Let My People Go:
Ethnic In-Group Bias in Judicial Decisions – Evidence from a Randomized Natural Experiment

Oren Gazal-Ayal[*] *& Raanan Sulitzeanu-Kenan*[**]

[*] Senior Lecturer, University of Haifa, Faculty of Law, Mount Carmel 31905, Israel, ogazal@law.haifa.ac.il.
[**] Lecturer, the Hebrew University of Jerusalem, Political Science Department, and Federmann School of Public Policy, Mount Scopus, Jerusalem 91905, Israel, rasulke@huji.ac.il.

Let My People Go:
Ethnic In-Group Bias in Judicial Decisions – Evidence from a Randomized Natural Experiment

A B S T R A C T

Does ethnic identity affect judicial decisions? We provide new evidence on ethnic biases in judicial behavior, by examining the decisions of Arab and Jewish judges in first bail hearings of Arab and Jewish suspects in Israeli courts. Our setting avoids the potential bias from unobservable case characteristics by exploiting the random assignment of judges to cases during weekends, and by focusing on the *difference* in ethnic disparity between Arab and Jewish judges. The study concentrates on the early-stage decisions in the judicial criminal process, controlling for the state's position, and excluding agreements, thereby allowing us to distinguish judicial bias from other sources of ethnic disparities. We find systematic evidence of in-group (same ethnic group) bias in detention decisions. However, in cases where the decision is to detain, no ethnic bias was found in the length of the detention. Possible interpretations and implications of these findings are discussed.

I. INTRODUCTION

Most judicial decisions should not be affected by the race or ethnicity of the parties, and all fair judicial decisions should be oblivious to the judges' identity. Still, many studies point to the possible effect defendants' race and ethnicity have on judicial decisions, given racial and ethnic disparities in various punitive treatments. Yet inferring judicial bias from racial and ethnic disparities has proven to be difficult. The question remains whether these differences reflect racial differences in criminal behavior, or the outcome of ethnic bias? More formally, a null hypothesis of "zero disparity in judicial outcomes" can be rejected only if *all* the relevant variables are controlled for. Indeed, differential control of contextual variables in various studies has produced conflicting results.

Moreover, most of these studies concentrated on the effect of race or ethnicity on sentencing. Yet sentences are affected by many procedural decisions that precede this final stage of the trial. Among others, sentences are influenced by the defendants' pretrial status – whether they were detained or not, the prosecution's position in the sentencing hearing, the defendants' representation, their response to the charges, and more. It is often difficult to control for all of these factors, and adequately show that it is the judge's differentiated treatment which leads to what appears to be a harsher treatment of minorities.

This study attempts to overcome these difficulties in three ways. First, rather than asking whether there is an ethnic gap in judicial treatment, we ask whether

there are systematic differences across the ethnic group of judges (see also: Abrams et al. 2008). Such a comparison allows us to exploit the random assignment of cases to judges, which solves the selection problem that stems from unobservable case and suspect characteristics (Angrist and Pischke 2009, Shadish et al. 2002). This design enables us to assess whether ethnic disparities in judicial decisions are the result of unobservable case and suspect variables or judicial ethnic bias. The second way this study attempts to overcome the aforementioned methodological pitfalls is by directing the attention to the earliest judicial decisions in the criminal process – the suspect's first bail hearing – usually within twenty-four hours from his/her arrest by the police. At this stage, judges have to make decisions based on very limited information, and those decisions are preceded, and thus potentially affected, by the least number of stages. Lastly, the Israeli bail hearing procedure requires the police to state the number of detention days it asks the court to order when applying for a detention order. This formally recorded number of requested days can serve to quantitatively control for the state's position, which is one of the major factors in any decision in criminal cases. Thus, the setting of the first bail hearing in Israeli courts presents a unique opportunity for examining the effect of ethnicity on *judicial* decisions, beyond its potential effect on the state's position in court.

Our findings show systematic evidence of ethnic in-group bias in the decision to detain or to release on bail. However, no ethnic in-group bias was found in the length of detentions. One inherent limitation of our study is that in

the absence of a known ethnic-neutral baseline, we cannot formally determine whether our findings reflect discrimination against Arab suspects by Jewish judges, discrimination in favor of Arab suspects by Arab judges, or a mixture of both. Still, the study importantly and compellingly shows that judges are not blind to ethnicity. Notably, judges from different ethnic groups vary in their treatment of ethnic groups in a way that is consistent with ethnic in-group bias.

The paper proceeds as follows. Section II reviews the literature on ethnic and racial bias in the criminal legal system, and presents the general approach adopted in this study: (1) addressing ethnic in-group bias; (2) focusing on the earliest judicial decisions of the criminal process; (3) utilizing the Israeli bail hearing process for qualitatively controlling for the state's position. Section III describes the initial bail hearing process in Israel. Section IV presents the research design, data, and measurement. It also includes tests of randomization (of judge's assignment to cases), provides descriptive statistics, and presents the statistical method employed. Section V presents the results of both analyses – detention probability and detention length, and Section VI concludes with a discussion on the empirical and theoretical implications of the findings.

II. Scientific Background

In nearly every Western society, members of disadvantaged ethnic or racial minorities are overrepresented in jails and prisons (Tonry, 1997). While the

reasons for this observable fact are numerous, the possibility that judicial discriminatory practices are partly responsible is troubling. If otherwise equal cases are treated differently due to the ethnic identity of the participants, the legitimacy of the judiciary is undermined. After all, one of the major justifications for an independent judiciary is its ability to protect minorities from prejudice and stereotypes and assure equality before the law.

Thus, the effect of race and ethnicity on judicial decisions in criminal cases has attracted much attention in the literature. Numerous studies have examined whether disadvantaged minorities are treated differently, mostly in sentencing outcomes. While many studies have found that minorities suffer from harsher sentences, the difficult question is not whether these disparities exist, but rather what their causes are. Are these disparities the result of discrimination in the criminal justice system? Do they result from differential criminal profiles of different ethnic groups? And/or do they stem from discrimination outside the criminal justice system? A central problem in providing a full answer to these questions draws on the lack of an available neutral baseline. Under these conditions, any statistical analysis of actual outcomes that employs a null hypothesis of equal outcomes across ethnic groups must assume control of all the relevant control variables. Hence, it is far from clear whether these disparate outcomes are the result of judges' ethnic and racial bias, or whether they result from omitted case characteristics that are correlated with ethnic identity (Abrams et al. 2008). Kleck (1981) analyzed the American research on racial bias in

sentencing and showed that the studies that appeared to find racial discrimination usually failed to adequately control for criminal record and other explanatory factors. More recent studies have led some commentators to conclude that there is little evidence of overt racial bias in sentencing (Klein et al. 1990, Sampson and Lauritsen 1997: 362, Warner 2000). This does not mean that racial discrimination has disappeared; rather that if it exists, it is more subtle and harder to observe (Daly and Tonry, 1997; Bushway and Piehl, 2001).

Other analyses are less conclusive. For example, Chiricos and Crawford (1995) argue that there is sufficient evidence to infer direct racial discrimination in sentencing. Similarly, Spohn's (2000) survey of the literature led her to conclude that "earlier refutations of the discrimination thesis were premature." She reports that a majority of studies suggest that race impacts the incarceration decision, but less than one-quarter of them provide evidence that race affects sentence length. Still, these conclusions do not rule out the possibility that it is prosecutors, defense attorneys or other non-judicial actors who are responsible for the biased result.

Research findings in other Western countries are equally ambiguous. It seems that wherever the issue has been thoroughly investigated, "neither proponents of the claim that bias is a cause of disparities nor proponents of the claim that bias has no influence will find empirical evidence to prove their claims" (Tonry, 1997). The research on sentencing of Jews and Arabs in the Israeli criminal justice system is no exception. Attempts to find judicial bias in

sentencing have achieved mixed results, with several studies having found harsher treatment of Arabs (Cohen and Palmor, 1985; Rattner and Fishman 1998, Fishman et al. 2006), while others have shown that evidence of differential treatment virtually disappears when pretrial detention and prosecutor's position in the sentencing hearings were controlled for (Hassin and Kremnitzer 1993; Haj Yahia, Rahav & Teichman, 1994).

Why is it, despite the numerous studies that have been conducted, still so hard to find a clear view regarding the existence of racial and ethnic discrimination in judicial sentencing? The difficulty is that race and ethnicity are often correlated with other factors which have been found to significantly influence sentences, such as criminal record (Kleck, 1981), pretrial detention (Demuth and Steffensmeier, 2004), unemployment (Chiricos and Bales, 1991), court appointed counseling (Holmes et al, 1996), crime type (Tonry¸ 1995), aggravated circumstances (Kleck, 1981) etc.; and when these factors are controlled for, the independent effect of race loses all or most of its explanatory power in some of the studies. When so many racially correlated factors influence sentencing, it is statistically difficult to determine whether race in itself influences sentencing directly. Determining whether these factors receive unjustified weight because they mainly affect minorities is even harder.

Even when race and ethnicity has a direct effect on sentencing, it is still questionable whether discrimination should be attributed to the sentencing judges. As Bushway and Piehl (2001) explain, the sentence outcome is the result of the

decision of multiple actors interacting in one system. Sentence disparity might be the result of differentiated treatment by the police, prosecutors (Humphrey and Fogarty 1987) probation officers (Bridges and Steen 1998; Ben-Zion and Palmor 1985) defense attorneys and other participants. In fact, it is likely that prosecutors have a more direct impact on racial disparities than judges (Humphrey and Fogarty 1987; Davis 1998). Their sentencing recommendations have a substantial impact (Englich and Mussweiler 2001; Englich, Mussweiler and Strack 2005; Kremnitzer and Hassin 1993), but even more importantly, prosecutors have direct control over the outcome through plea bargaining, which is the method of disposition in the vast majority of cases. If prosecutors are offering white defendants more lenient plea bargains than non-whites, the sentence disparity should be attributed to prosecutors, rather than to courts (Humphrey and Fogarty, 1987). The existing research on sentencing has made virtually no attempt to distinguish between prosecutors' and judges' differentiated treatment of minorities, and thus even if the studies that support the discrimination thesis are found more convincing than the contradicting studies, one cannot convincingly argue that this discrimination is the judges' fault.

One way of isolating judicial ethnic bias is by examining decisions of judges from different ethnic groups. While many studies have examined the effects of defendants' race on judicial decision making, few studies have examined both the judges' and the defendants' race. Most of these studies did not find that defendants' race had a different effect on the decisions of black and white judges.

Steffensmeier and Britt (2001), and Uhlman (1978) found that black judges sentence both black and white defendants somewhat more harshly than white judges, but the defendants' race had similar effects on white and black judges. Similarly, two studies of Spohn (1990a, 1990b) found no significant effect of judges' race on sentencing of white and black defendants. On the one hand, Welch et al. (1988) found that black judges treated black defendants more equally compared to white judges in the decision to incarcerate, and Johnson (2006) found that minority judges were less likely to incarcerate black and Hispanic offenders, but found no difference in the length of sentences. Price and Wolfers (2007), Parsons et al. (2007) and Larsen, Price and Wolfers (2008) found in-group racial bias in the decisions of officials in basketball and baseball leagues. Fishman et al. (2006) found that in Israeli courts Jewish judges were less likely to impose a prison sentence on Jewish defendants (compared to Arab defendants), whereas the ethnic identity of the defendants did not seem to play a role in the decisions of Arab judges.[1] In contrast, Zussman and Shayo (2010) found evidence of ethnic in-group bias in Israeli small claims court's rulings. Abrams et al.'s (2008) study of sentencing decisions in Cook County, Illinois found statistically significant between-judge variation in incarceration rates, although not in sentence lengths. All in all, the limited literature on the interaction between judges' and defendants' racial and ethnic background provides mixed evidence as to whether the harsher

---

[1] The unique importance of Fishman et al.'s (2006) study is in the inclusion of the ethnic identity of the victim of the crime, and its interactions with the ethnic identity of perpetrators and judges,

treatment of minorities in court results from judges' in-group bias.

This study differs from earlier works in three main respects, which together create a uniquely effective method of isolating judicial bias from other ethnicity related factors. First, in line with Abrams et al. (2008), rather than asking whether there is an ethnic gap in judicial treatment, we ask whether there are systematic differences across the ethnic group of judges. Such a comparison allows us to exploit the random assignment of cases to judges, which solves the selection problem that stems from unobservable case and suspect characteristics (Angrist and Pischke 2009, Shadish et al. 2002). This design enables us to assess whether ethnic disparities in judicial decisions are the result of unobservable case and suspect variables, or of judicial ethnic bias. If the former is the case, we should expect an overall difference in judicial treatment by suspects' ethnic group, but not a systematic variation in this difference across ethnic groups of judges, as random assignment ensures that each group receives the same case-mix. However, if Arab and Jewish judges are found to systematically treat Arab and Jewish suspects differently, this would provide strong evidence for judicial in-group bias.

The second way this research attempts to overcome the aforementioned methodological pitfalls is by directing our attention to the earliest judicial decisions in the criminal process – the suspect's first bail hearing – usually within twenty-four hours of his/her arrest by the police. At this stage, judges have to

---

but this is less relevant to the current study. Note also that their interest in the effects of the

make decisions based on very limited information, and those decisions are preceded, and thus potentially affected by the least number of stages. Each stage in the criminal process influences the result of the next stage. For example, it is well established that pretrial detention increases the probability of conviction and of imprisonment sentence (Rankin, 1964; Clarke & Kurtz, 1983; Frazier and Bishop, 1985; Williams, 2003). Similarly, the type of plea, the manner in which witnesses are questioned, the prosecution's positions at each step, and many other factors, can be correlated with race and ethnicity and thus affect the decisions in the following stages. By examining the practice of judges at the very beginning of the process rather than at the end, the effects of previous stages are minimized.

Focusing on early detention decisions also serves a theoretical consideration. It is often suggested that lack of information drive judges to rely on attributions linked to defendants' race as a mechanism to reduce uncertainty (Albonetti 1991; Bridges and Steen 1998; Schlesinger 2005; Demuth and Steffensmeier 2004). At the first bail hearing, which takes place only hours after the initial arrest, the judge has very limited information on the suspect and the offence. Thus, in this stage ethnicity might have a different, more pronounced role (Demuth 2003). Only a few studies have examined the effect of race on pretrial detention decisions, without an attempt to isolate initial decisions and later ones. Like sentencing research, these studies do not clearly support or refute the discrimination hypothesis. Several studies found indication of discrimination (Ayres and

---

victim's identity has led their study to look only at violent offences (*ibid*: p. 74).

Waldfogel 1994; Bynum 1982, Patterson and Lynch 1991; Petee 1994). Yet even in this relatively initial stage, some ethnically or racially correlated factors can influence the result, and thus the result is heavily influenced by the variables that were unavailable to the researchers.[2] The approach adopted in this study thus reduces these potential, unobservable case variables by focusing on pretrial detention decisions, and overcomes the remaining problem by looking at the variation across the ethnic group of randomly assigned judges.

Lastly, the Israeli bail hearing procedure allows us to control for the overall state's position. When applying for a detention order, the police are required to state the number of detention days it requests the court to order. This written number of requested days serves in our analyses to quantitatively control for the state's position, which is one of the major factors in any decision in criminal cases. To the best of our knowledge, none of the existing pretrial detention studies have used such a control. Disregarding the state's position in bail hearings is as problematic as disregarding plea bargaining in sentencing. To the extent that the state's position is systematically associated with ethnic characteristics, any potential disparity in the result might be attributed to the state and not the judge. Obviously, the state's agents might also be biased. Yet controlling for the state's position in the analyses enables us to infer *judicial* bias from racially differential treatment in the outcome of the judicial decision, separately from other sources of

---

[2] Ayres and Waldfogel (1994) concede that although their analysis controls for many variables they might still be missing one or more omitted variables that may account for the disparate outcome.

bias in the criminal justice system. Thus, the setting of the first bail hearing in Israeli courts presents a unique opportunity for examining the effect of ethnicity on judicial behavior, isolating judicial bias from other relevant and irrelevant factors already considered by the state.

III. THE INITIAL BAIL HEARING IN ISRAEL

The Israeli criminal procedure provides an excellent laboratory for examining the effect of ethnicity on pretrial decisions. First, it operates within a divided society in which criminal suspects and judges include both Arabs and Jews. Second, the pre-trial detention hearing is the earliest stage in the criminal process in which judicial decision is involved. Third, the prosecution's position is quantitatively presented in the hearing by the requested length of detention, and finally, judges and public defenders are randomly allocated for the different cases.

In this system Arabs play much the same role as other disadvantaged minorities elsewhere. While there are internal ethnic or faith-based divisions within the Jewish and the Arab groups, members of each group usually see themselves as part of one single ethnic and national identity. Since the establishment of Israel in 1948, a fragile and contested coexistence between the Jewish majority and the Arab minority has gradually been developed (for a recent review of this relationship see: Samooha 2005). Hostile undercurrents in these inter-group relations are further augmented by the ongoing Israeli-Arab conflict.

It is therefore not surprising that these social realities permeate the courts. Indeed, a sizable number of studies provide consistent evidence that there is a wide disparity in sentencing of Jews and Arabs in the Israeli justice system (Cohen 1985, Cohen and Palmor 1985, Hassin and Kremnitzer 1988, Kretzmer 1990, Rattner and Fishman 1998, Fishman et al. 2006). Like African-Americans in the United States or minority groups in the United Kingdom, Arabs are overrepresented in prisons; they suffer from negative stereotypes and are more likely to be socially associated with severe crimes (Fishman et al 1987). However, as in other countries, a strand of the literature suggests that when controlling for other major factors such as criminal record, pretrial detention and the prosecutors' and sentencing requests, little evidence of judicial discrimination remains (Hassin and Kremnitzer 1993; Haj Yahia et al. 1994).

As in many other jurisdictions, a suspect in Israel has to be brought before a judge within twenty-four hours of the arrest, unless a police officer has decided to release him/her earlier. The court can then detain the suspect for up to 15 days, though it usually orders a much shorter detention period. In the case that the suspect is detained, the police can ask the court to order the continuation for up to 15 days from time to time, as long as the total detention period does not exceed 30 days prior to the issuance of an indictment. After an indictment is issued, the case is handled by the prosecution, which can either agree to release the defendant on bail or request the court to remand the defendant for the full duration of the trial.

In the pre-trial hearing, the state is represented by a police detective (a

prosecutor represents the state only after an indictment is issued). In the application, the police officer asks the court either to release the suspect on bail or to detain him/her. In the latter case the officer also states how many days of detention are requested. The court reads the request and hears the suspect or his/her attorney, and then decides whether to detain the suspect, and if so, for how long. The factors considered at this stage are the suspect's dangerousness, the risk of flight, the need to prevent disruption of proceedings and investigation needs.

The suspect is brought before a magistrates' court judge, who has been designated in advanced to sit on all the bail hearings on that specific day. The public defender's office sends a lawyer to represent all the suspects that do not hire private lawyers.[3] Hence, judges and public defenders are randomly allocated for the different cases. With this in mind, we turn to the method of our study.

IV.   RESEARCH DESIGN

Data was drawn from the entire list of pretrial bail hearings taken on Fridays and Saturdays during 2004 at the Haifa, Nazareth and Acre magistrate courts ($N =$ 1852).[4] Fridays and Saturdays were selected for two reasons. First, while during working days particular judges handle this type of hearing, all the judges serve on pre-arranged rotation duty in these hearings over the weekend. This provides a

---

[3] At this stage suspects have a right to public representation only if they are found unable to hire a lawyer of their own, but since the hearing takes place soon after the arrest it is impossible to inquire about a suspect's financial situation and thus, in practice, every suspect is offered a public defender.

"naturally occurring" process of random allocation of judges to cases. Second, during these days, nearly all hearings deal with first detention decisions, i.e., regarding suspects that have first been apprehended during the previous twenty-four hours, while a larger proportion of continuation of detention decisions typifies other days. The number of judges in the sample is 59, out of whom 15 are Arabs and 44 are Jews. It should be noted that the three courts chosen include the majority (78.9%) of the Arab magistrate judges in Israel.[5] Table 1 provides initial evidence consistent with a random assignment of judges to cases, showing that none of the case variables are related to the ethnic identity of the judge. Appendix A also shows that none of these variables have any power in explaining the assignment of judges (Arabs or Jews) to cases.

TABLE ONE – ABOUT HERE PLEASE

The analysis included two stages. In the first stage, the judicial decision was analyzed as a dichotomous variable – release or detain. The second stage analyzed the decision as a count variable based on the length of detention in days, from one up to fifteen. Only those cases in which agreement was not reached among the parties regarding release or the length of detention, thereby resolved solely by a

---

[4] The weekend in Israel includes Friday and Saturday, while Sunday is a work-day.

[5] A total of nineteen Arab magistrate judges (excluding labor and family courts, which do not handle detention decisions) (based on the official Judicial Authority website: www.court.gov.il).

judicial decision, were included in the analyses.[6] Additionally, cases in which the police request was for bail (zero detention days) were omitted from the dataset, as in these cases the release was not in dispute.[7] Thus the effective sample used for the analyses included 1230 decisions.

We coded judges as Arabs or Jews based on their names and biographic information found on the Israeli Court System's official website. The same sources were used to code judges' gender.[8] Suspects were coded as Arabs or Jews based on their names as they appear in the protocol of the hearing. Other variables coded were the police request (1-15 days), whether the hearing took place on Friday or Saturday, and severity of the most severe offense mentioned in the detention procedure. Coding of offence severity followed a four scale ordinal variable which represented the maximum statutory penalty for the offence – (1) less than three years (misdemeanors under Israeli law); (2) three to seven years; (3) seven to fifteen years; (4) over fifteen years. Based on earlier research (Hasin and Kremnitzer 1993), a dummy variable indicating the identity of the court – Nazareth and Acre – was included (neither indication represents cases from the

---

[6] Agreed decisions are the result of three possible processes. One third of the agreed decisions are determined before the hearing, and the judge only approves the parties' agreement. The rest of the agreements are formed when the judge sends the parties to deliberate (typically in the corridor) and later approves their agreement, or suggests a possible outcome of the agreement and suggests to the parties to reach a proximate agreement. The exact proportions of these two situations are unknown. Clearly, consensual decisions follow a different dynamic from judicial decisions, and they are studied more closely by the authors elsewhere.

[7] Indeed in *all* of these cases suspects were not detained by judges.

[8] www.court.gov.il. The website also includes pictures of the individual judges. We can safely say that the ethnic identification of the judges is highly reliable. In the few cases in which we had any doubt, we verified the identity by contacting people who knew the judge in question. In one particular case we found that a certain judge was the son of an Arab father and a Jewish mother (which is quite rare in Israeli society). His decisions (only 9) were excluded from the dataset.

Haifa court). The type of representation provided in each case was coded (private, public defender, or unrepresented), and the type of the offense was coded following eight categories. In some of the cases, the police investigation is concluded by the time of the bail hearing. In these cases, detention is possible only if the prosecutor declares that the state intends to issue an indictment and requests that the defendant be detained for the full duration of the proceedings. Since the official legal justification for detention in these cases is slightly different, a variable for such prosecutorial declaration was also included in the analysis. Table 2 provides a list of the variables used in our analysis, as well as a comparison of the mean values between Arab and Jewish suspects. We do not have data on the defendants' prior record as this information was not included in the files. However, the control for the state's position in the hearings is likely to incorporate most of the relevant factors for the bail hearing, including prior record. More importantly, the random assignment of cases to judges eliminates the risk that prior criminal record will be able to account for the behavioral differences between Jewish and Arab judges.[9]

Table 2 shows that the police ask for slightly more detention days for Arab suspects, compared to Jewish suspects (5.106 vs. 4.925 days); Arab suspects do not differ from Jewish suspects in the extreme ranges of offence severity, yet they appear to be less likely to be detained by the police for mid-low severity offences

---

[9] The omission of this variable (like any other) may act to reduce the explained variance of the in the analyses, but in the context of an experimental design this does not derogate from reliability of

compared to Jewish suspects (.429 vs. .484), and more so for mid-high severity offences by a similar margin (.136 vs. .085). Unsurprisingly, Arab suspects are much more involved in immigration offences (.082 vs. .006)[10]; but are less frequently arrested for domestic violence (.167 vs. .230) and drug related offences (.095 vs. .139) compared to Jewish suspects. Arab suspects are less frequently represented by the public defender's office (.546 vs. .719), and significantly exceed the proportion of Jewish suspects in the magistrate court of Acre (.403 vs. .212). As to the dependent variables, Arab suspects have a lower proportion of releases compared to Jewish suspects (.234 vs. .300), yet they do not appear to differ in their mean length of detention (2.890 days vs. 2.816 days).

TABLE TWO – ABOUT HERE PLEASE

Statistical Analysis

In order to account for the data structure, in which judicial decisions are nested within individual judges, the following analyses use hierarchical modeling procedures, designed to account for the nested nature of detention decision within judges.[11] The models examining release versus detention were estimated with logistic and multilevel logistic regressions.

Detention length provides a number of complications that should be accounted for in the analysis. One may consider detention length as including

---

the estimate of the randomly assigned causal effect – in our case the ethnic identity of the judge (see: King 1991, pp. 1050-1).

[10] This offence is almost entirely committed by Arab Palestinians from the West Bank and Gaza.

[11] See also: Johnson 2006. For a concise review of the method see: Steenbergen and Jones 2002, and the references therein.

releases as 'zero-value detentions', or exclude releases from the range of detention length as categorically distinct. We tend to subscribe to the later conceptualization. Moreover, even if one considers release as a zero-value case of detention, this value occupies the lower limit of the range, potentially misrepresenting cases with theoretically varying values of underlying judicial valuation. For example, such cases may include those cases that only just fall short of sanctioning detention; those that clearly should be released; and model citizens that should not have been arrested by the police in the first place. Thus including zero detention cases as another value of the detention variable would violate the validity of lower value of this measure, and potentially inflate the number of zeros in such a variable. More formally, a strictly positive (excluding 'releases') variable of detention length is expected to follow a Poisson distribution,[12] and one that includes zero detention length is expected to follow a zero-inflated Poisson distribution.[13] Figure One graphically presents the histograms of detention length against the corresponding Poisson distributions. The two graphs confirm the theoretical expectations. It appears that the strictly positive variable approximates a Poisson distribution, while the variable that includes 'releases' as zero-value detentions exhibits an excessive number of

---

[12] The mean detention length of this variable is 2.85 days and the variance is 1.716; and the overdispersion parameter alpha is insignificantly different from zero ($p = 1.00$).

[13] The mean detention length of this variable is 2.08 days and the variance is 2.86; and the overdispersion parameter alpha is insignificantly different from zero ($p = 1.00$).

zeros.[14] For these theoretical and empirical reasons, our main analyses have used the strictly positive variable of detention length, and the models examining detention length were estimated with Poisson and multilevel Poisson regressions.

Notwithstanding this, three sets of robustness tests follow the main analyses. The first is intended to assess the robustness of the findings regarding detention length under alternative distributional assumptions. It includes analyses of detention length with linear and multilevel linear regressions. The second and third sets of analyses gauge the possibility of selection bias in estimating detention duration as a result of including only those cases in which suspects were detained. These employ the zero-inflated Poisson regression (ZIP), and the Heckman two-step selection model. The detailed analyses are reported in the following section.

FIGURE ONE – ABOUT HERE PLEASE

V.   RESULTS

The effect of in-group on the probability of release

Table Three presents the results of four logistic regression analyses with the log-odds of release as dependent variable. The second column of each model reports the odds-ratios after exponential transformations of the regression coefficients. We begin with a barebones specification in model (1), which is a basic difference-in-difference specification. Models (2) and (3) add controls for

---

[14] Indeed goodness-of-fit $\chi 2$ for the full Poisson regression model with strictly positive variable suggested a good fit (292.18, $p = 1.00$) , while this was not the case for the variable that includes

observable judge, court and case variables.[15] Model (4) presents a multilevel analysis thereby controlling for unobserved variation across individual judges.

TABLE THREE– ABOUT HERE PLEASE

The primary focus of our analysis is the interaction term *Arab judge×Arab suspect$_{ij}$* which indicates whether the effect of the ethnic identity of the suspect on the probability of release is different between Arab and Jewish judges. Across all of these specifications, we find that this difference is significant and positive. Given that the assignment of judges to cases is random, this statistical association enables us to infer that ethnic in-group increases the likelihood of release. In order to demonstrate this finding Figure Two presents the predicted probabilities of release for Arab and Jewish suspects across Arab and Jewish judges, holding all other variables at their modal values.[16] The significant difference between the effect of the suspect's ethnic identity on Arab and Jewish judges, inferred from the significant interaction term, is graphically depicted by the higher probability of release of Arab suspects for Arab judges, compared to the higher probability of release of Jewish suspects for Jewish judges. At the modal situation, Arab

releases (1649.33, $p < .001$).

[15] Initial VIF tests of multicolinearity have shown a mean VIF of 2.20, yet three variables exceeded VIF value of 4.0, with 'property offences' with the highest value of 5.06. Excluding 'property offences' from the analysis has led to a mean VIF of 1.59 and a maximum level of 2.77 (High severity). In order to avoid multicolinearity and given that 'property offences' were not found to have a significant association with any of the dependent variables, nor with the ethnic identity of suspects and judges; it was excluded from the final analyses.

suspects are 6.3% more likely than Jewish suspects to be released by an Arab judge, while Jewish suspects are 10.4% more likely than Arab suspects to be released by a Jewish judge.

FIGURE TWO – ABOUT HERE PLEASE

It should be noted that the significant *Arab judge×Arab suspect$_{ij}$* interaction term refers to the difference between the treatment of Arab and Jewish suspects in the comparison *between* Arab and Jewish judges (as the null hypothesis is: no difference in the treatment across judges' groups). However, analyses of release rates *within* each group of judges involve comparisons between the actual disparate rates of release (e.g. 24.1% and 30.4%) and an *equal* release rate of Arab and Jewish suspects. Despite our attempt to control for many case, judge, and court variables, the true ethnically neutral baseline of release probability is not known and not necessarily equal across suspects' ethnic group. Thus within-judges'-group comparisons cannot be used to infer whether Jewish, Arab or both groups of judges discriminate. Yet regardless of the ethnically neutral baseline of release probability, the differences in the treatment of Arab and Jewish suspects by the two groups of judges clearly shows that ethnic identity plays a role in these decisions, in a pattern that is consistent with in-group bias.

As could be expected, the police request was found to be negatively

---

[16] Male judge, police request = 5 days, at the Haifa Magistrate Court, mid-low severity of offence,

associated with the probability of release.[17] The severity of the offence is also negatively associated with release probability (45.0%, 37.8%, 28.2%, and 26.6%, for low, mid-low, mid-high and high severity levels, respectively).[18] Interestingly, male judges are more likely to release suspects on bail (37.8% vs. 28.4%). Suspects in the Nazareth Magistrate Court are less likely to be released (19.0% vs. 30.2% at Acre, and 37.8% at Haifa), and so are suspects who are represented by a public defender (37.8%) or unrepresented (38.0%) compared to those represented by a private attorney (46.9%). Within the types of offences, domestic violence and traffic violations were found to have a significant (positive) association with the probability of release (49.2%, and 63.3%, respectively, compared to the modal (violent) offence – 37.8%). It appears that accounting for the multilevel structure of the data resulted in some changes in the regression estimates. This result is consistent with the significant likelihood-ratio test of rho ($p = .015$), and suggests that inter-judge variance in release decisions is substantive and statistically significant.

The effect of in-group on the length of detention

Table Four presents the results of four Poisson regression analyses with detention length as dependent variable. This analysis was based on the subset of

---

violent offence, representation by a public defender, on a Friday.

[17] The predicted probability of release for the case in which the police request four detention days and other variables are at their modal value is 48.7%. An increase of one day in the police's request reduces the probability of release to 37.8%.

decisions to detain (no releases) ($N = 897$). In this analysis we also begin with a barebones specification in model (5). Models (6) and (7) add controls for observable judge, court and case variables, and model (8) presents a two level analysis thereby controlling for unobserved variation across individual judges.

No significant *Arab judge×Arab suspect$_{ij}$* interaction was found for any of the models nor any main effect of the judge or suspect's ethnicity on the length of detention. The police request was found to be significantly and positively associated with the length of the detention ordered, i.e., every additional day of detention requested by the police resulted in an increase of 12.7% in detention length (e.g. an increase from 5 to 6 days requested by the police is associated with and increase of 0.635 days of detention). While male judges were found to have a greater propensity to release suspects (compared to female judges), when they did decide to detain, they tended to order 10.3% longer detentions. Drugs related offences were found to result in 20.6% longer detentions, and detentions on Saturdays were 13.6% shorter than those ordered on Fridays. It appears that accounting for the multilevel structure of the data resulted in negligible differences in the regression estimates. This result is consistent with the insignificant likelihood-ratio test of alpha ($p = .499$), and suggests that inter-judge variance in determining detention length is statistically insignificant, and practically negligible.

---

[18] This and the following comparisons were computed when all other variables were held at their

modal values.

Robustness tests

In order to assess the robustness of the findings regarding detention length these relationships were also estimated with linear and multilevel linear regressions, which are reported in appendix B. Again, we begin with a barebones specification in model (9), adding controls for observable judge, court and case variables in models (10) and (11), and presenting a two level linear analysis in model (12). The findings were generally consistent with those found in the Poisson models.[19] More specifically, no significant *Arab judge×Arab suspect$_{ij}$* interaction was found for any of the models nor any main effect of the judge or suspect's ethnicity on the length of detention.

The second and third sets of analyses are intended to assess the possibility of selection bias in estimating detention duration as a result of including only those cases in which suspects were detained. First the zero-inflated Poisson regression (ZIP) was estimated using the entire set of decisions. The count model for detention length included the dependent variables regarding the ethnicity of the judges and suspects, judge gender, the police's request, offence type, and hearing day (Friday/Saturday), omitting variables that were not found to be significantly associated with detention length in model (8). The model predicting release included the dependent variables concerning the court's location, offense severity, and representation category. The analysis is reported in the left column in

---

[19] One exception is the findings of a significant association between 'Violent offence' and a positive increase in detention length, which was found in the multilevel linear regression analysis (model 12) and was not found in the Poisson analyses.

appendix C. The significant Voung test result ($p < .001$) indicates that the ZIP model fits this data better than standard Poisson. The estimates in the count model are consistent with those found in models (7) and (8), and the estimates predicting release conform to those found in model (4) (with the exception of the 'prosecutor's declaration).

Next the Heckman two-step selection model was estimated using the entire set of decisions. The allocation of dependent variables to the two models is identical to the one used in the ZIP analysis. This separation between the dependent variables follows Puhani (2000) in order to decrease multicollinearity in the data.[20] The results of this model are reported in the center column in appendix C. Finally, the inverse Mills ratio calculated by the Heckman two-step model is included as an additional dependent variable in the multilevel Poisson regression (see: Green 1994). Both linear and multilevel Poisson models with Heckman two-step selection correction conform to the results of models (7) and (8). We thus conclude that our findings in the analysis of the effects of ethnic in-group on detention length appear to be robust.

TABLE FOUR – ABOUT HERE PLEASE

VI. DISCUSSION

This is the first study to utilize a natural randomized experiment to assess the

---

[20] The full model R^2 of the regression of the Inverse Mills ration on the regressors of detention length is .67; mean VIF = 1.61, maximum VIF = 2.80.

effect of suspects' ethnic identity on pre-trial detention decisions across the ethnic identity of the deciding judges. We have found support for in-group bias in the likelihood of release, while we have found no support for such bias in the decision on the length of detention.

Comparing these findings to those of previous studies in the context of sentencing in the US presents a mixed result. They are contrary to the findings of Uhlman (1978), Spohn (1990b), and Steffensmeier and Britt (2001), but conform to the findings of Welch et al. (1988), and particularly Johnson (2006) who found that minority judges are less likely to incarcerate black and Hispanic offenders, yet found no difference in the length of sentences. Our findings also coincide with Abrams et al. (2008) who found statistically significant between-judge variation (yet not specifically across judge's ethnic identity) in incarceration rates, although not in sentence lengths. In the Israeli sentencing context our findings agree with those of Fishman et al. (2006) who found that Jewish judges were less likely to impose a prison sentence on Jewish defendants (compared to Arab defendants), whereas the ethnic identity of the defendants did not seem to play a role in the decisions of Arab judges. This inconsistent level of agreement between our findings and previous studies is not surprising given the conflicting results in the literature. We would, however, note the fact that both American and Israeli studies are included in those studies that are similar in their findings to ours, suggesting that these results are not limited to the Israeli context.

Judicial ethnic in-group bias as implicit bias

How should we account for the evidence for ethnic in-group bias in the decision to release or detain, and the lack of such bias in the decision on the length of detention, which also mirrors similar findings in the context of sentencing (Johnson 2006, Abrams et al. 2008)? The main difference between these two decisions is in the characteristic of the response alternatives – release or detain, as opposed to a range of detention lengths of from one to fifteen days. Judges are required to consider a series of varied details pertaining to the criminal behavior, its particular circumstances and consequences, future potential risks that the suspect poses if released, and a range of legal rules that apply to all these facts. While the first decision (detention/release) requires them to transform their overall impression from this elaborate set of information to one of two possible responses, the second one (detention duration) offers a more varied range of response choice.

Under these conditions one possible interpretation of the findings is that judicial ethnic in-group bias is an implicit bias. An implicit attitude is defined as an unconscious association between a target (e.g., an Arab suspect) and a given attribute (Bertrand et al. 2005, Greenwald and Krieger 2006). Such attitudes exist whether they are aligned or not with a person's explicit (conscious) attitudes, and thus affect judgment and behavior in unintentional and often unconscious ways. Rachlinski et al. (2009) have found that judges harbor similar implicit racial biases to those found in the general population, yet they can suppress their

implicit racial bias from influencing their decisions in some circumstances. While both implicit and explicit attitudes are at work in determining judgment, choice and behavior, it appears that implicit bias may have greater influence in less-controllable behavioral outcomes. Notably, three conditions are conducive to the influence of implicit bias: inattentiveness to task, time pressure or other cognitive load, and task ambiguity (Bertrand et al. 2005).[21] While the first two conditions are equal across the two judicial choices under question here, the third appears to be relevant.

Ambiguous tasks are those that do not provide a simple formula for their completion.[22] Deciding between releasing and detaining a suspect based on the information presented before a judge in a bail hearing forces her to make a choice between two very distinct outcomes even when the relevant considerations suggest an equivocal situation. Such a task thus imposes (at least in some cases) an implicit "rounding up" or "down" of case information evaluation in order to make a choice. This implicit "rounding" of evaluation is unguided and thus ambiguous. On the other hand, the decision on the length of a detention provides a range of ranked responses from 1 to 15 days, which facilitates matching a measured response to the case's specific circumstance. We therefore suggest that the finding of ethnic in-group bias in the decision to release (/detain) but not in

---

[21] One possible explanation for this effect of ambiguity on the expression of implicit bias is the process of 'attribution substitution', in which people, confronted with a difficult question, often answer an easier one instead, usually without being aware of the substitution (Kahneman and Frederick 2002, Jolls and Sunstein 2006).

the length of detentions can be understood as resulting from implicit (rather than explicit) bias. This bias is variably manifested in the judicial decisions based on their level of task ambiguity.[23] Further research is needed in order to assess this potential explanation, both in the context of detentions, and in sentencing.

From in-group bias to ethnic discrimination

The difficulties in studying ethnic discrimination in the legal system, discussed in the introduction, are mainly due to potential unobserved variables which entail uncertainty as to the true neutral baseline for treating members of different ethnic groups. Looking at *absolute* differences in the treatment of ethnic group members requires us to control all relevant variables in order to be able to equate the null hypothesis of 'zero difference' with the true neutral treatment. Indeed, differences in controls in previous studies, in Israel as in other countries, have resulted in mixed findings. Alternatively, replacing "ethnic bias" with "ethnic in-group bias" as the concept under investigation diverts our attention to *relative* difference in the (apparent) differential treatment, across the ethnic identity of judges, taking advantage of the random allocation of judges to cases, and provides an opportunity to avoid this problem. In this study, the null hypothesis is *not* that the rate of release or length of detention of members in the

---

[22] One notable example provided in the literature is the placing of résumés on the "yes" or "no" pile (Bertrand and Mullainathan 2004).
[23] It should be noted here that while the multilevel modeling in the analysis of detention/release was found to be substantively important, this is not the case for the analysis of detention length, suggesting that individual judges differ more with regard to the former decision.

two ethnic groups is equal, but rather that the treatment of these two groups (equal or not) is the same across the ethnic group of judges. Such a null hypothesis conforms to the normative expectations from judges in a system which upholds equality before the law, and its manifestation as zero difference between judges is the proper empirical expectation when judges are randomly assigned to cases.[24]

While we have statistically established that ethnic identity makes a difference in judicial decisions, one limitation of our study is that in the absence of a known ethnically-neutral baseline, we cannot formally determine whether our findings reflect discrimination against Arab suspects by Jewish judges, discrimination in favor of Arab suspects by Arab judges, or a mixture of both.

Still, in-group bias is not only a methodological bridge across the pitfalls of unobserved case characteristics. It is also theoretically important as it provides a mechanism for understanding discrimination of minorities (Greenwald and Grieger 2006). Since most judges belong to the ethnic and racial majority group in their jurisdiction, it is likely that the aggregate effect of in-group bias results in an overall discrimination of minorities in a national judicial system. In other words, systemic ethnic and racial disparities against minorities may simply result from many more majority judges practicing in-group bias compared to minority judges.

---

[24] It should be noted that another approach to uncovering judicial biases has been recently offered by Rachlinski et al. (2009), by directly testing for implicit racial bias and estimating its effect on simulated legal situations. The researchers administered the Implicit Association Test (IAT) to a sample of American trial judges, and then asked them to adjudicate a number of hypothetical cases. Their findings show that judges harbor implicit racial biases that are similar to those prevalent in the general US population, and that these biases can influence their judgment when they are unaware of a need to monitor their decisions for racial bias. Yet when judges are aware of this, they appear to be able to suppress that bias.

Further research is required to establish the characteristics of in-group bias in other legal systems, in other legal procedures, the determinants of these biases and their psychological mechanisms.

The existence of ethnic in-group bias in the early stages of the criminal process may also explain the difficulty to establish the existence of ethnic discrimination in sentencing, since these early decisions, such as those regarding detentions of suspects or defendants can indirectly affect the entire process. Thus, if Arabs are more likely to be detained compared to Jews – as a result of translating individual level in-group bias into aggregate level discrimination against minorities – and if the detention influences the severity of the final sentence, then research on sentencing may attribute the increased severity of sentences for Arabs to their greater likelihood of prior detention rather than to their being Arabs. It is possible that this is the reason that studies that did not control for prior detentions have found discrimination in sentencing among Arabs and Jews, while studies that controlled for it did not.[25] As noted earlier, the concentration on an early stage of the criminal process allows us to remove a considerable amount of "noise" from the analysis, and the potential to quantitatively control for the state's position by relying on the police request enables us to address this important factor. To conclude, our findings robustly support the existence of judicial ethnic in-group bias in the decision to detain or

---

[25] Compare Rattner and Fishmann (1998) and Fishmann et al. (2006), which found ethnic discrimination without controlling for prior detentions; to Hasin and Kremnitzer (1993) which

release, yet not in the decision on the length of detention.

References

Abrams, David, Marianne Bertrand, and Sendhil Mullainathan. 2008. Do Judges Vary in their Treatment of Race? *mimeo*, University of Chicago.

Albonetti, C.A. 1991. An Integration of Theories to Explain Judicial Discretion. *Social Problems* 38: 247-266.

Angrist, Joshua D. and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton and Oxford: Princeton University Press.

Ayres, I. and J. Waldfogel. 1994. A market test for race discrimination in bail setting. *Stanford Law Review* 46: 987–1047.

Berk, Richard A. 1983. An Introduction to Sample Selection Bias in Sociological Data, *American Sociological Review* 48(3): 386-398.

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. 2005. Implicit Discrimination. *American Economic Review* 95(2): 94-98

Bertrand, Marianne and Sendhil Mullainathan. 2004. Are Emily and Greg More Employable Than Lakisha and Jamal? *American Economic Review*, 94(4): 991–1013.

Bridges, G.S. and S. Steen. 1998. Racial Disparities in Official Assessments of

found that prior detentions accounted for a considerable part of the variance in sentencing, and did

Juvenile Offenders: Attributional Stereotypes as Mediating Mechanisms. *American Sociological Review* 63: 554 .

Bushway, S.D. and A.M. Piehl. 2001. Judging judicial discretion: Legal factors and racial discrimination in sentencing. *Law and Society Review* 35: 733

Bynum, T. 1982. Release on recognizance: Substantive or superficial reform? *Criminology* 20: 67–82.

Chiricos, T.G. and C. Crawford. 1995. Race and imprisonment: A contextual assessment of the evidence. In Ethnicity, race, and crime, Albany: State University of New York Press.

Chiricos, T.G. and W.D. Bales. 1991. Unemployment and Punishment: An Empirical Assessment. *Criminology* 29: 701-24.

Clarke, S.H. and S.T. Kurtz. 1983. The importance of interim decisions to felony trial court dispositions. *Journal of Criminal Law and Criminology* 74: 475–518.

Cohen, B. Z. 1985. Judicial discretion and sentencing disparity in adult felony courts in Israel. *Journal of Criminal Justice* 13: 99–115.

Cohen, B. Z. and S. Palmor. 1985. The impact of ethnicity on probation recommendations and dispositions in Israel. *Justice Quarterly* 2: 197-212.

Daly, K. and M. Tonry. 1997. Gender, Race and Sentencing. *Crime and Justice* 22: 201.

Davis, A.J. 1998. Prosecution and Race: The Power and Privilege of Discretion.

not find evidence for ethnic discrimination.

Fordham Law Review 67: 13.

Demuth, S. 2003. Racial and ethnic differences in pretrial release decisions and outcomes: A comparison of Hispanic, black, and white felony arrestees. *Criminology* 41: 873–908.

Demuth, S. and D. Steffensmeier. 2004. Ethnicity Effects on Sentence Outcomes in Large Urban Courts: Comparisons among White, Black, and Hispanic Defendants. *Social Science Quarterly* 85: 994 .

Demuth, S. and D. Steffensmeier. 2004. The Impact of Gender and Race-Ethnicity in the Pretrial Release Process. *Social Problems* 51:2 222-242.

Englich, B. and T. Mussweiler. 2001. Sentencing under uncertainty: Anchoring effects in the courtroom. *Journal of Applied Social Psychology* 31: 1535–1551.

Englich, B., T. Mussweiler and F. Strack. 2005. The Last Word in Court—A Hidden Disadvantage for the Defense. *Law and Human Behavior* 29: 705-722.

Fishman, G., A. Rattner and H. Turjeman. 2006. Sentencing Outcomes in a Multinational Society: When Judges, Defendants and Victims Can Be either Arabs or Jews. *European Journal of Criminology* 3: 69.

Fishman, G., A. Rattner and G. Weimann. 1987. The Effect of Ethnicity on Crime Attribution, *Criminology* 25: 507.

Frazier C.E. and D.M. Bishop. 1985. The Pretrial Detention of Juveniles and Its Impact on Case Dispositions. *Journal of Criminal Law and Criminology* 76:

1132.

Green, William H. 1994. Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models, *NYU Working Paper* No. EC-94-10, http://ssrn.com/abstract=1293115

Greenwald, Anthony G. and Linda Hamilton Krieger. 2006. Implicit Bias: Scientific Foundations, *California Law Review* 94(4): 945-967.

Haj Yahia, M.S., Rahav, G., & Teichman, M. 1994. The lower court and minorities in Israel. *Plilim: Israel Journal of Criminal Justice*, 4: 157-174 (In Hebrew).

Hassin, Y. and M. Kremnitzer. 1993. The Offence of Breking and Entering – Variables Affecting Consistency in Sentencing. *Mishpatim* 22: 571 (In Hebrew).

Heckman, James J. 1979. Sample Selection Bias as Specification Error, *Econometrica* 45: 153-161.

Holmes, M.D., H.M. Hosch, H.C. Daudistel, D.A. Perez, and J.B. Graves. 1996. Ethnicity, Legal Resources, and Felony Dispositions in Two Southwestern Jurisdictions. *Justice Quarterly* 13: 11-30.

Humphrey, J. and T.J. Fogarty. 1987. Race and Plea Bargained Outcomes: A Research Note. *Soc. F.* 66: 176.

Johnson, Brian D. 2006. The Multilevel Context of Criminal Sentencing: Integrating Judge-And County-Level Influences, *Criminology* 44(2): 259-298.

Jolls, Christine and Cass R. Sunstein. 2006. The Law of Implicit Bias, *California Law Review* 94(4): 969-996.

Kahneman, Daniel and Shane Frederick. 2002. Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.) *Heuristics and biases* (pp. 49–81) New York: Cambridge University Press.

King, Gary. 1991. Truth is Stranger than Prediction, More Questionable than Causal Inference. *American Journal of Political Science* 35(4): 1047-53.

Kleck, G. 1981. Racial Discrimination in Criminal Sentencing: A Critical Evaluation of the Evidence with Additional Evidence on the Death Penalty. *American Sociological Review* 46: 783.

Klein, Stephen, Joan Petersilia, and Susan Turner. 1990. Race and Imprisonment Decisions in California, *Science* 247(4944): 812-816.

Kretzmer, D. 1990. *The legal status of Arabs in Israel*, Boulder, CO: Westview Press.

Patterson, E. and M. Lynch. 1991. Bias in formalized bail procedures. In M. Lynch and E. Patterson (eds.), Race and Criminal Justice. Albany, NY: Harrow and Heston, pp. 36–53.

Petee, T. 1994. Recommended for release on recognizance: Factors affecting pretrial release recommendations. *The Journal of Social Psychology* 134: 375–382.

Puhani, Patrick A. 2000. The Heckman Correction for Sample Selection and its

Critique, *Journal of Economic Surveys* 14(1): 53-68.

Rachlinski, Jeffrey J., Sheri Lynn Johnson, Andrew J. Wistrich, and Chris Guthrie. 2009. Does Unconscious Racial Bias Affect Trial Judges?, *Notre Dame Law Review* 84(3): 1195-1246.

Rankin, A. 1964. The Effect of Pretrial Detention. *New York University Law Review* 39: 641.

Rattner, Arye and Gideon Fishman. 1998. *Justice for all? Jews and Arabs in the Israeli criminal justice system*. Westport, CT: Praeger.

Rawls, John. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.

Samooha, Sammy. 2005. *Index of Arab-Jewish Relations in Israel 2004*. Haifa: The Jewish-Arab Center, University of Haifa; Jerusalem: The Citizens' Accord Forum between Jews and Arabs in Israel; Tel Aviv: Friedrich Ebert Stiftung.

Sampson, R.J. and J.L. Lauritsen. 1997. Racial and Ethnic Disparities in Crime and Criminal Justice in the United States. *Crime and Justice* 21: 311.

Schlesinger, T. 2005. Racial and ethnic disparity in pretrial criminal processing. *Justice Quarterly* 22: 170 – 192.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, New York: Houghton Mifflin Company.

Spohn, C.C. 2000. Thirty Years of Sentencing Reform: The Quest for a Racially Neutral Sentencing Process. Criminal Justice 2000, Vol. 3. National Institute

of Justice.

Spohn, C.C. 1990a. Decision Making in Sexual Assault Cases: Do Black and Female Judges Make a Difference? *Women and Criminal Justice* 2: 83.

Spohn, C.C. 1990b. The Sentencing Decisions of Black and White Judges: Expected and Unexpected Similarities. *Law and Society Review* 24: 1197.

Steenbergen, Marco R. and Bradford S. Jones. 2002. Modeling Multilevel Data Structures, *American Journal of Political Science* 46(1): 218-237

Steffensmeier, D. and C.L. Britt. 2001. Judges' Race and Judicial Decision-Making: Do Black Judges Sentence Differently? *Social Science Quarterly* 82: 749-764.

Steffensmeier D. and S. Demuth, 2001, Ethnicity and Judges' Sentencing Decisions: Hispanic-Black-White Comparisons. *Criminology* 39: 145.

Tonry, Michael H.(ed.). 1997. *Ethnicity, crime, and immigration: Comparative and cross-national perspectives*. Chicago: University of Chicago Press.

Tonry, Michael H. (ed.). 1995. *Malign neglect--race, crime, and punishment in America*. New York: Oxford University Press.

Uhlman, T.M. 1978. Black Elite Decision Making: The Case of Trial Judges. *American Journal of Political Science* 22: 884.

Warner, D.E. 2000. Race and Ethnic Bias in Sentencing Decisions: A Review and Critique of the Literature. In Michael W. Markowitz & Delores D. Jones-Brown (eds.) *The System in Black and White: Exploring the Connection between Race, Crime and Justice*, Westport, Connecticut: Praeger.

Welch, S., M. Combs and J. Gruhl. 1988. Do Black Judges Make a Difference? *American Journal of Political Science* 32: 126.

Williams, M.R. 2003. The Effect of Pretrial Detention on Imprisonment Decisions. *Criminal Justice Review* 28: 299-316.

Zussman, A. and M. Shayo. 2010. Judicial Ingroup Bias in the Shadow of Terrorism, (unpublished manuscript).

Table One: Randomization tests

	Arab judge	Jewish judge	*P* value
Arab suspect	45.9%	44.0%	.579
Police request (in days)	5.007	5.005	.987
Low severity	26.0%	24.2%	.551
Mid-low severity	44.8%	46.3%	.658
Mid-High severity	10.8%	10.7%	.952
High severity	18.4%	18.8%	.889
Violent offences	24.5%	25.3%	.796
Illegal residence	2.5%	4.4%	.158
Sexual offences	3.6%	3.8%	.895
Property offences	23.8%	26.8%	.324
Domestic violence	22.0%	19.6%	.386
Drugs related offences	13.0%	11.7%	.547
Traffic violations	2.5%	2.2%	.753
Public defender	60.7%	65.3%	.158
Unrepresented	9.4%	9.3%	.981
Prosecutor's statement	9.8%	7.0%	.134
Day (Saturday)	36.5%	37.5%	.762

N = 1230

Table Two: Summary statistics

	Arab suspect		Jewish suspect		Difference
	Mean	SD	Mean	SD	
Police request (in days)	5.106	1.806	4.925	1.678	.181*
Low severity	.260	.439	.235	.425	.025
Mid-low severity	.429	.495	.484	.500	-.055*
Mid-High severity	.136	.342	.085	.279	.051***
High severity	.176	.381	.196	.397	-.020
Violent offences	.273	.446	.234	.424	.039
Illegal residence	.082	.275	.006	.076	.077***
Sexual offences	.037	.188	.038	.191	-.001
Property offences	.258	.438	.264	.441	-.005
Domestic violence	.167	.373	.230	.421	-.063***
Drugs related offences	.095	.294	.139	.346	-.044**
Traffic violations	.027	.164	.019	.137	.008
Public defender	.546	.498	.719	.450	-.173***
Unrepresented	.097	.296	.091	.287	.006
Prosecutor's statement	.079	.270	.075	.263	.004
Day (Saturday)	.366	.482	.377	.485	-.012
Nazareth	.291	.455	.268	.443	.024
Acre	.403	.491	.212	.409	.191***
Release	.234	.424	.300	.458	-.065**
Detention length[†]	2.890	1.322	2.816	1.300	.074
Judge's ethnicity (Arab)	.233	.423	.219	.414	.013
Judge's gender (male)	.676	.468	.639	.481	.037

N = 1230, † *N* = 897; *, ** and *** denote differences that are statistically significant at 10%, 5% and 1%, respectively.

Table Three: Estimating the effects of ethnic in-group on release rates – logistic and multilevel logistic regressions

	(1)		(2)		(3)		(4)	
	b	O.R.	b	O.R.	b	O.R.	b	O.R.
Arab judge × Arab suspect	.631* (.325)	1.880	.810** (.346)	2.248	.832** (.351)	2.298	.806** (.360)	2.239
Arab judge	-.618*** (.222)	.539	-.757*** (.241)	.469	-.796*** (.244)	.451	-.651** (.298)	.522
Arab suspect	-.459*** (.148)	.612	-.430*** (.160)	.651	-.476*** (.169)	.621	-.485*** (.173)	.616
Judge gender (male)			.532*** (.154)	1.702	.520*** (.158)	1.682	.428** (.206)	1.534
Police's request			-.415*** (.060)	.661	-.434*** (.062)	.648	-.446*** (.064)	.640
Court (Nazareth)			-.964*** (.191)	.381	-.946*** (.197)	.388	-.952*** (.243)	.386
Court (Acre)			-.256 (.164)	.774	-.280 (.170)	.756	-.342 (.244)	.710
Mid-low severity			-.142 (.161)	.868	-.310 (.213)	.733	-.298 (.216)	.742
Mid-high severity			-.628** (.282)	.534	-.726** (.298)	.483	-.733** (.303)	.480
High severity			-.708*** (.234)	.492	-.835*** (.303)	.434	-.813*** (.308)	.444
Violent offence					-.104 (.212)	1.109	.117 (.216)	1.124
Illegal residence					-.008 (.390)	.992	-.034 (.397)	.967
Sexual offence					.487 (.433)	1.627	.524 (.439)	1.689
Domestic violence					.584*** (.197)	1.793	.581*** (.200)	1.788
Drugs related offence					.203 (.346)	1.225	.165 (.351)	1.179
Traffic violation					1.108** (.463)	3.027	1.162** (.468)	3.196
Public defender					-.378** (.169)	.685	-.373** (.173)	.689
Unrepresented					-.374 (.266)	.688	-.364 (.272)	.695
Prosecutor's declaration					-.787** (.318)	.455	-.862*** (.325)	.422
Day (Saturday)					-.106 (.150)	.899	-.119 (.156)	.888
Constant	-.727*** (.092)		1.405*** (.322)		1.789*** (.409)		1.858*** (.431)	
Ln(sigma u)							-1.885 .679	
Likelihood-ratio test of ρ							4.72**	
N	1229		1229		1229		1229	

*, ** and *** denote significance levels of 10%, 5% and 1%, respectively.

Table Four: Estimating the effects of ethnic in-group on detention length –
Poisson and multilevel Poisson regressions

	(5)	(6)	(7)	(8)
	b	b	b	b
Arab judge × Arab suspect	.077	.080	.083	.083
	(.092)	(.093)	(.093)	(.093)
Arab judge	-.029	-.051	-.053	-.053
	(.063)	(.065)	(.065)	(.065)
Arab suspect	.007	-.021	-.015	-.015
	(.045)	(.047)	(.049)	(.049)
Judge gender (male)		.118***	.099**	.099**
		(.043)	(.044)	(.044)
Police's request		.119***	.120***	.120***
		(.009)	(.010)	(.010)
Court (Nazareth)		-.017	-.028	-.028
		(.048)	(.050)	(.050)
Court (Acre)		-.037	-.067	-.067
		(.052)	(.053)	(.053)
Mid-low severity		-.038	-.003	-.003
		(.055)	(.067)	(.067)
Mid-high severity		.060	.042	.042
		(.071)	(.076)	(.076)
High severity		.024	-.071	-.071
		(.064)	(.079)	(.079)
Violent offence			.062	.062
			(.059)	(.059)
Illegal residence			.028	.028
			(.128)	(.128)
Sexual offence			.113	.113
			(.108)	(.108)
Domestic violence			.011	.011
			(.064)	(.064)
Drugs related offence			.187**	.187**
			(.087)	(.087)
Traffic violation			-.015	-.015
			(.186)	(.186)
Public defender			-.026	-.026
			(.048)	(.048)
Unrepresented			.006	.006
			(.077)	(.077)
Prosecutor's declaration			-.009	-.009
			(.078)	(.078)
Day (Saturday)			-.146***	-.146***
			(.044)	(.044)
Inverse Mills ratio				
Constant	1.042***	.352***	.395***	.395***
	(.031)	(.074)	(.098)	(.098)
Ln(alpha)				-18.864
				(859.545)
Likelihood-ratio test of α				$6.9 \cdot 10^{-6}$
N	896	896	896	896

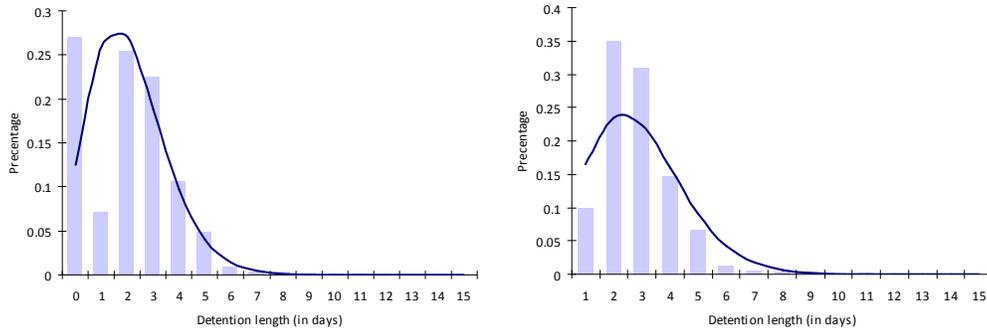*, ** and *** denote significance levels of 10%, 5% and 1%, respectively.

Figure One: Actual distribution of detention length including releases ($\lambda = 2.08$ days, N = 1230) (left); and excluding releases ($\lambda = 2.85$ days, N = 897) (right); and corresponding Poisson distributions
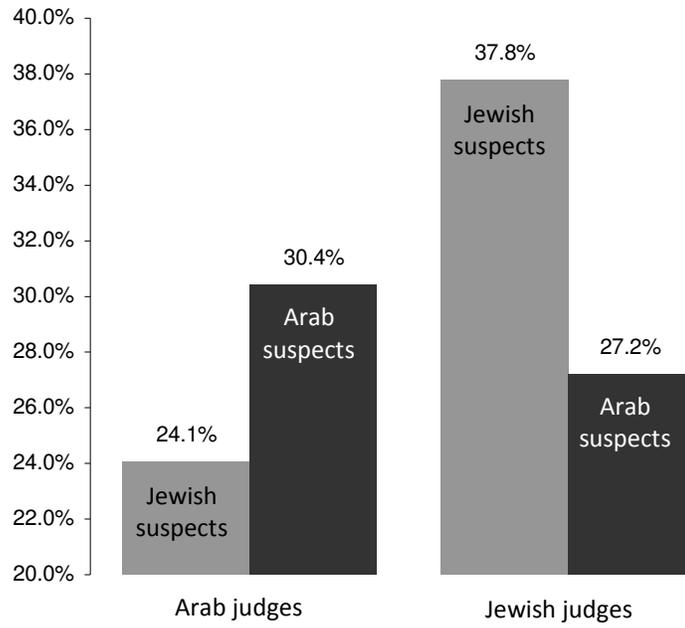
Figure Two: Predicted Probabilities of Release across the Ethnic Identity of
Judges and Suspects (at the modal case)

Appendix A: Further randomization tests

Dependent variable: logit(Arab judge)
Each cell reports p-values from Wald test

	(1)	(2)	(3)	(4)	(5)
Arab suspect	.579	.604	.324	.494	.481
Police request (in days)	.990	.931	.854	.951	.832
Mid-low severity		.554	.129	.126	.136
Mid-High severity		.785	.499	.441	.463
High severity		.674	.266	.230	.223
Violent offences			.757	.741	.727
Illegal residence			.077	.093	.079
Sexual offences			.981	.938	.889
Domestic violence			.217	.206	.216
Drugs related offences			.549	.503	.522
Traffic violations			.950	.941	.815
Public defender				.139	.176
Unrepresented				.475	.557
Prosecutor's statement					.162
Day (Saturday)					.966
Constant	<.001	<.001	<.001	.002	.002
χ^2	.31	.67	6.88	9.05	10.98
-2 Log likelihood	1311.894	1311.529	1304.810	1302.636	1300706
Pseudo R-square	.000	.000	.005	.007	.008

$N = 1230$

Appendix B: Estimating the effects of ethnic in-group on detention length – linear and multilevel linear regressions

	(9)	(10)	(11)	(12)
	b	b	b	b
Arab judge × Arab suspect	.223	.166	.192	.117
	(.206)	(.161)	(.159)	(.156)
Arab judge	-.080	-.103	-.113	-.094
	(.139)	(.111)	(.110)	(.159)
Arab suspect	.021	-.066	.056	.009
	(.101)	(.081)	(.084)	(.081)
Judge gender (male)		.317***	.276***	.208*
		(.073)	(.073)	(.121)
Police's request		.424***	.419***	.417***
		(.020)	(.020)	(.019)
Court (Nazareth)		.019	-.013	-.018
		(.083)	(.084)	(.134)
Court (Acre)		-.064	-.143	-.100
		(.088)	(.089)	(.164)
Mid-low severity		-.148	-.057	-.019
		(.091)	(.110)	(.105)
Mid-high severity		.090	.047	.014
		(.125)	(.132)	(.126)
High severity		.050	-.195	-.174
		(.111)	(.137)	(.131)
Violent offence			.164	.213**
			(.100)	(.096)
Illegal residence			.135	.103
			(.204)	(.196)
Sexual offence			.308	.320*
			(.193)	(.184)
Domestic violence			.024	-.006
			(.105)	(.100)
Drugs related offence			.492***	.489**
			(.153)	(.148)
Traffic violation			-.099	-.020
			(.300)	(.286)
Public defender			-.065	-.020
			(.083)	(.080)
Unrepresented			.016	.040
			(.133)	(.130)
Prosecutor's declaration			-.018	.010
			(.128)	(.123)
Day (Saturday)			-.356***	-.292***
			(.073)	(.074)
Inverse Mills ratio				
Constant	2.836***	.520***	.661***	.574***
	(.069)	(.134)	(.172)	(.190)
Likelihood-ratio test of sigma				30.64***
N	896	896	896	896

*, ** and *** denote significance levels of 10%, 5% and 1%, respectively.

Appendix C: Estimating the effect of ethnic in-group on detention length –
robustness tests

	Zero inflated Poisson	Heckman two-step selection model – linear model	Multilevel Poisson with Heckman selection correction
Detention length			
Arab judge × Arab suspect	-.011	.173	.075
	(.097)	(.157)	(.093)
Arab judge	.008	-.095	-.047
	(.068)	(.108)	(.064)
Arab suspect	.035	-.060	-.016
	(.045)	(.079)	(.047)
Judge gender (male)	.071	.265***	.096**
	(.045)	(.072)	(.043)
Police's request	.144***	.418***	.119***
	(.009)	(.019)	(.009)
Violent offence	.042	.139	.049
	(.055)	(.090)	(.053)
Illegal residence	-.009	.145	.014
	(.126)	(.182)	(.119)
Sexual offence	.071	.266	.099
	(.109)	(.182)	(.102)
Domestic violence	-.048	-.006	.002
	(.067)	(.099)	(.062)
Drugs related offence	.134**	.366***	.130*
	(.067)	(.116)	(.067)
Traffic violation	-.363*	-.104	-.031
	(.189)	(.285)	(.178)
Day (Saturday)	-.168***	-.355***	-.145***
	(.045)	(.071)	(.043)
Inverse Mills ratio			.064
			(.144)
Constant	.119	.521***	.332***
	(.076)	(.183)	(.107)
Release			
Court (Nazareth)	-2.66***	.531***	
	(.937)	(.104)	
Court (Acre)	-.345	.148	
	(.220)	(.092)	
Mid-low severity	-.341	.179*	
	(.234)	(.094)	
Mid-high severity	-1.502***	.627***	
	(.527)	(.154)	
High severity	-1.294***	.616***	
	(.380)	(.128)	
Public defender	-.681***	.095	
	(.235)	(.093)	
Unrepresented	-.491	.081	
	(.356)	(.149)	
Prosecutor's declaration	-14.191	.245	
	(669.76)	(.166)	
Constant	-.088	.105	
	(.266)	(.115)	
N	1229	1229	896

*, ** and *** denote significance levels of 10%, 5% and 1%, respectively.