Defense Use of Digital Discovery in Criminal Cases: A Quantitative Analysis

Michael Braun
Southern Methodist University

Jenia I. Turner
Southern Methodist University

Ronald F. Wright
Wake Forest University

April 16, 2024

Recent criminal court reforms have required prosecutors to provide defense attorneys with broader and earlier discovery of evidence. For these discovery reforms to fulfill their aims of improved fairness and efficiency, defense attorneys must take advantage of the evidence disclosed by the prosecution. Prior studies suggest, however, that a range of factors, including low pay and high caseloads, impede effective defense representation. If similar factors hinder defense attorneys from reviewing discovery, discovery reforms would fail to meet their goals, and defendants would receive sub-standard representation.

The recent adoption of digital evidence platforms by local jurisdictions allows us to study whether defense attorneys consistently fulfill their duty to review discovery. Analyzing data from digital evidence platforms used in felony cases in four Texas counties between 2018 and 2020, we examine whether and when defense attorneys fail to access evidence disclosed by the prosecution. We find that a substantial number of defense attorneys never access the discovery. The access rate varies by county, year, offense seriousness, attorney category, attorney experience, and file type. Drawing on review of prior scholarship and Bayesian analysis of the data, we discuss plausible interpretations of these variations.

Keywords: discovery, disclosure, digital evidence, criminal defense, effective assistance of counsel

Introduction

In criminal cases, access to the evidence gathered by the prosecution is increasingly regarded as an essential precondition to a fair proceeding and a just outcome. Over the last two decades, states across the country have moved toward earlier and broader discovery in criminal cases (Moore, 2012; Turner & Redlich, 2016). As part of this trend, the Texas Legislature in 2013 mandated open-file discovery in criminal cases, requiring prosecutors to disclose to the defense virtually all evidence relevant to the case (Texas Code of Criminal Procedure art. 39.14).

Yet preliminary evidence on open-file discovery laws suggests that they have had no discernible effect on case dispositions (Grunwald, 2017). Prior research on defense attorney constraints offered one possible explanation: defense attorneys do not take full advantage of discovery in some cases, whether because of lack of time and resources or for other reasons (Grunwald, 2017). The possibility that defense attorneys are not reviewing evidence disclosed to them by the prosecution requires close examination for several reasons. First, the failure would frustrate the promise of open-file discovery laws to deliver fairer and more efficient dispositions. Second, the failure may constitute neglect of the attorneys' professional duty and may raise constitutional questions about the validity of any convictions following such neglect.

Digital evidence platforms allow prosecutors to upload their required discovery in electronic format. Those systems can then track defense attorney activity on the system as they view and download those files. The recent adoption of digital evidence platforms in many district attorney's offices around Texas allowed us to investigate whether—and if so, under what circumstances—defense attorneys fail to view or download discovery after the prosecution makes it available. To pursue this inquiry, we obtained and analyzed platform data from felony cases (n = 69,356) in four Texas counties from 2018 to 2020.

We find that a substantial number of defense attorneys never access the electronic discovery that prosecutors made available to them in felony cases. We analyze variations in the rate of discovery by county, offense seriousness, defense attorney experience level, defense attorney category (retained, public defender, appointed counsel), volume of discovery per case, and types of files uploaded. Our results can help explain why more liberal discovery rules do not appear to have made a significant impact on criminal case dispositions. The results also help identify factors that stand in the way of effective defense representation. These findings have implications for ineffective assistance of counsel litigation, strategies for defense attorney supervision, and criminal justice reform.

Discovery and defense attorney performance

Legal framework for discovery

Defense attorney access to evidence in criminal cases has long been considered a fundamental element of procedural fairness. The U.S. Supreme Court held in *Brady v. Maryland*, 373 U.S. 83 (1963), and subsequent cases that due process requires prosecutors to disclose all material exculpatory and impeachment evidence to the defense before trial. The *Brady* framework has gaps and flaws, however, and prosecutors' failure to disclose relevant evidence has contributed to numerous wrongful convictions over the years. Considering the limited effects of constitutional doctrine, many states have expanded discovery obligations through statutes and rules of criminal procedure (Brown, 2017; Turner & Redlich, 2016).

These broader discovery obligations were introduced to promote fairness in criminal cases. Legislation such as the Michael Morton Act in Texas and the open-file discovery law in North Carolina were adopted in response to wrongful convictions that had occurred as a result of discovery failures (Turner & Redlich, 2016). These statutes were expected to help prevent miscarriages of justice and make the process more just and efficient. Defense attorneys who see the strengths and weaknesses of the prosecutor's evidence could take realistic positions during plea negotiations and better prepare to answer that evidence at trial (Bibas, 2004). Proponents of the law also hoped that the free flow of information between the parties before trial would promote more efficient criminal proceedings as defense attorneys would not have to request specific items of evidence, and disputes over discoverable evidence would decrease (Turner & Redlich, 2016; Burke, 2009; Medwed, 2010; Uphoff, 1992).

Because discovery is central to fair dispositions in criminal cases, lawyers have professional obligations in most cases to review discovery as part of their broader duty to "investigate and engage investigators" (American Bar Association, 2017, § 4-4.1). An attorney's failure to review discovery may also violate a defendant's Sixth Amendment right to effective assistance of counsel (Strickland v. Washington, 466 U.S. 668 (1984)).

Studies of defense attorney performance

While states have adopted broader discovery laws over the last two decades, it remains an open question whether these laws have delivered greater fairness. An early empirical study of open-file discovery statutes in North Carolina and Texas found that these laws had not benefited defendants "in terms of charging, plea bargaining, and sentencing" (Grunwald, 2017, p. 777). The study's author hypothesized that heavy caseloads and a lack of resources might explain defense attorneys' failure to take advantage of broader discovery (Grunwald, 2017).

Other studies point to similar explanations for defense attorneys' failure to represent clients effectively (Brown, 2005). Relevant factors include overwhelming caseloads and financial incentives that reward quick disposition of cases. For example, prior studies have found less effort by defense attorneys in appointed cases, particularly when work is reimbursed at a flat rate (Agan, Freedman, & Owens, 2021; Anderson & Heaton, 2012; Cohen, 2014; Iyengar, 2007; Lee, 2021; Schwall, 2017). Research has also found that high caseloads, a significant problem for public defenders and appointed counsel in many jurisdictions, might hamper defense attorneys' investigative and legal efforts (Gottlieb & Arnold, 2021; Iyengar, 2007; Klein, 1986). In addition, some studies have concluded that attorney characteristics, such as the quality of the law school from which the attorney graduated and the experience level of the attorney, may affect outcomes for defendants (Abrams & Yoon, 2007; Iyengar, 2007; Roach, 2014).

Prior scholarship has also highlighted reasons why defense performance may be inferior in less-serious cases. Flat-rate payments are lower and caseloads higher for attorneys who represent defendants in misdemeanor and low-level felony cases (Texas Indigent Defense Commission, 2024), and these factors can hamper defense performance (Lee, 2021; Roberts, 2011). Appointed counsel report devoting less time to case-related tasks, including discovery review and investigation, as the seriousness of offenses charged (and the associated flat rate payment) diminishes (Carmichael et al., 2014).

The desire of some detained defendants to take an attractive early plea offer can also discourage defense discovery efforts in less serious cases. Recent scholarship has highlighted that detained defendants in minor cases are significantly more likely to plead guilty because offers for time served or probation would mean a prompt release from pretrial detention (Heaton, Mayson, & Stevenson, 2017; Roberts, 2011; Smith & Maddan, 2020). Some in-custody defendants who feel pressure to respond to plea offers quickly reportedly instruct their attorneys to forego review of discovery (Anonymous, 2024).

Another factor that increasingly burdens defense representation is the rapid growth of voluminous digital evidence in criminal cases (Brown, 2021; Kimpel, 2021; Turner, 2019). Storing and processing large digital files present defense attorneys with significant challenges, particularly when the files are disclosed in unfamiliar formats that require specialized and expensive proprietary software to access (Turner, 2019). Criminal defense attorneys are often solo practitioners who lack the resources to hire technology experts to help them handle complex or voluminous discovery (Turner, 2019). They may be "slow to adapt to new strategies," including to new technologies (Anderson & Heaton, 2012, p.198).

Based on interviews with criminal defense attorneys, Anonymous (2024) present qualitative analysis of factors that may predict failure to view or download discovery. First, difficulties with using the digital evidence platforms, especially in the early years after the platforms' introduction, and the lack of technological expertise by some attorneys (particularly older attorneys) were mentioned as reasons for the failure to access discovery in some cases. Second, attorneys suggested that the gravity of the charges might affect the decision whether to access the discovery, with attorneys in some lower-level cases being less likely to review discovery. Third, interviewees pointed to inadequate pay and high caseloads to explain why some attorneys—especially appointed counsel in counties with low flat rates of compensation—may not access discovery.

Data collection in the current study

A recent development allows us to study defense attorney discovery practices empirically. Court systems around the country have begun installing digital case management platforms that allow prosecutors to upload evidence and share it with the defense through the click of a button (Urban Institute, 2018).

To examine whether and when defense attorneys fail to access discovery in criminal cases, we analyzed data collected directly from one such digital platform. We submitted Public Information Act requests to seven prosecutor's offices for data concerning digital evidence in felony cases closed by these offices between January 1, 2018, and December 31, 2020. We received useable data from four counties, all of which use the same digital evidence platform. To maintain confidentiality of the data, we refer to these counties, in descending order of population, as Pentagon, Rectangle, Triangle, and Circle.[1] In the counties we studied, these digital platforms are the *only* mechanism by which defense attorneys can access discovery evidence.[2]

Generally speaking, the discovery platform records certain characteristics of the electronic files, the charges and cases to which the files are associated, when the files were uploaded and made available to the defense, and a log of the attorney's efforts to view or download each file. However, each county configures the platform differently and decides which data it retains and is willing to share. Given both the total volume of data, and the variation across counties in terms of which variables were available, we pre-processed the data to construct a common data structure that lets us focus on the most critical research questions.

We define an *evidence file* to be a computer file containing evidence related to a case. The evidence file is associated with a *charge record* that corresponds to a single incident, charge, or count. A typical charge record includes the incident date, an offense code or description that

1. Confidentiality also requires that we withhold the name of the platform.

2. There are limited exceptions where evidence must be viewed in a secure setting (e.g., sexually explicit images involving children).

connects the charge to a section of the Texas Penal Code, a unique identifier for the defense attorney, and the *attorney category*. The dataset also includes charge records that are not directly related to the offense itself (e.g., a subsequent probation violation or an administrative note).[3]

In all but Pentagon County, the attorney category indicates whether defense counsel was appointed or retained. Pentagon County also employs public defenders, but their implementation of the platform does not retain attorney category information in the charge records. Although the Pentagon County implementation of the platform does not include this information, it is available through court records. However, there was no way to automatically merge attorney categories from the court records database into the platform data. Instead, we manually looked up this information for a random sample of 24% of the charge records. In the tables in this paper, Pentagon refers to all cases in the platform database, and Pentagon* refers to this subset.[4] More details are in section 1.1 of the Web Appendix.

A *case* is comprised of one or more charges with a common state tracking number (TRN). The TRN is established at the time of initial arrest, so all related incidents, charges, and counts fall under it. Thus, it is the level of aggregation that most closely corresponds to how the courts would handle a single legal case. For example, a domestic violence case might consist of incidents on multiple dates, with multiple offenses (e.g., assault, protective order violation), resulting in multiple charges under the same TRN. We only considered *closed* cases that included at least one felony charge, and in which at least one evidence file was made available to the defense, from 2018 to 2020. More details about how the raw charge and case records were cleaned and processed appear in section 1.2 of the Web Appendix.

3. The dataset did not include accessible information on defendant race or gender.
4. This means that for the analyses of Pentagon County that follow, some are conducted on all available cases, and others on the random subset, depending on whether the attorney category variable is used.

Construction of variables and descriptive statistics

Case access as outcome variable

The outcome variables in our analyses depend on whether an evidence file is downloaded. A *download* includes any attempt (successful or not) by defense counsel to save an evidence file to local storage or to view it through the online platform.[5] None of the counties provided a complete log of all attempts to download a file. Instead, we have dates for only the most recent download attempt in each case. Triangle County provided download dates for individual evidence files, but Pentagon, Rectangle, and Circle Counties provided download dates only at the case level.

Given the limited information available about when specific files are downloaded, we say that a case is *accessed* if the defense attorney downloaded all, some, or one of the files. For example, if a case in Rectangle County has 10 discoverable files, that case is accessed if defense counsel downloads one or more of those files, regardless of the precise number.

Table 1 shows the number of cases in our dataset by county, the proportion of those cases that were not accessed, the number of distinct defense attorneys represented among those cases, and the proportion of those attorneys with at least one unaccessed case. The bottom row of Table 1 shows that a significant proportion of attorneys in the four counties, ranging from 36% to 61%, had at least one felony case in 2018-2020 for which they never downloaded any of the files dis-

5. The system cannot distinguish between a successful download from a download that was initiated but interrupted. In both cases, the system categorizes the file as downloaded. Nor can the system confirm if the defense physically viewed the contents of the file after saving it locally. Because of these limitations of the access log data, we consider streaming or viewing a file online as equivalent to downloading the file and saving it to local storage.

closed by the prosecution. Further, the second row shows that depending on the county, the percentage of felony cases in which no evidence was accessed by the attorney ranged from 4% to 27%. Differences among the counties in their practices regarding data retention and release prevent us from performing rigorous statistical analyses on variation across counties.

Table 1: Aggregate Data by County

	Pentagon	**Pentagon***	**Rectangle**	**Triangle**	**Circle**
Total cases	20,705	4,943	25,755	15,236	2,717
Percentage of cases *not* accessed	27%	26%	4%	19%	5%
Unique attorneys	840	547	785	511	86
Proportion of Attys with ≥1 unaccessed case	61%	61%	36%	61%	40%

Explanatory variables

Drawing on prior scholarship, while working within the constraints of the data, we identified several potential factors that could influence case access. The first factor we consider is *attorney category*, which indicates whether the attorney on the case is appointed, retained, or (for Pentagon County only), a public defender. For Pentagon, Rectangle, and Circle Counties, the unique identifier for defense counsel is that attorney's State Bar ID. By cross-referencing this ID with the State Bar of Texas membership database, we were able to collect limited demographic data about each attorney: gender, law school, and graduation data, and license date. We converted license date into a "years of experience" variable by computing the number of years elapsed until the start of our observation period. We do not have attorney demographics for Triangle County because the attorney identifier in that dataset is anonymized.

We also considered the nature of the offense for each case. The charge records in the dataset include 485 distinct textual descriptions of felony offenses. To manage the complexity of our analysis, we collapsed these codes into 15 *offense types*, and ordered them based on the severity of punishments ordinarily imposed in such cases, while adjusting upward some offenses with strong reputational effects for defendants. These offense types first appear in Table 2.

When a case contains charges with more than one offense, the most serious offense is controlling. Table 2 shows the offense types in descending order of seriousness. We classified homicide and sexual offenses as the "most serious," followed by crimes against individuals (e.g., assault, robbery, and burglary), DWI and felony-level traffic offenses (because they threaten injury and can carry substantial penalties), property crimes (including theft and fraud), and other offenses that are often secondary to more serious ones. Thus, a property damage case does not involve burglary or theft (which are both ranked as more serious) but may include weapons or drug charges (which we rank as less serious). This hierarchy focuses on the offense in a case that would be most likely to influence the decisions of a prosecutor or judge. For example, in the case of a sexual assault during a burglary, the sexual assault would drive a prosecutor's or judge's decisions more than the burglary charge. More information on how we assigned offenses to levels of the offense type variable are in section 2 of the Web Appendix.

Table 2 shows the proportion of unaccessed cases, and the proportion of attorneys with at least one unaccessed case, for each county, broken down by year of offense, attorney category, offense type, and range of years of attorney experience.

Table 2: Downloads by Offense and Attorney Characteristics

	Proportion of cases *not* accessed					Proportion of attorneys with ≥ 1 unaccessed case				
	Pent.	**Pent.***	**Rect.**	**Tri.**	**Circ.**	**Pent.**	**Pent.***	**Rect.**	**Tri.**	**Circ.**
Offense Year										
2018	30%	30%	5%	20%	7%	58%	54%	27%	56%	34%
2019	29%	28%	4%	19%	4%	61%	53%	30%	55%	27%
2020	22%	20%	3%	17%	3%	56%	46%	28%	54%	30%
Attorney Category										
Appointed		28%	3%	18%	5%		66%	54%	75%	42%
Retained		27%	6%	20%	4%		38%	25%	46%	16%
PD		19%					48%			
Offense Type										
Homicide	4%	9%	0%	4%	0%	6%	11%	0%	5%	0%
Sex offenses, children	8%	7%	1%	6%	3%	12%	9%	3%	9%	6%
Other sex offenses	17%	22%	1%	9%	8%	17%	24%	2%	10%	10%
Agg Crimes vs persons	21%	22%	3%	11%	4%	47%	35%	14%	27%	17%
Robbery	20%	19%	3%	10%	4%	42%	27%	9%	20%	9%
Burglary	27%	23%	4%	14%	5%	48%	32%	13%	34%	11%
Crimes vs persons	25%	26%	2%	13%	4%	50%	40%	11%	37%	16%
DWI/traffic	22%	24%	2%	8%	3%	30%	24%	8%	21%	8%
Theft or fraud	30%	29%	4%	22%	7%	67%	49%	30%	54%	38%
Property damage	29%	31%	4%	7%	0%	36%	36%	6%	11%	0%
Weapons violations	25%	27%	3%	21%	5%	40%	36%	8%	35%	10%
Drug violations	32%	29%	5%	24%	4%	63%	51%	31%	63%	25%
Evidence tampering	26%	29%	5%	24%	0%	28%	32%	9%	33%	0%
Evading arrest	31%	34%	5%	22%	8%	47%	41%	11%	32%	17%
Other offenses	28%	26%	5%	27%	6%	39%	26%	8%	32%	10%
Attorney experience										
0-4 years	21%	19%	3%		1%	44%	39%	31%		14%
4-10 years	21%	22%	4%		4%	53%	54%	34%		45%
10-20 years	28%	27%	3%		7%	65%	70%	35%		41%
20-60 years	29%	27%	5%		4%	65%	61%	38%		41%
Attorney gender										
Male	30%	28%	4%		5%	60%	59%	36%		40%
Female	22%	22%	4%		2%	66%	65%	36%		36%
Law School Rank										
Ranks 1-76	27%	26%	3%		5%	59%	58%	36%		45%
Ranks 77-148	27%	27%	5%		4%	64%	65%	36%		38%

File downloads as outcome variable

Triangle County is the only county that configured its discovery platform to retain download information for original files. Given the large number of files that prosecutors make available for discovery, our dataset is a 5% subsample, stratified by case (TRN) and file type. From the 15,236 Triangle County cases in our dataset, there are 508,532 distinct computer files containing digital evidence. For each file, we observe a file name, the date the file was made discoverable, and the date when the file was most recently accessed by the defense (if at all). The file naming scheme is not standardized, so we cannot reliably determine the contents of the file from its name. The best we can do is to infer the type of file using its extension (e.g., .docx, .jpg). We reduced the 126 file extensions to six file types: Document, Image, Video, Audio, Archive, and Other. Examples of the file extensions and typical content for each file type appear in section 3 of the Web Appendix. Table 3 shows the average number of files per case, and Table 4 shows the proportion of cases with at least one file of each type.

Table 3: Average Number of Files per Case, by File Type and Case Characteristics (Triangle County)

	Cases	Document	Image	Video	Audio	Other	Archive	All
All cases	**15,236**	**20.7**	**2.3**	**4.4**	**1.6**	**1.0**	**3.2**	**33.4**
Homicide	26	135.1	31.4	34.5	32.6	17.3	45.8	296.7
Sex offenses, children	122	41.4	7.2	5.8	5.2	2.2	6.3	68.0
Other sexual offenses	58	37.8	1.9	7.3	6.6	4.2	6.6	73.4
Agg crimes vs persons	1,279	32.7	8.5	9.8	4.8	1.6	6.5	63.9
Robbery	470	41.7	3.9	15.1	6.9	2.5	11.4	81.4
Burglary	767	26.8	4.1	8.5	3.2	1.8	4.7	49.1
Crimes vs persons	2,189	26.1	2.7	3.5	2.2	1.0	2.9	38.4
Property damage	118	22.3	5.7	4.5	2.4	2.5	3.8	41.2
Theft or fraud	2,557	17.6	1.5	3.4	1.0	1.1	2.2	26.8
DWI and other traffic	1,055	23.8	.4	4.0	.9	.9	3.1	33.2
Weapons violations	321	24.1	1.3	8.6	1.1	1.3	3.7	40.2
Drug violations	5,426	13.6	.8	2.5	.3	.5	2.0	19.6
Evidence tampering	198	12.1	2.5	2.1	.7	.6	1.9	19.9
Evading arrest	455	12.1	.3	3.3	.4	.5	2.2	18.9
Other offenses	195	19.9	6.7	1.6	1.4	1.1	1.4	32.1
Appointed counsel	10,389	20.0	1.7	4.5	1.6	1.0	3.1	31.9
Retained counsel	4,847	22.3	3.6	4.3	1.7	1.2	3.5	36.6
2018	3,708	16.1	1.0	.9	1.1	.6	1.6	21.4
2019	7,131	22.1	3.0	3.6	1.8	.9	3.1	34.6
2020	4,397	22.4	2.3	8.7	1.8	1.5	4.8	41.5

Table 4: Percentage of Cases with at Least One File of Each Type (Triangle County)

	Cases	Document	Image	Video	Audio	Other	Archive
All cases	**15,236**	**100%**	**16%**	**39%**	**45%**	**42%**	**70%**
Homicide	26	100%	85%	96%	100%	92%	96%
Sex offenses, children	122	99%	52%	66%	76%	60%	93%
Other sexual offenses	58	100%	45%	55%	71%	57%	84%
Agg. crimes vs persons	1,279	100%	31%	53%	88%	56%	88%
Robbery	470	100%	41%	76%	91%	70%	96%
Burglary	767	100%	32%	59%	79%	65%	85%
Crimes vs persons	2,189	100%	28%	46%	80%	45%	82%
Property damage	118	100%	27%	52%	67%	63%	81%
Theft or fraud	2,557	100%	12%	35%	39%	39%	56%
DWI and other traffic	1,055	100%	9%	49%	40%	60%	89%
Weapons violations	321	100%	15%	46%	37%	57%	72%
Drug violations	5,426	100%	5%	28%	20%	28%	60%
Evidence tampering	198	99%	9%	27%	19%	37%	53%
Evading arrest	455	100%	7%	34%	22%	38%	63%
Other offenses	195	100%	16%	23%	20%	36%	32%
Appointed counsel	10,389	100%	15%	39%	45%	40%	66%
Retained counsel	4,847	100%	18%	41%	46%	45%	78%
2018	3,708	100%	9%	14%	41%	31%	49%
2019	7,131	100%	17%	35%	46%	37%	69%
2020	4,397	100%	20%	68%	49%	58%	89%

In general, the total number of files is greatest for the most serious offenses, and the number of files per case increased from 2018 to 2020. Because the document file type includes procedural documents, it is present in nearly every case. But not all file types are represented in all cases. For example, audio files (possibly recordings of 911 calls) are present for most robbery and crimes against persons cases, but not for cases of evading arrest or evidence tampering.

Empirical analysis

Our empirical analysis consists of two Bayesian hierarchical logistic regressions. The outcome variables are binary indicators: case access for Model 1 and file download in Model 2. The first levels of the hierarchical models are similar to classical logistic regressions, where the log *odds* of the outcome variable is a linear function of covariates.[6] The second levels model heterogeneity in attorney effects, and for Model 2, the influence of case characteristics on the marginal effects of file type.[7] The marginal effect of a covariate is interpretable as an *odds ratio* relative to a reference group: an expected multiple of accessed cases in one group over the baseline group when both groups have the same number of unaccessed cases. An odds ratio of 1 is a "null result," indicating no difference in the propensity to access evidence between the focal and baseline cases.[8]

There are several reasons why Bayesian data analysis is appropriate for our analysis (Gelman & Hill, 2007, §11.5). The first is that the prior distribution on attorney-specific parameters also represents the mixture of attorneys' heterogeneous latent propensities to access evidence for a case. A typical econometric approach to incorporating heterogeneous attorney effects would be to define "fixed effects" for all attorneys as separate parameters. But these parameters would not be well-identified for attorneys who appear in only a very small number of observations. Those attorneys are involved in too few cases to isolate the effect of, say, offense type, from the unob-

6. The odds of case access is the expected multiple of accessed cases over unaccessed cases, and is related to the probability of access by $O_{ij} = p_{ij}/(1 - p_{ij})$. For example, odds of $O_{ij} = 3$ (i.e., "3-to-1 odds in favor") means that there is a $p_{ij} = .75$ probability that the case is accessed (since 0.75 / 0.25 = 3). There would be three times as many accessed cases as unaccessed cases.

7. For a description of hierarchical modeling, including hierarchical logistic regression, see chapters 11-14 of Gelman & Hill (2007). Our modeling approach is similar to that found in Braun et al. (2018).

8. Even though the logistic regression model is nonlinear, the odds ratio does not depend on the other parameters in the model. Therefore, we will report the marginal effects of our case characteristics in terms of odds ratios, rather than raw, untransformed parameters. More details about the specification and estimation of our models appear in the Web Appendix.

served tendencies of the attorney. In essence, attorneys who are sparsely represented in our dataset create multiple "small sample size" problems for their corresponding parameters. A potential remedy would be to drop all cases with attorneys whose case counts fall below some arbitrary threshold. But 26% of attorneys appear in only one case. Dropping their cases means throwing away valuable information and injecting selection bias into the estimates (because the dropped cases would not be randomly sampled).

Instead, our Bayesian approach induces dependence across heterogeneous parameters, where posterior estimates of parameters are "shrunk" toward a common mean for all attorneys in the population. That is, the posterior distribution of an attorney-specific effect draws on information contained not only in cases handled by that attorney, but also in the distribution of those effects across all attorneys in the dataset. This additional structure allows estimates for attorneys who are sparsely represented in the data to borrow information from the rest of the attorney population without resorting to an unwieldy number of attorney-level fixed effect parameters and interaction terms (Gelman & Hill, 2007). This partial pooling of information is a fundamental feature of Bayesian hierarchical models and lets us retain data from cases handled by low-volume attorneys that would otherwise be discarded.

Another reason we prefer Bayesian analysis in hierarchical models is interpretability of the results. Our quantitative analysis is meant to be descriptive, so we are not testing sharp null hypotheses of causal effects. Rather than focus on arbitrary cutoffs on *p*-values to assess the statistical significance of an effect, the Bayesian paradigm lets us measure the strength of an effect by considering the posterior probability that a value is greater than the null effect (i.e., an odds ratio is greater than 1), which we denote as $Pr>1$. In a Bayesian sense, the closer $Pr>1$ is to 0 or 1, the stronger our posterior beliefs should be that the odds ratio points in a particular direction

(Gelman & Tuerlinckx, 2000; Gelman & Carlin, 2014). When *Pr>1*=0.5, the posterior median is exactly 1, and is analogous to a null result.

Rather than ask the reader to decipher results from detailed numerical tables, we will present our results graphically as posterior intervals of marginal effects. The purpose of these figures is not to precisely estimate a probability of counsel viewing a case, or even to demonstrate statistical significance of an effect. Instead, we want to illustrate the relative odds ratios across levels of categorical variables. When viewing the upcoming figures, readers should focus on the proportions of the box plots that fall on either side of 1 (the odds ratio of a "null effect") and on the extent to which various box plots overlap. The *Pr>1* appear next to each box plot in the figures. Estimates based on more extensive or highly informative data (as in the larger counties) will be more precise with narrow posterior intervals, while estimates from Circle County, where prosecutors closed significantly fewer felony cases, will have wider intervals.

Model 1: Case Access

Equation 1 is the first level of Model 1, where Y_{ij} is a Bernoulli random variable that indicates whether case i of attorney j is accessed. Access occurs with probability p_{ij}. The log odds of case access (logit of p_{ij}) is a linear function of three categorical variables, each expressed as vectors of dummy variables: attorney category ($\mathbf{A_{ij}}$), discovery year ($\mathbf{W_{ij}}$) and offense type ($\mathbf{X_{ij}}$).[9] More detailed definitions of the mathematical symbols are in Section 4.1 of the Web Appendix.

9. $\mathbf{A_{ij}}$ includes an intercept, and dummy variables for retained counsel. For Pentagon County, $\mathbf{A_{ij}}$ also includes a dummy variable for public defenders. The reference level is appointed counsel. $\mathbf{W_{ij}}$ is a vector of dummy variables indicating a discovery year of 2019 or 2020 (2018 is the reference level). $\mathbf{X_{ij}}$ is a vector of 14 dummy variables for offense type, with robbery as the reference level. We chose

$$
\begin{aligned}
Y_{ij} &\sim \text{Bernoulli}(p_{ij}) \\
\text{logit}\, p_{ij} &= \boldsymbol{\beta}_j \mathbf{A}_{ij} + \boldsymbol{\Lambda} \mathbf{W}_{ij} + \boldsymbol{\Gamma} \mathbf{X}_{ij}
\end{aligned}
\tag{1}
$$

In Eq. 1, the coefficient $\boldsymbol{\beta_j}$ is heterogeneous across attorneys, while $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$ are homogeneous. The variation in $\boldsymbol{\beta_j}$ allows us to capture attorney-specific propensities to access cases for different attorney categories. Equation 2 describes this heterogeneity as the second level of our hierarchical model.

$$
\boldsymbol{\beta}_j \sim \text{MVN}\left(\boldsymbol{\Delta} \mathbf{Z}_j,\ \ \text{diag}(\tau) \boldsymbol{\Sigma}\, \text{diag}(\tau)\right)
\tag{2}
$$

In Eq. 2, $\mathbf{Z_j}$ is a vector of dummy variables indicating attorney experience: 4 to 10 years, 10 to 20 years, and more than 20 years (fewer than 4 years is the reference level). Thus, $\Delta \boldsymbol{Z_j}$ represents the expected case access propensity for an attorney with a given experience level. A more formal definition of the hierarchical model, including the covariance structure, and the hyper-priors for all model parameters (e.g., $\boldsymbol{\Delta}$, τ and $\boldsymbol{\Sigma}$) is in section 4.1.3 of the Web Appendix. Details on model estimation are in section 4.1.4 of the Web Appendix.

robbery as the reference level because it falls roughly in the middle of our hierarchy of severity. More details are in the Web Appendix.

Figure 1. Posterior intervals of odds ratios of case access, relative to appointed counsel

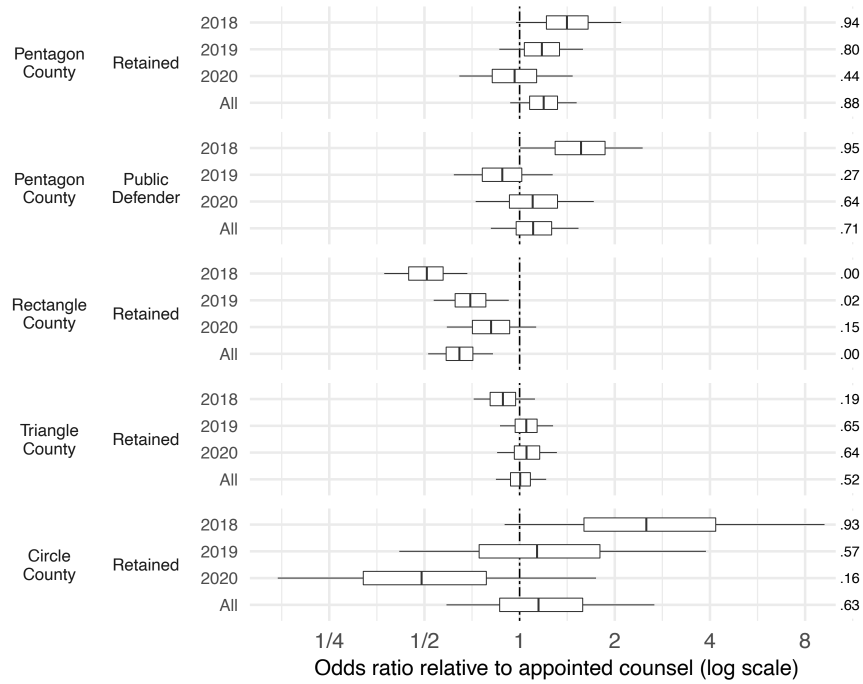


Figure 1 shows the posterior intervals of the odds ratio for attorney category, relative to the reference level of appointed counsel, broken down by year. In the top panel of Figure 1, Pentagon County retained counsel were more likely to access case evidence than appointed counsel in 2018, with a posterior median odds ratio of 1.41 (*Pr>1*=.94). That is, given an equal number of unaccessed cases between retained and appointed counsel in 2018, we estimate 41% more of the accessed cases will have been handled by retained than appointed counsel. But this difference disappeared by 2020 (posterior median odds ratio of 0.96; *Pr>1*=.44). In Rectangle County, we find a stronger effect in the opposite direction: retained counsel were *less* likely than appointed counsel to access case evidence, although that difference also attenuated over time. We consider possible explanations for these conflicting results in the Discussion section below.

Figure 2 shows the associations between case access and offense type. The box plots show posterior intervals of the odds ratios (marginal effects) relative to the robbery reference level.[10] The width and positioning of the intervals allows for comparisons across the remaining 14 offense types. We observe a general pattern that discovery is most likely to be accessed in cases concerning the most serious offenses—homicide and sexual offenses.

But there are some exceptions to this pattern. In Pentagon, Rectangle, and Triangle Counties, DWI and other traffic cases are more likely to be accessed than adjacent offenses in our ranking. This is consistent with qualitative findings in Anonymous (2024) that outcomes in DWI cases often turned on discovery such as lab results and videos of the field sobriety test.

When the number of cases within a particular offense type is small, the posterior interval for that type is wide enough that we cannot make sharp distinctions between offense types of similar seriousness. This is particularly true in Circle County, where the number of cases is small to begin with. But we can still observe general patterns without testing specific null hypotheses that compare offense types.

10. For example, in Pentagon County the posterior median odds ratio for sexual offenses involving children, relative to robbery, is 3.24. This means that for an equal number of unaccessed child sex offenses and robbery cases, we predict there would be more than three times as many child sex offenses would be accessed than robbery. In Triangle County, case access for DWI and Robbery cases are nearly equal.

Figure 2. Posterior intervals of odds ratios of case access, relative to robbery cases.

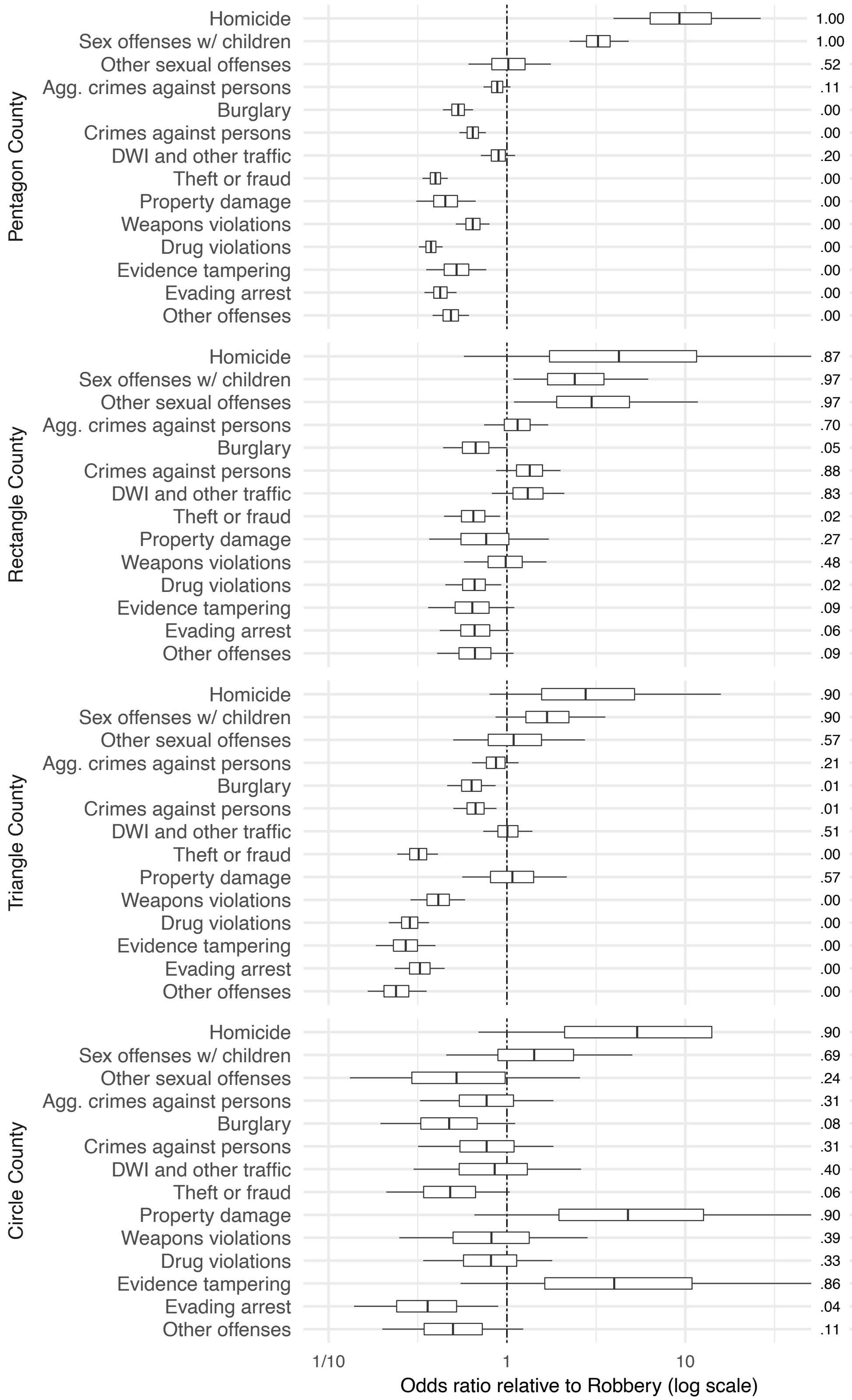
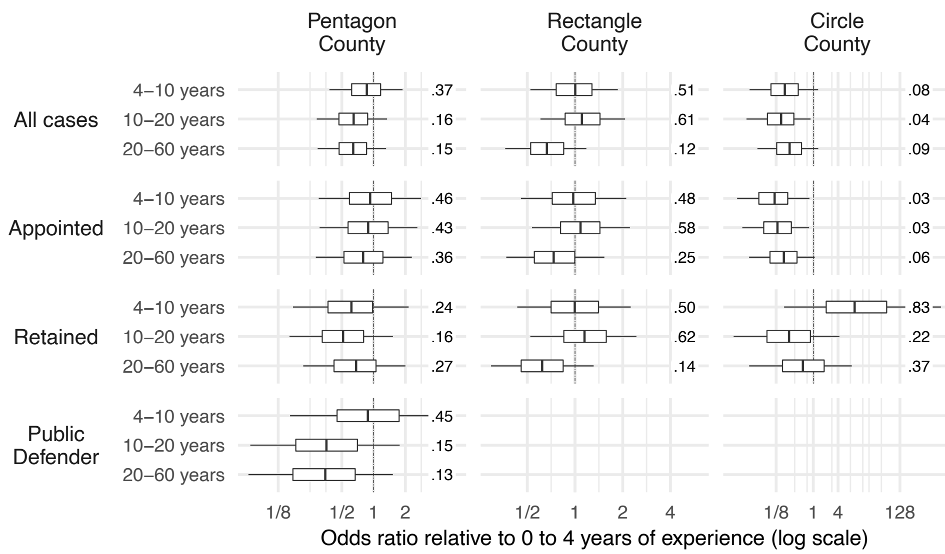


Figure 3 plots posterior distributions over odds ratios for attorney experience, relative to the 0-to-4-year reference level. We see that in two out of the three counties, the rate of case access is lower for attorneys with 20-60 years of experience. In all three counties, the rate of access is highest for attorneys with the least experience, 0-4 years. In other words, our analysis reveals an interesting pattern of more experienced attorneys (more than 4 years) being less likely to access evidence (with one exception being *retained* counsel in Circle County with 4 to 10 years of experience). We examine possible explanations for this pattern in the Discussion section.

Figure 3. Posterior intervals of odds ratios of case access, relative to attorneys with 4 or fewer years of experience.



Model 2: Evidence file access

Model 2 is a Bayesian hierarchical model that captures attorneys' propensities to download specific evidence files. It is applicable only to Triangle County, which was the only county to provide download data regarding specific evidence files. Equation 3 presents the first level of the model, where D_{kij} is a Bernoulli random variable, with probability q_{kij}, indicating whether file k for case i handled by attorney j is downloaded. The log odds of download (logit of q_{kij}) depends on α_j, an attorney specific random effect, and $\mathbf{T_{ij}}$, a vector of dummy variables for the file types: Image, Video, Audio, Archive, and Other (Documents is the reference level). Examples of which kinds of files are categorized into which file types appear in section 3 of the Web Appendix.

$$\begin{aligned} D_{kij} &\sim \text{Bernoulli}(q_{kij}) \\ \text{logit}\, q_{kij} &= \alpha_j + \mathbf{\Theta}_{ij}\mathbf{T}_{ij} \end{aligned} \tag{3}$$

Equation 4 shows the second level of the hierarchical model. The case-specific coefficients in Θ_{ij} depend on the same case characteristics as in Model 1: A_{ij} indicates appointed counsel, $\mathbf{W_{ij}}$ indicates year and $\mathbf{X_{ij}}$ indicates offense type. We also include the log of the number of files associate with the case (S_{ij}). The attorney effect, α_j, is a normal random variable with zero mean; because Triangle County de-identified defense attorneys, we cannot include attorney demographics in this model.

$$\begin{aligned} &\mathbf{\Theta}_{ij} \sim \text{MVN}\left(\bar{\mathbf{\Theta}},\ \ \text{diag}(\sigma)\mathbf{\Omega}\,\text{diag}(\sigma)\right) \\ &\bar{\mathbf{\Theta}}_{ij} = \delta A_{ij} + \mathbf{\Lambda W}_{ij} + \mathbf{\Gamma X}_{ij} + \rho \log S_{ij} \end{aligned} \tag{4}$$

The complete hierarchical model and information about the estimation process are described in section 4.2. of the Web Appendix.

Figure 4 presents estimated posterior intervals of expected probabilities that a file of each type is downloaded, after controlling for attorney effects and averaging over case characteristics and the number of files per case. On average, individual image files are more likely to be downloaded than documents, while video files are significantly less likely.

Figure 4. Posterior Intervals of Probabilities of Downloading Evidence Files

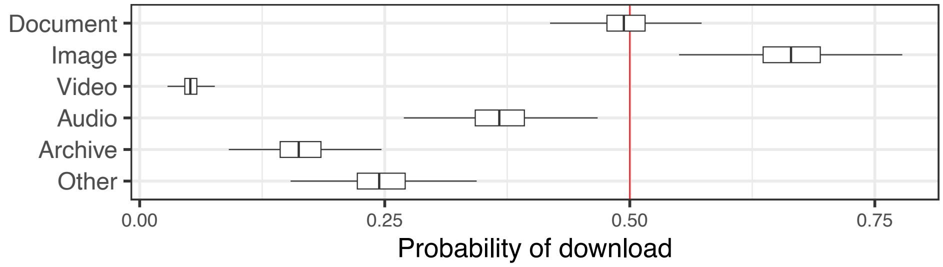
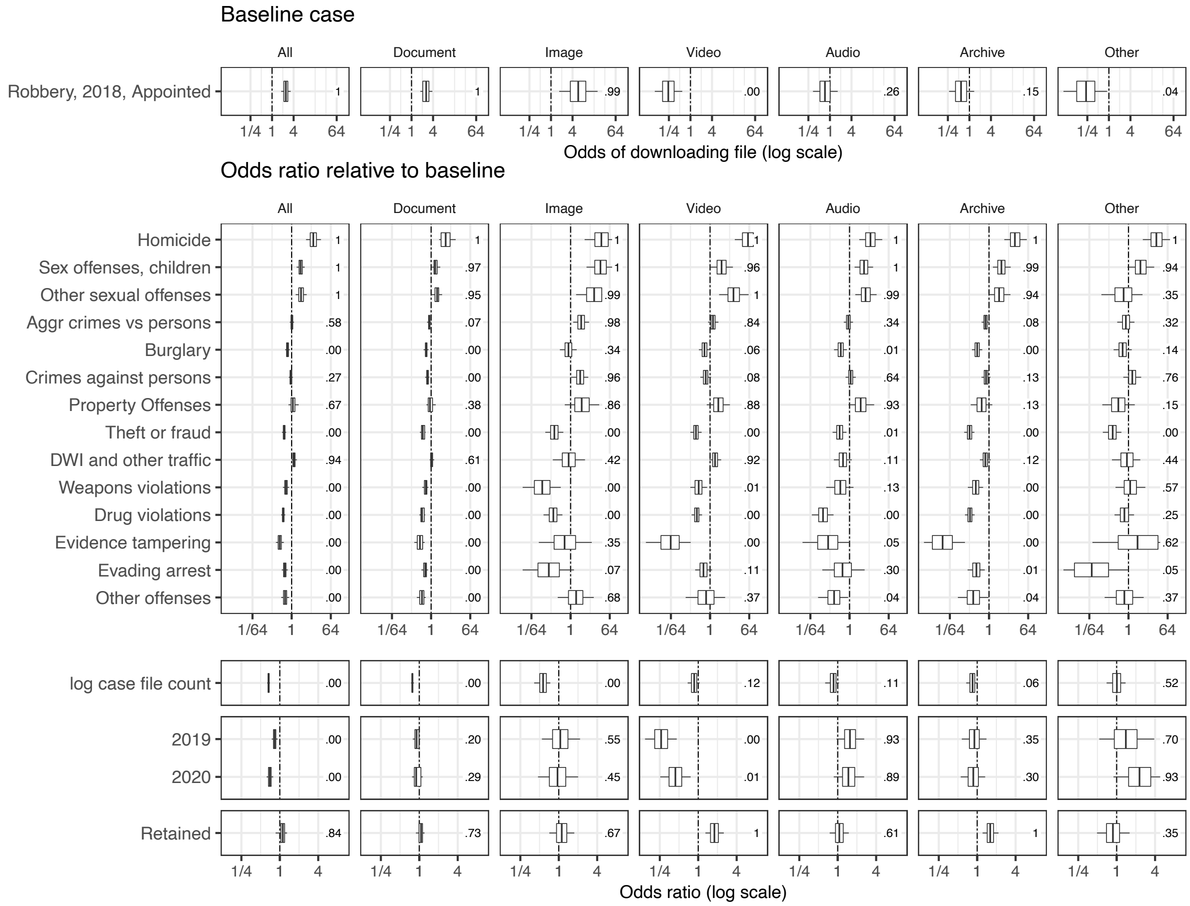


Figure 5 illustrates the estimated effects of the case covariates on file downloads. In the top row of plots, the posterior intervals are odds of download for the same reference levels as Model 1 – a robbery offense from 2018 defended by appointed counsel – along with a reference level of one file for the case. The posterior median odds of download of any file for the reference level is 2.43, which corresponds to a download probability of 0.81 (top left plot in Figure 5). But these posterior median odds vary by file type, from 5.77 for image files to 0.24 for video files. The remaining rows show the posterior distributions of odds *ratios* for file downloads, relative to the reference level. For example, the posterior median odds ratio for video files for burglary is 0.55. Suppose we had the same number of undownloaded files for burglary and robbery cases. This odds ratio for burglary means that we should expect a bit more than half the number of downloaded files for burglary than robbery cases.

We find that when prosecutors make more files discoverable, defense attorneys are less likely to download any one of those files. As Figure 5 indicates (in the first entry of the “log case file count” row), cases with a larger the number of files present lower odds that counsel will open and review any single file, compared to the baseline robbery case. This is consistent with the hypothesis that attorneys are less likely to view individual files in cases with voluminous evidence (Turner, 2019).

Figure 5. Posterior Intervals of odds and odds ratios for downloads of individual evidence files



In the aggregate (the left column of Figure 5), files were less likely to be downloaded in 2019 and 2020 than in 2018, which is the opposite effect that we observed for case access overall. However, that pattern appears to be driven mostly by video files. This could happen because of an increase in the availability of video evidence, such as police bodycam footage and recordings from home security systems, as reported in the literature (Turner, 2019).

Differences in horizontal positions within rows of Figure 5 represent statistical interactions between file type and case characteristics. Retained counsel are more likely to access video and archive files, compared to appointed counsel and document files. This is consistent with the hypothesis that counsel who are paid per hour are more likely to invest time in downloading and reviewing voluminous discovery than those paid a fixed fee.

Image and video files are even more likely to be accessed than document files for the most serious offenses. This is again consistent with the hypothesis that defense attorneys will be more diligent in downloading discovery when the stakes of the case are higher, and this holds even when the discovery is time-consuming to review.

Discussion

Our quantitative analysis of the novel dataset from discovery platforms yields several noteworthy findings about defense attorney use of digital evidence. Overall, we find that attorneys fail to access evidence files in a substantial portion of cases, with non-access rates ranging from 4 to 27 percent of felony cases in the counties we studied (see Table 1). Furthermore, a significant proportion of attorneys (36 to 61 percent) failed to download any evidence in at least one felony case.

In the counties we studied, prosecutors used the digital platform as the exclusive means of disclosing evidence in the vast majority of cases during 2018-20 (Anonymous, 2024). Therefore, whenever we see that attorneys failed to access the evidence on the platform, this means that they did not review it. In turn, this likely hampered their ability to investigate leads, negotiate with the prosecution, and prepare for trial. Conversely, downloading a file does not necessarily mean that an attorney reviewed the file. Therefore, the number of cases in which defense attorneys failed to read or view discovery was likely higher than our findings on access of discovery suggest.

Case access and offense type

There is a clear association between different access rates and the seriousness of the offense charged. As shown in Figure 2, attorneys engaged in higher levels of discovery activity in cases with more serious criminal charges. This finding is consistent with scholarship highlighting several factors that hinder effective representation in lower-level cases. Specifically, high caseloads (Gottlieb & Arnold, 2021; Iyengar, 2007; Klein, 1986) and low flat-rate pay (Lee, 2021; Schwall, 2017) pose challenges for defense lawyers who handle less serious cases.

Defendants who remain in jail while their charges are pending face strong pressure to plead guilty early in lower-level felonies, because a guilty plea leads to prompt release in those cases (Heaton, Mayson, & Stevenson, 2017; Roberts, 2011; Smith & Maddan, 2020). In fact, some prosecutors' offices offer misdemeanor charges or diversion out of criminal court entirely for defendants facing certain low-level felony charges – but only for defendants who accept the offer and end their cases promptly. It is therefore likely that some defendants in less serious cases instruct their attorneys to spend less time on discovery review (Anonymous, 2024). Such instructions from the client might justify the attorney's decision to move forward with guilty plea negotiations even before accessing the discovery files.

This association between case access and seriousness of the criminal charge did not hold true for all crime categories. Some charges that carry lesser statutory punishments (such as DWI) still produce unusually high rates of discovery access, possibly because the electronic files are more likely to hold relevant evidence (Anonymous, 2024).

Case access and attorney experience

We find that the attorneys with the fewest years of experience tended to be the most diligent in accessing discovery. As Figure 3 indicates, access rates go down, by and large, for attorneys with more than four years of practice experience.

The lower rate of case access among experienced lawyers might reflect less technological skill among older lawyers. When an attorney hits a technological barrier during an attempted access of discovery, the younger attorneys with greater fluency in electronic data management and online document platforms might persist until they succeed. Alternatively, more experienced attorneys may have greater confidence that they can evaluate the cases and represent some clients' interests without opening the electronic files at all. Whether or not this confidence among more experienced attorneys is justified, their years in practice might convince these attorneys that they can identify the cases where discovery is likely to add value to a case, based just on an interview with the client and an initial review of the charge. While this strategy is risky and draws disapproval from other defense attorneys (Anonymous, 2024), it might explain the mindset of some attorneys with more than four years of experience.

This finding is potentially at odds with prior research showing that less practiced criminal defense attorneys produced worse outcomes for their clients (Abrams & Yoon, 2007). But we did not test the association between discovery and case disposition, so it is possible that more experienced attorneys held other comparative advantages over junior attorneys, which allowed them to deliver better results for defendants even without reviewing discovery.

Case access and attorney compensation categories

Our findings – reflected in Figure 1 – are mixed for the potential association between access rates and the category of defense attorney assigned to a case: retained, appointed, or public defender. Prior literature suggests that retained attorneys have an incentive to provide more careful and complete service for their clients, who might take their cases to a competitor. Challenges such as low pay, flat fees, and high caseloads might cause appointed counsel to fail more often to access the evidence (Agan, Freedman, & Owens, 2021; Anderson & Heaton, 2012; Cohen, 2014; Gottlieb & Arnold, 2021; Iyengar, 2007; Klein, 1986; Lee, 2021; Schwall, 2017).

Consistent with this theory, retained attorneys in Pentagon County – and to a lesser extent Circle County – were more likely than appointed counsel to access case evidence.[11] Likewise, retained counsel were generally more likely to access video and archive files, compared to appointed counsel and document files. This is consistent with the hypothesis that counsel who are paid per hour are more likely to invest time in downloading and reviewing voluminous discovery than those paid a fixed fee.

In Rectangle County, on the other hand, we find a stronger effect in the opposite direction, with retained counsel *less* likely than appointed counsel to access case evidence. In Triangle County, we do not see a significant difference in the rate of access between appointed and retained counsel.

The appointed attorneys in Rectangle County may access discovery in their cases more often than retained attorneys because appointed lawyers in Rectangle are paid an hourly rate, creating incentives to invest more time reviewing discovery. By contrast, in Pentagon, Triangle, and Circle Counties, the presumptive payment method in appointed felony cases is a fixed fee. Fixed fees have been shown to disincentivize defense efforts (Agan, Freedman, & Owens, 2021; Lee, 2021; Schwall, 2017). Further research with data from additional counties would be helpful to determine whether the type of payment is indeed associated with lesser discovery efforts. Since

11. In Pentagon County, we also have data on access rates of public defenders. Public defenders are more likely than appointed counsel to access discovery. This is contrary to predictions that high caseloads for public defenders would discourage discovery access. On the other hand, better training and better technology infrastructure in a public defender's office may explain why public defenders perform better than appointed counsel (Cohen, 2014).

fixed-fee payments are merely presumptive in many Texas counties, data showing the type of payment in individual cases could also be used to explore this question. Such data were not available to us.

The hourly rate of compensation for appointed lawyers might also make a difference. In Triangle County, even when attorneys obtain the hourly rate, that rate is substantially lower than the hourly rate for surrounding counties (Anonymous, 2024). This may explain the relatively infrequent discovery access in Triangle County, as reflected in Table 1. Conversely, in Circle County, which also sets a presumptive flat rate, the flat rate is higher than the rate used by comparable counties in Texas. This may encourage the higher discovery access rate we observe in Circle.

Circle is also significantly smaller than the three urban counties we examined. The tight-knit nature of the courtroom community may help explain discovery diligence despite a presumptive fixed-fee compensation scheme. In a small legal community, criminal justice actors will interact with each other regularly, and reputation is likely to play a greater role in motivating behavior than it does in large urban areas (Battle, 1971; Schneider, 2007). Because the platform allows prosecutors (and other defense attorneys who inherit a case) to see if an attorney has failed to download discovery, defense attorneys in small legal communities may care about the reputational effects that failure to view discovery might have on their practice. Finally, analysis of caseloads across Texas found that in Circle and Rectangle counties, a smaller proportion of appointed attorneys are overburdened with cases than in Pentagon and Triangle (Davis et al., 2018). This could help explain why Circle and Rectangle County attorneys are more diligent in accessing evidence files for more cases.

In sum, our findings do not support a broad claim that retained attorneys perform better in discovery than appointed attorneys. Appointed lawyers did not systematically access discovery at a lower rate than other attorneys. Instead, the mixed results across different counties are better understood to support a narrower hypothesis: flat-fee payments for appointed lawyers produce lower rates of discovery access, while hourly payments for appointed lawyers lead to rates of access closer to the performance of retained lawyers. Consistent with this more targeted explanation, appointed attorneys did perform slightly worse than retained attorneys in two counties (Pentagon and Circle) where the courts relied mostly on flat-fee payments. On the other hand, in one county that paid appointed lawyers on an hourly basis (Rectangle), appointed attorneys performed better. While we do not have adequate case-level data about flat-fee versus hourly compensation to confirm this finding in a robust way, the differences among counties raise the question of whether the use of flat-fee payments has a negative impact on discovery performance.

Case access and file downloads over time

As shown in Table 2, our data confirm that the number of cases showing at least some discovery access improved over time from 2018 to 2020 in all but one county. Early technological problems apparently faded as defense attorneys improved their own skills and the technology infrastructure in their offices.

On the other hand, as Figure 5 shows, the number of individual files that attorneys left unopened increased over the years. This probably happened because video files became more common each year, and those files are the most burdensome type for defense attorneys to review. This is consistent with scholarship highlighting the special challenges of voluminous digital evidence such as video files (Brown, 2021; Kimpel, 2021; Turner, 2019).

File downloads and file types

Figure 4 indicates that attorneys downloaded individual image files more readily than they downloaded document files; conversely, they were significantly less likely to download video files than document files. We propose three possible explanations for this finding. First, video files are often too large and therefore difficult to download and store. Second, body camera videos tend to be repetitive and therefore at least some of them are seen as irrelevant to the attorney's efforts and the case outcome (Anonymous, 2024). Finally, because of their size, videos are also the most time-consuming to download, watch, and analyze, which is especially likely to discourage downloads for attorneys who are carrying heavy caseloads, are working on a short timeline to respond to an exploding plea offer, or are paid a flat rate.

File downloads and overall file volume

Our results, as presented in Figure 5, also reveal that defense attorneys are less likely to download any files at all in cases with the largest number of discoverable files, after controlling for offense type and other covariates. When a defense attorney encounters two cases with similar features, the attorney is less likely to complete a download of any files at all if the overall list of files listed on the discovery platform is longer and more daunting. The platform offers limited clues about the content of the files available for discovery. Thus, the sheer number of uploaded files in the case might offer the attorney the only basis for a preliminary guess about the amount of time required to complete discovery.

Implications

Our analysis spotlights the failure by a significant number of defense counsel to perform a critical task in representing their clients. This failure may violate defendants' constitutional right to effective assistance of counsel (Anonymous, 2024), and it undercuts the promise of discovery reform to promote fairness in the criminal process. It should prompt policymakers to consider reforms that remove structural barriers to effective representation and better deter discovery neglect in individual cases. These reforms might include technological improvements to the discovery platforms, intensified training of defense attorneys and their support staff, revisions to attorney compensation methods, and increased monitoring of defense attorney compliance with their discovery obligations (Anonymous, 2024).

But to fully understand when and why defense attorneys disregard their discovery duties, and to shape reform efforts accordingly, further research would be useful. As digital evidence platforms become more common, researchers could expand the number of jurisdictions to study. Future research would especially benefit from studying vouchers documenting individual payments to defense attorneys within counties that rely on both flat-rate and hourly payments. Scholars could also obtain relevant data to test the associations between caseloads and access rates, while noting that many criminal defense attorneys work in multiple counties, handle both state and federal cases, and work on civil as well as criminal cases. If researchers could obtain the relevant data, they could also analyze whether there is an association between access rates and the defendant's detention status and guilty plea timing.

Finally, scholars with access to more detailed and comprehensive data might be able to examine whether discovery access is correlated with better outcomes for the defendant. While legal doctrine and scholarship presume that discovery is an essential element of effective representation, this question deserves closer empirical examination. In circumstances where an early guilty plea results in more favorable treatment, foregoing discovery to obtain the advantages of such a plea bargain may be advantageous to the average defendant (Anonymous, 2024).

Conclusion

Scholarship on defense representation in criminal cases has documented various factors that hinder defense attorneys from performing adequately in criminal cases. Our analysis of digital evidence platform data uncovers a previously undocumented failing of defense attorneys to review the evidence disclosed by the prosecution. Reviewing discovery is a critical function of criminal defense attorneys, and failure to do so amounts in some cases to constitutionally deficient representation.

Analysis of the data suggests that attorneys are more likely to neglect their duties to access discovery in cases featuring less serious offenses. Low pay for appointed counsel in flat-fee jurisdictions, high caseloads for public defenders, and a deluge of (often repetitive) digital discovery also appear to limit attorneys' capacity to review evidence. Contrary to expectations, more experienced attorneys were less likely to access discovery.

Some of these problems appear to become less serious over time. But the failure to access rates are significant and persistent enough in some counties that state courts, legislatures, and bar associations ought to analyze their own data and consider reforms to deliver for criminal defendants their rights to full discovery and effective assistance of counsel.

Acknowledgements

The authors thank Mary Liz King and Leela Orbidan for their excellent research assistance, D.A.'s Offices, and the Texas Indigent Defense Commission for providing data for the project. We are also grateful to participants in the Duke Empirical Criminal Law Roundtable for valuable comments.

References

Abrams, D. & Yoon, A. (2007). The luck of the draw: Using random case assignment to investigate attorney ability, *University of Chicago Law Review, 74*(4), 1145–1177.

Agan, A., Freedman, M. & Owens, E. (2021). Is your lawyer a lemon? Incentives and selection in the public provision of criminal defense, *The Review of Economics and Statistics, 103*(2), 294–309.

American Bar Association. (2017). Criminal justice standards: Defense function.

Anonymous. (2024). Publication details withheld to safeguard blind peer-review process.

Anderson, J. M. & Heaton, P. (2012). How much difference does the lawyer make? The effect of defense counsel on murder case outcomes, *Yale Law Journal, 122*(1), 154–217.

Battle, J. B. (1971). In search of the adversary system: The cooperative practices of private criminal defense attorneys, *Texas Law Review, 50*(1), 60–118.

Bibas, S. (2004). Plea bargaining outside the shadow of trial, *Harvard Law Review, 117*(8), 2463–2547.

Braun, M., Rosenthal, J. & Therrian, K. (2018). Police discretion and racial disparity in organized retail theft arrests: evidence from Texas, *Journal of Empirical Legal Studies 15*(4), 916–950.

Brown, D. K. (2021). Disclosure, security, technology. In R. F. Wright, K. L. Levine & R. M. Gold (eds.), *The Oxford handbook of prosecutors and prosecution* (101–116). Oxford University Press.

Brown, D. K. (2017). Discovery. In E. Luna (ed.), 3 *Reforming criminal justice: Pretrial and trial processes* (147–170). Arizona State University.

Brown, D. K. (2005). The decline of defense counsel and the rise of accuracy in criminal cases, *California Law Review, 93*(6), 1585–1645.

Burke, A. (2009). Revisiting prosecutorial disclosure, *Indiana Law Journal, 84*(2), 481–518.

Carmichael, D., Clemens, A., Caspers, H., Marchbanks, M.P., & Wood, S. (2014). *Guidelines for indigent defense caseloads: A report to the Texas Indigent Defense Commission*. Public Policy Research Institute.

Cohen, T. H. (2014). Who is better at defending criminals? Does type of defense attorney matter in terms of producing favorable case outcomes, *Criminal Justice Policy Review, 25*(1), 29–58.

Davis, N. T., Naufal, G., Caspers, H., & Burkhart, G. T. (2018). Indigent defense caseloads in Texas: Assessing the extent of high-volume practice. [unpublished manuscript]. http://www.nicholastdavis.com/wp-content/uploads/2018/08/caseload_draft.pdf.

Gelman, A. & Carlin, J. (2014). Beyond power calculations: Assessing Type S (Sign) and Type M (magnitude) errors, *Perspectives on Psychological Science*, 9(6), 641–651.

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A. & Tuerlinckx, F. (2000). Type S error rates for classical and Bayesian single and multiple comparison problems, *Computational Statistics, 15*, 373–390.

Gottlieb, A. & Arnold, K. (2021). The effect of public defender and support staff caseloads on incarceration outcomes for felony defendants, *Journal of the Society for Social Work and Research, 12*(3), 569–589.

Grunwald, B. (2017). The fragile promise of open-file discovery, *Connecticut Law Review*, *49*(3), 771–836.

Heaton, P., Mayson, S. G. & Stevenson, M. (2017). The downstream consequences of misdemeanor pretrial detention, *Stanford Law Review, 69*(1), 711–794.

Iyengar, R. (2007). An analysis of the performance of federal indigent defense counsel, NBER Working Paper No. 13187.

Kimpel, A. (2021). Violent videos: Criminal defense in a digital age, *Georgia State University Law Review, 37*(2), 305-426.

Klein, R. (1986). The emperor Gideon has no clothes: The empty promise of the constitutional right to effective assistance of counsel, *Hastings Constitutional Law Quarterly, 13*(4), 625–693.

Schneider, A. K. (2007). Cooperating or caving in: Are defense attorneys shrewd or exploited in plea bargaining negotiations?, *Marquette Law Review, 91*(1), 145–162.

Lee, A. J. (2021). Flat fee compensation, lawyer incentives, and case outcomes in indigent criminal defense. [Unpublished manuscript]. Department of Economics, University of Texas at Austin.

Medwed, D. S. (2010). *Brady*'s bunch of flaws, *Washington and Lee Law Review, 67*(4), 1533–1567.

Moore, J. (2012). Democracy and criminal discovery reform after *Connick* and *Garcetti*, *Brooklyn Law Review, 77*(4), 1329–1388.

Roach, M. A. (2014). Indigent defense counsel, attorney quality, and defendant outcomes, *American Law and Economic Review, 16*(2), 577–619.

Roberts, J. (2011). Why misdemeanors matter: Defining effective advocacy in the lower criminal courts. *University of California Davis Law Review, 45*(2), 277–372.

Schwall, B. (2017). More bang for your buck: How to improve the incentive structure for indigent defense counsel, *Ohio State Journal of Criminal Law, 14*(2), 553–578.

Smith, A. & Maddan, S. (2020). Misdemeanor courts, due process, and case outcomes, *Criminal Justice Policy Review, 31*(9), 1312–1339.

Texas Indigent Defense Commission. (2024). *Current attorney fee schedules*. http://tidc.tamu.edu/public.net/Reports/FeeDocuments.aspx.

Turner, J. I. (2019). Managing digital discovery in criminal cases, *Journal of Criminal Law and Criminology, 109*(2), 237–312.

Turner, J. I. & Redlich, A. D. (2016). Two models of pre-plea discovery in criminal cases: An empirical comparison, *Washington and Lee Law Review, 73*(1), 285–408.

Uphoff, R. J. (1993). Criminal Discovery in Oklahoma, *Oklahoma Law Review, 46*(3), 381–414.

Urban Institute (2018). Collecting and using data for prosecutorial decisionmaking: Findings from 2018 National survey of state prosecutors' offices.

Web Appendix

Defense Use of Digital Discovery in Criminal Cases: A Quantitative Analysis

1 Data Processing

1.1 Pentagon County Attorney Categories

The digital discovery platform for Pentagon County does not contain information about attorney type: appointed, retained, or public defender. The data are present in the county court system database, indexed by internal case number. We had to look up this information manually. We randomly selected 24% of charge records and asked research assistants to look up the attorney category information in the county court database. The internal case number was also available in the digital evidence platform data, so we cross-referenced based on that number and merged into our existing data set the attorney category information that our RAs had found.

1.2 Charge Record Clean-Up

The following charge records were dropped from the analysis:

1. "Phantom" records with missing TRN numbers (these charges cannot be grouped together as distinct cases), offense types, or defendant identifiers.
2. Status changes for probation or diversion programs (e.g., probation is completed or revoked). These records do not constitute a new charge, even though they are linked to the original case. The original charge records were retained.
3. Likely duplicates or non-cases (e.g., disposition flags like "Administrative closure," "Consolidated," "Transferred," "Case Refiled")
4. Charges rejected (e.g., no-billed by a grand jury, on the request of a complaining witness or a law enforcement agency, "interest of justice," "prosecutorial discretion," defendant is unapprehended or deceased, "expired," or "other").

Additionally, we dropped cases where the defense attorney category was ambiguous (e.g., cases with multiple charges where counsel for some charges are retained and others are appointed), or where the TRN appears in more than one county.

2. Categorizing Offense Types

In Texas, all arrests (for crimes that are Class B misdemeanors or higher) are recorded in the state Computerized Criminal History (CCH) system, managed by the Department of Public Safety (DPS). The CCH documentation provides a list of offense descriptions and the corresponding statute and section to which the offense refers. Because statutes are amended from year to year, we merged four versions of the offense lists that were active from 2018 to 2020.

We filtered this list to identify unique felony offenses that appear in our dataset. For records that did not match the CCH list exactly, we manually identified the appropriate statute, chapter, and section. We then arranged the list by statute and section to ensure that all similar offense descriptions were grouped together. This process gave us a list of 485 distinct descriptions of felonies that appear in at least one charge record.

Next, we assigned an offense type to each offense description. Despite the large number of distinct offense descriptions, many were similar enough that offense types could be determined in blocks (e.g., we classified 110 distinct offense descriptions related to the Texas Controlled Substances Act as "drug violations"). To manage the complexity of our analysis, we collapsed these codes into 15 offense types and ordered them based on the severity of punishments ordinarily imposed in such cases, while adjusting upward some offenses with strong reputational effects for defendants. The classification of offenses to offense types were based on our professional judgment as law professors and/or former prosecutors, and it was guided by how an offense would be handled by a prosecutor or judge.

We then assigned offense types to charge records based on the offense description.

3. File types

For Model 2, file types were inferred by the file name extension, according to the classification in Table A1. Overall, there were 126 different file name extensions represented in the Triangle County dataset. Many of these are associated with executable applications, such as a viewer for a video file in a proprietary format. We did not include those files in our analysis.

Table A1: Triangle County Evidence File Types, Based on File Name Extension

File Type	**Example file extensions**	**Typical content**
Document	.docx, .pdf, .txt, .ppt, .xlsx	Printed evidence, warrants, notices, motions, call logs
Image	.jpg, .tiff, .png	Evidence (e.g., photos of injuries)
Video	.mp4, .mov, .avi, proprietary formats	Surveillance video, police body cameras, interviews
Audio	.mp3, .wav	911 calls, witness interviews, jail calls
Archive	.zip	Multiple files of any type
Other	.xml, .eml, .asx	Applications, metadata, cell phone logs, music playlists.

4. Bayesian hierarchical model notation and specifications

Scalar values (including indicator values) are in italics. Vectors and matrices are in bold.

4.1 Model 1

4.1.1. Data

Y_{ij}	Binary indicator of access to case i for attorney j	
$\mathbf{A}_{ij}$	$= \begin{cases} (1, A_{ij}^{\text{Ret}}, A_{ij}^{\text{PD}}) & \text{for Pentagon County} \\ (1, A_{ij}^{\text{Ret}}) & \text{for all other counties} \end{cases}$	Intercept and dummy variables for attorney category (reference: appointed)
$\mathbf{W}_{ij}$	$= (W_{ij}^{(19)}, W_{ij}^{(20)})$	Dummy variables for discovery year (reference: 2018)
$\mathbf{X}_{ij}$	$= (X_{ij}^{(1)}, \dots, X_{ij}^{(14)})$	Dummy variables for offense type (reference: robbery)
$\mathbf{Z}_j$	$= (Z_j^{(4-10)},\ Z_j^{(10-20)},\ Z_j^{(20+)})$	Dummy variables for attorney experience (reference: <4 yrs)

4.1.2. Parameters

p_{ij}		Probability of case access
$\boldsymbol{\beta}_j$	Coefficient vector on A_{ij}	2×1 vector (3×1 for Pentagon County)
$\boldsymbol{\Lambda}$	Coefficient vector on $\mathbf{W}_{ij}$	2×1 vector
$\boldsymbol{\Gamma}$	Coefficient vector on $\mathbf{X}_{ij}$	14×1 vector
$\boldsymbol{\Delta}$	Coefficient matrix on $\mathbf{Z}_j$	2×3 matrix (3×3 for Pentagon County).
$\boldsymbol{\Sigma}$	Correlation matrix of $\boldsymbol{\beta}_j$	2×2 matrix (3×3 for Pentagon County)
$\boldsymbol{\tau}$	Scaling vector on $\boldsymbol{\Sigma}$	2×1 vector (3×1 for Pentagon County)

4.1.3. Hierarchical model specification

$$
\begin{aligned}
Y_{ij} &\sim \text{Bernoulli}(p_{ij}) \\
\text{logit}\, p_{ij} &= \boldsymbol{\beta}_j \mathbf{A}_{ij} + \boldsymbol{\Lambda}\mathbf{W}_{ij} + \boldsymbol{\Gamma}\mathbf{X}_{ij} \\
\boldsymbol{\beta}_j &\sim \text{MVN}\left(\boldsymbol{\Delta}\mathbf{Z}_j,\ \text{diag}(\tau)\boldsymbol{\Sigma}\text{diag}(\tau)\right) \\
\text{vec}\boldsymbol{\Delta} &\sim \text{MVN}(\text{vec}\boldsymbol{\Delta}_0, 16) \\
\boldsymbol{\Lambda} &\sim \text{MVN}(0,16) \\
\boldsymbol{\Gamma} &\sim \text{MVN}(0,16) \\
\tau &\sim \text{Half-MVN}\left(0, \sqrt{\frac{\pi}{2}}\right) (\text{so that } E(\tau) = 1) \\
\boldsymbol{\Sigma} &\sim \text{LKJ}(1)
\end{aligned}
$$

where

$$
\boldsymbol{\Delta}_0 = \begin{cases} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} & \text{for all counties except Pentagon} \\ \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \text{for Pentagon County} \end{cases}
$$

The LKJ(1) distribution (Lewandowski, Kurowicka and Joe 2009) is a uniform prior on the correlation matrix on β_j (see Stan Development Team 2024, section 1.13). The use of the half-MVN distribution and the scaling vector τ is described in Gelman (2006).

4.1.4. Estimation

We sampled from posterior distributions using the default "No-U-Turn" sampler (Hoffman and Gelman 2014), as implemented by the Stan modeling and computational platform (Stan Development Team 2024). We ran 32 independent Hamiltonian Monte Carlo (HMC) chains in parallel, each for 1,000 warmup and 1,000 post-convergence iterations, and retained 10% of the post-convergence iterations for inference. Raw parameter samples were then transformed into odds and odds ratios. The empirical distributions of these transformed samples comprise the plots in the results figures.

4.2. Model 2

4.2.1. Data

D_{kij}		Binary indicator of download of file k for case i of attorney j
r_{ij}		Binary indicator of whether attorney j is retained for case i
$\mathbf{W}_{ij}$	$= \left(W_{ij}^{(19)}, W_{ij}^{(20)}\right)$	Dummy variables for discovery year
$\mathbf{X}_{ij}$	$= \left(X_{ij}^{(1)}, \ldots, X_{ij}^{(14)}\right)$	Dummy variables for offense type
$\mathbf{T}_{kij}$	$= \left(T_{kij}^{(1)}, \ldots, T_{kij}^{(5)}\right)$	Dummy variables for file type
S_{ij}	Log number of files for case i of attorney j.	
$\boldsymbol{\tau}$	Scaling vector on Ω	XXX

4.2.2 Parameters

q_{kij}	Probability of file download	
$\boldsymbol{\Theta}_{ij}$	Coefficient vector on $\mathbf{T}_{ij}$	5×1 vector
$\boldsymbol{\Lambda}$	Coefficient vector on $\mathbf{W}_{ij}$	2×1 vector
$\boldsymbol{\Gamma}$	Coefficient vector on $\mathbf{X}_{ij}$	14×1 vector
$\boldsymbol{\Omega}$	Correlation matrix of $\boldsymbol{\Theta}_{ij}$	5×5 matrix
ρ	Coefficient on S_{ij}	scalar
$\boldsymbol{\sigma}$	Scaling vector on Ω	5×1 vector

4.2.3. Hierarchical model specification

$$
\begin{aligned}
D_{ij} &\sim \text{Bernoulli}(q_{ij}) \\
\text{logit}\, q_{kij} &= \alpha_j + \mathbf{\Theta}_{ij}\mathbf{T}_{ij} \\
\alpha_j &\sim N(0,9) \\
\mathbf{\Theta}_{ij} &\sim \text{MVN}(\bar{\mathbf{\Theta}},\ \text{diag}(\sigma)\mathbf{\Omega}\text{diag}(\sigma)) \\
\bar{\mathbf{\Theta}}_{ij} &= \delta r_{ij} + \mathbf{\Lambda W}_{ij} + \mathbf{\Gamma X}_{ij} + \rho \log S_{ij} \\
\sigma &\sim \text{Half-MVN}\left(0, \sqrt{\frac{\pi}{2}}\right) (\text{so that } E(\sigma) = 1) \\
\mathbf{\Omega} &\sim \text{LKJ}(1) \\
(\mathbf{\Lambda}, \mathbf{\Gamma}, \delta, \rho) &\sim \text{MVN}(0,16)
\end{aligned}
$$

Similar to Model 1, the LKJ(1) distribution (Lewandowski, Kurowicka and Joe 2009) is a uniform prior on the correlation matrix on Θ_{ij} (see Stan Development Team 2024, section 1.13). The use of the half-MVN distribution and the scaling vector σ is described in Gelman (2006).

4.2.4. Estimation

Given the large number of file records in the dataset, we estimated the model on a 20% random sample of cases. The model was estimated using the same method and tools as Model 1. We ran 30 independent Hamiltonian Monte Carlo (HMC) chains in parallel, each for 100 warmup and 100 post-convergence iterations, and retained all post-convergence draws for inference. Raw parameter samples were then transformed into odds and odds ratios. The empirical distributions of these transformed samples comprise the plots in the results figures.

Web Appendix References

Gelman, Andrew (2006). Prior Distributions for Variance Parameters in Hierarchical Models. Bayesian Analysis, 1(3): 515–533.

Hoffman, M. and Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research, 15: 1593–1623.

Lewandowski, D., Kurowicka, D., and Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. Journal of Multivariate Analysis, 100(9): 1989–2001

Stan Development Team (2024). Stan Modeling Language Users Guide and Reference Manual. Version 2.34. url: https://mc-stan.org.