Defending Your Life: A Computational Analysis of California Parole Suitability Hearings

Hannah Laqueur[*]   Anna Venancio[†]

Abstract

Parole boards hold tremendous discretionary power over offenders' lives and can have a significant impact on rates of incarceration. Yet parole represents a largely invisible dimension of the criminal justice system. This chapter analyzes over 8,000 transcripts from parole suitability hearings in California from 2011-2014 to make visible a decision-making process that is largely hidden from the public eye. The chapter demonstrates techniques for extracting structured data from semi-structured transcripts to offer both a quantitative and qualitative account of the Parole review process in California. We assess factors that may impact the parole decision including inmate risk score, the commitment offense, and whether victims are present at the hearing. We assess decision-maker consistency, both within individual decision-makers and across decision-makers. Finally, we parse the spoken words of inmates during their parole hearings to assess the narrative aspects of the process, and whether this may make a difference beyond the facts on file.

---

[*]Hannah Laqueur is Assistant Professor of Emergency Medicine at the University of California, Davis. hslaqueur@ucdavis.edu

[†]Anna Venancio is a senior data scientist at Primer AI.

Introduction

There are over 35,000 inmates in California serving life sentences with the possibility of parole, over a quarter of the entire state prison population. For these inmates, the sentencing phase does not end with a judge in the courtroom. Instead, their freedom depends on an affirmation by the Board of Parole Hearings ("the Board") that they are suitable for release from prison. The Board holds tremendous discretionary power over inmates' lives, determining whether or not they will die in prison. Yet relatively little is known about the decision-making process. Final deliberations among the parole commissioners are held in private, with only the final decision revealed. Hearings are only open to the victim's family and, with permission, members of the press. However, the *transcripts* from California parole hearings are publicly available. This chapter uses computer text processing to analyze over 8,000 transcripts (the population of transcripts from hearings held between 2011-2014) to provide a window into this relatively invisible and understudied process.[1]

The central question before the Board is whether the inmate poses an unreasonable risk of danger to society if released from prison. The Board has extensive latitude in making this determination. They may consider "all relevant, reliable information available to the panel" (Cal. Code Regs., tit. 15, §2281). The Board's assessment includes paper files indicating the offense for which the inmate was committed, the inmate's psychological risk assessment score, prison record, and letters of support. Beyond these facts on paper, the Board also holds a hearing, typically

---

1. There have been two empirical reports on California Lifer parole, but analyzing only a sample of 754 hearing transcripts from suitability hearings conducted between October 1, 2007, and January 28, 2010. Our analysis, by contrast, we have the population of hearings from 2009-2014, over 12,000 transcripts, which allows for more current and more extensive analyses(Weisberg, Mukamal, and Segall 2011). We focus on 2011 forward because this represents the current parole administration era.

lasting two to three hours. In the words of the Board of Parole Hearings Executive Officer Jennifer Shaffer, this is the Board's opportunity to try to understand: "Who were you then? Who are you today and what's the difference?" (Ulloa 2017).

Inmates and their attorneys view the hearing to be a critical performance. The Post-Conviction Justice Project's guide for California parole attorneys describes the work of preparing the client to testify as "the most important thing you can do to assist them in being found suitable for parole."[2] The hearing is seen an opportunity for the inmate to narrate a story of redemption: to demonstrate "insight" into the commitment offense, express remorse, and describe the ways in which he has reformed.

In this chapter, we use natural language text processing to explore inmate speech during the hearing and assess whether the narration may make a difference. We also identify non-verbal factors that are predictive of parole outcomes, such as the inmate's psychological risk score, and analyze potential cognitive biases among the parole decision-makers. We find the psychological risk score is by far the most important predictor of whether or not an inmate will be granted parole: inmates evaluated to be "high risk" of future violence and offending are virtually certain to be denied parole, whereas over half of the inmates identified as "low risk" are granted parole. Given the importance of risk, we focus much of our analyses on hearings for low risk inmates, a setting in which the decision is more likely to be affected by other factors.

The chapter proceeds as follows. Section provides background on the California parole system for inmates serving life sentences with the possibility of parole ("lifers"). This includes a discussion of the legal, political and administrative changes over the last decades. Section offers a description of the current parole suitability hearing process. Section describes the construction of the dataset built using computer text pars-

---

2. http://uscpcjp.com

ing of the universe of transcripts of California Parole Board hearings held from 2011-2014. Section 0.3 describes the variables that can be consistently pulled from the transcript text and that are predictive of a grant or denial of parole. We also use an instrumental variables approach to estimate the causal effect of risk score on the likelihood of a parole grant. Section 0.5 examines inter and intra decision-maker inconsistency in the parole hearing process. This includes estimation of the extent to which key system actors — presiding commissioners, deputy commissioners, and psychologists affect an inmate's chances of parole. Additionally, the section analyzes within decision-maker variation driven by cognitive biases or extraneous factors that should have no bearing on the decisions. Section 0.7 turns to an analysis of inmate speech during the suitability hearing. We use a bag-of-words approach, focusing on unigrams and bigrams for classification. Section 0.11 concludes.

Trends in California Parole

The system of indeterminate sentencing and release through parole reached a high point in California, and nationally, in the 1970s. At the time, more than 95% of prisoners in California were released by the discretionary decision of a parole board (Petersilia 2003).[3] In 1977, California overhauled its indeterminate sentencing system replacing it with a determinate system for almost all crimes. Indeterminate sentences remain only for those convicted of the most serious crimes and given a life sentences with the possibility of parole.[4] This shift in sentencing reflected broader nationwide changes in criminological thinking and criminal justice policy-making. It was a moment of pessimism about the possibilities of rehabilitation, conservative concerns criminals were being coddled,

---

3. More than 70 percent of inmates in the US were released by parole.
4. Over 90% of inmates serving life terms with the possibility of parole are in prison for first or second degree murder.

and progressive worries that the discretionary system of indeterminate release led to discrimination and inequality.

Under the current system, the determination of parole suitability for those serving life sentences with the possibility of parole is made by the Board of Parole Hearings ("the Board"), an executive branch agency within the California Department of Corrections and Rehabilitation (CDCR). For decades, the possibility of parole on the books was meaningless in practice—almost no one faced an actual possibility of parole release. Throughout the mid-1980s until the early 2000s, on average, the Board granted parole in only 2% of scheduled hearings. In the rare instances in which the parole board determined an inmate suitable for parole, the Governor almost always reversed the decision. During Governor Gray Davis' tenure, from 1999 to 2003, he vetoed 98% of the parole recommendations resulting in a total of only two inmates paroled; between 2003 and 2011, Governor Arnold Schwarzenegger vetoed an average of 73% of the parole recommendations that came before him (Sarosy 2013). This refusal to grant parole contributed to the tripling of the lifer population in California's prisons since the late 1980s: in the late 1980s, an inmate sentenced to a life term with the possibility of parole for second-degree murder served an average of five years; two decades later, he or she would serve an average of 24 years (Mullane 2012). There are now roughly 35,000 inmates serving life sentences with the possibility of parole, over 30% of the population in prison.[5]

There has been a radical shift in the rate of parole release in recent years. As shown in Figure 1, the number of grants by the Board has been increasing steadily since the early 2000s, with a substantial rise since 2007. The *rate* of release by the Board has also increased significantly. Because hearings may be scheduled but then postponed or stipulated by

---

5. California has the highest proportion of inmates serving life sentences with the possibility of parole than any other state in the nation (Weisberg, Mukamal, and Segall 2011).

the inmate, the best measure of the rate of release is the percentage of cases granted out of the number of hearings actually conducted. Data on this is available since 2007 and is presented in Figure 2. In 2007, 8% of hearings resulted in a grant by the Board; in 2014, the rate of release was 36%.[6]

Figure 1. *Number of Hearings Resulting in a Grant: 1978-2015*
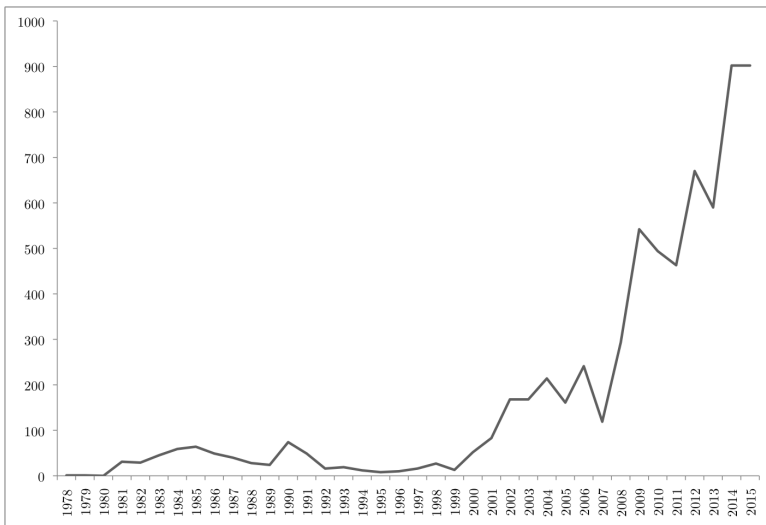
The recent change in the Board's grant rate tells only part of the story. There has also been a substantial shift in the governor's rate of approving his appointed Board's release decisions. California is one of only four states — Louisiana, Maryland, and Oklahoma, the others — in which the governor has this power of reversal. This additional layer of administrative review was instituted in 1988 with the passage of proposition 89. The Governor has the power to reverse Board decisions in all murder cases.

---

6. Data from 2007 - 2008 are from the Stanford Criminal Justice Center *Life in Limbo* Report; 2009-2014 are from this paper's parole hearing transcript dataset.

Figure 2. *Rate of Parole Grant: 2007-2014*
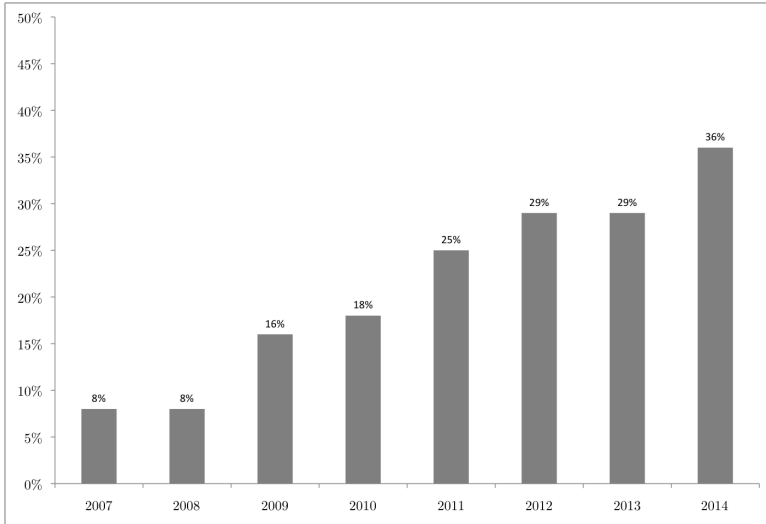
For offenders convicted of a crime other than murder, the Governor can request an En Banc Board review a suitability determination. The current Governor, Jerry Brown, has departed from his predecessor's practice of routine reversals, affirming over 80% of the Parole Board's release decisions since he took office in 2011. Thus, Governor Brown, rather than compensating for the Board's increasingly leniency, has re-enforced and affirmed the trend in the Board decision-making. The Board and governor appear to work in tandem rather than a true two-stage process.

What has lead to the substantial change in the rate of parole release? And how do we explain the change in the governors' position towards his Board's decisions? At least in some part, these changes may be explained by recent changes in the law brought about by the California Supreme Court. In 2008, the Court held that the Board and the governor must reach the release decision based on an inmate's "current dangerousness." Simply citing the heinousness of the inmate's commitment offense is an insufficient basis on which to deny parole (*In re Lawrence* 2008). This new

standard, in demanding the Board and governor establish some rational nexus between the evidence and a conclusion the inmate still poses a criminal risk, may make blanket denials less likely. At the very least, even if the law does not actually demand a higher threshold, it may offer political cover. Indeed, Brown has explained his lower reversal rate by stating: "I'm bound to follow the law" (Egelko 2011).

The increasing rate of release also came in the shadow of the state budget crisis and the U.S. Supreme Court's 2011 *Brown v. Plata* decision, which mandated that California reduce its prison population. Despite speculation by many that prison overcrowding encouraged more leniency in parole grants, the Brown administration nonetheless has insisted lifer releases are unrelated to any efforts to reduce the prison population numbers. On the other hand, Brown has recognized changes in public attitudes on crime. "There's still public safety (as a concern), but there's different dominating issues" (Egelko 2011).

Whatever the causes of these changes in California parole, we are at a moment in which, for the first time in decades, the possibility of parole means a real possibility. With this possibility may come actual incentives that encourage the convicted to reform, the putative purpose of parole. Engaging in productive in-prison programming, avoiding in-prison infractions, and narrating an understanding of past wrongs and future ambitions may have a substantive impact on the chances of release. In short, parole hearings now serve a purpose and represent a real process. In what follows, we describe the details of this process, and then turn to an analysis of relevant decision variables and inmate speech.

The Parole Hearing Process

California lifer inmates becomes eligible for parole consideration a year before their minimum eligible parole date, a date set at the time of sentencing. As a general rule, the initial and all subsequent hearings are presided over by a commissioner and deputy commissioner. There are twelve full-time commissioners, each appointed by the Governor and

confirmed by the Senate for staggered three-year terms with eligibility for reappointment. Deputy commissioners are civil servants who need only to have a "broad background in criminal justice."(Hopkins and King 2010) There are approximately 70 deputy commissioners across the state. Hearings are usually conducted at the prison where the inmate resides and usually take between two and five hours. Commissioners travel across the state to preside over hearings although are more frequently assigned to prisons near their residences. The identity of the commissioners assigned to hearings at a given prison in a given week is kept confidential until the week's hearings begin.

Before every hearing, the Board receives the central file, which includes the inmate's behavioral record in prison, vocational and education certificates, and the results of the psychological evaluations assessing recidivism risk. The hearing itself generally proceeds in three parts: first with a discussion of the commitment offense and the inmate's pre-conviction behavior and circumstances; then a consideration of post-conviction factors including prison behavior, programs and the risk assessment score; and finally, a review of post-release parole plans and, if applicable, any statements of support for or against the inmate's release.

As the California Supreme Court stated in the 2008 *In re Shaputis* decision "[T]he paramount consideration for both the Board and the Governor under the governing statutes is whether the inmate currently poses a threat to public safety."[7] Prior to the 2008 decision, the Board, rarely granting parole, would often cite the incarcerating offense as the sole justification for the denial. The Board may no longer make decisions based on mere retributive impulses; the central question before the Board is whether the inmate poses a risk of future violence and recidivism. In making this determination the Board considers a host of factors, perhaps most importantly the psychological risk assessment score. Until 2013,

---

7. In re Shaputis, 190 P.3d 573 (Cal. 2008)

psychologists from the Board of Parole Hearings' Forensic Assessment Division (FAD) generated a risk score using a combination of instruments, known as the HCR-20, LS/CMI, and PCL-R and expressed that score in terms of one of five risk levels: low, low/moderate, moderate, moderate/high, and high. Since 2014, the Board has revised its evaluation procedures slightly by ceasing to use the LS/CMI and expressing the risk rating in terms of one of three levels: low, moderate and high.[8] (LSA, Vol. 5-4, April 2014). The psychologist's evaluation may incorporate consideration of the commitment offense, historic risk factors, institutional programming, and the inmate's past and present mental state.[9]

In addition to the risk score, the Board may consider factors including whether the inmate has a violent criminal history, exhibits signs of remorse, has plans for the future, has been engaged in institutional activities, vocational and educational certificates, participation in self-help groups such as a Alcoholics Anonymous or Narcotics Anonymous, and the extent of his or her misconduct in prison, such as infractions for fights, use of drugs, or possession of a cell phone. The Board also considers any letters of support or opposition to the inmate's release and testimony from victims or victims' kin.

If an inmate is found unsuitable for parole, the law requires that a subsequent hearing be set three, five, seven, ten or fifteen years in the future. At each subsequent hearing the Board reviews transcripts from previous hearings along with an updated record. Despite the fact that fifteen years is the default by law, few subsequent hearings are set for fifteen years: in more than 90% of cases in our dataset the subsequent hearing is

---

8. Our data mostly does not reflect the shift from a five-level score to a three-level score. It will be some time before the change will manifest itself since most inmates will still be assessed in the hearing based on the risk scores from previous years.

9. An evaluation completed after 2009 is valid for five years. If an inmate's petition to advance is granted they are given a new subsequent risk assessment before the next scheduled hearing.

set for three years from the time of denial. Under Marsy's Law, the Board automatically reviews and may advance a three-year denial one year after the denial was issued. At the same time, an inmate may also request an advanced date regardless of the denial period. If the Board determines there has been a "change in circumstances or new information" and there exists a "reasonable likelihood that consideration of public and victim's safety does not require the additional incarceration," a hearing date will be advanced. (California Penal Code section 3041.5(b)(3)).

If the commissioners find the inmate suitable for parole, the decision is subject to review by the full Board. This rarely happens in practice. Absent action by the full Board, the grant of parole becomes final 120 days after the hearing and then goes to the Governor for his review. Again, in California, the Governor has the unusual power to reverse the decision of the parole board for inmates serving sentences for murder, the vast majority of lifers.

A parole candidate scheduled to go before the Board may seek to waive, stipulate or postpone the hearing. Waivers may be requested for anywhere between one and five years. A request for a short-term delay (postponement) for "exigent circumstances" such as emergencies, illness, or incomplete files may also be made. Both waivers and postponements are granted as a matter of course if made 45 days before the scheduled hearing, but there are limits as to the number of waivers and postponements an inmate may request. Alternatively, an inmate may stipulate to unsuitability for a period of three, five, seven, ten or fifteen years in accordance with the current denial periods (California Penal Code, 3041.5(b)). A stipulation is similar in nature to plea bargains, and the Board may choose to deny a stipulation. In practice, however, the Board appears to grant the vast majority.

Finally, all inmates are entitled to a state-appointed attorney at the parole hearing if they have not retained an attorney. These "panel" attorneys are compensated $50 an hour with a cap at $400, which means they spend a maximum of eight hours per case. Advocates and private attor-

neys with whom we have spoken contend that, not unlike plea bargaining in criminal court, the panel attorneys often try to persuade parole candidates to stipulate to a denial and defer the hearing rather than have to spend the time going through the hearing process (DeBacco 2010).

The Dataset

Anatomy of a transcript

Through a public records request, we obtained all parole suitability hearings conducted between 2011-2014. [10] The records are obtained as pdfs, which we converted to text files. They have an intrinsic page layout, which is flattened in the text file. This is strongly reflected in the INDEX at the beginning of the file (see an example in the following table).

The presence of the INDEX underlines the predetermined flow of the parole hearing. This can be very useful in order to extract structured pieces of the information from the raw text (for instance, parole outcome), but also implies some text cleaning to be able to highlight only transcribed speeches. Beyond these architectural elements, the transcripts have another very distinctive attribute to take into account during text analysis: a dialog structure. In order to focus on the inmate's part of the dialog, the text must be parsed to selectively retain only the inmate's contribution.

---

10. 2009 is the first year in which the universe of transcripts are available electronically. However, we focus our analyses on 2011 forward because this marks the beginning of the current Board of Parole Hearings administration under Governor Jerry Brown and Executive Officer Jennifer Shaffer, and the period after the important *In re Shaputis* California Supreme Court decision (2011)

Table 1. *Example of an INDEX*

From raw text to structured data

We used Python regular expressions to pull key information from each hearing transcript. These variables include the result of the parole hearing, the commitment crime, the psychological risk assessment score as well as the identity of the evaluating psychologist, the minimum eligible parole date, the inmate's lawyer, the district attorney if present at the hearing, the number of victims or victim's next of kin present at the hearing, whether or not an interpreter was present at the hearing, the results of any previous suitability hearings, the inmate's date of entry into prison, the presiding and deputy commissioners, the date and time of the hearing, and the results of the immediately preceding hearings.

Figures 3 and 4 below present snapshots of hearing transcripts, highlighting variables that can be consistently retrieved from the raw text and that might impact the parole board's decision.

From raw text to inmate dialog

During the Parole Hearing, several key figures participate, asking questions to prompt the inmate and the inmate's attorney to shine light on the

Figure 3. *Structured data in the presiding commissioner's opening statements (example of extracted information).*

PRESIDING COMMISSIONER CHRONES: All right. Good morning, everybody. The time is 9:00 a.m. , and this is a Subsequent Parole Consideration Hearing for Leo Robles, CDC Number B-00842. Today is March 13th, 2012. We are located at the California Men's Colony . Mr. Robles was received on February 18th, 1966, from Santa Clara and Marin Counties. The controlling offense for which he was committed include Murder in the First Degree , case number 41303. Two counts

of Penal Code Section 187, Assault With Intent to Commit Murder , case number 2780. And a – Penal Code Section 217, and an additional count of Assault in Prison under the same case number as the previous, I believe, with an additional count of 12022 (b) PC. Mr. Robles has a minimum eligible parole date of August 12th, 1977 . This hearing is being recorded, and for purpose of voice identification.

Figure 4. *Structured data in the rest of the text (example of available information : chronos, discipline 115s, risk assessment...).*

DEPUTY COMMISSIONER CASSADY: Okay.   PRESIDING COMMISSIONER PRIZMICH: So let me go through a brief description of 115s . February 21, 1975 Attempted Mail – Subverting the Mail I guess. He was found guilty on that. February 21, 1975, Possession of a small amount of marijuana, found guilty; (...)
–
DEPUTY COMMISSIONER TURNER: Okay.   Did you get any kind of, like laudatory chrono from your supervisor saying, you know, you're a good worker and all that?

–
Dr.   Twohy   formed   the   opinion   that   you present a low to moderate risk for violence in the free community.

inmate's past and current behavior. Because we not only want to extract key variables but also analyze the inmate's participation in the dialog, this means building a parser able to selectively extract parts of the text, as shown in Figure 5.

Pre-processing steps relied on *spaCy*, free open-source library in Python

Figure 5. *Parsing the text to obtain only the parts of the dialog spoken by the inmate (examples of sentences to extract).*

DEPUTY COMMISSIONER ZARRINNAM: Ali Zarrinnam, Z- A-R-R-I-N-N-A-M, Deputy Commissioner.
INMATE ROBLES: I'm Robles, R-O-B-L-E-S. Number B-as in boy-00842.

ATTORNEY EISENSTAT: Your first name, sir. INMATE ROBLES: Pardon?
ATTORNEY EISENSTAT: Your first name. INMATE ROBLES: Leo, L-E-O.
PRESIDING COMMISSIONER LABAHN: And Mr. Robles, it will be helpful –
INMATE ROBLES: Oh, okay. I'm sorry.

for Natural Language Processing Toolkit. We describe them in more detail in Section 0.7.

Non-Verbal Factors that Influence Parole

Risk Score

By far, the strongest predictor of the release outcome is the inmate's risk score. A rating of high or high moderate essentially forecloses the possibility of parole: less than 3% of the 1,193 of the inmates in our dataset with a high or high moderate risk score are granted parole. On the other hand, 56% of the 1,673 inmates given a low risk score are granted parole.

We estimate the causal impact of the risk assessment score using an instrumental variables approach (Angrist, Imbens, and Rubin 1996). Risk assessments are completed by one of 46 psychologists in the Board of Parole Hearings Forensic Assessment Division (FAD). The FAD psychologists work exclusively on conducting these assessments for use in suitability hearings. Evaluations are conducted in-person using a combination of instruments – HCR-20, LS/CMI, and PCL-R – as well as "clinical judgment." The process of pairing inmates with psychologists is based solely on geography. Thus, within a given prison and year, an inmate's assignment to one of the 46 psychologists is essentially random. It is thus possible to use the systematic differences in psychologists' penchants for

15

assigning higher (or lower) risk ratings to instrument for risk measured as a cardinal variable from a score of 1 - 5, with 1 being "low risk" and 5 high risk. This allows use to estimate the causal effect of risk on the likelihood of parole grant. The results are presented in table 2, which shows the estimate of the average effect of moving down a risk score level on the likelihood of parole release. The instrumental variables estimate is slightly smaller and the confidence interval wider than a simple regression estimate, but the two estimates are remarkably similar: the average of moving up each risk assessment score category (i.e. from low to low-moderate) results in a 15% reduction prison release chances. We ran several robustness checks including models with controls for prison and year, and found estimates of similar magnitude and significance.

Table 2. *Effect of Risk Score on Grant Chances*

	IV Estimate	Naive OLS Estimate
Risk (1-5)	-0.154 (0.023)	-0.161 (0.004)

Other Potentially Relevant Variables

Table 3 below presents other variables that we can consistently computationally retrieve from the hearing transcripts that could impact the parole decision: the inmate's commitment offenses, whether the inmate used an interpreter, and the presence of victims at the hearing. We presents the simple difference in mean release rate as well as adjusted estimates controlling for other relevant variables including the prison where the inmate is housed, their attorney, and the number of previous hearings. The adjusted estimates are made using Targeted Maximum Likelihood Estimation (TMLE), a double robust estimation technique that incorporates data-adaptive modeling ad has been shown to be less statistically biased and more efficient than conventional parametric method (Van der Laan and Rose 2011). At the same time, even after adjusting for covariates in a

rigorous manner, there are too many possible unmeasured confounders — variables that may affect both the variable of interest and the outcome — to allow for these estimates to be interpreted causally. Instead, the measures offer a means of noting important associations that are not easily explained away and thereby raise research questions. For example, we find the presence of an interpreter is strongly associated with release, even after adjusting for an assortment of other variables. Does speaking through an interpreter actually result in a sharp drop in one's chances of getting parole because it is harder for the inmate to connect personally with the commissioners? Or is speaking through an interpreter simply associated with other variables we can't easily extract from transcripts, such as inmate ethnicity?

Table 3. *Likelihood of Parole Grant: Variable Importance Measures*

	Bivariate Estimate	**Adjusted Estimate**
Murder	0.045 (0.015)	0.004 (0.014)
Robbery	-0.001 (0.02)	0.031 (0.019)
Kidnap	-0.033 (0.019)	-0.004 (0.019)
Attempted	-0.012 (0.018)	0.013 (0.018)
Sex Crime	-0.186 (0.028)	-0.125 (0.02)
Interpreter	-0.032 (0.016)	-0.066 (0.015)
0 vs 1 Victim	-0.015 (0.04)	-0.001 (0.044)
0 vs 2 Victims	-0.008 (0.03)	-0.02 (0.033)
0 vs 3 Victims	-0.058 (0.027)	-0.057 (0.028)
0 vs 4 Victims	-0.073 (0.031)	-0.06 (0.031)
0 vs 5 Victims	-0.085 (0.039)	-0.073 (0.038)
0 vs More than 5 Victims	-0.139 (0.033)	-0.141 (0.027)

In most cases, the grant rate does not appear to vary with the commitment offense. One clear exception are offenses involving a sex crime. After controlling for other variables, we find the grant rate is an estimated 13% lower when the inmate's commitment offense included a sex

crime. This finding comports with evidence presented in previous research (Weisberg, Mukamal, and Segall 2011) and conventional wisdom among attorneys and advocates, who suggest the Board is particularly unlikely to release an inmate whose crime involved sexual violence, especially if the crime involved a child. As we see in the text analysis in section 0.7, the word "sex" is also negatively associated with the chances of a grant of parole.

The role victims play in the parole suitability hearing has been the subject of considerable debate since the passage of Marsy's Law (Proposition 9) in 2008. Marsy's Law, among many things, gave victims and their next of kin expanded rights to receive notice and testify at suitability hearings. It was also under Marsy's Law, ostensibly in an effort to relieve the hardship placed on victims of attending hearings, that the deferral lengths between hearings was extended to the current three to fifteen years period (DeBacco 2010). The Stanford Criminal Justice Center Report (Weisberg, Mukamal, and Segall 2011), the first study to offer an empirical portrait of the California parole process, used a sample of hand-coded transcripts from 2007-2010 and found the overall grant rate when victims attended hearings was 5% as compared to 14% when victims did not attend. The estimates presented here look at the influence of victim attendance at the hearing ranging from zero in attendance to more than five. This analysis shows that victim attendance at the parole hearing is an important predictor of an inmate's parole chances. While there is minimal difference when one or two victims are at a hearing as compared to zero, the chances of parole drop precipitously as more victims are in attendance: hearings with more than five victims result in a grant roughly 14% less often than hearings with no victims in attendance, even after adjusting for highly predictive variables.

The estimate of the rate of release at hearings conducted through an interpreter is 7% lower than at hearings in which an interpreter was not present. Again, it cannot be concluded that speaking through an interpreter actually causes the Board to be less likely to grant parole. It may

be that unmeasured factors, such as ethnicity, are associated with having an interpreter and with lower chances of parole. Or perhaps the need for an interpreter represents a broader set of disadvantages faced by inmates who do not speak English and therefore have a harder time engaging in prison programming and planning for post-prison release.

Given that essentially no inmate with a psychological risk score of high or high moderate is granted parole, whereas over half of low risk inmates are granted parole, we perform the same set of variable analyses on the subset of low risk inmate hearings. It is among low risk inmates that there is theoretically the most room for other factors, such as victim or interpreter presence, to impact the parole decision.

Table 4 below shows a simple difference in mean grant rate among low risk inmates for the important variables discussed above, next to the estimates for the full population. Table 5 shows the adjusted estimates for low risk inmate versus the adjusted estimates for the full population.

Table 4. *Likelihood of Grant Among Low Risk vs. All Inmates (Bivariate Estimates)*

	Low Risk	All Inmates
Murder	0.036 (0.033)	0.045 (0.015)
Robbery	0.006 (0.048)	0.045 (0.015)
Kidnap	-0.061 (0.045)	-0.033 (0.019)
Sex Crime	-0.215 (0.071)	-0.186 (0.028)
Interpreter	-0.112 (0.031)	-0.032 (0.016)
Victim Present	-0.155 (0.027)	-0.063 (0.014)
0 vs 1 Victim	-0.089 (0.076)	-0.015 (0.040)
0 vs 2 Victim	-0.045 (0.059)	-0.008 (0.030)
0 vs 3 Victim	-0.119 (0.049)	-0.058 (0.027)
0 vs 4 Victim	-0.200 (0.054)	-0.073 (0.031)
0 vs 5 Victims	-0.216 (0.069)	-0.085 (0.039)
0 vs More than 5 Victims	-0.267 (0.060)	-0.139 (0.033)

Table 5. *Likelihood of Grant Among Low Risk vs. All Inmates (Adjusted Estimates)*

	Low Risk	All Inmates
Murder	-0.054 (0.047)	0.004 (0.014)
Robbery	-0.010 (0.059)	0.031 (0.019)
Kidnap	-0.108 (0.067)	-0.004 (0.019)
Sex Crime	-0.165 (0.088)	-0.125 (0.02)
Interpreter	-0.113 (0.031)	-0.066 (0.015)
Victim Present	-0.174 (0.026)	-0.127 (0.013)
0 vs 1 Victim	-0.129 (0.075)	-0.001 (0.044)
0 vs 2 Victim	-0.065 (0.058)	-0.02 (0.033)
0 vs 3 Victim	-0.144 (0.048)	-0.057 (0.028)
0 vs 4 Victim	-0.224 (0.054)	-0.06 (0.031)
0 vs 5 Victims	-0.248 (0.068)	-0.073 (0.038)
0 vs More than 5 Victims	-0.248 (0.059)	-0.141 (0.027)

As was the case among the full population, crime type appears to generally make little difference in the parole outcome among low risk inmates, with the exception of sex crimes. In this case, among low risk inmates, serving time for a sex crime is associated with a 17% lower chance of parole, controlling for other variables, as compared to 13% among the full population. Both the presence of an interpreter and the presence of victims appear to be even *more* important in the parole decision for low risk inmates. An interpreter is twice as important among the low risk group. Hearings with five or more victims present result in a grant roughly 25% less often than hearings with no victims in attendance, after adjusting for relevant covariates. This is approximately 1.8 times greater difference than among the full population.

Inconsistency in the Parole System

"Justice," the trope goes, "is what the judge ate for breakfast."(Kozinski 1992) The problem of inconsistency in judicial decision-making has been documented in a number of contexts. Recent research indicates, for example, that the outcome of a football game (Chen and Spamann 2014), the results of the immediately preceding case (Chen, Moskowitz, and Shue 2014), and the time of day (Danziger, Levav, and Avnaim-Pesso 2011) can substantially affect legal decisions. Also startling are the wide between-judge disparities found in domains including immigration asylum (Ramji-Nogales, Schoenholtz, and Schrag 2007; Fischman 2014), social security disability (Nakosteen and Zimmer 2014), and criminal setencing (Abrams, Bertrand, and Mullainathan 2012).

In what follows, we offer estimates of the extent to which differences between major decision-makers — presiding commissioners, deputy commissioners, and FAD psychologists — affect parole outcomes. The results suggest a process that contains some real element of chance. Who an inmate happens to get as the presiding commissioner or deputy commissioner can alter whether they are released or remain in prison. On the other hand, there is no evidence of *intra* commissioner inconsistency, at least with respect to the time of the decision, which the highly publicized Israeli parole board study (Danziger, Levav, and Avnaim-Pesso 2011) found to be important, and "gambler's fallacy," a psychological effect documented in other judicial decision-making contexts. In sum, the findings, although they cannot speak to the inherent quality of the release decisions themselves or whether the current release rate of roughly a third is the "right" rate, do suggest that the parole board appears to be operating in a relatively consistent and professional manner.

Inter Decision-Maker Inconsistency

The following analysis of inter decision-maker inconsistency provides two measures of decision-maker differences: average and extreme.[11] The average estimate provides an overall measure of how inconsistency affects outcomes. This estimate of average difference may be interpreted as the probability that a hearing would be decided differently by two randomly selected decision-makers due to one of them being systematically more likely to grant parole than the other. The estimate of extreme inconsistency is the difference in grant rates between the harshest and most lenient commissioner and deputy commissioner. This can be thought of as the percentage of cases that could come out differently if they were assigned to the most lenient decision-maker rather than the harshest.

Estimation Procedure

Estimates are made using year-prison fixed effects with dummy variables for actor identities. Commissioners are assigned to hearings based largely on geography (they are more likely to decide hearings at prisons near their residence), and different security level prisons will have inmates with different characteristics. Commissioners in one year may be hearing very different types of cases than a commissioner hearing cases in another year. Within prison and year, however, commissioners should see the same type of inmates on average, and thus, at least with an infinite sample, estimates should provide reasonable quantification of actors' relative impacts on the probability that an inmate will be granted parole.[12]

---

11. The method for estimating average inconsistency was developed by Joshua Fischman (Fischman 2014).

12. We conduct a randomization check and find support for the assumption that decision-makers are randomly assigned to inmates once prison and year are controlled for. We use the presence of a district attorney at the hearing and whether a hearing is an initial hearing or subsequent as test outcomes. If cases are randomly assigned, district attorney attendance

A second stage is needed in the estimation procedure to correct for finite sample bias and to properly estimate confidence intervals. In a finite sample estimates will be biased upwards because there will almost always be differences between decision-makers due to chance even if they would make the same decision in each case. For the measure of average difference in grant rates, we estimate and correct for finite sample bias via subsampling (Fischman 2014).[13] Measures of extreme differences in grant rates are especially susceptible to finite sample bias because we would expect the extremes to regress toward the mean with larger samples. Thus, taking a particularly conservative approach to the estimates of extreme differences, we permute outcomes within prison-year combinations to estimate bias under the null distribution that assumes no actual differences between decision-makers. For both the extreme and average measures of inconsistency, we also guard against finite sample bias by limiting our analysis to decision-makers with at least 100 observations.

Estimates of Inter Decision-Maker Differences

It is worth stressing that these estimates represent lower bounds for inconsistency in decision-making. Insofar as commissioners grant parole differently in different types of cases, the grant rate differentials may understate the degree of inconsistency. For example, one commissioner may decide to grant in 20% of hearings and another in 30% of hearings. While we know that the commissioners decide at least 10% of cases differently,

---

and hearing number should not be related to the decision-maker. We find no evidence of such a relationship.

13. Our estimates of finite sample bias are slightly responsive to the percentage of the dataset we subsample. For Presiding Commissioner inconsistency, as the percentage subsampled increases from 25% to 55%, our estimate of bias decreases by about .5%. For Deputy Commissioner inconsistency, the estimate of bias decreases by almost 1%. We report bias corrected estimates derived from a subsampling percentage of 40%.

the number could be considerably higher. Imagine, for example, that the first commissioner's grants are all first-degree murder cases while the second commissioner's grants are all second-degree murder cases and she denies parole in all first murder cases. The inconsistency between the two commissioners would actually be 30%, but our estimate would only be 10%.

Nonetheless, even the lower bound estimates of inconsistency in the decision-making are illuminating. As the estimates presented in Table 6 show, at least 11% of cases could be decided differently based simply on the presiding commissioner that happens to be assigned to an inmate's case and at least 15% of cases could be decided differently depending only on the deputy commissioner assigned to hear the case. An evaluating psychologist can affect an inmate's chances of parole by at least 6% despite the fact that they are not involved in the actual parole decision. As shown the preceding section, an inmate's risk score is an important determinant of the likelihood they will be released; consistent with *In re Lawrence*, risk is the primary consideration for the Board. This raises two questions. First, given the importance of risk, how accurate are the current clinical risk assessment instruments used by the Board? Second, are the risk tools being applied consistently? If two clinicians administer a tool to an inmate, the score should not reflect the idiosyncratic tendencies of a given clinician. At the same time, the average differences in parole rates are, as expected, substantially smaller and relatively small: at least 3% of cases would be decided differently if they were randomly re-assigned to a different psychologist, at least 6% of cases would be decided differently if they were randomly re-assigned to a different Presiding Commissioner, and at least 7% of cases would be decided differently if assigned to a different Deputy Commissioner.

Given a substantial literature devoted to the study of how judicial decisions differ by judicial ideology, one might surmise commissioner characteristics could account for the differences in the grant rate. One might expect, for example, that commissioners appointed by Governor

Table 6. *Inconsistency in Decision Making: Differences in Grant Rate*

Decision-Maker	Average Difference	Extreme Difference
Presiding Commissioners	6.1% (5.0% - 7.2%)	11.3% (7.8% - 14.8%)
Psychologists	3.3% (1.2% - 5.4%)	5.6% (3.3% - 9.7%)
Deputy Commissioners	7.0% (5.7% - 8.3%)	14.6% (10.3% - 18.7%)

Schwarzenegger, a governor known for being "tough on crime," would be less likely to grant parole than commissioners appointed by Governor Brown (Ball 2011). Yet despite the well-powered dataset, there is surprisingly little difference in grant rates between commissioners with different backgrounds and characteristics: appointing governor, gender, prior employment with the CDCR, military experience, nor prior service as parole panel attorney appear to matter in the likelihood the hearing will result in a grant.

Intra-Commissioner Inconsistency

A growing body of work in the judicial decision-making literature has documented extraneous factors and cognitive biases that can make judges internally inconsistent in their decision-making. Most relevant to the present study of California parole, a highly publicized study of Israeli judges found the time of day in which the decisions were made in relation to breaks and meals had a significant impact on the parole decision (Danziger, Levav, and Avnaim-Pesso 2011). Danzinger et al. found the rate of parole grants dropped from roughly 65% to almost zero as a session neared its end, and then rose again to 65% after a session break. The authors speculate that the effect is likely driven by mental depletion. However, more recent work has challenged the results, arguing cases are not randomly distributed throughout the day and because favorable rulings take longer than unfavorable ones, judges may be strategic in the order in which they hear cases(Weinshall-Margel and Shapard 2011).

The California Parole Board does not take such clear and scheduled recesses. But it is possible to capture and assess the effect of the time of day on release outcomes. Table 7 shows the differences in grants for each start hour relative to 8:00 a.m., the earliest time in which hearings are scheduled. Start times are rounded to the nearest hour. The model includes prison-commissioner fixed effects. That is, for each presiding commissioner within a given prison, the model measures differences in release decisions that depend on the time in which the hearing is conducted. The results suggest no such significant differences in the outcome regardless of the time of day the hearing begins. One possible exception is late-starting hearings. There is some evidence that hearings starting between 4:00 pm and 6:00 pm are more likely to end with a grant, but the sample sizes at those times are too small to allow for robust conclusions.

Table 7. *Effect of Hearing Start Time on Grant Rate*

Start Time	Grant Rate (relative to 8:00 a.m.)
9:00 AM	-0.029 (0.018)
10:00 AM	0.005 (0.022)
11:00 AM	-0.010 (0.018)
12:00 PM	-0.010 (0.010)
1:00 PM	0.007 (0.017)
2:00 PM	-0.002 (0.021)
3:00 PM	0.070 (0.028)
4:00 PM	0.062 (0.043)
5:00 PM	0.102 (0.088)
6:00 PM	-0.029 (0.018)

The "gambler's fallacy" — the mistaken idea that the chances of something occurring increases (or decreases) depending on recent occurrences, despite the fact that the probability of the occurrence is fixed — is another cognitive bias scholars have speculated could impact judicial decision-

making (Tversky and Kahneman 1974).In the parole context, this would be the tendency to respond to streaks of grants (or denies) by becoming more likely to deny (or grant) in the next hearing. Chen, Moscowitz, and Shue (2014) document such negative auto-correlation in the decisions made by asylum court judges and find growing effects as the length of a streak of decisions in one direction or another increases. We find no evidence for such effects in the California parole hearing process. Running a model of lagged release decisions regressed on the present hearing outcome does generate statistically significant lagged coefficients, prima facie evidence that a streak of previous grants (or denies) increases the probability that a commissioner will grant (or deny) the next inmate's request for parole. For every sequential grant, commissioners are about 3% less likely to grant parole to the next inmate. However, recent work suggests that standard techniques for analyzing the gambler's fallacy are subject to finite sample bias (Miller, Sanjurjo, et al. 2015). Indeed, in our dataset, further analysis reveals the result is *not* driven by commissioner psychology but is instead an artifact of finite sample bias. Permuting hearing outcomes, thereby randomly reassigning a grant or deny to each hearing, and repeatedly running the same lagged models on the permuted data generates estimates of negative correlation that are essentially equivalent to the initial estimate.

In summary, there is no evidence that the gambler's fallacy affects parole decisions. The current administration has made efforts towards increasing professionalization — commissioners are sent to national judicial college for training, and efforts have been made to reduce workloads to guard against decision fatigue and to allow for more deliberate decision making.[14] It is possible that this has had an impact on parole decisions. Another explanation is simply that the cognitive biases doc-

---

14. This comes from personal communication with Howard Mosely and Jennifer Shaffer, December 2014.

umented in previous studies have been overstated. In fact, statistical properties of the underlying data may explain phenomenon that have been attributed to decision-fatigue or the the gambler's fallacy.

Inmate Expression Analysis

Many factors, explicit or not, are naturally taken into account when the Board makes parole determination. Here we propose to examine the relationship between the parole outcome and the words the inmates used during the parole suitability hearing.

Pre-Processing

The parser described in Section extracts out the parts of the dialog spoken by the inmate. The resulting text is then tokenized to obtain unigrams (single words) and bigrams (two consecutive words), a standard process in the 'bag of words' approach to text analysis. This approach discards word order and uses a simple count of the number of times the word appears in the text. The tokenization is performed using *spaCy* in the Python Natural Language Processing (NLP) toolkit and includes partial pre-processing of the text such as removal of a subset of punctuation and of all *spaCy* stopwords - i.e. common words such as "a," "and," "like", "this", "that", "and", "so", "on". Unigrams are also normalized further by applying *spaCy*'s lemmatisation algorithm, which reduces all words to their base form. For example, "talk", "talked", "talks", "talking" are reduced to the verb lemma "talk."[15]

Finally, we zoom into word functions within sentences by relying on *spaCy*'s Part of Speech (POS) tagger to identify verbs, nouns, adjectives,

---

15. Research indicates for simple and common tasks like measuring sentiment, topic modeling, and classification, n-grams do little to enhance performance (Hopkins and King 2010).

adverbs, and interjections. This allows us to pay particular attention to more coherent subcategories of the vocabulary, for instance by selectively retaining only verbs or nouns used by the inmates.

Text Based Prediction

To examine the relationship between the parole outcome and the words the inmates used, we established a vocabulary (a list of words) and transformed each transcript into a sparse vector by looking at the counts of each word in the pre-processed text. To investigate the impact of words on the outcome of the hearing, we favored explanatory models and found logistic regression to satisfy our needs. Discounting interaction effects, which are are hard to capture for words, logistic regression produces coefficients whose magnitude reflects the importance of the features in the model.

We focus the analyses on suitability hearing transcripts for inmates identified as low risk. As described in Section 2, almost no inmate with a psychological risk score of high or high moderate is granted parole, while over half of low risk inmates are granted parole. We therefore focus our attention on this subset for whom the hearing itself could theoretically make a difference in the parole decision. It is among low risk inmates that there is the most room for commissioner discretion and room for subtle and subjective factors, such as inmate expression, to impact the decision.

Finally, we restricted the dataset to parole hearings resulting in a clear "deny" or "grant" outcome. Although several outcomes, including "postpone", "waive" or "stipulate" are possible[16], retaining deny or grant results gives us a clearer signal. The positive class is taken to be the grant outcome.

---

16. See http://www.cdcr.ca.gov/BOPH/pshResults.html for a detailed description of the various hearing results

We begin with an analysis of lemmatized unigrams with stopwords removed. The vocabulary is taken to be the top 1,000 words from both granted and denied hearings. The features for each transcript are a straight forward word count vector. The dataset is split into training/validation sets, with a 80/20 ratio. A logistic regression classification model is trained using Scikit Learn's LogisticRegressionCV, with the L2 regularization constant determined with cross validation to avoid overfitting.

The accuracy of the model, obtained on the withheld validation set, is 66%. The baseline accuracy, obtained with a simple prediction of majority label[17] is 56%. Diving a little deeper into the performance of the model, we investigated the overall precision and recall, as well as the performance on a per class level. Overall precision is at 65% and recall is at 66%, giving an f1-score of 65%. [18] This underlines that there is some signal captured by the model. Looking at per label performance shows that recall for the minority class is what is keeping the overall performance metrics down, as it is at 47%. The precision of minority class is at 61% and that of the majority class is at 68%.

Table 8 presents the top 10 most important unigram features of grants and denials. "Thank" shows up as the most important feature for parole grants. This is consistent with the interpretation that inmates who are polite and show respect and deference are more likely to be granted parole. As the Post-Conviction Justice Project's manual for attorneys states: "it is crucial that they not seem hostile, disrespectful, or defensive. The

---

17. Predicting the majority label in all cases is one of the simplest way of obtaining a baseline model for any classification problem.

18. Precision gives and idea of the quality of the label, it is computed as the ratio of correctly predicted positive observations to the total predicted positive observations. Recall gives an idea of quantity of data points of each class captured and is computed as the ratio of correctly predicted positive observations to the all observations in actual class. The F1-score is used to capture both measurements, and is computed by obtaining the harmonic average of precision and recall.

Panel generally looks most favorably on clients who are respectful, penitent and truly appear to be remorseful." [19]

Table 8. *Top 10 Most Important Unigram Features For Parole Grants & Denials*

Grant		Denial	
thank	0.107	**would've**	-0.045
sponsor	0.032	**guess**	-0.043
service	0.032	**mistake**	-0.042
situation	0.032	responsibility	-0.040
turn	0.032	**115**	-0.036
commissioner	0.030	respect	-0.036
emotion	0.027	sell	-0.032
world	0.027	class	-0.032
alcoholic	0.027	cell	-0.031
care	0.026	**lie**	-0.031

Much of what we see are words suggestive of factors that we would expect to influence the parole decision, as compared to words that represent inmate expression. "Sponsor" is likely an indicator of an inmate who has participated in substance abuse programming such as Alcoholics Anonymous, which the Board looks upon favorably as an indicator of rehabilitation and reform. "Alcoholic" is also among the top ten most important features of a parole grant among low risk inmates.

With respect to denials, the conjunction "would've" is the most important unigram. To fully understand its meaning, we would of course need to look at complete transcript text. But often, would've is part of the expression of past unreal conditionals—phrases used to express wishes about the past, often expressions of regret. "had I known A, I *would have*

19. http://uscpcjp.com/wp-content/uploads/2016/11/Parole-Manual-and-Case-Law-Chart.pdf

done B." We know the parole board looks for expression of remorse and regret, but perhaps this past hypothetical conditional appears to avoid full responsibility.

"Would've" is followed in importance by the word "guess". This may point to the parole board's desire for the inmate to claim ownership of their lives. For example, as shown in the snapshot below of a suitability hearing transcript, the commissioner pushes the inmate on the question of his honesty and ownership of his past transgressions - dealing drugs and pimping. "Were you a pimp at any time in your life?" The commissioner asks. "I guess - I guess I was" the inmate responds.

Figure 6. *Transcript Excerpt: "Guess"*

```
And you would deal drugs. Do you recall selling drugs?
INMATE MITCHELL: Yes, ma'am.
PRESIDING COMMISSIONER FRITZ: You do. Okay. I guess with the clinician
you initially denied selling drugs. Then noted, "I think I may have
sold drugs or something like that." And you were vague about having
pimped. Were you a pimp at any time in your life?
INMATE MITCHELL: I guess — I guess I was. I was, like, what do you
call a watcher, you know.
PRESIDING COMMISSIONER FRITZ: Now what does that mean?
INMATE MITCHELL: When you —- if you got somebody working, and you look
out for them and stuff like that.
PRESIDING COMMISSIONER FRITZ: Okay. So someone, maybe a prostitute
working for you.
INMATE MITCHELL: Yeah.
```

"Lie", also an important feature of denial, points to the centrality of truthfulness in the Board's assessment of insight, remorse, and rehabilitation. "If you're *lying* to the Panel today, what do you think that says about your rehabilitation?" a commissioner asked inmate # P-05816 at his suitability hearing. In another case, arguing against parole, the District Attorney states: "He is still trying to present his crime as an insurance fraud that turned into a robbery without his prior knowledge. The prisoner either has no idea why he committed the robbery or he's hiding the truth. Either way, that would make him a risk for future violence if paroled....The prisoner has committed perjury before on numerous panels and some have expressed concern about his truthfulness."

Finally, we also find a number of words predictive of denial that are indicative of negative behaviors."115" represents a rules violation report. This can be for a serious violations (e.g. fighting, possessing escape paraphernalia), or an administrative violations. The snapshot below shows an example on dialogue around an inmate's 115 for a relatively minor infraction of cursing to a guard after being asked to tuck in his shirt.

Figure 7. *Transcript Excerpt: "115"*

```
PRESIDING COMMISSIONER LABAHN: Okay. So the 115 indicates that after being told to tuck
your shirt in, that you responded fuck you; is that true?
INMATE COLEMAN: That was a while back.
PRESIDING COMMISSIONER LABAHN: That was ten years ago.
INMATE COLEMAN: Yes.
PRESIDING COMMISSIONER LABAHN: I'll tell you what, I'm a little younger than you, not tha
much, but if I told somebody in uniform fuck you, I'd probably remember it even ten years
ago. I think I'd remember saying that. Did you say that?
INMATE COLEMAN: If they said I said it, I said it, sir.
PRESIDING COMMISSIONER LABAHN: You don't remember it.
INMATE COLEMAN: Sometimes my mind gets a little foggy.
PRESIDING COMMISSIONER LABAHN: All right.
```

Examining Parts of Speech

We also examine building classification models on parts of speech - nouns, verbs, adverbs, and adjectives. Table 9 presents the model accuracy, precision, recall, and F1 score for the four classification models. We again use a logistic regression model with cross-validation; performance metrics are obtained on the withheld validation set. We find verbs and nouns perform much better than do adjective or adverbs. Adjective and adverbs are often taken in natural language processing to be indicators of subjectivity; they are part of speech used to convey sentiment. That they perform much less well in our classification task offers some indication that the parole decision is indeed grounded in the facts as the Board sees them, rather than in the more subjective way in which the inmate narrates them.

Table 9. *Comparing Parts of Speech*

NOUNS: 64% Accuracy

	Precision	Recall	f1-score
0	0.57	0.45	0.5
1	0.68	0.77	0.72
Avg/total	0.63	0.64	0.63

VERBS: 64% accuracy

	Precision	Recall	f1-score
0	0.57	0.49	0.53
1	0.68	0.75	0.71
Avg/total	0.64	0.64	0.64

ADVERBS: 60 % accuracy

	Precision	Recall	f1-score
0	0.54	0.27	0.36
1	0.6	0.83	0.7
Avg/total	0.58	0.59	0.55

ADJECTIVES: 57 % accuracy

	Precision	Recall	f1-score
0	0.45	0.1	0.17
1	0.58	0.91	0.71
Avg/total	0.53	0.57	0.48

Bigrams

The inherent drawback of the unigram model is its inability to capture relationships between two words (e.g. a word and its modifier, a word and its negation, etc), because it treats each word in isolation. In what follows, we consider bigram features. Table 10 presents the top ten features again using a logistic regression classification model

Many of the bigrams most predictive of grants are expressions of thanks, as we saw with the unigram analysis: "thanks yes", "thanks sir", "sir thank." There are also several affirmations: "thanks yes", "yeah yes". This again

Table 10. *Top 10 Most Important Bigram Features of Grant & Denial*

Grant			Denied		
Thank	Thank	0.105	I'm	trying	-0.042
Thank	Yes	0.041	don't	remember	-0.039
's	going	0.034	cell	phone	-0.036
's	house	0.031	don't	recall	-0.031
feel	like	0.030	don't	don't	-0.028
Yeah	Yes	0.027	don't	know	-0.028
Thank	sir	0.025	know	going	-0.028
sir	Thank	0.025	Yes	don't	-0.027
don't	care	0.025	I'm	saying	-0.025
don't	need	0.025	ask	forgiveness	-0.024

comports with advice given in the Post-Conviction Justice Project guide for Parole attorneys: "it can be extremely effective to have on record something as simple as your client testifying "Yes, I did the crime, I'm responsible and I'm sorry.""

On the other hand, we see a number of negative bigrams are predictive of denial: "don't remember", "don't recall", "don't don't"", "don't know." This difference in positive versus negative sentiment in grants versus denials is illustrative of one of the advantages of using bigrams: we are able to see this negative (and positive) syntatic structure, which would not have appeared when looking at unigrams alone.

Besides these bigrams expressing negative sentiment, several of the expressions are ones regarding memory: "don't remember", "don't recall", "don't know." This is consistent with the idea that the Board is looking for the inmate to have "insight" into their past misdeeds. Insight is of course impossible if you can't remember the set of events in question.

The snapshots below offers two examples from suitability hearings of the inmate telling the Board he doesn't remember. In the first case, he

tells the Board he doesn't remember committing the crime for which he is incarcerated: "I probably did it...I *don't remember* picking her back up again.... No, I have no idea why I did it, if I did it." Studies of parole

Figure 8. *Transcript Excerpt: "Don't Remember"*

```
I probably did it. I remember being with her the whole evening. I remember taking her home
and the child, too. I don't remember picking her back up again. No, we didn't have an
argument. No, I have no idea why I did it, if I did it.
```

release decision-making have consistently demonstrated that an inmate's willingness to acknowledge their culpability and express remorse for the commitment offense is a vital component of the parole decision calculus (Medwed 2007). Worse than not remembering, claims of innocence are essentially never looked upon favorably by the board, and substantially diminish the chances of parole. The attorney manual referenced above advises: "The Panel generally looks most favorably on clients who are respectful, penitent and truly appear to be remorseful...If your client maintains their innocence, it is possible for them to be remorseful about what happened to the victim even if they are not responsible for it." While the parole board cannot require the inmate to admit their guilt as a condition of parole (Cal. Penal Code 5011(b).), it is a much more challenging task to express remorse for an event that the inmate claims he had no part.

Figure 9. *Transcript Excerpt: "Don't Remember"*

```
INMATE STEPHENS: I don't remember saying I still have the occasional violent fantasies.
Now, as far as, you know, objectification of females, sometimes, yes. I mean, Playboy
magazine objectifies females, but I try to look at them as --
DEPUTY COMMISSIONER CHAMBERS: (Inaudible).
INMATE STEPHENS: Yeah, I try to look at them as humans and people because I got a little
thing that I do. I imagine them with a little stuffed animal, and all of a sudden, they
quit being a thing or an object and you can see the person, you can see the warmth, you
can see the glow.
DEPUTY COMMISSIONER CHAMBERS: So you view them as objects of sexual, for your sexual
gratification.
INMATE STEPHENS: Sometimes, long ago in the past, yeah.
```

In the snapshot of a transcript shown in figure 9, the inmate is not claiming innocence, or even denying the commitment offense. The ex-

change is with regard to his memory of an admission of sexual and violent fantasies.

Conclusion

Parole suitability hearings for those sentenced to prison in California for life with the possibility of parole was for decades a process without a purpose. Hearings were held, but the outcome was almost never in doubt: virtually no one was released. Things have changed. Parole now represents an important prison release valve in the California criminal justice system, yet it has received relatively little attention.

With a unique dataset, built using computer text processing, this chapter has offered a rigorous account of release decisions. The results suggest a system that contains some element of chance. Who an inmate happens to get as an evaluating psychologist or commissioner can alter whether they are released or remain in prison for many more years. Further, the analysis of variables suggests that, at least to some degree, factors that by law should not matter in the parole decision, such as the presence of victims at the hearing, may well have a substantial impact on how the Board decides. On the other hand, the psychological risk score, a measure of an inmate's dangerousness and risk of re-offending, which is the central criteria upon which the Board is required to make its decision, is indeed of great importance in the decision. And there is no evidence that the California parole commissioners exhibit the psychological effects that researchers have discovered in other adjudicators: the gambler's fallacy and decision fatigue appear to play little to no role in the California Parole Board's decision-making.

Our analysis of inmate speech during the suitability hearing helps to illustrate and confirm some of the conventional wisdom with respect to what matters at the hearing: the idea that the parole board looks most favorably on clients who are respectful, clearly state admission to the crime, and appear rehabilitated, for example. Our classification models do capture some signal, suggesting inmate narration at the hearing may

be of some importance. On the other hand, that the models are not more predictive suggests that inmate narration is likely not a critical determinant of the parole outcome. This is as we would hope for a system that strives to be rational, consistent, and objective.

References

. 2011. S188655. December.

Abrams, David, Marianne Bertrand, and Sendhil Mullainathan. 2012. "Do Judges Vary in Their Treatment of Race?" *Journal of Legal Studies* 41 (2): 347–383. Accessed October 7, 2015. `http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=1800840`.

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of causal effects using instrumental variables." *Journal of the American statistical Association* 91 (434): 444–455.

Ball, W David. 2011. "Tough on Crime (on the State's Dime): How Violent Crime Does Not Drive California Counties' Incarceration Rates-And Why It Should." *Ga. St. UL Rev.* 28:987.

Chen, Daniel L., Tobias J. Moskowitz, and Kelly Shue. 2014. "Decision-Making Under the Gambler's Fallacy: Evidence from Asylum Judges, Loan Officers, and Baseball Umpires." *Loan Officers, and Baseball Umpires (November 4, 2014).* Accessed October 5, 2015. `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2538147`.

Chen, Daniel L., and H. Spamann. 2014. *This morning's breakfast, last night's game: Detecting extraneous factors in judging.* Technical report. Working paper, ETH Zurich.

Danziger, Shai, Jonathan Levav, and Liora Avnaim-Pesso. 2011. "Extraneous factors in judicial decisions." *Proceedings of the National Academy of Sciences* 108 (17): 6889–6892. Accessed October 5, 2015. `http://www.pnas.org/content/108/17/6889.short`.

DeBacco, James. 2010. *Bridges to Freedom.* `https://bridgestofreedom.files.wordpress.com/2014/10/a-b2f-front-pages.pdf`.

Egelko, Bob. 2011. *Brown paroles more lifers than did predecessors*, April. `https://m.sfgate.com/crime/article/Brown-paroles-more-lifers-than-did-predecessors-2373800.php`.

Fischman, Joshua B. 2014. "Measuring inconsistency, indeterminacy, and error in adjudication." *American law and economics review* 16 (1): 40–85. Accessed October 5, 2015. `http://aler.oxfordjournals.org/content/16/1/40.short`.

Hopkins, Daniel J, and Gary King. 2010. "A method of automated non-parametric content analysis for social science." *American Journal of Political Science* 54 (1): 229–247.

In re Lawrence. 2008.

Kozinski, Alex. 1992. "What I ate for breakfast and other mysteries of judicial decision making." *Loy. LAL Rev.* 26:993.

Medwed, Daniel S. 2007. "The Innocent Prisoner's Dilemma: Consequences of Failing to Admit Guilt at Parole Hearings." *Iowa L. Rev.* 93:491.

Miller, Joshua B., Adam Sanjurjo, et al. 2015. *Surprised by the Gambler's and Hot Hand Fallacies? A Truth in the Law of Small Numbers.* Technical report. Accessed October 5, 2015. `ftp://ftp.igier.unibocconi.it/wp/2015/552.pdf`.

Mullane, Nancy. 2012. *Life after murder: Five men in search of redemption.* Public Affairs.

Nakosteen, Robert, and Michael Zimmer. 2014. "Approval of social security disability appeals: analysis of judges' decisions." *Applied Economics* 46 (23): 2783–2791. Accessed October 5, 2015. `http://www.tandfonline.com/doi/abs/10.1080/00036846.2014.914147`.

Petersilia, Joan. 2003. *When prisoners come home: Parole and prisoner reentry.* Oxford University Press.

Ramji-Nogales, Jaya, Andrew I. Schoenholtz, and Philip G. Schrag. 2007. "Refugee roulette: Disparities in asylum adjudication." *Stanford Law Review:* 295–411. Accessed October 5, 2015. `http://www.jstor.org/stable/40040412`.

Sarosy, Charlie. 2013. "Parole Denial Habeas Corpus Petitions: Why the California Supreme Court Needs to Provide More Clarity on the Scope of Judicial Review." *UCLA L. Rev.* 61:1134.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under uncertainty: Heuristics and biases." *science* 185 (4157): 1124–1131.

Ulloa, Jazmine. 2017. "More California inmates are getting a second chance as parole board enters new era of discretion." *Los Angeles Times* (July). `http://www.latimes.com/politics/la-pol-ca-parole-board-proposition-57-20170727-htmlstory.html`.

Van der Laan, Mark J., and Sherri Rose. 2011. *Targeted learning: causal inference for observational and experimental data.* Springer Science & Business Media. Accessed October 5, 2015. `https://books.google.com/books?hl=en&lr=&id=RGnSX5aCAgQC&oi=fnd&pg=PR3&dq=Targeted+Learning:+Causal+Inference+for+Observational+and+Experimental+Data&ots=FPb9cuAT7B&sig=P-RCh8efs-QtJYp1duNsq2W7Tbw`.

Weinshall-Margel, Keren, and John Shapard. 2011. "Overlooked factors in the analysis of parole decisions." *Proceedings of the National Academy of Sciences* 108 (42): E833–E833.

Weisberg, Robert, Debbie A. Mukamal, and Jordan D. Segall. 2011. *Life in Limbo: An Examination of Parole Release for Prisoners Serving Life Sentences with the Possibility of Parole in California.* Stanford Law School, Stanford Criminal Justice Center. Accessed October 5, 2015. `https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=264072`.