Racial Bias and In-group Bias in Virtual Reality Courtrooms

Samantha Bielen Hasselt University

Wim Marneffe Hasselt University

Naci Mocan Louisiana State University, NBER and IZA

August 2020

We shot videos of criminal trials using 3D Virtual Reality (VR) technology, prosecuted by actual prosecutors and defended by actual defense attorneys in an actual courtroom. This is the first paper that utilizes VR technology in a non-computer animated setting, which enables us to replace white defendants with individuals who have Middle Eastern or North African descent in a real-life environment. We alter only the race of the defendants in these trials, holding all activity in the courtroom constant, creating arguably perfect counterfactuals (http://proficient.ninja/splitscreen/). Master's level Economics and Law students, and undergraduate economics students are randomly assigned to watch, with VR headsets, the trials that differed only in defendants' race. Background information obtained from these evaluators allowed us to identify their cultural heritage. Evaluators made decisions on guilt/innocence as well as prison sentence and fine in accordance with the guidelines provided by the relevant law. By design, the race of the defendant is uncorrelated with the characteristics of both the prosecutors and the defense attorneys, as well as with any activity in the courtroom. Defendant race is also uncorrelated with evaluator attributes. We find that both white and minority evaluators are harsher towards minority defendants during the conviction decision. In the sentencing phase defendants receive favorable treatment from evaluators of their own race. This pattern of behavior leads to significant bias against minorities at all stages: conviction, prison sentence, and fine, which is partly the reflection of the fact that the numerical majority of the evaluators are white. Evaluators' concerns about terrorism do not impact the racial biases in these decisions. The same racial bias is observed in the decisions of practicing attorneys. Adding a small number of prosecutors and judges to the sample of attorneys generates similar results as those obtained from the attorney sample.

We thank Christine Jolls, Ozkan Eren, Shawn Kantor, Carl Kitchens, Marc Isaac, Bulent Unel, Stefan Voigt, Barton Willage, Richard Boylan, Vivian Frances, and participants of seminars at the University of Iowa, Florida State University, the European Law and Economics Association Meetings in Milan, Italy, the Economic Consequences of Legal Development Workshop at UCLouvain, WEA Meetings in Tokyo, Japan, ASREC Europe Conference in Lund, Sweden and the Economic Analysis of Litigation Workshop in Granada, Spain for helpful comments. An anonymous referee and the editor Sam Peltzman provided constructive comments. Funding from the Faculty of Business Economics of Hasselt University and the Special Research Fund (BOF) of Hasselt University is gratefully acknowledged. We thank Johan Ackaert and Sofie Hennau for allowing us to engage their law students in the study, and to the Office of the District Attorney of Limburg, Belgium and in particular the District Attorney, Guido Vermeiren for sharing their expertise and for allowing us to use their main courtroom. We also thank prosecutors Bruno Coppin and Pieter Strauven, and attorneys Sophia Lycops, Tom Van Overbeke and Mieke Denissen for participating in the trials. We are grateful to Jeff Luyten and Michel Séquaris from Uncanny Virtual Experiences for making this project possible by providing technical support in creating the VR videos and to Sem Metten for creating the website that communicated with the VR videos and that enabled the participants to register their decisions. Bielen acknowledges financial support from Research Foundation Flanders, grant number 12S3117N.

Racial Bias and In-group Bias in Virtual Reality Courtrooms

I. Introduction

The investigation of racial bias in decision-making is important for both scientific inquiry and for public policy. A particularly important decision is judgment about others with consequential outcomes. For example, although judicial decisions are expected to be made blindly, whether such impartiality exists in practice has long been debated, and a great deal of research has focused on the investigation of biases against minorities in conviction, sentencing and clemency decisions made by judges, juries or state governors (Alesina and La Ferrara 2014, Rehavi and Starr 2014, Abrams et al. 2012, Argys and Mocan 2004, Glaeser and Sacerdote 2003). This research is important because these are high-stake decisions with economic consequences, and they are made with deliberation. A related, and more nuanced line of research has recently emerged to investigate the existence of in-group bias (decision-maker's preferential treatment of the members of his/her own group).¹

The analysis of judicial decisions is complicated by a number of inherent empirical problems related to omitted variables and selection. Although random assignment of defendants to judges alleviates some of the selection problems (Eren and Mocan 2018, Depew et al. 2017, Abrams et al. 2012, Shayo and Zussman 2011), random assignment does not resolve all selection issues. Even if defendants are randomly assigned to judges, there are a number of intervening steps taken by prosecutors and defense attorneys before defendants appear in front of judges, and this process may confound the inference obtained from the analysis of judicial decisions.² Similarly, the composition of cases that are adjudicated can also change because of the strategic behavior of defense attorneys.³ The impression of judge racial bias may also be a reflection of omitted variables. For example, if prosecutors are more diligent and aggressive towards a certain type of defendant in comparison to others, this differential effort in the

¹ Examples include Depew et al. (2017), Grossman et al. (2015), Anwar et al. (2012), Shayo and Zussman (2011), and Gazal-Ayal and Sulitzeanu-Kenan (2010).

 $^{^2}$ After cases are (randomly) assigned to judges, prosecutors enter the process and they have the discretion to determine the charges levied against the defendants. The attitude of the prosecutor regarding charging a particular defendant more or less aggressively has an impact on the eventual outcome of the case.

³ For example, after the case is assigned to a judge the defense attorney may be more or less open to plea bargaining based on the attributes of the judge, such as the race and the reputation of the judge.

courtroom can have a systematic effect on judge decisions. The effort and effectiveness of defense attorneys may also systematically differ between groups of defendants. To the extent that such effectiveness cannot be measured, it creates an omitted variable bias.⁴ Thus, racial differences in judicial decisions could be driven by the actions of prosecutors and attorneys, not by racial bias of judges.⁵ These pre-trial decisions, as well as unobservable (to the researcher) actions of various parties in the courtroom (ranging from the style of language used by the defense and the prosecutor to the body language) are nearly impossible to control for.

In this paper we eliminate biases caused by these confounders. Using Virtual Reality (VR) technology we record trials in a real courtroom, with actual prosecutors and actual defense attorneys. This is the first paper that uses 3D Virtual Reality technology with actual people in these videos, as opposed to computer-animated scenes. The details are provided in Sections II and III, and the Online Appendix. The VR technology allows us to *replace the defendants* in the 3D VR videos of the trials, holding constant every spoken word and every action in the court, enabling us to create arguably perfect counterfactuals to identify the effect of defendant's race. A glimpse of one of the trials with two different defendants can be seen at http://proficient.ninja/splitscreen/.

Two-hundred and one evaluators made decisions on conviction, prison sentence and fine in accordance with the guidelines provided by the relevant penal code. To guard against sample selection bias in the sentencing phase, we implement a trimming procedure to remove the marginal convicted defendants from the analysis of sentencing, who were likely discriminated against in the conviction phase. The evaluators are Master's level law, Master's level economics, or undergraduate economics students, practicing criminal attorneys, prosecutors or judges. By design, the race of the defendant is uncorrelated with the characteristics of both the prosecutor and the defense attorney and with all courtroom activity. Defendant race is also uncorrelated with evaluator attributes.

⁴ For instance, minority defendants may not be able to afford high-quality attorneys, and ineffective defense provided by low quality attorneys will translate into worse outcomes for minorities such as higher conviction probabilities and longer sentences (Anderson and Heaton 2012). Part of the negative outcomes associated with court-assigned attorneys to low-income defendants is attributable to attorney effort, motivated by the compensation structure of these attorneys (Agan et al. 2018).

⁵ Also, each particular case which is brought to trial has its own set of mitigating and aggravating circumstances, its own features about the background and the criminal history of the defendant, and the circumstances of the victim. Not being able to adjust for these factors constitutes another omitted variable problem.

We hold the prosecutor and the defense attorney as well as their actions in the courtroom constant. We do not reveal the names of the defendants to evaluators. The defendants speak only three times during the trial to answer the questions posed by the judge, as they would typically do in such trials. These answers are: "Yes," "Not guilty", and "No." All actor-defendants are born in Belgium and they speak Dutch fluently with no accent. This implies that "foreignness" or the minority status of the defendants can only be inferred from their darker skin complexion, but not from their names or accents. As explained in Online Appendix, a separate group of 89 students, who were not involved in the experiment, confirmed that the defendants are easily identified as being white or minority by the pictures of their faces.

We find that verdicts about the defendants' guilt vs. innocence and the decisions about sentence severity are not race-blind. There are statistically and economically significant overall racial bias against minorities at all stages: conviction, prison sentence, and fine. We also uncover behavioral differences between white and minority evaluators. White evaluators reveal positive in-group bias during all stages of decision-making; that is, they treat white defendants favorably in comparison to minority defendants during the conviction decision, and also when assigning a prison sentence and fine. A more nuanced picture emerges in case of minority evaluators, who treat defendants of their own race more *harshly* during the conviction decision, but more *leniently* during sentencing. Thus, minority defendants are treated more harshly by both white and minority evaluators during conviction. In contrast, defendants receive favorable treatment from evaluators of their own race during sentencing. When we investigate whether disparate treatment of defendants is impacted by evaluators' concerns about terrorism we find no evidence that such concerns significantly impact racial biases in these decisions.

The analysis of attorney decisions reveals the same pattern of racial bias we find in the sample of economics students and law students. We also have a sample of judges and prosecutors⁷ who made conviction and sentencing decisions on these cases. Although the decisions of judges and prosecutors cannot be analyzed separately because of their small sample size, when we add judges and prosecutors to the sample of attorneys we obtain similar results as those obtained from the attorney sample. The student sample and the attorney sample have different strengths. The student sample is larger and it allows us to investigate the in-group bias. The attorney sample is smaller, and in-group bias cannot be analyzed in this sample because all attorneys are white. On the other hand, attorney decisions have more external validity.

⁷ The group of prosecutors does not include those who have participated in the shooting of the VR trials.

II. General Idea and the Contribution of the Paper

Imagine we are watching a trial. The district attorney presents his case the way he sees appropriate. The defense attorney defends her client with the knowledge and the skill she possesses. The facts of the case are presented and debated by both parties. The defendant, who sits in front of the judge, listens to the arguments and counter-arguments made by the prosecutor and by his defense attorney. The defendant is a white man. At the end of the trial the judge renders a verdict about conviction/acquittal; and if she finds the defendant guilty, she assigns a prison sentence and/or a fine.

Now imagine we go back in time and watch *the same exact trial*. The same prosecutor, the same defense attorney, the very same arguments, the same words, the same body language. Everything that took place in the courtroom is precisely the same, with one difference: the defendant is someone else. He is not a white man, but he has dark skin. Would the decision of the judge about the verdict, the prison sentence length and the fine be different? Are these judicial decisions influenced by whether or not the judge and the defendant have similar racial backgrounds? If so, are these influences stronger for white judges or minority judges? If there are race-driven effects on these decisions, are they impacted by whether the evaluator believes that terrorism is a major problem in the country? These are the questions we try to answer in this paper.

Of course, if the judge were to watch the same trial the second time (the only difference between the two versions being the defendant's race), she would immediately recognize that this was the same case she adjudicated before. Therefore, while the first evaluator watches the trial with the white defendant, we ask another evaluator to watch the same trial with the minority defendant. Each of the six criminal trials analyzed has a different defendant who faces a different burglary or assault charge with different circumstances. Each of these six trials has two versions. The only difference between the two versions is that while the defendant in Version 1 of each trial is a young white man, the defendant in Version 2 of the same trial is a young man with darker skin color. We randomly assign these trials to 201 evaluators, making sure that everyone watches all six trials in random order, and that everyone watches three trials with white defendants and three other trials with minority defendants. Half of the evaluators watch Version 1, while the other half watch Version 2 of each VR video.

The VR technology allows us to video-record the courtroom activity in 3D. It also allows us to replace the defendant in the VR video with another individual who is video-recorded separately. In the end, this production generates two identical VR videos with one difference: Version 1 contains the white defendant, and Version 2 contains the minority defendant.

Figure 1 displays a side-by-side image of Version 1 and Version 2 of a particular scene from one of the trials. The version in the top frame involves a white defendant and the version in the bottom frame involves a minority defendant. In each frame, the person on the right is the prosecutor who is presenting his case. The defendant is sitting in the middle; and the person on the left, behind the defendant, is the defense attorney. This picture depicts how the judge (the evaluator) observes the trial, although watching the Virtual Reality videos using 3D headsets produces a much more realistic image of the actions and sounds of the courtroom.

Two short segments of this trial (shown in Figure 1) titled "A Snapshot of the Virtual reality Videos of the Same Trial" are at the link: <u>http://proficient.ninja/splitscreen/.</u> This particular clip presents a 40-second segment of the fourth trial shown (Video No 4), but it displays two versions of the same trial simultaneously on the split screen. The top segment of the screen displays the trial with the white defendant and the bottom segment displays the same trial with the minority defendant. Every detail in both videos is identical with the exception of the defendants. The second video at this link presents the full version of the same trial with the minority defendant (Video No 4, Version 1), and the clip below it (Video No 4, Version 2) displays the same trial with the white defendant.

Another short video-clip at the link <u>http://proficient.ninja/uhasselt/</u> provides a glimpse of a different trial (note that the prosecutor and the defense attorney as well as the defendant are different). This video is titled "Video No 1, Version 1" to indicate the first trial (Video No 1) with the white defendant (Version 1). Two short segments are included in the clip. The first segment shows a section from prosecutor's opening remarks; and the second part shows a segment in which the defense attorney speaks. Pictures from other trials can be found at this link by scrolling down.

The prosecutors and the defense attorneys in all videos are actual prosecutors and actual defense attorneys who practice criminal law in the province of Limburg, Belgium. We used a real courtroom in Hasselt, Limburg. The videos are shot from the bench of the judge so that the evaluators could observe the trial from the vantage point of the judge. At the conclusion of each trial, each evaluator rendered a verdict on guilt vs. innocence, and assigned a sentence if s/he found the defendant guilty. We compare the convictions and the prison sentence lengths and

fines between the evaluators who watched the same exact trial but faced a white vs. minority defendant. This design enables us to identify the impact of the race of the defendant on these decisions, holding all else constant. Furthermore, because we have information on evaluators' race, we investigate whether evaluators are more lenient toward defendants who are of the same race.⁹

Previous experimental research, that investigated the impact of defendant characteristics (e.g. race, mental illness) on trial verdicts, provided mock trial transcripts to be read by the evaluators. In these studies the race of the victim or defendant is manipulated using photographs and names (Mossiere et al. 2018; Maeder et al. 2012), or evaluators were given vignettes that described the offender and the circumstances of the case (Mercado et al. 2006). Some studies showed videos of mock trials where the trial conditions, such as eyewitness evidence, has been manipulated by the experimenters (Jones et al. 2017).

It is also informative to compare our design to that of audit studies, where pairs of individuals (auditors) with similar/identical characteristics with the exception of one attribute (e.g. race, gender, criminal record) are sent to apply for a job (Pager et al. 2009, Pager 2003, Neumark et al. 1996), to buy a car (Ayres and Siegelman 1995), a housing unit (Ondrich et al. 2000), and so on. These matched auditors are trained to respond to and interact with the decision-maker in the same manner. The difference in outcomes between the matched pairs is attributed to discriminatory behavior of the decision-makers, related to the attribute that differs between the pairs. A primary weakness of audit studies is the difficulty to match the auditors so that they are identical in all characteristics that might matter for the decision-maker.¹⁰ Importantly, auditors are typically aware of the purpose of the study, and as highlighted by Riach and Rich (2002), Heckman (1998) and Turner et al. (1991), they may consciously or unconsciously behave and interact with the decision-maker in a manner that may generate evidence for discrimination. These concerns are alleviated in our study because unlike typical audit studies in which the actors interact and have conversations with the decision-makers in order to buy or rent something or to interview for a job, our actor-defendants don't interact with

⁹ This component could not be implemented in the sample of attorneys because all attorneys were white. The same was true for judges and prosecutors who evaluated these cases.

¹⁰ As described by Heckman (1998) and Heckman and Siegelman (1993), auditors may also differ in dimensions that are unobservable to the designer of the audit study, but that are at least somewhat visible to the decision-maker and that are relevant for the decision. These unobservables as well as their distribution between groups can produce differential outcomes may cause differences in outcomes.

the evaluators. This is organic and natural in our setting, because as explained earlier, the defendant's representation is provided by his attorney who does all the talking on behalf of the defendant in the courtroom. Our actor-defendants speak only three times during their trial to say "Yes," "Not guilty", and "No." Also, unlike audit studies, everything that should matter for the outcome is identical between white and minority defendants, including the details of the case and the behaviors of the prosecutor and the defense attorney in the courtroom.¹¹

There exists research that aims to differentiate between theories and mechanisms of racial bias. For example, List (2004) conducted an audit study with buyers and sellers of collectable baseball cards and then combined it with dictator games conducted in the lab with these card dealers. Levitt (2004) analyzed the behavior of the contestants of the game show *Weakest Link*, and Bertrand and Mullainathan (2004) analyzed in a correspondence study the extent to which call-backs of job applicants are responsive to the race information contained in their names and the amount of information provided about the qualifications and credentials of the applicants. In a recent empirical study Arnold, Dobbie and Hull (2020) analyze judges' bail decisions in New York City. Their data set contains extraordinary detail on defendants, including criminal history and past pre-trial misconduct (failing to appear in court, or being arrested for a new crime after having been released on bail) and the conduct of the defendant after the bail has been granted by the judge. Levering the detail of the data and imposing a structure on the manner in which judicial decisions are made, Arnold, Dobbie and Hull (2020) shed light on the theoretical channels driving the racial discrepancy in bail decisions.

Our paper does not attempt to tease out the extent to which the bias is driven by taste-based or statistical discrimination because our research design does not permit a meaningful analysis along these dimensions. Nevertheless, in an effort to provide more insight into the source of the identified bias, we expand the analysis. For example, it could be argued that racial biases identified in the paper might be driven by certain perceptions about individual defendants that are visible to evaluators, such as defendants' perceived trustworthiness or their perceived proclivity for delinquency. We show, however, that accounting for perceived criminal

¹¹ A related stand of literature is based on correspondence studies, where pairs of fictitious resumes, letters of interests, or applications are sent to prospective employers (or other decision-makers) where the applications between the pairs differ in one attribute, such as the name (signaling the race), age, or sex of the applicants. Differences in the call-back rates are attributed to the difference in the attributes. A detailed discussion of the correspondence studies is provided by Bertrand and Duflo (2017).

proclivity or perceived honesty of individual defendants (assessed by a different group of students who are not involved in the experiment) does not alter the results.

Nine to 15 days after the experiment the evaluators completed an online survey of 40 questions on a variety of subjects, including whether they consider terrorism as an important issue. The results are similar between those who think that it is very important that the Belgian society pays attention to terrorism, and those who were concerned less.

III. The Design of the Study and the Data

We are granted permission by the district attorney of Limburg, Belgium to use actual criminal cases as well as actual prosecutors and an actual courtroom. Upon discussions with prosecutors we decided to focus on burglary and assault cases because such cases provide substantial discretion to the judge about sentencing, and no special background or additional expert information is needed to deliver a verdict. Three actual burglary and three actual assault cases are selected from the court archives or from the docket.

Two prosecutors who specialize in burglary and assault cases volunteered to participate. We contacted a number of law firms specializing in criminal cases to be part of the experiment. After interviews, we selected three lawyers (two females and one male) who have experience in both burglary and assault cases. The lawyers and the prosecutors were given the case files one week before the shooting of the videos and were asked to prepare these cases similar to an actual court trial. The lawyers did not talk to the prosecutors before the shooting to keep the trial environment as realistic as possible. Both the prosecutors and lawyers tried and defended the cases the way they would normally do in real-life trials. There was no script to follow. They presented their case with no instructions from anyone.

In Belgium the judge is in control of the court hearing although his/her actual role is limited. The judge talks very little during the trial; he/she speaks just to direct the proceedings. The judge first asks the defendant if he understands the charges against him and whether he pleads guilty or not guilty. In each of our videos, the defendants indicate that they understand the charges and that they plead not guilty. The judge then allows the prosecutor to start with his statement. Once the prosecutor is finished, the defence attorney launches his/her pleading.¹²

¹² In an actual hearing, the judge sets the time frame for her verdict (normally four weeks after the conclusion of the hearing) and she could potentially ask the defendant some questions during the trial. Our set up (the evaluators watching the VR videos with headsets) does not allow the evaluators to

For the sake of realism in the videos, we needed an actual judge to control the flow of proceedings. We selected a retired judge from the region of Hasselt to be in control of the cases, as he would normally do. The VR camera was located before the judge's bench. Thus, the evaluators who watched the videos could only hear his voice in the VR video, but could not see him. This way the setting remained realistic and the evaluators viewed the cases from the perspective of the judge. We used one of the courtrooms in the main court house in Hasselt.

We employed three white defendants and three non-white defendants for our six trials.¹³ We limited the number of trials to six because on average a trial took 12 minutes, which implied that watching the six VR videos and making decisions on these six trials lasted for two hours. Asking the evaluators to watch trials for more than two hours was infeasible for a number of reasons. There was a five-minute break after watching the first three videos and making judgments on these cases. Interactions with other participants were not allowed. The evaluators were not primed in any way. The pictures of the actors are provided in Figure 2. Snapshots of them during the trials can be found by scrolling down at the link http://proficient.ninja/uhasselt/.

One hundred and fifty three students from the Faculty of Business Economics, and Faculty of Law of Hasselt University participated as evaluators. The Economics students were a mixture of juniors and master's students that were enrolled in the Policy Evaluation course, consisting of 86 students. We also recruited practicing attorneys from the Limburg Bar Association, and 36 attorneys decided to participate. Additional information is provided in the Appendix.

The students participated in the experiment on November 22-23, 2017. The attorneys participated during a 10-day period in July 2018. There were no unusual news or events in Belgium during this period related to immigration, terrorism or asylum seekers. It should be remembered, however, that the terrorist attacks in Brussels that targeted the airport and the subway system, took place on March 22, 2016, which is 20 (28) months before the students (attorneys) evaluated the cases in this experiment. In addition, the terrorist attacks in Paris, where most of the perpetrators were residents of Brussels, took place on November 13, 2015.

interrogate the defendant. It is, however, not uncommon in Belgian criminal proceedings for the judge not to ask any questions to the defendant.

¹³ We selected male actors who lived outside of the region where the evaluators reside in order to avoid the possibility of an evaluator recognizing one of the actors. We also made sure that all actors were about the same age, and we gave them instructions on wardrobe so that the clothing worn by the pairs of actors was similar.

Thus, it is not unreasonable to assume that these major events were still remembered by the evaluators.

The evaluators were granted anonymity during the experiment and in the follow-up survey, which they completed online 9 to 15 days after the experiment (see the Online Appendix for details). After reading the case folder, evaluators could click on a button on the screen to start the VR video. They had 30 seconds to put on the headsets to watch the video. After watching the video, they took off the headsets, and clicked on a button on the screen to go to the deliberation phase. First, they were asked to decide on conviction or acquittal. If they decided to convict, they had to assign the sentence (prison sentence and/or fine). As is the case with real judicial decision, they also had to make a decision on whether and how much to suspend the prison sentence and the fine. All of these decisions were made by clicking and typing on the screen. Details are provided in the Online Appendix,

After watching and rendering decisions on all six cases, participants were presented with an overview of their decisions for all the six cases and they could alter any decisions they made (conviction, prison term and fine) on any of the six cases before they were finished with the experiment. In the analyses we use the final decisions made on each case.¹⁴ Nine days after the experiment, the evaluators were asked to complete an online survey, which is used to elicit information about their cultural background and attitudes towards social and economic issues.

If the defendants in the VR videos are not recognizable as being a member of a minority or majority group in Belgium, then our design will provide null results.¹⁵ Put differently, if the evaluators could not infer correctly the racial background of our defendant-actors by looking at them, then the evaluators would not assign differential verdicts or sentences even if the evaluators had racial biases.¹⁶ The pictures of the defendants are provided in Figure 2. To analyze whether the defendants can be identified as white or minority, we used a different group of 89 students who were shown pictures of the defendants. As detailed in the Online Appendix, the students identified the defendants' race correctly with 99 percent accuracy.

¹⁴ To investigate if the evaluators took their task seriously, we conducted a number of tests such as the analysis of the decisions by the speed of the decisions made by the evaluators, by the order in which the decisions were made, and so on. These are detailed in the Extensions and Robustness section of the Appendix.

¹⁵ Online Appendix, Section 4 provides a discussion on race vs. ethnicity.

¹⁶ To put bluntly, the question is whether the evaluators can tell the difference between a Belgian person of European heritage and another individual who is a racial/ethnic minority with dark skin (e.g. a Belgian with Moroccan or Turkish origin)? Or, would the evaluators think that the latter person is a white European, but happens to have sun tan?

Because we have six different cases (three burglaries and three assaults) and two versions of each case (one with a white defendant and one with a non-white defendant) we have a total of 12 videos in the experiment, and each defendant-actor participated in two videos. We made six sets with different sequences of the videos and three different combinations to make sure that each evaluator would see each defendant only once. The sets are listed in Appendix Table A1. The details are discussed in the footnote to Appendix Table A1.

IV. Descriptive Statistics

Table 1 presents the descriptive statistics related to student evaluator attributes. Information on evaluators' sex and whether they are law students or economics students were obtained during the experiment. Using an online survey that was registered nine days after the experiment, we measured evaluators' beliefs on a number of social issues, and elicited additional personal background information from them. *Minority Evaluator* takes the value of one if either the mother or the father of the evaluator was born outside of Belgium, and if that country is outside of the European Union (EU).¹⁷ Thus, minority evaluators are first- or secondgeneration immigrants whose ancestral origins are in Turkey (10 evaluators), Morocco (4 evaluators), and one evaluator each from Iraq, Zimbabwe, and Armenia. Alternatively, we determined if the evaluator was a minority in Belgium based on the answers to the following question. "Do you or anyone in your household speak a language other than Dutch?"¹⁸ In this alternative definition, we coded the evaluator as a minority if he/she indicated that either Arabic, Armenian, Turkish or Shona is spoken at home.¹⁹ The dummy variable *Terrorism is a very* Important Problem in Belgium takes the value of 1 if the individual assigned a value of 7 to the question "On a score of 1-7, to what extent do you think it is important that Belgian society pays attention to terrorism?"

Table 2 presents the descriptive statistics of the decisions made by evaluators. Row (1) displays information about the conviction decision by the race of the defendant [Columns (1)

¹⁷ If the parent was born in a country which is part of the EU, such as Germany or France, the evaluator is not considered as a minority in Belgium. There was one evaluator with parent(s) from Japan, one evaluator with parent(s) from South Korea, and another one from India. These three individuals of Asian heritage are excluded from the analysis, but considering them as minorities did not alter the results.

¹⁸ Hasselt University is located in Flanders, the Dutch-speaking part of Belgium.

¹⁹ We did not consider an evaluator as a minority in Belgium when for example French, Polish or Italian is spoken at home.

and (2)], as well as by race-matching between the defendant and the evaluator [Columns (3) to (6)]. The stars in the cells indicate statistically significant differences in the mean values in a particular row between columns (1) and (2), between columns (3) and (4), and columns (5) and (6). For example, columns (1) and (2) reveal that the conviction rate among minority defendants is 12 percentage points higher in comparison to the conviction rate of white defendants (0.79 vs. 0.67) and that the difference is statistically different from zero. Columns (3) and (4) reveal that when minority defendants are matched with white evaluators, the conviction rate is higher in comparison to situations where white defendants face white evaluators (0.78 in column (3) vs. 0.68 in column (4)). Interestingly, the same relatively unfavorable outcome for minority defendants is observed even when they are matched with minority evaluators. Minority evaluators convict minority defendants at a higher rate in comparison to white defendants (0.82 in column (5) vs. 0.65 in column (6)).

Row (2) of Table 2 displays the average prison sentence imposed on defendants if they were found guilty. Part of the prison sentence can be suspended at the discretion of the evaluator. Row (3) presents the average suspended prison sentence among various groups. The difference between the initial prison sentence and the suspended sentence is the actual, effective prison term faced by the convicted defendants, displayed in row (4). For example, column (1) shows that convicted minority defendants receive an average of 10.15 months prison time (row 2), and that 5.81 months of this initial sentence is suspended (row 3). Thus, row (4) and column (1) show that the effective prison term among minority defendants is 4.34 months, on average.

Columns (5) and (6) and row (4) show that minority evaluators assign prison sentences to minority defendants that are about 2.7 months shorter than they assign to white defendants. This indicates that minority evaluators treat minority defendants more favorably in assigning a prison sentence, although minority evaluators are harsher toward defendants of their own race during the guilt/innocence decision. Columns (3) and (4) of row (4) reveal that white evaluators also exhibit positive in-group bias and assign white defendants prison terms that are about one month shorter (3.61 months vs. 4.54 months) in comparison to sentences they assign to minority defendants. Thus, row (4) indicates that evaluators provide preferential treatment in assigning prison sentences to defendants who are of the same race, which signifies positive in-group bias in effective prison term.

Rows (5) to (7) of Table 2 display information on fines. Columns (3) to (6) of row (7) reveal that there are racial differences in fines by race-matching between the defendant and the

evaluator, although the differences are not statistically significant. The details of this picture will be revealed in the empirical analysis section.

Columns (1) and (2) of row (4) show that minority defendants receive longer sentences in comparison to their white counterparts. Columns (1) and (2) of row (7) reveal the same picture about fines, although the differences are not different from zero. This statistical insignificance is misleading because racial bias in the conviction decisions generates a selected sample of defendants who are sentenced. Note in Table 2 that while 361 minority defendants are convicted and sentenced, there are only 309 white defendants who are sentenced. These 52 "excess" minority defendants end up in the sentencing phase because of the 12 percentage point difference in the conviction rates between the races (0.79 vs. 0.67).²¹ Thus, we present the empirical results both ways: i) that accounts for this selection bias and ii) that ignores the selection bias.

V. Econometric Analysis

Overall Racial Bias

To investigate the existence of overall racial bias in conviction and sentencing decisions we run regressions of the following type:

(1) $Y_{cj} = \alpha_1 + \beta_1$ Minority Defendant_{cj} + $X_j \Phi + v_c + \varepsilon_{cj}$

where Y_{cj} stands for the outcomes related to case (trial) *c*, evaluated by individual *j*. The first outcome is an indicator to represent whether the defendant is convicted by evaluator (*j*). The second and third outcomes are the prison term and the amount of fine imposed on the defendant if he is found guilty. *Minority Defendant* equals one if the defendant in case *c* (in the VR video of trial *c*) is a racial minority. Recall that a particular case is identical across all evaluators who watch the VR video of that case, with the exception of the race of the defendant. Six courtroom hearings (c=1, 2,...6) are watched and judged by each evaluator (j). Three of these trials involve burglary, and the other three are assault cases. In some specifications we include trial fixed effects (v_c), in others we control for the type of crime on which the trial was based (assault vs. burglary). Vector X includes attributes of the evaluators watch the same

²¹ The same point is made by Depew et al. (2017). Because the sequential decision of conviction and sentencing of the defendants by the same judges generates sample selection in the sentencing phase, the authors analyzed the sentencing outcomes of those defendants who pled guilty (as opposed to those who are convicted and then sentenced by the same judge).

trials, the order in which these trials are watched is different by design.²³ The estimated coefficient β_1 identifies whether minority defendants are treated differently in comparison to their white counterparts in conviction and in sentencing decisions.

In-group bias

To test for the existence of in-group bias we make use of information obtained from the evaluators regarding their cultural heritage. This information allows us to estimate

(2) $Y_{cj} = \alpha_2 + \beta_2 \text{Minority Defendant}_c + \gamma_2 \text{Minority Evaluator}_j$

+ δ_2 Minority Defendant_c*Minority Evaluator_j + $X_j\Theta$ + μ_c + ω_{cj}

In Equation (2) *Minority Evaluator* is a dichotomous indicators which equals 1 if the evaluator, who makes the conviction and decisions, is a minority him/herself in Belgium. The coefficient γ_2 captures the difference in outcomes between white defendant-minority evaluator pairs and white defendant-white evaluator pairs, and $\gamma_2+\delta_2$ captures the differential impact on sanctions assigned on minority defendants by minority versus white evaluators. Thus, δ_2 is the difference-in-difference estimate, which signifies the differential decisions of minority evaluators versus white evaluators in their treatment of minority defendants over white defendants. Put differently, if δ_2 is different from zero, this reflects in-group bias.

While specification (2) identifies in-group bias, it does not provide insights into the source of the bias. One of the contributions of this paper is its ability to unbundle the in-group bias. Because the race of the evaluators is, by design, uncorrelated with all courtroom attributes and with all case characteristics, and because it is also uncorrelated with the race of the defendant, we can run the regressions below, conditional on evaluator race, to investigate the source of the in-group bias.

(3A) $Y_{cj} = \alpha_3 + \gamma_3$ White Defendant $+X_j\Theta + C_c\Lambda + \pi_c + u_{cj}$ in the sample of White Evaluators (3B) $Y_{cj} = \alpha_4 + \gamma_4$ Minority Defendant $+X_j\Pi + C_c \boldsymbol{\Xi} + \theta_c + \tau_{cj}$ in the sample of Minority Evaluators

²³ As Appendix Table A1 demonstrates, evaluators who were randomly assigned to Set 1 or Set 2 watched these six cases in a particular sequence (Sequence 1), while those who were randomly assigned to Sets 3 or 4 watched the same videos in a different order (Sequence 2); and Sequence 3 is the third sequence in which the videos are watched by evaluators. Vector X contains Sequence1 and Sequence2 which are two dummy variables that control for the sequence in which the videos are watched.

The coefficient γ_3 reveals whether white defendants are treated differently in comparison to minority defendants by white evaluators. The coefficient γ_4 provides the same information for minority evaluators.²⁴ Note that in Equation (2), which analyzes in-group bias, $(\alpha_2 + \beta_2)$ represents the impact of white judges on minority defendants, and α_2 stands for the impact of white judges on white defendants. Thus, β_2 represents the differential treatment of minority defendants vis-à-vis white defendants by white judges. Similarly, $(\beta_2 + \delta_2)$ stands for the differential treatment of minority defendants by minority judges. If *p* stands for the proportion of white judges in the sample, then the overall differential treatment of minority defendants by all judges is equal to $[\beta_2+(1-p)*\delta_2]$, which should be equivalent to the estimated coefficient of *Minority Defendant* in Equation (1).

VI. Results

Conviction Decisions

The results, reported in Table 3, are based on 153 law students or economics students, who made decisions on conviction vs. acquittal on each of the six cases they watched, generating a sample of 918 decisions. The analyses of attorney decisions are reported later in the paper. The first three columns of Table 3 present the results of estimating Equation (1) which investigates the overall racial bias in conviction decisions; and columns (4)-(6) display the results of the investigation of in-group bias, based on Equation (2). In columns (1) and (4) we control for whether the case was an assault or burglary as well as whether the decision on the case was made during the first half of the experiment or in the second half (Early Trial). Columns (2) and (5) present the results from another specification which includes trial fixed-effects, the sex of the evaluator and whether the decision was made during the first half of the experiment. Columns (3) and (6) control for the sequence type (the order in which videos are watched), in addition to trial fixed effects and evaluator's sex.

Columns (1) to (3) show that minority defendants are about 12 percentage points more likely to get convicted in comparison to white defendants. This indicates that two defendants, whose cases are identical in every respect except for their race, receive different resolutions.

²⁴ Note that γ_3 in Equation (3A) represents the differential treatment of white defendants by white evaluators, and is equivalent to (- β_2) in Equation (2). Similarly, γ_4 in Equation (3B) stands for the differential treatment of minority defendants vis-a-vis white defendants by minority evaluators, which is represented by $\beta_2+\delta_2$ in Equation (2). Thus, the in-group bias coefficient δ_2 of Equation (2) can be constructed by the sum of γ_4 and γ_3 .

Specifically, the defendants face the same exact criminal charge with the same exact mitigating and aggravating circumstances. They are charged by the same prosecutor and are defended by the same attorney in the same exact way (same spoken language and same body language in the courtroom). Nevertheless, minority defendants are 12 percentage points (about 16 percent) more likely to get convicted in comparison to the white defendant.

Columns (4) to (6) of Table 3 present the estimation results of Equation (2). The coefficient of the interaction term (*Minority Defendant x Minority Evaluator*) represents the impact of in-group bias in conviction decisions. The point estimate is about 0.085, although not statistically significant at conventional levels. This provides suggestive evidence for negative in-group bias: defendants are 8.5 percentage points (about 12 percent from the mean conviction rate) more likely_to get convicted if they face an evaluator who is of the same race. Estimating the model by including evaluator fixed effects provided very similar results.

Table 4 sheds light into the source of the statistically insignificant in-group bias reported in Table 3. The two columns of the table present the estimation results of Equations (3A) and (3B). Column (1) shows that white evaluators are 11 percentage points (15 percent) less likely to convict white defendants, indicating positive in-group bias. On the other hand, column (2) displays *negative* in-group bias in minority evaluators' conviction decisions. Minority defendants are 20 percentage points (27 percent) more likely to get convicted by minority evaluators. Thus, white evaluators exhibit positive in-group bias, while minority evaluators display negative in-group bias, and the latter effect is stronger in magnitude. This indicates that the statistically insignificant in-group bias reported in columns (4)-(6) of Table 3 is the result of opposing in-group biases between racial groups.

Preferential treatment of in-group members is intuitive and unsurprising, and there is evidence of such behavior (Chen and Li 2009, Shayo and Zussman 2011). Negative in-group bias, however, has also been reported recently in naturally-occurring data by Depew et al. (2017), who found that Louisiana juvenile court judges assign longer sentences to defendants who are of the same race as them. This could be because judges are elected in Louisiana and therefore they may be trying to avoid creating the impression to the voters of favoring their own race, and thus they may be harsher as a result towards these defendants. This is, of course, not the case in our setting because our evaluators do not have such a concern. Another explanation could be that minority evaluators might consider a defendant of that minority group as a representative of the group. In that case, the evaluator might be harsher towards the defendant can

translate into convicting the defendant even if the case against him in the court hearing may not have been very strong. Negative in-group bias that can be attributed to this effect has been detected in lab experiments where in-group members have violated a social norm (Mendoza et al. 2014, Goette et al. 2006). Similarly, List and Price (2009) reported that minority donors are less likely to make a contribution for a charity during a door-to-door fundraising drive if the solicitor is also a minority, and that the size of the gift is lower, conditional on the decision to give, suggesting lack of trust.

Sentencing Decisions: Prison Term and Fine

Those defendants who are found guilty are sentenced to a prison term and/or fine. Because there is overall racial bias in the conviction decision (columns 1 to 3 of Table 3) this bias contaminates the sample of defendants who are sentenced. More specifically, the existence of racial bias in convictions implies that some minority defendants, who should not have been convicted, are nevertheless found guilty because of their race. If these individuals represent borderline cases in the conviction-acquittal decision, they may receive lenient punishment in the sentencing phase. In this case, OLS estimates of racial bias in sentence length and fine would be biased downwards. To account for such selection, we follow Hoffman and Oreopoulos (2009) and Angrist et al. (2006) and trim the sample that contains those who are found guilty. This procedure assumes that evaluators use a higher standard to determine guilt in case of white defendants, which implies that the marginal convicted minority defendant is "less guilty" than the marginal convicted white defendant. Therefore, we trim out "excess" minority defendants (with lowest levels of punishment) to obtain equal-size groups by race at the sentencing phase. For comparison purposes we also report the results based on the entire sample of convicted defendants, without trimming the sample.

As displayed in Table 2, white evaluators made 408 decisions on minority defendants and they made another 408 decisions on white defendants (row 1, columns 3 and 4). These white evaluators convicted minority defendants 78 percent of the time, while they convicted white defendants with 68 percent probability. This generated 319 convicted and sentenced minority defendants, but only 276 white defendants who are found guilty and then sentenced (see rows 2-7 of columns 3 and 4). Put differently, the 10 percentage point difference in the conviction rates between white and minority defendants (0.78 vs. 0.68), which is due to racial bias of white evaluators, generated 43 excess minority defendants to be sentenced by white evaluators (319-276=43). As shown in row (1) and columns (5) and (6) of Table 2, minority evaluators too are biased against minority defendants in their conviction decisions (conviction rates of 0.82 vs.

0.65). This has created 9 excess minority defendants (42-33=9) sentenced by minority judges. The trimming procedure allows us to determine these marginal defendants and drop them from the sample. This is achieved by analyzing the sentence distribution of all 319 minority defendants who are convicted by white evaluators and by dropping 43 of these who received the lowest sentences from these white evaluators. Among the 319 minority defendants who are found guilty by white evaluators, there are exactly 43 defendants for whom both the effective prison sentence is zero and the effective fine is zero. That is, white evaluators assigned prison terms and fines, and then suspended all of the prison term and the entire fine in these 43 cases. We applied the same procedure to drop the 9 minority defendants (to bring down the number of sentenced minority defendants to the number of sentenced white defendants: from 42 to 33) who were convicted and then sentenced by minority evaluators: those who are at the far left tail of the sentence distribution.²⁶

Table 5 presents the estimation results of both the overall bias (columns 1-3) and in-group bias (columns 4 -6) where the outcome is the effective prison sentence. Recall that effective prison sentence is the actual prison sentence imposed on the defendant, which is the difference between the initial sentence handed down by the evaluator and the suspended sentence (see Table 2). Panel A of Table 5 presents the results based on the trimmed sample, and Panel B displays the results obtained from the whole sample (the entire group of convicted defendants). Columns 1 to 3 display the estimation results of the overall racial bias in the assigned effective prison term. As expected, the coefficient of *Minority Defendant* is larger when the regressions use the trimmed sample in Panel A of Table 5. The results indicate that minority defendants receive about 0.7 months longer prison terms when the racial selection bias in conviction is not adjusted for (panel B of Table 5), but that minority defendants receive 1.3 months longer sentences (about 32% from the mean) if selection is accounted for.²⁷ Models with evaluator fixed effects provided the same inference. Columns (4)-(6) of Table 5 reveal the existence of

²⁶ Six minority defendants who are convicted by minority evaluators had zero effective prison time and zero effective fine assigned to them. These are the minority defendants with no effective sentences, and they are dropped. Among the remaining convicted minority defendants, we dropped those who had zero effective prison time coupled with lowest effective fines (which were 100 and 300 Euros).

²⁷ As discussed earlier, the racial bias coefficient in Table 6 (the coefficient of *Minority Defendant* in columns 1-3) is also recoverable from the two coefficients obtained from the in-group bias regressions of Table 6. For example, using column 6 of the trimmed sample of Table 6, $1.684 - (3.766) \times 0.11$ is equal to 1.269 (where 0.11 is the proportion of minority evaluators in the sample). The coefficient of *Minority Defendant* in column (3) is 1.285.

positive in-group bias in prison sentencing. The estimated coefficient of the interaction term *Minority Defendant * Minority Evaluator* is negative and significantly different from zero in every specification. The magnitude of the coefficients indicates that if the defendants are sentenced by an evaluator of their own race, they receive prison sentences that are almost 4 months shorter. This is a big impact as the average sentence length is 4 months.

Table 6 displays the results of the models where the dependent variable is the logarithm of the fine.²⁸ Columns (1) to (3) display the regression results investigating the existence of overall racial bias in fines. The results in Panel B, that are based on the entire sample, indicate that minority defendants receive fines that are 53 percent higher than those assigned to white defendants. Panel A, which reports the results based on the trimmed sample, reveal that correcting sample selection (stemming from the bias in conviction decision), more than doubles the estimated coefficient of the Minority Defendant dummy. Columns (4) to (6) of Table 6 reveal that the coefficient of the interaction term (*Minority Defendant x Minority Evaluator*) is negative in all specifications and that the point estimate is around -0.6, indicating that defendants who have the same race as the evaluator receive fines that are 54 percent lower, although the effect is not statistically significant.^{29 30}

Table 7 presents the results that unbundle the in-group bias effect in prison sentence and fine. Column (1) shows that convicted white defendants receive prison sentences that are about 1.7 months shorter in comparison to their minority counterparts if they are judged by a white person. Column (2) reveals that minority defendants receive sentences that are 1.7 months shorter if the evaluator is also a minority (the effect is significant at the 11 percent level). Columns (3) and (4) display the results based on untrimmed sample, which reveal the same inference. Thus, evaluators of both races treat defendants of their own race equally favorably in the assignment of prison sentencing. Put differently, convicted defendants receive shorter sentences if they are matched with an evaluator of their own race and this bias is driven equally by both white and minority evaluators.

²⁸ Because some effective fines are 0, we added one Euro to effective fines assigned by the evaluators.

²⁹ The impact is calculated as $exp\{\beta-0.5Variance(\beta)\}-1$, where β is the estimated coefficient, and Var(β) is its variance (Kennedy 1981).

³⁰ We also used an alternative method of correcting sample selection. We assigned zero punishment (zero months of prison term and zero fine) to all defendants who are found not guilty. We then used all defendants, regardless of whether or not they were convicted, in the sentencing regressions. The estimated impacts were somewhat smaller, but statistically significant.

Column (5) of Table 7 shows that white evaluators fine white defendants more leniently in comparison to minority defendants. Specifically, white evaluators assign fines to convicted white defendants that are 68 percent lower in comparison to those they assign to minority defendants (the coefficient of Minority Defendant is -1.128). As shown in column (6), minority evaluators too favor defendants of their won race. Minority evaluators assign minority defendants 39 percent lower fines, although this effect is statistically not different from zero. Columns (7) and (8), which present the results obtained from the untrimmed sample, display a similar picture. Thus, columns (4) to (6) of Table 6 and columns (5) to (8) of Table 7 reveal positive in-group bias in the assignment of fines, indicating that convicted defendants face lower fines when the person who fines them is of the same race, and that this effect is stronger for white evaluators. Models with evaluator fixed effects provided the same results.

In summary, the results reveal positive in-group bias in conviction and sentencing for white evaluators. That is, white evaluators are biased in favor of white defendants (or against minority defendants) at all stages of the decision-making. The decisions of minority evaluators are nuanced. They exhibit negative in-group bias in conviction, indicating that minority evaluators are harsher towards defendants of their in-group during the guilt-innocence decision. On the other hand, minority evaluators treat defendants of their own race more leniently when assigning prison sentence and fine, although the latter effect is not statistically significant. This is the first paper that identifies such opposing effects within the same decision sequence. Because the majority of the evaluators in the sample are white, these in-group biases translate into overall racial biases in conviction, prison time and fines against minorities.³¹ A number of extensions and robustness analyses did not alter the conclusions. For example, the results are similar between the first three and last three trials. The results are also similar between fast decisions and slow decisions, as well as decisions that have, and have not been altered by evaluators at the end of the experiment. These and other sensitivity analyses are summarized in the Appendix.

³¹ Recall that, as described at the end of section V, the racial bias coefficient β_1 (the coefficient of *Minority Defendant*) in Equation (1) is equal to β_2 + $p\delta_2$, where β_2 and δ_2 are the relevant coefficients from the in-group bias regression (2), and p is the proportion of white evaluators.

Is it Race or Something Else?

Could these findings be an artifact of some other attribute of the defendants? For example, could it be the case that the body language of the defendants is influenced by their race and that it is the body language, rather than the race of the defendants, that triggers the response of the evaluators? There are a number of answers to this question. First, in our six trials and 2 versions of each trial (generating 12 versions with six white and six minority defendants), body language is very similar between white and minority defendants (see the pictures of the trials with white and minority defendants by scrolling down at http://proficient.ninja/uhasselt/). This is because after shooting a particular trial with the white defendant, that video was shown to the minority defendant (who would replace the white defendant in that particular video) so that he could mimic the body language of the white defendant.

The body language of the defendants did not project any disrespect towards the judge or the prosecutor either.³² The only possible difference was the white defendant in trial 1, who occasionally crossed his legs, while other defendants never crossed their legs. If crossing legs while sitting in front of the judge is considered as disrespectful behavior which should be punished, this would imply that this white defendant received harsher punishment than what was appropriate, and it indicates that our results are potentially an underestimate of the racial bias we detect against minorities.

One can argue that minorities would be more timid and nervous during a trial because of cultural reasons, while white defendants would be more self-confident. Self-confidence could signal innocence or trustworthiness, while being timid and nervous could suggest the implication of guilt, and this could be the reason why the evaluators were biased against minorities. As discussed earlier, minority defendants were told to mimic the body language of the white defendants. Second, recall that minority defendants are more likely to get convicted by *both* white and minority evaluators. Given that minority evaluators are less likely to fall into such a trap of cultural misunderstanding of the body language of their own in-group, our findings cannot be attributed to possible differences in body language.

³² The spoken language did not differ between the defendants. They did not speak during the trial with the exception of their statements regarding their understanding of the charges, their statement about a "not guilty" plea and their statement about not adding anything else to their attorney's defense. These statements were: "Yes," "Not guilty" and "No."

What is the Driver of the Bias?

It can be conjectured that white evaluators hold minorities responsible (explicitly, or implicitly) for undesirable social phenomena. For example, as shown in Table 1, 48 percent of the evaluators are very concerned about terrorism in Belgium. If those evaluators who think that terrorism is a very important problem also believe that minorities are more likely to be associated with terrorism, then these beliefs may translate into evaluators' propensity to treat minority defendants relatively harshly. To shed some light into this mechanism, we used the question that was posed to evaluators about terrorism.³³ We ran the models by dividing the sample into two groups of evaluators: those who think that it is very important that terrorism is a very important problem in Belgium, and those who are less concerned. Table 8 summarizes the results. The first two columns pertain to the results obtained from the former group, and columns (3) and (4) pertain to the latter one. Column (1) of Panel A shows that minorities are 11 percentage points more likely to get convicted in comparison to whites if the evaluators think that terrorism is a very important problem. But column (3) mirrors this result and shows that the minority impact is 13.3 percentage points among those evaluators who are not as concerned about terrorism.³⁴ Thus, there is no compelling evidence that evaluators' concerns about terrorism is the driver of the differential conviction rates between white and minority defendants. The estimated coefficients are also similar between columns (2) and (4) of Table 8 (which display the results of in-group bias effects) reflecting similar patterns of behavior between the two groups of evaluators. Consequently, the results displayed in Panel A of Table 8 are similar to those obtained from the entire sample, presented in Table 3.

Panels B and C of Table 8 summarize the results pertaining to sentence length and fine. Once again, the results are similar between evaluators who are very concerned about terrorism (columns 1 and 2) and those who are less concerned (column 3 and 4). Thus, the result displayed in Table 8 indicate that the racial bias against minority defendants are unlikely to be driven by

³³ Recall that evaluators answered this question via an online survey nine to 15 days after the experiment, and that the terrorism question was part of a set of questions including personal characteristics of the respondent and perceptions on institutions, jobs, income, and human values.

³⁴ The average conviction rate in the former group is 0.78, while it is 0.69 in the latter, indicating that the probability of getting convicted for minorities is 19 percent higher than that of a white defendant in the latter group of evaluators, while being minority increases the probability of conviction by 14 percent in the former group.

evaluators' concerns about terrorism.³⁶ Of course, as described earlier in the paper, the terrorist attacks in Brussels and Paris, which took place about two years before the experiment, may have been ingrained in the psyche of the evaluators regardless of how they revealed their concerns about terrorism in our survey. As a result, dividing the sample between those who are more or less concerned about terrorism may not lead to a significant difference between these two groups. On the other hand, it should be recognized that the trials in the experiment are not related to acts of terrorism and therefore judgements on the merits of these cases should, in principle, not be impacted by anxieties about terrorism.

Another concern might be that one or more of the defendants, who happen to be white, might look more trustworthy to the evaluators in comparison to other defendants. If people who look more trustworthy are generally treated more favorably, then the racial bias reported in the paper may be attributable to differences in perceived trustworthiness between white and minority defendants. We did not ask the evaluators about their assessment of the defendants' trustworthiness because such an inquiry could have made the evaluators suspicious about the intent of the experiment. Instead, we showed the pictures of the defendants to a different group of 49 students, who were juniors of Hasselt University. These 49 students were asked to rate each defendant, based on their pictures (see Figure 2), on a scale from 1 to 7 regarding their trustworthiness and criminal proclivity. These questions were: "How likely is it that this person is involved in a small crime such a shoplifting (1: Very likely- 7: Very unlikely)" and "How likely is it that this person would return a wallet to its owner after finding it on the street? (1: Very likely- 7: Very unlikely)".³⁸ Accounting for these perceived honesty scores of the defendants did not impact the results in a systematic way, indicating that the racial bias reported in the paper is unlikely to be driven by perceptions of honesty/dishonesty of the defendants.

³⁶ Using different cutoffs to classify evaluators to groups that are more or less concerned about terrorism did not later the results.

³⁸ White defendants, as a group, are perceived as more trustworthy. The average value among the three white defendants to the question about returning the wallet to its owner (being unlikely to commit a minor crime) is 4.49 (4.47), while the average among the three minority defendants is 4.03 (3.94). There is, however, variation between defendants. For example, Minority Defendant-2 is perceived more trustworthy than White Defendant-1 and about as trustworthy as White defendant-3. The same minority Defendant-1 is deemed to be less likely to commit a crime than White Defendants 1 and 3.

VIII. The Analysis of Attorney Decisions

As described earlier, 36 attorneys also participated in the experiment. Table 9 displays the summary statistics related to the conviction and sentencing decisions made by attorneys. They convict at a lower rate in comparison to students, but consistent with the behavior of students, attorneys too convict minority defendants at a higher rate (69 percent vs. 55 percent). Average effective prison sentence is similar between attorneys and students, and average fine assigned by attorneys is only slightly lower (420 Euros vs. 468 Euros).³⁹

Because there is no racial variation within attorneys (they are all white), in-group bias in their decisions cannot be analyzed. Instead, we focus on the investigation of overall racial bias. Columns (1) and (2) of Table 10 present the results regarding racial bias in conviction decisions, and show that minority defendants are about 14 percentage points more likely to be found guilty by practicing attorneys, holding constant everything else about the attributes of the case and trial environment. It is interesting to note that the size of the racial bias coefficient estimated in the sample of attorneys (14 percentage points as shown in Table 10) is almost the same as the one estimated among the sample of law students (14.6 percentage points as shown in Appendix Table A2, column 10).

Columns (3)-(4) and (5)-(6) of Table 10 present the regression results that analyze the impact of defendant race on prison sentence and fine, respectively, assigned by attorneys.⁴⁰ Columns (3)-(4) indicate that minority defendants receive sentence lengths that are on average 0.7 month longer, although this effect is not statistically different from zero at conventional levels. On the other hand, columns (5)-(6) reveal that attorneys assign 131 percent larger fines to convicted minority defendants (implied by the coefficient of 0.927). Thus, the analysis of attorneys' decisions reveal a pattern that is similar to those observed in the behavior of law students and economics students. Minority defendants are more likely to get convicted for the

³⁹ These are weighted averages of fines assigned to minority and white defendants shown in row (7) of Tables 2 and 9.

⁴⁰ Each of the 36 lawyers evaluated each of the six cases, yielding to 216 decisions on conviction versus acquittal (108 cases of white defendants and 108 cases of minority defendants). Fifty-nine of the white defendants are convicted (55 percent), while conviction rate among minority defendants was 69 percent (74 minority defendants are convicted). This implies the existence of 15 "excess" convicted minority defendants. Ranking of the 74 convicted minorities by sentence length and fine showed that 12 defendants received zero effective prison term and zero effective fine, despite their conviction. These are the marginal convicted defendants, who are arguably not guilty despite being convicted. Of the remaining convicted defendants with effective prison terms of zero months, we dropped three people who are sentenced to zero months of effective prison term and the lowest initial fines assigned as punishment (two people with 100 Euro fines, and one person with 104 Euros).

same exact crime, based on the same arguments made by the prosecutors and for the same defense by their attorney, and they are more likely to receive stiffer punishment upon conviction.

We also recruited judges and prosecutors to evaluate these cases. The small size of the judge and prosecutor sample does not allow us to analyze them separately, but we merged them with the attorney sample to investigate the sensitivity of the results. Adding 12 judges/prosecutors to the sample of 36 attorneys generated similar results to those obtained from the sample of attorneys, suggesting that judge and prosecutor decisions on these trials are not substantially different from those of the attorneys.⁴¹

IX. Summary and Conclusions

In this paper we investigate whether the decision about guilt vs. innocence of an individual, who is being accused of a crime, is race-blind. We also ask whether the extent of the punishment depends on race. Do these decisions depend on whether or not the defendant and the evaluator are of the same race? If so, are these influences stronger in case of white or minority evaluators? If there exist race effects on these decisions, are they impacted by whether the evaluator believes terrorism is a major problem in the country?

Some of these are old questions, and all of them are important for both scientific inquiry and public policy. These questions, however, have been difficult to answer because of a number of inherent omitted variables and endogeneity issues. At the heart of the issue lies the near impossibility to create a 'counterfactual scenario' in a trial, which involves the defense attorney, the prosecutor, and the defendant. We create a design which holds constant everything that takes place during a trial, with the exception of the race of the defendant. We shot 3D Virtual Reality (VR) videos of six criminal trials in Belgium, prosecuted by actual prosecutors and defended by actual defense attorneys in an actual courtroom. Only the defendants in the courtroom are actors. The prosecutors and the defense attorneys are given the case files one week before the shooting of the trials so that they could do the background work and prepare

⁴¹ There is no statistically significant racial bias in prison sentencing in either the attorney sample or in the attorney&prosecutor&judge sample. The point estimate of Minority Defendant in the conviction decision regression is 0.139 in the former sample, while it is 0.103 in the latter sample, and both are significantly different from zero, although the difference in the magnitudes is not inconsistent with the null effect in a judge& prosecutor only sample. The point estimate of Minority Defendant in the fine regression is 0.927 in the attorney sample and it is 0.860 in the attorney&prosecutor&judge sample (both significantly different from zero), and the difference in the magnitudes is inconsistent with the hypothesis that the judge&prosecutor sample would produce a nil effect.

their case and their defense. The prosecutors and the defense attorneys presented their cases orally in the court room, as they would normally do. The VR technology enabled us to replace white defendants in the courtroom with individuals who have Middle Eastern or North African descent. This allowed us to alter only the race of the defendants in these trials, holding all activity in the courtroom constant, including every word spoken by the prosecutor and the defense attorney, and all the body language in the courtroom. A short clip of two videos can be seen here http://proficient.ninja/splitscreen/. This paper is the first one that utilizes 3D Virtual Reality technology that uses actual people, rather than computer-animated scenes. Scrolling down at the link http://proficient.ninja/uhasselt/ allows one to observe scenes from all six trials. Full versions of one of these trials can also be watched (in 2D) at this link. Two versions of this particular trial are titled Video No 4 - Full Version 1 (with the minority defendant), and Video No 4, Full Version 2 (with the white defendant).

This design allows us to bypass the identification challenges faced by previous empirical research, and it enables us to create arguably perfect counterfactuals. In our case, the race of the defendant in each trial is uncorrelated with the characteristics of the prosecutors, with the characteristics of defense attorneys, and with any activity in the courtroom. Defendant race is also uncorrelated with evaluator attributes. The only variation in a given trial is obtained from the race of the defendants. We hold all actions (body language, spoken words, and so on) of the prosecutor and the defense attorney in the courtroom constant. Evaluators watched one of the two versions of a particular trial. They observed identical courtroom activity with only one difference: half of the evaluators saw a white defendant, the other half saw a minority defendant. The evaluators didn't know the names of the defendants; thus names cannot be used as signals of minority status (Bertrand and Mullainathan 2004). As is the usual situation in such trials, defendants barely spoke during the trial: they spoke only three times during the trial to answer three questions of the judge about whether they understand the charges against them (they answered "Yes"); whether they plead guilty or not guilty (they answered "Not guilty"), and whether they had anything else to add to what their lawyer said during the trial (they answered "No").⁴² Finally, all actor-defendants were instructed to wear similar clothes. Thus, the differentiation between white and minority defendants is obtained from the variation in their skin color. This point is verified by providing pictures of the defendants to another group of 89

⁴² All actor-defendants are born in Belgium and they speak fluent Dutch. Because they had no accent and because they spoke only a few words during the trial their minority status could not be inferred from the way they spoke.

freshmen students of Hasselt University. These students identified the race of each defendant correctly with 99 percent accuracy. A total of 153 Master's degree law students and undergraduate and Master's degree economics students are randomly assigned to watch, from the view point of the judge, the VR videos of these trials and make decisions on conviction as well as prison sentence and fine in accordance with the guidelines provided by the relevant law.

The results provide suggestive evidence for *negative in-group bias* during the conviction phase. Evaluators are more likely to convict a defendant if he is of the same race as the evaluator, although this effect is not statistically significant. Because evaluator race is, by design, uncorrelated with all other variables including the race of the defendant, in-group bias can be decomposed to determine its sources. This analysis shows that the statistically insignificant in-group bias in convictions emerges because of two opposing forces. White evaluators exhibit positive in group bias, while minority evaluators exhibit negative in-group bias; i.e. both white and minority evaluators treat white defendants favorably relative to minority defendants during the conviction stage. These effects produce the overall racial bias against minorities in conviction decision: minority defendants are 12 percentage points more likely to get convicted in comparison to white defendants even though they are tried for the same exact case and even though everything that went on during the trial is the same.

According to the relevant criminal law, convicted defendants can be assigned a prison sentence and/or a fine. In this sentencing stage, we find *in-group favoritism*. Convicted defendants receive prison sentences that are about 4 months shorter if the evaluator is of the same race. We show that this *positive* in-group bias in prison sentences is driven equally by white and minority evaluators. Because most evaluators are white, this behavior translates into racial bias in prison sentencing with minority defendants being sentenced on average 1.3 months longer prison terms, which corresponds to a 32 percent increase in prison time.

Positive in-group bias also exists in the assignment of fines, but this in-group bias is stronger in case of white evaluators, who assign fines to white defendants that are 68 percent lower in comparison fines for minority defendants. This translates into racial bias in fines where minority defendants receive stiffer fines than their white counterparts.

Using the subjective assessments of 49 individuals who are unrelated to the study, we show that controlling for perceived criminal proclivity or perceived honesty of defendants does not alter the results. When we analyze whether disparate treatment of defendants depends on the extent to which evaluators are concerned about terrorism in the country, we find no discernable behavioral difference between those evaluators who are very concerned about terrorism and those who are less concerned. This may suggest that the source of the racial bias may be deeprooted. Alternatively, it may be that the terrorist attacks in Brussels and Paris, which took place about two years earlier, generated increased animosity towards minorities which has not yet dissipated and that they impact the decisions on minority defendants even though the alleged offenses in these trials are unrelated to any terrorist act.

We repeated the same experiment with a group of 36 practicing attorneys, most of whom specialize in criminal law. Although this sample is smaller, it provides stronger external validity. Because all the attorneys in the sample are white, we could not investigate in-group bias in their decisions. The results were similar to those obtained from law students and economics students, both qualitatively and in magnitude. A small group of judges and prosecutors, who were not involved in the experiment, also participated as evaluators of these VR trials. Their small sample size did not permit us to analyze their decisions separately, but when we merged the sample of judges and prosecutors with the sample of lawyers, the analyses provided similar results to those obtained from the lawyer sample, suggesting that judge/prosecutor decisions are not substantially different from those of the lawyers.

A large number of additional analyses confirm the robustness of the results. For example, the speed with which the evaluators made their decisions, or whether the evaluators altered their original conviction or sentencing decisions have no impact on the results. Similarly, the decisions made during the first half of the experiment (the first three trials) are no different from those made in the second half, and the decisions of men and women are similar.

It is important to underline that for minorities the in-group bias in the conviction phase is negative, but that the in-group bias in the sentencing phase is positive. In other words, minority evaluators are harsher towards defendants of their own race when it comes to the guilt-innocence decision, but they favor same-race defendants during the sentencing phase. This finding is important because it reveals a changing pattern of bias (negative-then-positive in-group bias) in the same decision sequence by the same group of evaluators, and this is the first paper that identifies this pattern.

Figure 1 A Snapshot of the Virtual Reality Videos of the same Trial



This scene can be watched at http://proficient.ninja/splitscreen/

Figure 2 Six Defendants in Six Trials



Minority Defendant-1 (M1)



Minority Defendant-2 (M2)



Minority Defendant-3 (M3)



White Defendant-1 (WH1)



White Defendant-2 (WH2)



White Defendant-3 (WH3)

1			
	Evaluator Attributes		
Minority Evolution	0.11		
Minority Evaluator	(0.32)		
Mala Evaluator	0.47		
Iviale Evaluator	(0.50)		
Low Student	0.44		
Law Student	(0.50)		
	Evaluator Beliefs		
Terrorism is a very Important	0.48		
Problem in Belgium	(0.50)		
N	153		

Table 1Descriptive Statistics of Evaluators

Information on whether evaluators were law students or economics students and their sex were obtained when the evaluators participated in the experiment. Information on the belief regarding the importance of terrorism as a problem was obtained through an online survey completed 9 to 15 days after the experiment. It was part of a survey containing other questions ranging from job markets, to income, to human values.

			White E	White Evaluator		Evaluator
	Minority	White	Minority	White	Minority	White
	defendant	defendant	Defendant	Defendant	Defendant	Defendant
	(1)	(2)	(3)	(4)	(5)	(6)
(1) Conviction	0.79***	0.67***	0.78***	0.68***	0.82**	0.65**
Rate	(0.41)	(0.47)	(0.41)	(0.47)	(0.39)	(0.48)
Ν	459	459	408	408	51	51
(2) Initial Prison	10.15	9.55	10.47	9.51	7.69	9.82
Sentence (months)	(8.62)	(7.57)	(8.85)	(7.42)	(6.15)	(8.91)
Ν	361	309	319	276	42	33
(3) Suspended Prison	5.81	5.73	5.93	5.90	4.88	4.33
Sentence (months)	(5.20)	(5.57)	(5.26)	(5.62)	(4.72)	(4.94)
Ν	361	309	319	276	42	33
(4) Effective Prison	4.34	3.81	4.54*	3.61*	2.81**	5.48**
Sentence (months)	(6.73)	(5.11)	(7.03)	(4.93)	(3.48)	(6.31)
Ν	361	309	319	276	42	33
(5) Initial Fine	782.62	696.70	776.63	685.74	828.10	788.36
(Euros)	(1192.96)	(1088.64)	(1218.33)	(1101.48)	(990.45)	(985.24)
Ν	361	309	319	276	42	33
(6) Suspended Fine	290.87	278.10	288.28	272.96	310.55	321.03
(Euros)	(559.83)	(519.81)	(557.29)	(528.46)	(585.33)	(445.62)
Ν	361	309	319	276	42	33
(7) Effective Fine	491.75	418.61	488.35	412.78	517.55	467.33
(Euros)	(999.70)	(889.49)	(1027.55)	(906.93)	(764.96)	(737.55)
Ν	361	309	319	276	42	33

 Table 2

 Descriptive Statistics of Conviction Decision, Sentence Length and Fine Imposed by Evaluators

Cells contain means and (standard deviations). * signifies difference in the means between two columns of a group (columns 1 and 2, columns 3 and 4, or columns 5 and 6) at the 10% level. ** indicates difference at the 5% level, and *** represents difference at the 1% level or better.

		•					
	Ove	erall Bias		In-group Bias			
	(1)	(2)	(3)	(4)	(5)	(6)	
Minority Defendant	0.111***	0.120***	0.120***	0.103***	0.110***	0.110***	
	(0.028)	(0.025)	(0.024)	(0.029)	(0.026)	(0.026)	
Minority Evaluator				-0.025	-0.041	-0.039	
				(0.083)	(0.076)	(0.076)	
Minority Evaluator ×				0.070	0.089	0.085	
Minority Defendant				(0.102)	(0.087)	(0.084)	
Law Student	0.016	-0.011	-0.012	0.017	-0.011	-0.012	
	(0.028)	(0.028)	(0.028)	(0.029)	(0.029)	(0.029)	
Male Evaluator		-0.088***	-0.095***		-0.088***	-0.095***	
		(0.028)	(0.029)		(0.029)	(0.029)	
Early Trial	-0.127***	-0.101***		-0.127***	-0.101***		
	(0.028)	(0.030)		(0.028)	(0.030)		
Sequence 1			-0.073**			-0.073**	
			(0.034)			(0.034)	
Sequence 2			-0.037			-0.036	
			(0.031)			(0.031)	
Assault Case	-0.105***			-0.105***			
	(0.029)			(0.029)			
Observations	918	918	918	918	918	918	
Trial FE	No	Yes	Yes	No	Yes	Yes	

Table 3Overall and In-group Racial Bias in Conviction Decisions

Early Trial dummy = 1 if the trial is one of the first three trials watched by that evaluator. Sequence 1 and Sequence 2 are dichotomous indicators of the order in which the videos are watched. Standard errors are clustered at the evaluator level. *** p<0.01, ** p<0.05, * p<0.1

	0	
	(1)	(2)
	White Evaluators	Minority Evaluators
White Defendant	-0.110***	
	(0.026)	
Minority Defendant		0.204**
		(0.076)
Law Student	-0.011	-0.007
	(0.029)	(0.139)
Male Evaluator	-0.087***	-0.092
	(0.029)	(0.118)
Early Trial	-0.121***	0.072
2	(0.031)	(0.071)
Observations	816	102
Trial FE	Yes	Yes

Table 4 Decomposing In-group Racial Bias in Conviction Decisions: Conviction Regressions Conditional on Evaluator Race

Early Trial dummy = 1 if the trial is one of the first three trials watched by that evaluator. Standard errors are clustered at the evaluator level. *** p<0.01, ** p<0.05, * p<0.1

	A: Regressions using trimmed sample					
-	С	verall Bias		In-group Bia		
-	(1)	(2)	(3)	(4)	(5)	(6)
Minority Defendant	1.256**	1.300***	1.285**	1.634***	1.689***	1.684***
	(0.514)	(0.498)	(0.511)	(0.556)	(0.545)	(0.556)
Minority Evaluator				1.914	2.103*	2.080*
				(1.266)	(1.205)	(1.226)
Minority Evaluator ×				-3.544***	-3.671***	-3.766***
Minority Defendant				(1.136)	(1.211)	(1.240)
Law Student	-0.874	-0.387	-0.372	-0.844	-0.336	-0.327
	(0.654)	(0.666)	(0.670)	(0.663)	(0.672)	(0.678)
Male Evaluator		1.666**	1.751**		1.696**	1.776**
		(0.716)	(0.754)		(0.713)	(0.750)
Early Trial	-1.498***	-1.112*		-1.492***	-1.087*	
	(0.457)	(0.563)		(0.457)	(0.564)	
Sequence 1			0.232			0.227
-			(0.966)			(0.966)
Sequence 2			-0.298			-0.314
*			(0.647)			(0.648)
Assault Case	-3.188***			-3.199***		
	(0.428)			(0.424)		
Observations	618	618	618	618	618	618
Trial FE	No	Yes	Yes	No	Yes	Yes
		B∙ F	Regressions i	using untrimmed s	ample	
-		D. 1		untillined a	sample	
-	С	verall Bias		Ising untrimined s	n-group Bias	
-	(7)	Verall Bias (8)	(9)	<u>In In I</u>	a-group Bias (11)	(12)
Minority Defendant	(7) 0.605	0verall Bias (8) 0.708	(9) 0.692	<u>In In I</u>	n-group Bias (11) 1.117**	(12) 1.114**
Minority Defendant	(7) 0.605 (0.470)	0verall Bias (8) (0.708 (0.451)	(9) 0.692 (0.454)	$\frac{\frac{1}{(10)}}{1.011^{**}}$ (0.508)	<u>a-group Bias</u> (11) 1.117** (0.493)	(12) 1.114** (0.495)
Minority Defendant Minority Evaluator	(7) 0.605 (0.470)	0verall Bias (8) (0.708 (0.451)	(9) 0.692 (0.454)	$ \frac{\frac{\text{Ir}}{(10)}}{1.011^{**}} \\ (0.508) \\ 1.911 $	1-group Bias (11) 1.117** (0.493) 2.085*	(12) 1.114** (0.495) 2.077*
Minority Defendant Minority Evaluator	(7) 0.605 (0.470)	0.708 (0.451)	(9) 0.692 (0.454)	$ \frac{\frac{Ir}{(10)}}{1.011^{**}} \\ (0.508) \\ 1.911 \\ (1.265) $	(11) 1.117** (0.493) 2.085* (1.202)	(12) 1.114** (0.495) 2.077* (1.219)
Minority Defendant Minority Evaluator Minority Evaluator ×	(7) 0.605 (0.470)	0.708 (0.451)	(9) 0.692 (0.454)		<u>ample</u> <u>1-group Bias</u> (11) 1.117** (0.493) 2.085* (1.202) -3.720***	(12) 1.114** (0.495) 2.077* (1.219) -3.832***
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant	(7) 0.605 (0.470)	0.708 (0.451)	(9) 0.692 (0.454)		ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202)	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student	(7) 0.605 (0.470) -0.848	-0.345	<u>(9)</u> 0.692 (0.454) -0.312		ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student	(7) 0.605 (0.470) -0.848 (0.617)	-0.345 (0.630)	<u>(9)</u> 0.692 (0.454) -0.312 (0.634)		Interpretended Interpr	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator	(7) 0.605 (0.470) -0.848 (0.617)	-0.345 (0.630) 1.607**	(9) 0.692 (0.454) -0.312 (0.634) 1.655**	$\begin{array}{r} & \\ \hline \\ \hline$	ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629**	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674**
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator	(7) 0.605 (0.470) -0.848 (0.617)	-0.345 (0.630) 1.607** (0.680)	$\begin{array}{r} (9) \\ \hline 0.692 \\ (0.454) \\ \end{array}$ $\begin{array}{r} -0.312 \\ (0.634) \\ 1.655^{**} \\ (0.712) \end{array}$		ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676)	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674** (0.707)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial	(7) 0.605 (0.470) -0.848 (0.617) -1.402***	-0.345 (0.630) 1.607** (0.680) -1.089**	$\begin{array}{r} \hline (9) \\ \hline 0.692 \\ (0.454) \\ \hline \\ 0.634) \\ 1.655^{**} \\ (0.712) \end{array}$	$\frac{1}{(10)}$ 1.011** (0.508) 1.911 (1.265) -3.638*** (1.155) -0.846 (0.627) -1.385***	ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676) -1.049**	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674** (0.707)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial	(7) 0.605 (0.470) -0.848 (0.617) -1.402*** (0.427)	-0.345 (0.630) 1.607** (0.680) -1.089** (0.521)	$\begin{array}{r} (9) \\ \hline 0.692 \\ (0.454) \\ \end{array}$ $\begin{array}{r} -0.312 \\ (0.634) \\ 1.655^{**} \\ (0.712) \end{array}$	$\frac{Ir}{(10)}$ 1.011^{**} (0.508) 1.911 (1.265) -3.638^{***} (1.155) -0.846 (0.627) -1.385^{***} (0.427)	ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676) -1.049** (0.523)	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674** (0.707)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1	-0.848 (0.617) -1.402*** (0.427)	-0.345 (0.630) 1.607** (0.680) -1.089** (0.521)	(9) 0.692 (0.454) -0.312 (0.634) 1.655** (0.712) -0.021	$\begin{array}{r} & \\ \hline \\ \hline$	ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676) -1.049** (0.523)	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674** (0.707) -0.008
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1	(7) 0.605 (0.470) -0.848 (0.617) -1.402*** (0.427)	-0.345 (0.630) 1.607** (0.680) -1.089** (0.521)	$\begin{array}{r} (9) \\ \hline 0.692 \\ (0.454) \\ \end{array}$ $\begin{array}{r} -0.312 \\ (0.634) \\ 1.655^{**} \\ (0.712) \\ \end{array}$ $\begin{array}{r} -0.021 \\ (0.886) \end{array}$	$\begin{array}{r} & \\ \hline \\ \hline$	1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676) -1.049** (0.523)	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674** (0.707) -0.008 (0.890)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2	<u>(7)</u> 0.605 (0.470) -0.848 (0.617) -1.402*** (0.427)	-0.345 (0.630) 1.607** (0.680) -1.089** (0.521)	$\begin{array}{r} (9) \\ \hline 0.692 \\ (0.454) \\ \hline 0.634) \\ 1.655^{**} \\ (0.712) \\ \hline -0.021 \\ (0.886) \\ -0.278 \end{array}$	$\begin{array}{r} & \\ \hline \\ \hline$	ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676) -1.049** (0.523)	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674** (0.707) -0.008 (0.890) -0.292
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2	(7) 0.605 (0.470) -0.848 (0.617) -1.402*** (0.427)	-0.345 (0.630) 1.607** (0.680) -1.089** (0.521)	$\begin{array}{r} (9) \\ \hline 0.692 \\ (0.454) \\ \hline 0.634) \\ 1.655^{**} \\ (0.712) \\ \hline -0.021 \\ (0.886) \\ -0.278 \\ (0.627) \end{array}$		ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676) -1.049** (0.523)	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674** (0.707) -0.008 (0.890) -0.292 (0.625)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2 Assault Case	(7) 0.605 (0.470) -0.848 (0.617) -1.402*** (0.427) -3.193***	-0.345 (0.630) 1.607** (0.680) -1.089** (0.521)	$\begin{array}{r} (9) \\ \hline 0.692 \\ (0.454) \\ \end{array}$ $\begin{array}{r} -0.312 \\ (0.634) \\ 1.655^{**} \\ (0.712) \\ \end{array}$ $\begin{array}{r} -0.021 \\ (0.886) \\ -0.278 \\ (0.627) \end{array}$	Image Ir (10) 1.011** (0.508) 1.911 (1.265) -3.638*** (1.155) -0.846 (0.627) -1.385*** -1.385*** (0.427)	1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676) -1.049** (0.523)	(12) $1.114**$ (0.495) $2.077*$ (1.219) $-3.832***$ (1.237) -0.297 (0.642) $1.674**$ (0.707) -0.008 (0.890) -0.292 (0.625)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2 Assault Case	(7) 0.605 (0.470) -0.848 (0.617) -1.402*** (0.427) -3.193*** (0.411)	-0.345 (0.451) -0.345 (0.630) 1.607** (0.680) -1.089** (0.521)	$\begin{array}{r} (9) \\ \hline 0.692 \\ (0.454) \\ \hline 0.634) \\ 1.655^{**} \\ (0.712) \\ \hline -0.021 \\ (0.886) \\ -0.278 \\ (0.627) \end{array}$	$\begin{array}{r} & \\ & \\ \hline \\ \hline$	1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676) -1.049** (0.523)	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674** (0.707) -0.008 (0.890) -0.292 (0.625)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2 Assault Case Observations	(7) 0.605 (0.470) -0.848 (0.617) -1.402*** (0.427) -3.193*** (0.411) 670	-0.345 (0.630) 1.607** (0.680) -1.089** (0.521)	(9) 0.692 (0.454) -0.312 (0.634) 1.655** (0.712) -0.021 (0.886) -0.278 (0.627) 670	$\begin{tabular}{ c c c c c c c c c c c c c c c c c c c$	ample 1-group Bias (11) 1.117** (0.493) 2.085* (1.202) -3.720*** (1.202) -0.323 (0.636) 1.629** (0.676) -1.049** (0.523)	(12) 1.114** (0.495) 2.077* (1.219) -3.832*** (1.237) -0.297 (0.642) 1.674** (0.707) -0.008 (0.890) -0.292 (0.625)

Table 5Overall and In-group Racial Bias in Prison Sentencing

Effective Prison Sentence is equal to initial prison sentence minus suspended sentence. Standard errors are clustered at the evaluator level. *** p<0.01, ** p<0.05, * p<0.1. See notes to Table 3.

	A: Regressions using trimmed sample					
	(Overall Bias		Ι	n-group Bia	S
	(1)	(2)	(3)	(4)	(5)	(6)
Minority Defendant	1.082***	1.061***	1.071***	1.151***	1.131***	1.135***
	(0.198)	(0.194)	(0.196)	(0.211)	(0.206)	(0.209)
Minority Evaluator				1.138*	1.213**	1.216**
				(0.599)	(0.577)	(0.568)
Minority Evaluator ×				-0.655	-0.664	-0.616
Minority Defendant				(0.571)	(0.585)	(0.576)
Law Student	-0.746***	-0.689**	-0.695**	-0.674**	-0.599**	-0.604**
	(0.270)	(0.281)	(0.280)	(0.275)	(0.288)	(0.286)
Male Evaluator	× ,	0.084	0.053	· · · ·	0.117	0.084
		(0.284)	(0.279)		(0.281)	(0.277)
Early Trial	0.975***	0.627**	()	0.962***	0.619**	
j	(0.239)	(0.249)		(0.240)	(0.252)	
Sequence 1	(0.20))	(0.2.12)	0.017	(01210)	(0.202)	-0.010
sequence i			(0.334)			(0.319)
Sequence 2			0.078			0.081
Sequence 2			(0.309)			(0.309)
Assault Case	-0.098		(0.50))	-0.116		(0.30))
Assault Case	(0.231)			(0.232)		
Observations	618	618	618	618	618	618
Trial FE	No	Yes	Yes	No	Yes	Yes
	110			110		105
		B. Regn	essions using	o untrimmed	sample	
	(B: Regro	essions using	g untrimmed	sample In-gr	oup Bias
	(7)	B: Regro Dverall Bias (8)	(9)	$\frac{\text{g untrimmed}}{(10)}$	sample In-gro	oup Bias
Minority Defendant	(7) 0.461**	B: Regro Dverall Bias (8) 0.436**	$\frac{(9)}{0.443^{**}}$	$\frac{10}{0.550**}$	sample In-gr (11) 0 523**	oup Bias (12) 0 522**
Minority Defendant	(7) 0.461** (0.200)	B: Regro Dverall Bias (8) 0.436** (0.198)	(9) 0.443** (0.199)		sample In-gro (11) 0.523** (0.210)	oup Bias (12) 0.522** (0.211)
Minority Defendant	(7) 0.461** (0.200)	B: Regro Dverall Bias (8) 0.436** (0.198)	(9) 0.443** (0.199)		sample In-gro (11) 0.523** (0.210) 1.220**	oup Bias (12) 0.522** (0.211) 1.237**
Minority Defendant Minority Evaluator	(7) 0.461** (0.200)	B: Regro Dverall Bias (8) 0.436** (0.198)	(9) (0.443** (0.199)		sample In-gro (11) 0.523** (0.210) 1.220** (0.576)	0.522** (0.211) 1.237** (0.576)
Minority Defendant Minority Evaluator	(7) 0.461** (0.200)	B: Regro Dverall Bias (8) 0.436** (0.198)	(9) (9) (0.443** (0.199)	(10) 0.550** (0.212) 1.152* (0.600) -0.872	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendent	(7) 0.461** (0.200)	B: Regro Dverall Bias (8) 0.436** (0.198)	(9) 0.443** (0.199)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant	(7) 0.461** (0.200)	B: Regro Dverall Bias (8) 0.436** (0.198) 0.681**	(9) 0.443** (0.199)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) 0.661**	sample In-grv (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) 0.500*	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) 0.612*
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student	(7) 0.461** (0.200) -0.725** (0.202)	B: Regro <u>Overall Bias</u> (8) 0.436** (0.198) -0.681** (0.204)	<u>(9)</u> 0.443** (0.199) -0.699** (0.202)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.200)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.212)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.211)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student	(7) 0.461** (0.200) -0.725** (0.292)	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112		(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.986
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator	(7) 0.461** (0.200) -0.725** (0.292)	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.205)	<u>(9)</u> 0.443** (0.199) -0.699** (0.303) 0.057 (0.201)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.204)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.200)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator	(7) 0.461** (0.200) -0.725** (0.292)	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.62***	<u>(9)</u> 0.443** (0.199) -0.699** (0.303) 0.057 (0.301)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial	(7) 0.461** (0.200) -0.725** (0.292) 0.945***	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.653*** (0.242)	<u>(9)</u> 0.443** (0.199) -0.699** (0.303) 0.057 (0.301)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299) 0.934***	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304) 0.645***	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial	(7) 0.461** (0.200) -0.725** (0.292) 0.945*** (0.234)	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.653*** (0.242)	<u>(9)</u> 0.443** (0.199) -0.699** (0.303) 0.057 (0.301)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299) 0.934*** (0.235)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304) 0.645*** (0.244)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1	(7) 0.461** (0.200) -0.725** (0.292) 0.945*** (0.234)	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.653*** (0.242)	<u>(9)</u> 0.443** (0.199) -0.699** (0.303) 0.057 (0.301) -0.230	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299) 0.934*** (0.235)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304) 0.645*** (0.244)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1	(7) 0.461** (0.200) -0.725** (0.292) 0.945*** (0.234)	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.653*** (0.242)	<u>(9)</u> 0.443** (0.199) -0.699** (0.303) 0.057 (0.301) -0.230 (0.353)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299) 0.934*** (0.235)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304) 0.645*** (0.244)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2	(7) 0.461** (0.200) -0.725** (0.292) 0.945*** (0.234)	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.653*** (0.242)	<u>(9)</u> 0.443** (0.199) -0.699** (0.303) 0.057 (0.301) -0.230 (0.353) 0.035	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299) 0.934*** (0.235)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304) 0.645*** (0.244)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2	(7) 0.461** (0.200) -0.725** (0.292) 0.945*** (0.234)	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.653*** (0.242)	<u>(9)</u> 0.443** (0.199) -0.699** (0.303) 0.057 (0.301) -0.230 (0.353) 0.035 (0.343)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299) 0.934*** (0.235)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304) 0.645*** (0.244)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300)
Minority Defendant Minority Evaluator Minority Evaluator Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2 Assault Case	(7) 0.461** (0.200) -0.725** (0.292) 0.945*** (0.234) -0.255	B: Regro <u>Overall Bias</u> (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.653*** (0.242)	<u>(9)</u> 0.443** (0.199) -0.699** (0.303) 0.057 (0.301) -0.230 (0.353) 0.035 (0.343)	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299) 0.934*** (0.235)	sample In-gr (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304) 0.645*** (0.244)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2 Assault Case	(7) 0.461** (0.200) -0.725** (0.292) 0.945*** (0.234) -0.255 (0.222)	B: Regro <u>Overall Bias</u> (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.653*** (0.242)		(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299) 0.934*** (0.235) -0.270 (0.223)	sample In-grv (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304) 0.645*** (0.244)	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300)
Minority Defendant Minority Evaluator Minority Evaluator × Minority Defendant Law Student Male Evaluator Early Trial Sequence 1 Sequence 2 Assault Case Observations	(7) 0.461** (0.200) -0.725** (0.292) 0.945*** (0.234) -0.255 (0.222) 670	B: Regro Dverall Bias (8) 0.436** (0.198) -0.681** (0.304) 0.112 (0.305) 0.653*** (0.242) 670	(9) 0.443** (0.199) -0.699** (0.303) 0.057 (0.301) -0.230 (0.353) 0.035 (0.343) 670	(10) 0.550** (0.212) 1.152* (0.600) -0.872 (0.606) -0.661** (0.299) 0.934*** (0.235) -0.270 (0.223) 670	sample In-grv (11) 0.523** (0.210) 1.220** (0.576) -0.854 (0.617) -0.599* (0.313) 0.142 (0.304) 0.645*** (0.244) 670	oup Bias (12) 0.522** (0.211) 1.237** (0.576) -0.781 (0.605) -0.612* (0.311) 0.086 (0.300) -0.256 (0.343) 0.038 (0.342)

Table 6Overall and In-group Racial Bias in Fines

See notes to Table 5.

	Prison Sentence			Fine				
	Trimmed	l Sample	Untrimmed Sample		Trimmed Sample		Untrimmed Sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	White	Minority	White	Minority	White	Minority	White	Minority
	Evaluators	Evaluators	Evaluators	Evaluators	Evaluators	Evaluators	Evaluators	Evaluators
White	-1.689***		-1.118**		-1.128***		-0.521**	
Defendant	(0.546)		(0.494)		(0.206)		(0.210)	
Minority		-1.729		-2.317**		0.468		-0.302
Defendant		(1.034)		(1.057)		(0.531)		(0.513)
Law Student	-0.084	-3.029**	-0.112	-2.395*	-0.645**	0.154	-0.673**	0.479
	(0.706)	(1.276)	(0.672)	(1.197)	(0.294)	(1.408)	(0.320)	(1.451)
Male Evaluator	1.535**	3.072*	1.482**	2.977*	0.055	0.790	0.104	0.724
	(0.774)	(1.752)	(0.740)	(1.667)	(0.302)	(0.810)	(0.326)	(0.854)
Early Trial	-1.068*	-1.056	-1.006*	-1.415	0.570**	1.019	0.623**	1.011*
	(0.621)	(1.101)	(0.587)	(0.939)	(0.270)	(0.679)	(0.268)	(0.544)
Observations	552	66	595	75	552	66	595	75
Trial FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7
Decomposing In-group Racial Bias in Prison Sentencing and Fines
Prison and Fina Pagrossions Conditional on Evaluator Paga

Early Trial dummy = 1 if the trial is one of the first three trials watched by that evaluator. Standard errors are clustered at the evaluator level. *** p<0.01, ** p<0.05, * p<0.1

	$\Lambda \cdot Overall a$	nd In group Dagie	Bias in Conviction	Decisions
-	Terrorism is a v	very important	Terrorism is a	very important
-	problem = 1		proble	m = 0
	(1)	(2)	(3)	(4)
Minority Defendant	0.111***	0.100***	0.133***	0.127***
	(0.031)	(0.032)	(0.038)	(0.040)
Minority Evaluator		-0.038		-0.032
		(0.140)		(0.084)
Minority Evaluator ×		0.123		0.045
Minority Defendant		(0.118)		(0.124)
Observations	438	438	480	480
_	B: Overall	and In-group Rac	ial bias in Prison Se	ntencing
	Terrorism is a v	ery important	Terrorism is a	very important
_	problei	n = 1	proble	em = 0
_	(1)	(2)	(3)	(4)
Minority Defendant	1.004**	1.368***	1.579*	1.942**
	(0.491)	(0.491)	(0.824)	(0.898)
Minority Evaluator		2.330		1.463
		(2.150)		(1.451)
Minority Evaluator ×		-3.874**		-2.926**
Minority Defendant		(1.917)		(1.369)
Observations	310	310	308	308
	C: C	Verall and In-grou	p Racial Bias in Fir	ies
_	Terrorism is a v	ery important	Terrorism is a	very important
_	problei	m = 1	proble	em = 0
	(1)	(2)	(3)	(4)
Minority Defendant	1.003***	1.223***	1.094***	0.992***
	(0.275)	(0.286)	(0.287)	(0.318)
Minority Evaluator		2.564***		0.139
		(0.482)		(0.789)
Minority Evaluator \times		-2.335***		0.790
Minority Defendant		(0.579)		(0.752)
Observations	310	310	308	308

Table 8
Overall and In-group Racial Bias in Conviction, prison Sentencing and Fines
Subsamples by Evaluators' Concern about Terrorism

Columns (1) and (3) are based on the same specification as in column (2) of Table 3. Columns (2) and (4) are based on the same specification as in column (5) of Table 3. Standard errors are clustered at the evaluator level. *** p<0.01, ** p<0.05, * p<0.1

Attorney Sample:
Descriptive Statistics of Conviction Decision, Sentence Length and Fine

	Minority	White
	Defendant	Defendant
	(1)	(2)
(1) Conviction Rate	0.69**	0.55**
	(0.47)	(0.50)
N	108	108
(2) Initial Prison Sentence	11.2	11.1
	(6.5)	(7.5)
N	74	59
(3) Suspended Prison Sentence	6.2	6.5
-	(4.2)	(4.1)
N	74	59
(4) Effective Prison Sentence	5.0	4.5
	(6.5)	(7.5)
Ν	74	59
(5) Initial Fine	711	701
	(840)	(779)
Ν	74	59
(6) Suspended Fine	286	284
	(460)	(364)
N	74	59
(7) Effective Fine	425	417
	(554)	(577)
N	74	59

Cells contain means and (standard deviations). * signifies difference between the means in the corresponding row at the 10% level. ** indicates difference at the 5% level, and *** represents difference at the 1% level or better.

Attorney Sample							
	Conviction		Pri	son	Fine		
	(1)	(2)	(3)	(4)	(5)	(6)	
Minority	0.139**	0.139**	0.707	0.716	0.925**	0.927**	
Defendant	(0.060)	(0.061)	(0.890)	(0.862)	(0.430)	(0.421)	
Male	-0.106*	-0.108*	2.138	1.840	0.569	0.431	
Evaluator							
	(0.061)	(0.055)	(1.717)	(1.425)	(0.656)	(0.644)	
Early Trial	-0.104		-0.298		0.167		
	(0.063)		(1.379)		(0.457)		
Sequence 1		0.008		1.241		0.648	
		(0.068)		(2.447)		(0.758)	
Sequence 2		-0.149**		0.202		0.246	
_		(0.061)		(1.022)		(0.897)	
Observations	216	216	118	118	118	118	
Trial FE	Yes	Yes	Yes	Yes	Yes	Yes	

Table 10 Overall Racial Bias in Conviction Decisions, Prison Sentencing and Fines: Attorney Sample

Early Trial dummy = 1 if the trial is one of the first three trials watched by that evaluator. Sequence 1 and Sequence 2 are dichotomous indicators of the order in which the videos are watched. Standard errors are clustered at the evaluator level. *** p < 0.01, ** p < 0.05, * p < 0.1

References

Abrams, David S, Marianne Bertrand and Sendhil Mullainathan. 2012. "Do Judges Vary in Their Treatment of Race?" The Journal of Legal Studies, 41:2, 347-383

Agan, Amanda, Matthew Freedman and Emily Owens.2018. "Is Your Lawyer a Lemon? Incentives and Selection in the Public Provision of Criminal Defense" Forthcoming in The Review of Economics and Statistics.

Alesina, Alberto and Eliana La Ferrara. 2014. "A Test of Racial Bias in Capital Sentencing." The American Economic Review, 104:11, 3397-3433

Anderson, James M and Paul Heaton. 2012. "How Much Difference Does the Lawyer Make? The Effect of Defense Counsel on Murder Case Outcomes." Yale Law Journal, 122:154, 154-217

Angrist, Joshua, Eric Bettinger and Michael Kremer. 2006. "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." The American Economic Review, 96:3, 847-862

Arnold, David, Will Dobbie, and Peter Hull Hull, P. (2020). *Measuring racial discrimination in bail decisions*. National Bureau of Economic Research Working paper No: 26999.

Anwar, Shamena, Patrick Bayer and Randi Hjalmarsson. 2012. "The Impact of Jury Race in Criminal Trials." The Quarterly Journal of Economics, 127:2, 1017-1055

Argys, Laura M and H Naci Mocan. 2004. "Who Shall Live and Who Shall Die? An Analysis of Prisoners on Death Row in the United States." The Journal of Legal Studies, 33:2, 255-282

Ayres, Ian and Peter Siegelman 1995. "Race and gender discrimination in bargaining for a new Car." *The American Economic Review*, 85(3): 304-321.

Bertrand, Marianne, and Esther Duflo. 2017. "Field Experiments on Discrimination." in Abhijit Banerjee and Esther Duflo (es.) *Handbook of Fields Experiments*. North Holland.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." American Economic Review, 94 (4): 991-1013.

Chen, Yan and Sherry Xin Li. 2009. "Group Identity and Social Preferences." American Economic Review, 99:1, 431-457.

Depew, Briggs, Ozkan Eren and Naci Mocan. 2017. "Judges, Juveniles and In-group Bias." The Journal of Law and Economics, 60:2, 209-239

Eren, Ozkan and Naci Mocan. 2018. "Emotional Judges and Unlucky Juveniles." American Economic Journal: Applied Economics, 10:3, 171-205

Gazal-Ayal, Oren and Raanan Sulitzeanu-Kenan. 2010. "Let My People Go: Ethnic In-Group Bias in Judicial Decisions—Evidence from a Randomized Natural Experiment." Journal of Empirical Legal Studies, 7:3, 403-428

Glaeser, Edward L and Bruce Sacerdote. 2003. "Sentencing in Homicide Cases and the Role of Vengeance." The Journal of Legal Studies, 32:2, 363-382

Goette, Lorenz, David Huffman and Stephan Meier. 2006. "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence using Random Assignment to Real Social Groups." American Economic Review, 96:2, 212-216

Grossman, Guy, Oren Gazal-Ayal, Samuel Pimentel and Jeremy Weinstein, 2015. "Descriptive Representation and Judicial Outcomes in Multiethnic Societies." American Journal of Political Science. 60: 1, pp. 44-69.

Heckman, james. 1998. "Detecting Discrimination." Jornal of Economic Perspectives. 12(2): 101-116.

Heckman, James and Peter Siegelman 1993. "The Urban Institute audit studies: Their methods and findings" in Michael Fix and Raymond Struyk (eds.) *Clear and Convincing Evidence: Measurement of Discrimination in America*, Urban institute press, Washington, pp. 187-258

Hoffman, Florian and Philip Oreopoulos. 2009. "A Professor Like Me: Influence of Professor Gender on University Achievement." Journal of Human Resources. 44:2, 479-494.

Jones, Angela, Amanda Bergold, Marlee Kind Dillon and Steven Penrod. 2017. "Comparing the Effectiveness of Henderson Instructions and Expert Testimony: Which Safeguard Improves Jurors' Evaluations of Eyewitness Evidence?" Journal of Experimental Criminology, 13, 29-52.

Kennedy, Peter E. 1981. "Estimation with Correctly Interpreted Dummy varaibles in Semilogarithmic Equations." American Economic Review. p. 801

Levitt, Steve D. 2004. Testing theories of discrimination: evidence from Weakest Link. The Journal of Law and Economics, 47(2), 431-452.

List, John and Michael Price.2009. "The role of social connections in charitable fundraising: Evidence from a natural field experiment." Journal of Economic Behavior and Organization, 69:2, 160-169

List, John. 2004. "The nature and extent of discrimination in the marketplace: Evidence from the field." The Quarterly Journal of Economics, 119(1), 49-89.

Maeder, Eveltn, Annik Mossieere and Liann Cheung. 2012. "Candian Mock Juror Attitudes and Decisions in Domestic Violence Cases involving Asisna and White Interracial and Intraracial Couples." Journal of Interpersonal Violence. 28:4, 667-684.

Mendoza, Saaid A, Sean P Lane and David M Amodio. 2014. "For Members Only: Ingroup Punishment of Fairness Norm Violations in the Ultimatum Game." Social Psychological and Personality Science, 5:6, 662-670

Mercado, Cynthia Calkins, Brian Bornstein and Robert Schopp. 2006. "Decision-Making about Volitional Imapirment in Sexually Violent Predators." Law and Human Behavior. 3, 587-602.

Mossiere, Annik, Evelyn Maeder and Emil Pica. 2018."Racial Composition of Couples in Battered Spouse Syndrome Cases: A Look at Juror Perceptions and Decisions." Journal of Interpersonal Violence. 33:18, 2867-2890.

Neumark, David, Roy Bank, and Kyle D. Van Nort, 1996." Sex discrimination in restaurant hiring: An audit study." The Quarterly journal of economics, 111(3): 915-941.

Omi, Michael and Howard Winant, 1994. Racial Formation in the United States. New York: Routledge.

Ondrich, Jan, Stephen Ross, and John Yinger, 2000. How common is housing discrimination? Improving on traditional measures. Journal of Urban Economics, 47(3): 470-500.

Pager, Devah, Bruce Western, and Naomi Sugie. (2009. "Sequencing disadvantage: Barriers to employment facing young black and white men with criminal records." The ANNALS of the American Academy of Political and Social Science, 623(1):195-213.

Pager, Devah. 2003. "The mark of a Criminal Record." American Journal of Sociology, 108: 937-975.

Rehavi, M Marit and Sonja B Starr. 2014. "Racial Disparity in Federal Criminal Sentences." Journal of Political Economy, 122:6, 1320-1354.

Riach, P. A., and J. Rich, 2002. Field experiments of discrimination in the market place. The economic journal, 112(483), F480-F518.

Sadowski, W., & Stanney, K.M. (2002). Presence in virtual environments. In K. M. Stanney (Eds.). Handbook of virtual environments: Design, implementation and applications (pp. 791-806). Mahwah : IEA

Shayo, Moses and Asaf Zussman. 2011. "Judicial Ingroup Bias in the Shadow of Terrorism." The Quarterly Journal of Economics, 126:3, 1447-1484.

Shayo, Moses and Asaf Zussman. 2017. "Conflict and the Persistence of Ethnic Bias." American Economic Journal: Applied Economics, 9:4, 137-65.

Turner, Margery Austin, Michale Fix and Raymond J. Struyk. 1991. *Opportunities denied, opportunities diminished: Racial discrimination in hiring.* The Urban institute Press, Washington DC.

Appendix

<u>1. Bench Trials in Belgium</u>

Jury trials do not play an important role in the Belgian judiciary, and they are mostly reserved for murder cases. Furthermore, jury trials were abolished in 2016, although they have been reinstalled one year later because the abolishment was deemed unconstitutional by the Belgian constitutional court. Belgium is a civil law country and therefore it relies on an inquisitorial criminal justice system. Consequently, judges play a more important role in trials than they do in adversarial systems dominated by plea bargaining. Plea bargaining did not exist in Belgium until recently. Its implementation started in 2016, and it is applicable for a limited number of crimes (those punishable with fewer than 5 years of prison time). Some crimes, such as rape, sex crimes against minors and manslaughter, are excluded from the possibility of plea bargaining regardless of the sentence. Unlike in the U.S., Belgian law requires the prosecutor's plea bargain to be endorsed by a judge.

2. Details of the Experiment

Cases

Although we did not change any details of the actual cases used in the experiment (obtained from the office of the prosecutor), we altered the names of the victims and witnesses to guarantee anonymity of the real cases. Synopses of actual case files are provided to the evaluators before the beginning of the experiment so that the evaluators could assess the background of each case, the forensic evidence, the background of the defendant, and so on, as the judge would do. We did not provide the names of the defendants to evaluators in these case files to avoid any signaling about the race of the defendants. In the videos, the defendant is always referred to as 'my client' (by the lawyer) or as 'the defendant' (by the prosecutor). For burglary cases we changed the location of the event to eliminate the possibility of an evaluator being familiar with a particular burglary incident in a particular neighborhood. Given that the actual defendants were in some cases older than our actor-defendants, we changed the age of the defendant in the case files that are provided to the evaluators to match the age of the defendant-actors in our courtrooms.

Evaluators

We involved 153 students from the Faculty of Business Economics, and Faculty of Law of Hasselt University to act as judges in these trials. The Economics students were a mix of juniors and master's students that were enrolled in the Policy Evaluation course. The Economics group consisted of 86 students. They were randomly subdivided into four groups (because we had 25 headsets to watch the VR videos) to participate in the

experiment.¹² The evaluators watched the VR videos using VR headsets (Oculus Rift headsets). The lab hosted 25 students at one time (See Appendix Figures, Figure 1).

The evaluators were told that they were participating in an experiment which was an exercise in the adjudication of criminal acts. We organized a random lottery (with 20 movie tickets) for evaluators who participated in both the experiment and the follow up survey.³ A follow up survey, which is completed by the evaluators over the internet, is used to gather background information on the evaluators. This information is collected not on the same day of the experiment but nine days later to minimize concerns about incorrect information being provided.⁴

We recruited practicing attorneys from the Limburg Bar Association. By email and by phone we contacted 250 lawyers who were listed under the category of "criminal law" on the web site of the Association. Thirtysix lawyers agreed to participate. The attorney received the same information as the student participants and the flow of the proceedings was identical.⁵ Because the overwhelming majority of the lawyers were white we could not investigate in-group bias in their decisions, but we were able to analyze the existence of racial bias in lawyers' decisions.

We also tried to recruit judges and prosecutors, with limited success. We asked the court presidents of multiple courts and the head of one District Attorney's office for permission to conduct the experiment.⁶ Only eight judges and four prosecutors agreed to participate. The VR experiment with judges and prosecutors was conducted partly in the court buildings, and partly at Hasselt University for those who weren't available on the dates the VR equipment was set in the court buildings. Participating judges and prosecutors are guaranteed anonymity.

¹ All four groups watched the videos on the same day, and we made sure there was no interaction between any of the four groups of students. To avoid interaction between the groups, we put them in separate rooms and let them watch a movie (or be lectured by an instructor) while one of the other groups was involved in the experiment. They were also instructed not to inform each other via mail or phone. A professor was always supervising the group of students who were not in the lab.

 $^{^{2}}$ Law students are enrolled in Master's degree program at Hasselt University. They participated in the experiment as part of a course on Research Methods, which is mandatory class for all law students. There were 67 law students enrolled in this course, and they all took the experiment the day after the Economics students. Law students were randomly divided into three groups and again we made sure they could not interact with the other groups as the experiment was ongoing for one of the groups.

³ The experiment took place in November 2017. There was no unusual news or events in Belgium during this period related to immigration, terrorism or asylum seekers

⁴ If we had asked the evaluators questions about their cultural background, about their concern regarding crime and the judicial system etc. immediately after they completed the experiment, they might have been cognizant about the purpose of the survey, and may have provided untruthful responses. The survey contained 40 questions, and the key questions were sprinkled throughout. For example, we asked innocuous questions about concerns regarding unemployment, importance of family, and so on, before asking whether they think terrorism is a major problem in Belgium.

⁵ The lawyers took the experiment individually during a ten-day period in July 2018.

⁶ Obviously, the prosecutors who participated in the VR trials did not take part in the evaluations. The courts and District Attorney's office sent an email to all judges and prosecutors and asked them if they wanted to take part in a study on criminal trials that uses virtual reality technology.

3. VR Production

A professional Virtual Reality production firm was hired, and a cameraman and a technician were in charge to ensure the technical quality of the video and the audio. A static 360 camera (OZO) was used to shoot the cases. The camera was positioned before the bench of the judge and was on the same height as the judge, which allowed the participant in the experiment to observe the court room from the exact same angle as the actual judge. All videos are shot in one court room in the main court building in Hasselt, Belgium.

We developed a website that guided the participants through the experiment. The website contained instructions for the evaluators on what to do (when to put on the headsets, when to take them off, when to read a case folder, etc.). The website also communicated with the headset so that the videos would play automatically in the correct order, and provided a platform which allowed the evaluators to submit their decisions in each case (conviction/acquittal and sentences in case of conviction). As a real judge would do, the evaluators had to first read the case folder pertaining to that particular case. The case folders were color-coded so that each of the six cases were assigned a different color folder. The case folders contained the relevant background information about the cases, including the summary of the police report, and the criminal history of the defendant (See Online Appendix, Section 5 for the details of each of the six cases, the way they were presented to the subjects).

4. Extensions and Robustness

We used an alternative measure to determine the cultural background of evaluators. Instead of making use of information on parents' country of origin, we determined whether the evaluator is a minority in Belgium by using information on the language used in the household (see the Descriptive Statistics Section for details). Using this alternative indicator of minority status provided very similar point estimates.

The unconditional mean of the prison sentence and the fine assigned by the evaluators are smaller than their respective variances and the distributions are left-skewed. Estimation of the prison term and fine regressions using negative binomial models, which provided the same inference. During the experiment the evaluators had the option (as the actual judicial procedure allows for) to reflect and to revise their original decisions on conviction, prison sentence and fine. There are 153 evaluators who made 2,258 decisions (918 decisions on whether to convict, and 670 decisions on prison sentence and fine on those who are found guilty). We divided the sample into two groups: those evaluators who never altered their first decisions (88 evaluators) and those who made at least one change in their decisions (65 evaluators). Changing a decision may indicate that the evaluator contemplated more carefully about the case and therefore felt the need to revise his/her original decision. Alternatively, if an evaluator was very deliberate in watching the presentations of the prosecutor and the defense attorney and if the evaluator read the case file carefully, he/she did not have the need to go back and revise the original decisions about conviction and/or sentencing. It is also possible that changing a decision at the end of the experiment may indicate that the evaluator has realized the purpose of the experiment, and as a result he/she went back to revise at least one of the decisions he/she made during the experiment. In this case, any statistically significant effect of racial bias would disappear or would be smaller in the sample of evaluators who changed at least one of their decisions. As shown in columns (4) and (5) of Table A2, the result were similar between these two groups of evaluators.

To investigate whether evaluators took their task seriously, we analyzed the time they spent in making their decisions. Figure 2 in Appendix Figures displays the distribution of total time spent on 6 trials by evaluators

for decision-making. For example, 500 seconds means that that the evaluator took on average 1 minute and 23 seconds to decide on a case after he/she completed watching the VR video of the case.⁷ Recall that the evaluators read the case files prior to watching the VR videos of the case. The case file contains information about the case (police report, etc., as well as the sentencing guidelines for that particular crime (See the online Appendix). A quick decision after reading the case files and after watching the trial may imply that the evaluator watched the trial carefully and formed an opinion during the trial, and did not have to think long about the verdict and punishment. Alternatively, a quick decision. To investigate the sensitivity of the results to decision time, we dropped from the sample 25 percent of the fastest evaluators and 25 percent of the slowest evaluators and re-estimated the models, which provided similar results (see columns 6 and 7 of Table A2). We repeated the exercise by dropping the slowest and fastest 30 percent, 20 percent, and 15 percent of the distribution of cases, and obtained very similar results.

Similarly, the evaluators may have gotten fatigued during the experiment and they may have lost their concentration towards the end. If this is the case, the decisions made later during the experiment should be less careful and more noisy. Alternatively, they may have realized the purpose of the experiment, and may have made their decisions accordingly. To investigate this point, we analyzed separately the first three decisions and the last three decisions made by the evaluators, which showed no difference between these groups. Finally, we re-estimated the models using the sample of male or female evaluators and using only law students or only economics students. Appendix Table A2 summarizes the results of these exercises and demonstrates the robustness of the results.

Punishment decisions (prison sentence and fine) are made on those who are convicted. Recall that the results reveal racial bias in the conviction decisions against minorities (columns 1-3 of Table 3). This means that the sample of convicted defendants include some minorities who are found guilty because of their race. Thus, in the analyses pertaining to prison term and fine we focus on the trimmed samples which drop marginally innocent minority defendants. Alternatively, instead of trimming the sample to eliminate marginally not guilty defendants, we used all defendants regardless of their conviction status, but assigned a prison sentence of zero and a fine of zero to those who were found not guilty. We then used the entire sample to run prison sentence and fine regressions. The results were consistent with those reported in Tables 6 and 7. The estimated impacts were smaller in magnitude but they were sizable and statistically significant.

Finally, we created a dummy variable which takes the value of one if the defendant is convicted and received an effective prison sentence greater than zero months. The variable equals zero if the defendant is acquitted, or if the defendant is convicted but was assigned no prison term (entire sentence suspended, making effective prison term zero). Using this variable on the whole sample of students we found that the outcome is 7.8 percentage points more likely to take place for minority defendants (0.078, se=0.028). The in-group bias coefficient was -0.141 (se=0.086) in this specification. In the attorney sample, the coefficient of minority defendant was 0.120 (se=0.065)

⁷ This is total time spent to make the decision on guilt/innocence, and on prison sentence and fine.

Appendix Table A1

Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
(Sequence1,	(Sequence1,	(Sequence2,	(Sequence2,	(Sequence3,	(Sequence3,
Version 1)	Version 2)	Version 1)	Version 2)	Version 1)	Version 2)
WH1 SOBUR ¹	^{M1} SOBUR ¹ BUR ¹	${}^{M2}_{TO}AS^3_P$	$^{WH2}_{TO}AS^3_P$	WH2 TOBUR ³	^{M2} TOBUR ³
M1 SOBUR ² BUR ²	WH1 SOBUR ²	$^{WH3}_{M}AS_{P}^{2}$	${}^{M3}_{M}AS^2_{P}$	${}^{M3}_{M}AS^{1}_{P}$	WH3 MAS ¹
WH2 TOBUR ³	${}^{M2}_{TO}BUR^3_B$	$^{M3}_{M}AS^{1}_{P}$	WH3 MAS ¹	${}^{M1}_{SO}BUR^2_B$	BUR ² _B
$^{M3}_{M}AS^{1}_{P}$	WH3 MAS ¹	WH2 TOBUR ³	${}^{M2}_{TO}BUR^3_B$	$^{WH3}_{M}AS_{P}^{2}$	$^{M3}_{M}AS^2_P$
$^{WH3}_{M}AS^2_P$	$^{M3}_{M}AS^2_P$	$_{SO}^{M1}BUR_B^2$	WH1 SOBUR ²	WH1 SOBUR ¹ BUR ¹	^{M1} SOBUR ¹ BUR ¹ B
$_{TO}^{M2}AS_{P}^{3}$	$^{WH2}_{TO}AS^3_P$	WH1 SOBUR ¹ BUR ¹	${}^{M1}_{SO}BUR^1_B$	${}^{M2}_{TO}AS^3_P$	$^{WH2}_{TO}AS^3_P$

The Sequence of Trials Watched by Evaluators, and the Identity of Defendants, Prosecutors and Defense Attorneys in Each Trial

Each evaluator was randomly assigned to one of these six sets. Each set includes six cases (trials), that are watched in the order listed in each column. Each cell, such as ${}^{WH_1}_{S0}BUR^1_B$, summarizes the characteristics of that particular trial. BUR and AS stand for a burglary case, and an assault case, respectively. BUR¹ means the first burglary case, BUR² means the second burglary case, AS³ stands for the third assault case, and so on.

A subscript to the right identifies the prosecutor: BUR_B^1 means that the first burglary case is prosecuted by Bruno (B). There are two prosecutors in the experiment: Bruno (B) and Pieter (P). The subscript to the left identifies the defense attorney. There are three defense attorneys: TO, SO and M.

The superscript to the left identifies the defendant. There are three minority defendants: M1, M2, and M3; and there are three white defendants: WH1, WH2 and WH3. See Figure 2 for the pictures of all six defendants.

The entry ${}^{\text{WH1}}_{\text{SO}}\text{BUR}^{1}_{\text{B}}$ represents the first burglary case (BUR¹), where the defendant was WH1. The case was prosecuted by B, and the defense attorney was SO.

The first VR video of Set 1 is ${}^{WH1}_{S0}BUR^1_B$ and the first video of Set 2 is ${}^{M1}_{S0}BUR^1_B$. These two videos are identical in all respects but one: the race of the defendant (WH1 vs. M1); See Figure 1.

Each evaluator watched one of the six sets depicted in Table 1. Thus, each one of the three burglary cases (BUR¹, BUR², BUR³) as well as each of the three assault cases (AS^1 , AS^2 , AS^3) is watched by each evaluator.

	Whole Sample (1)	First 3 decisions (2)	Last 3 decisions (3)	Change (4)	No change (5)	Drop Slow 25% (6)	Drop Fast 25% (7)	Males (8)	Females (9)	Law Students (10)	Econ Students (11)
Racial bias in conviction	0.120*** (0.024)	0.115*** (0.041)	0.123*** (0.033)	0.061* (0.031)	0.164*** (0.035)	0.123*** (0.029)	0.107*** (0.029)	0.129*** (0.040)	0.115*** (0.030)	0.146*** (0.037)	0.098*** (0.033)
In-group bias in conviction	0.085 (0.084)	0.148 (0.114)	0.027 (0.114)	-0.016 (0.072)	0.221 (0.138)	-0.037 (0.106)	0.016 (0.091)	0.159 (0.135)	0.025 (0.107)	0.199 (0.142)	0.061 (0.102)
Racial bias in prison sentence	1.285** (0.511)	1.093 (0.907)	1.482*** (0.562)	0.859 (0.530)	1.692** (0.837)	1.325** (0.655)	0.462 (0.324)	1.830** (0.910)	0.722 (0.446)	1.013** (0.480)	1.479* (0.837)
bias in prison sentence	-3.766*** (1.240)	-3.055* (1.580)	-4.340** (1.953)	-3.946** (1.591)	-3.158** (1.455)	-4.045** (1.809)	-3.050** (1.190)	-5.894*** (1.955)	-2.163 (1.496)	-2.529** (1.066)	-4.305** (1.706)
Racial bias in fine	1.071*** (0.196)	0.999*** (0.305)	1.125*** (0.268)	1.113*** (0.269)	1.037*** (0.285)	1.074*** (0.219)	1.032*** (0.226)	1.076*** (0.289)	1.214*** (0.250)	1.105*** (0.294)	1.047*** (0.266)
In-group bias in fine	-0.616 (0.576)	-0.273 (0.567)	-1.172 (0.927)	-1.101 (0.791)	-0.012 (0.825)	-0.235 (0.852)	-0.829 (0.587)	-1.740*** (0.625)	0.400 (0.776)	-0.942 (0.826)	-0.450 (0.759)

Table A2Results of Sensitivity Analyses

Table A2 summarizes the results obtained from the sensitivity analyses. Overall Bias indicates the estimated coefficient of Minority Defendant in the relevant regression. In-group bias pertains to the coefficient of Minority Defendant x Minority Evaluator. Column (1) replicates the estimates obtained from the entire sample as presented in Tables 4B, 4A, 6B, 6A, 7B and 7A. Columns (2) and (3) report the results based on the first three and last three decisions, respectively, of each evaluator. Column (4) presents the estimates related to cases in which the evaluators have modified at least one of their initial decisions. Column (5) pertains to the sample of cases where the initial decisions are not modified. Finally, columns (6) and (7) present the estimates where the slowest 25 percent and fastest 25 percent of evaluators are dropped from the estimation sample. As Table A2 reveals, the estimates are highly consistent across various sub-samples, which indicates that fastness or slowness of decision-making, altering or not altering the initial decisions, or decisions made earlier or later during these six trials have no significant impact on the results. Similarly, there is no appreciable difference between law students and economics students (columns (10) and (11)) and between male and female evaluators (columns (8) and (9)), with one difference: racial in-group bias is not significantly different from zero in case of females

APPENDIX FIGURES

Appendix Figure 1 Evaluators Watching the VR Videos of the Trials



Appendix Figure 2 Distribution of Time Taken to Make Decisions (in Seconds)



ONLINE Appendix

1. Details of the Experiment

1.1. Anonymity of Evaluators and the Set-up of the Experiment

The participants are granted anonymity during the experiment and in the follow-up survey (see Section 2 of this Appendix). Upon entering the computer lab, students had to randomly draw a three digit number from a bowl. After drawing the number, the students could randomly choose one of the 25 computers in the lab. Before they could start with the experiment they had to type in the three digit number in a custom made online tool.

For the purpose of the experiment we developed a website that would guide the participants through the experiment. The website contained clear instructions for participants on what to do (when to put on the headsets, when to take off the headsets, when to read a case folder, etc.). The website also communicated with the headset so that the videos would play automatically in the correct order. The experiment took 2 hours, with a five minute break after watching three videos.

Participants first had to read the case folder of a certain case. The folders were color coded to avoid confusion. After reading the case folder students could click on a button to start the video. They had 30 seconds to put on the headsets. After watching the video they could take off the headset and click on the button to go to the deliberation phase.

1.2. Sentence Assignments

After watching each case video, the students were faced with the choice to either acquit or convict the defendant. The picture below shows the screen students were shown during the deliberation phase.



(*Translation: Conviction decision: Please indicate whether you want to convict or acquit the defendant. Note: You can read the pink case folder again if you want. Buttons: Convict; Acquit)*

If the evaluators acquitted the defendant, they would automatically proceed to the next case folder and they were instructed to read the next case folder. If they convicted the defendant they would go to the next screen to assign the sentences (which could be a prison sentence, a fine or a combination of both, depending on the circumstances of the case). For each case the evaluators could assign a sentence that was in accordance with the range prescribed by the penal code (for instance, a prison sentence between 12 months and 60 months). The Belgian criminal code also allows for the suspension of sentences (both for prison sentence and for fine). So the evaluators also had the option to assign a certain sentence of 12 months and suspend 4 months of this sentence. This would mean that the defendant would go to prison for only 8 months, and the other 4 months were suspended. Both the law and econ students were given a small lecture (by the same lecturer) on the assignment of sentences before they participated in the experiment. If the penal code prescribes that the actual judge in a certain type of crime has to assign both a prison sentence and a fine or either one of them, we made sure the participants were faced with the same option. See the picture below.

We also made sure the students could not type in values for the postponed sentences that were higher than the total sentences. We also made sure that students could not go outside of the range prescribed by the penal code. The information on the sentence range allowed by the penal code was on the case folders (see the case folders at the end of the annex for the ranges of the sentences for each case).

After making the decision on acquittal or conviction and assigning the sentences for each of the six cases, the students were presented an overview of all their conviction decisions and sentences. They were then informed that they could make changes to their answers. We registered both the initial answers of the students as well as their changed answers.

UHasselt	
	Strafmaat
	U heeft geoordeeld dat de beklaagde schuldig is. Gelieve een straf toe te wijzen.
	U vult eerst de totale gevangenisstraf en geldboete in, en vervolgens het deel daarvan dat wordt uitgesteld. U bent uiteraard niet verplicht om (een deel van) de straf uit te stellen. U moet én een gevangenisstraf én een geldboete uitspreken.
	Totale gevangenisstraf in maanden Waarvan uitgesteld (in maanden)
	Totale geldboete in euro Waarvan uitgesteld (in euro)
	Verder

(Translation of text: You have decided that the defendant is guilty. Please assign a sentence. You first have to fill in the total prison sentence and fine, and next you fill in the part of the sentence that is suspended. It is of course not mandatory to suspend (part of) the sentence. You have to fill in a prison sentence and a fine. Translation of boxes: Total prison sentence in months; postponed prison sentence (in months); Total fine in euro; postponed fine (in euro))

Overzicht veroordelingsbeslissingen

٠	Zaak	Becordeling	Gevangenisstraf	Waarvan uitgesteld	Boete	Waarvan uitgesteld
ţ.	Roze zaek	Verbordelen Vrijspreken	10	5	500	20
2	Cranje zaak	Veroordelen Vrijspreken	30	. 0	1000	0
3	Blauwe zook	Verbordelen Vijspreken	0	0	0	0
4	Groene zaak	Veroordelen Vrijspreken	4	3	500	0
5	Gele zaak	Veroordelen Vijppreken	Ø	0	0	0
6	Witte zaak	Veroordelen Vrijspreken	10	10	600	0

(*Translation: Overview conviction decision; # Case; Decision ; Prison sentence; Postponed prison sentence; Fine; Postponed fine*)

1.3. Pilot Testing the Experiment

Two weeks prior to the experiment, we had an extensive testing phase. We tested the entire experiment with two groups of respectively 8 and 15 participants (who had no connections to the university). They tested both the functionality and the user friendliness of the online website and the headsets. This lead to the inclusion of a 1 minute instruction video on how to use the headsets in the beginning of the actual experiment. The participants also tested the clarity and terminology of the case folders which lead to minor changes in the wording.

2. <u>The Follow-up survey</u>

2.1.Drafting and Testing the Follow-up Survey

We decided to organize a follow up survey for the participants. For the questions of the survey we made use of questions from the European Social Survey rounds 5 and 8. The survey contained groups of questions on the following topics: personal characteristics of the respondent, crime and migration, institutions, income and human values. We tested the survey on a group of 15 participants to make sure the questions were clear and the total duration of the survey took no more than 12 minutes. The survey was programmed in Qualtrics and was available via an online link. In the beginning of the survey the participants had to fill in the same three digit number they used in the experiment, this way we were able to match their answers and guarantee their anonymity.

2.2.Organization of the Follow-up Survey

We conducted the survey nine days to two weeks after the experiment during a lecture in the respective courses of the economics and law students. All students were informed to bring their laptops to fill in the survey. We sent an email with the link to students who were not in class so they could also fill in the survey. In total 153 of the 165 students that participated in the experiment filled in the survey. These 153 students constitute the sample used in the paper.

The survey contained 40 questions, and the key questions were sprinkled throughout. For example, we asked innocuous questions about concerns regarding unemployment, importance of family, and so on, before asking whether terrorism is a major problem in Belgium.

For the lawyer participants we registered the survey immediately after the experiment. Race questions were not included in the survey given to lawyers because all lawyers were white.

3. <u>Testing the photos of defendants: Race of the Evaluators</u>

We used a different group of 89 students who were enrolled in a freshman Microeconomics course of Business Engineering degree at Hasselt University to test whether the students can correctly identify the race of the actors used in the experiment. These 89 students have not participated in the experiment and presumably have no knowledge of the experiment. The students were in an auditorium and a large screen in the auditorium displayed defendants' pictures one at a time. Under each picture there were options to choose from regarding the racial/ethnic heritage of the person in the picture. The options were: Western European descent, Middle Eastern or North African descent, and Asian descent. Students were told that examples of Western European descent would be countries such as Belgium, Netherlands, Germany and France; Middle Eastern and North African descent would include such countries as Turkey, Morocco, Syria and Iraq, and examples of Asian descent would include countries such as China, Japan and South Korea. Each student had a hand-held device to enter their choices within 30 seconds after each picture was shown. The students were not allowed to talk to each other during this process. In addition to the six defendants in our trials, we added the picture of a young Chinese man for the students to evaluate.

The results showed almost unanimous consensus regarding the racial/ethnic heritage of the individuals in the pictures; and their heritage was identified correctly. Specifically, 97 percent of the students correctly guessed that the first minority defendant was of Middle Eastern/North African descent. All students who participated in this exercise guessed correctly the heritage of the second and third minority defendants. The students guessed with 98 percent accuracy that the first and the second White defendants were of Western European origin (Belgian), and they guessed with 100 percent accuracy that the third white defendant-actor had in fact Western European descent. Also, all students correctly identified the Chinese person as someone of Asian descent. Overall the minority defendant-actors have been identified as being a minority with almost 99 percent accuracy, and white defendants are identified correctly as being of Western European descent with 98.5 percent accuracy.

4. <u>Race vs. Ethnicity</u>

Ethnicity refers to common ancestry, a perception of common history and culture. Race typically refers to a group of people defined by physical characteristics. Race classifications are not established by a set of natural or biological factors but they are human constructs where skin color, eye shape, height, hair type are used as markers to define races (Omi and Winant 1994). The European Court of Human Rights states that "Whereas the notion of race is rooted in the idea of biological classification of human beings into subspecies on the basis of morphological features such as skin colour or facial characteristics, ethnicity has its origin in the idea of societal groups marked in particular by common nationality, religious faith, shared language, or cultural and traditional origins and backgrounds".

Race and ethnicity often overlap, and groups of people may move between categories. For example, as explained by Omi and Winant (1994), for the British, Irish were considered as a different race, although they had similar physical attributes as the British. In the U.S. Irish used to be treated as similar to the Black race. Now they are considered "white" but they constitute an ethic group. U.S. Census Bureau listed five races in 1870: White, Colored (Blacks), Colored (Mulattoes), Chinese, Indian. In 1950 there were three race classifications: White, Black, Other. In 1990 the classifications became White, Black, Asian or Pacific Islander, American Indian or Alaska Native, and in 2010 they consisted of White, Black, Asian, Hawaiian or Pacific Islander, American Indian or Alaska Native. The Census Bureau contemplated using Arab as a new racial category for the 2020 census, although the idea was not implemented.

5. <u>Case descriptions</u>

5.1.Burglary-1

<u>Charge</u>

Burglary, to the detriment of Frederik Smits and Bram Rogiers (victims). The following goods have been taken: a vehicle BMW 320cd, 100,00 EUR cash, a men's watch with a value of 50,00 EUR and spare keys of the BMW 320.

<u>Defendant</u>

- 20 years old
- Unmarried
- Unemployed

Criminal record of the defendant

- Refusal drug test and driving without driver's license: fine 2,400 EUR, 3 months driving ban.
- Driving without insurance and without driver's license, with vehicle that does not meet technical requirements of vehicles (motor vehicle inspection) and is not registered: fine of 2,400 EUR, 3 months driving ban.
- Burglary: 3-month prison sentence.
- Burglary by breaking, climbing in, or false keys: 8-month prison sentence.

Police report summary

- Victim Frederik Smits reported a burglary with breaking in his house.
- The forensics investigation for trace evidence in the house does not yield any results with the DNA and fingerprint databank.
- The burglar has gained access by ramming down a wooden door.
- 2 days after the burglary the victim informs the police that a friend has found the missing BMW.
- Forensics conducts an investigation of trace evidence on the car, and finds a fingerprint and DNA on exterior door.
- The fingerprint turns out to be the defendant's.
- House search (with consent defendant) at defendant's does not yield results (stolen goods not found).
- During interrogation the defendant denies having something to do with these facts, and does not know how his fingerprint ended up on the car. He does not know the victim.
- Victim says that he does know the defendant, however, the defendant has never been in his car with him.

Sentence range allowed by the penal code

- In case of conviction the defendant can be sentenced with a prison sentence from 1 month till 60 months and a fine between 156 EUR and 6,000 EUR.

5.2. Burglary-2

<u>Charge</u>

Burglary, to the detriment of John Peeters and Elke Deferm (victims). The following goods have been taken: a music installation of Bose, a golden watch from Ferrari, parfum of Mugler and Burberry, a HP laptop and 1,000 EUR cash.

<u>Defendant</u>

- 22 years old
- Unmarried
- Unemployed

Criminal record of the defendant

- Threatening to attack persons: 3-month prison sentence (suspended) and a fine of 600 EUR (suspended).
- Document fraud: 100 hours of community service.
- Indecent exposure: 2-month prison sentence (suspended).

Police report summary

- Victim John Peeters reported a burglary with breaking in his house.
- The forensics investigation includes taking trace evidence in the house and on an outside window. The DNA on the outside of the window matches with the DNA of the defendant.
- The defendant denies the facts and has no explanation why his DNA is on the window. He also claims to be physically not capable of breaking into the house due to a broken toe and a malfunctioning knee.
- The defendant admits to have been in the area of the break-in frequently at the time because he had a relationship with someone in the same area.

Sentence range allowed by the penal code

- In case of conviction the defendant can be sentenced with a prison sentence from 1 month till 60 months and a fine between 156 EUR and 6,000 EUR.

5.3. Burglary-3

Charge

Burglary, to the detriment of Tom Bamps (victim). The following goods have been taken: 2 HP laptops, 1 Apple tablet, 1 Rolex watch, 2 golden necklaces and 2,500 EUR cash.

Defendant

- 23 years old
- Unmarried
- Unemployed

Criminal record of the defendant

- No criminal record

Police report summary

- Victim Tom Bamps reported a burglary with breaking in his house. The victim came home from a night out and finds that the alarm system is not turned on. Upon entering the house

he noticed that there had been a burglary. The neighbor indicates she noticed a black Volkswagen Passat standing in front of the house in the evening but she did not notice the number plate.

- The victim claims to have turned on the alarm system before he left the house. The alarm system appears to be working and there is no sign that the alarm system has been tampered with.
- The burglar(s) gained access through a small window in the back of the house that was not protected by the alarm system.
- The alarm system was only installed in the week prior to the break-in. The victim states that one of the technicians was asking a lot of questions on their social activities in the coming weeks. The victim saw him in their street a few days before the burglary.
- The forensic investigation turns up one finger print on the outside of the window that was used by the burglar(s) to gain access to the house.
- Both technicians that installed the alarm system are interrogated. One of them has an alibi for the night of the burglary. The second technician was fired from the company the day before the burglary under suspicion of stealing material from the company.
- The technician denies the allegations of breaking in. He states to have no financial problems and that his parents would support him if he had financial problems.
- The technician owns a black Volkswagen Passat (same model as seen in front of the house the night of the burglary). He claims not to have been in the street at the time of the facts.
- The fingerprint found on the window matches the fingerprint of the technician. He claims that the fingerprint was there because he installed the alarm system the week before the burglary.

Sentence range allowed by the penal code

- In case of conviction the defendant can be sentenced with a prison sentence from 1 month till 60 months and a fine between 208 EUR and 8,000 EUR.
- -

5.4 Assault -1

<u>Charge</u>

Intentional assault of partner.

<u>Defendant</u>

- 25 years old
- Unmarried
- Unemployed

Criminal record of the defendant

- No criminal record

Police report summary

- Victim and defendant have a relationship for five years and have a 3-year old son together.
- There was an argument between the victim and the defendant, because he was not allowed to take the son to a family gathering.
- The victim tried to film the argument with her smartphone and the defendant grabbed the smartphone and smashed it. The victim states that the defendant then grabbed her and threw her on the table and on the ground. The defendant says she lost her balance and fell but that he was not behaving violently.

- The victim says she managed to get upstairs and call the police with another cell phone.
- The police see no apparent injuries on the victim. The police also see no signs on the table of a fight.
- The victim goes to the hospital on the same day and the medical record shows that she has pain in the neck, nausea and a tingling feeling in both her arms. The defendant claims that she had these complaints for an entire year and it has nothing to do with the argument.
- The victim claims that he has been aggressive before and hit her on multiple occasions in front of their son. The police was called to the house a year ago, but no police report was filed.
- The victim and the defendant decide to live in the same house for financial reasons. Once they sell the house, they will each go their own way.

Sentence range allowed by the penal code

In case of conviction the defendant can be sentenced with a prison sentence from 4 months till 24 months and a fine between 300 EUR and 1,200 EUR.

5.5. Assault-2

<u>Charge</u> Intentional assault of partner.

<u>Defendant</u>

- 24 years old
- Married
- Employed

Criminal record of the defendant

- No criminal record

Police report summary

- Victim calls the police and claims that the husband has locked her in the house.
- Police enter the house through a window.
- Victim says that she and her husband have been problems for 6 months.
- Victim claims that her husband destroyed her clothes three weeks ago and strangled her in front of their 2 year old son. She shows the police pictures of her injured neck but she did not file it.
- Victim claims that her husband hit her and locked her in the house.
- The defendant claims that he did not hit his wife, and that she has a key so that she could have left the house. He claims to have never hit his wife and that she destroyed her own clothes three weeks ago.
- The defendants sister is interrogated and claims that the victim lived with her for a while because of the problems in her marriage. The sister claims that she could not believe that her own brother would use violence and states that the victim was aggressive herself.

Sentence range allowed by the penal code

In case of conviction the defendant can be sentenced with a prison sentence from 3 months till 24 months and a fine between 208 EUR and 1,600 EUR.

5.6 Assault-3

<u>Charge</u> Intentional assault of partner.

Defendant

- 23 years old- Unmarried - Employed

Criminal record of the defendant

- Driving under influence: loss of driver's license for a month and 1,100 EUR fine.
- Driving under influence: loss of driver's license for 4 months.
- Document fraud: 7 months prison sentence (suspended) and fine (962.5 EUR).
- Online fraud: 6 months prison sentence and fine (600 EUR).

Police report summary

- Police are called to the house for a domestic dispute. The police have been at the house before for domestic disturbances while both partners are drunk.
- Police find a broken glass on the kitchen sink and hair of the victim on the stairs.
- The victim goes to the doctor to document the injuries and provides the police with previous reports of injuries which she claims have been caused by the defendant.
- The argument started the previous day during a night out at a bar, when the defendant claimed that the victim was hanging around other men. The defendant left irritated and his partner stayed there. She did not come home that night, but only around noon the next day. The victim claimed to spent the night at a friend's place, while the defendant claims that she spent the night with another man. When she arrived at home, the couple started to have the argument.
- The victim claims that the defendant pulled her by her hair throughout the house. After that she claims he pushed her down the stairs. The defendant denies to have hit her. He claims that she fell from the stairs herself (without him pushing her) and that he did aggressively tried to help her up by pulling her hair.
- Both of them admit to drink too much. The defendant claims that the victim has a real problem.
- The DA's office has tried mediation in this case to avoid it coming to court, but the mediation was not successful.
- The couple is back together at the moment of the trial.

Sentence range allowed by the penal code

In case of conviction the defendant can be sentenced with a prison sentence from 1 month till 12 months and a fine between 156 EUR and 600 EUR.