Developing High-Quality Data Infrastructure for Legal Analytics: Introducing the Israeli Supreme Court Database

Keren Weinshall* and Lee Epstein

Driving discovery in the study of law and legal institutions often requires infrastructure in the form of databases and other tools. The challenge is how to build the infrastructure. For obvious reasons, transplanting coding rules and variables from one dataset to the next is perilous; specialized knowledge of local conditions is necessary before one piece of datum is collected. Also required is adherence to a universal set of principles that distinguish high-quality infrastructure; namely, that the tool is capable of addressing real-world problems, accessible, reproducible and reliable, sustainable and updatable, and foundational. These principles guided construction of the Israeli Supreme Court Database, new and original infrastructure encoding information from all panel cases opened between 2010 and 2018 in the Israeli Supreme Court.

I. Introduction

The past decade has witnessed dramatic growth in the empirical analysis of apex courts worldwide, from Argentina (Muro et al. 2018) and Brazil (Arguelhes & Hartmann 2017) up to Canada (Alarie & Green 2017) and from Taiwan (Chen et al. 2015) across the globe to Israel (Weinshall-Margel 2016) and most of Europe (Hanretty forthcoming).[1] This is welcome news because studies of judicial behavior add to the store of knowledge on law and legal institutions, provide guidance to policymakers, educate the public about their courts, help lawyers develop strategies, and even prompt judges to rethink their choices (Posner 2008; Epstein et al. 2013; Wistrich et al. 2015). The less-welcome news is that data

*Address correspondence to Keren Weinshall, Vice Dean, Katia & Hans Guth-Dreyfus Chair in Conflict Resolution and Law at The Hebrew University of Jerusalem, Mt. Scopus, Jerusalem 9190501 Israel; email: keren.weinshall@mail.huji.ac.il.; Epstein is Ethan A. H. Shepley Distinguished University Professor at Washington University in St. Louis.

[1]The cites in parentheses are examples; for each high court many more papers and books have been published. The Israeli Supreme Court alone has generated scores of studies (e.g., Sommer 2009; Eisenberg et al. 2010, 2012, 2013; Weinshall-Margel 2011; Dotan 2013; Gliksberg 2014; Rosenthal et al. 2018; Anidjar et al. forthcoming).

infrastructure designed to advance knowledge and drive innovation, discovery, and invention for the analysis of judges and their courts has not kept pace with the accelerating interest.[2]

Why? One answer is that the field has mostly eschewed high-quality infrastructure in the form of public multi-user databases designed to capture a range of foundational information in favor of hand-coded datasets aimed at answering particular research questions (see generally Epstein et al. forthcoming).[3] The "one-off" approach has its benefits; chiefly, the resulting dataset is precisely tailored to the researchers' theoretical framing, definitions, and hypotheses. However, it also has substantial costs. Because encoding characteristics of courts and judges can be expensive, many tailored datasets consist of a small number of observations, decreasing statistical power and negating the combinatorial advantage. For the same reason, they are rarely updated, limiting their capacity to address contemporary problems. Finally, even when scholars include the same cases and covariates in their studies, conflicting results can, and do, emerge because of different data-collection procedures and practices.[4] Taken collectively, these costs impede the drive to discovery.

Law and courts scholars worldwide acknowledge the problems of weak data infrastructure—not to mention the challenges in making headway (Honnige & Gschwend 2010; Kapiszewski & Ingram forthcoming). One is the lack of consensus in the community over what form the data should take and for what purposes infrastructure should be developed. Some scholars favor quantitative (numerical) data and a selection process that allows for statistical inference; others are more interested in non-numerical data that they can interpret, organize into categories, and use to identify patterns. Much hand-wringing also ensues over how to define and measure concepts of interest (e.g., judge ideology, judicial independence, case subject matter).

Frankly, these divisions should not obstruct forward movement. Important breakthroughs have and will continue to follow from data infrastructure that relies on randomness or intention to select observations, that encodes data with numbers or archives text, or that permits for causal inference or deep description. Data are data, methods are methods (Patty 2015). As long as infrastructure can advance knowledge and accelerate discovery, these are differences without meaning.

---

[2]Sometimes, infrastructure is a method, procedure, or application that makes our work easier, faster, and better. No doubt such advances in law and the social sciences have been made, but apps and the like are not foremost on the minds of most scholars in the field; their core concern rather lies with products designed to capture data generated by courts, judges, lawyers, and other legal and political actors. For this reason we use the term "data infrastructure."

[3]Existing multi-user databases relevant to the study of judicial behavior include the *Biographical Directory of U.S. Federal Judges* (Federal Judicial Center 2020), the Comparative Constitutions Project (Elkins et al. 2020), the German Federal Courts Dataset (see Hamann 2019), the Norwegian Supreme Court Database (see Grendstad et al. 2015), the U.S. Supreme Court Database (Spaeth et al. 2020), the U.S. Supreme Court Justices Database (Epstein et al. 2020), and V-Dem (2020). The European Court of Human Rights Database (Cichowski & Chrun 2014), the International Criminal Tribunals Database (Meernik 2014), the National High Courts Database (Haynie et al. 2003), and U.S. Courts of Appeals Database (Songer 1996; Kuersten & Haire 2002) are also public multi-user databases but have not been updated for at least five years. See also note 5.

[4]To provide a simple example: Shamir (1990) and Dotan (1999) identify conflicting trends in the win rate of Palestinians at the Israeli Supreme Court because of the researchers' different definitions (and coding) of litigant success.

A more difficult challenge is how to build data infrastructure. The perils of simply transplanting coding rules and variables from one dataset to the next are obvious; specialized knowledge of local conditions is necessary before one piece of datum is collected. But even that knowledge is insufficient. Building high-quality data infrastructure also requires adherence to a universal set of principles; namely, that the tool is (1) aimed at solving or developing implications for real-world problems, (2) open and accessible, (3) reproducible and reliable, (4) sustainable and updatable, and (5) built to serve as a foundation for present and future research needs.

These principles guided development of the Israeli Supreme Court Database: a new data infrastructure designed to advance knowledge and encourage innovation by encoding information from the final decisions of all 16,109 panel cases opened by the Supreme Court of Israel between 2010 and 2018 (with annual updates). The database, including an online codebook and an analysis tool, is available at http://ISCD.huji.ac.il.

The sections that directly follow introduce the database, as well as describe relevant features of the Israeli judicial system. The balance of the article fleshes out the principles for creating high-quality data infrastructure with the hope of providing useful guidance to developers of the many nascent projects worldwide.[5]

II. A Very Short Primer on the Israeli Judicial System

The Israeli Supreme Court Database (the Database) is designed to capture the characteristics of resolved cases that made their way to or started in the Court between 2010 and 2018. Both routes are possible, as Figure 1 shows, because the Court is something of a hybrid between a U.S.-styled Supreme Court and a European Constitution Court. Like the Supreme Courts in the United States, Canada, India, and Norway, among others, the Israeli Court sits as the country's highest appellate court, hearing cases from lower courts (usually district courts[6]) through one of three legal procedures: civil appeals, criminal appeals, and administrative appeals. Mostly, these appeals are a matter of right, not discretion as they are in other Supreme Courts.[7]

Between 2010 and 2018, 8,690 panel cases came to the Court sitting as the highest appellate court; the other 7,419 were "High Court of Justice" (HCJ) cases. When the Court

---

[5]For example, at a 2019 conference at the European University Institute on multi-user databases for the study of judicial behavior (October 11–12), developers presented plans for building datasets on the Inter-American Court of Human Rights, the Court of Justice of the European Union, the Costa Rican Supreme Court, the German Federal Constitutional Court, and the Swedish Supreme Court.

[6]About 83 percent of all appeals come from district courts sitting as general courts and 16 percent from district courts sitting as administrative affairs courts.

[7]Then again, as Figure 2 shows, only 61 percent of the appeals result in a decision on their merits; the rest are withdrawn (usually on the justices' recommendation) or settled. Further, the Court does exercise discretion over appeals in civil and criminal cases that originated in the magistrate courts and were then appealed to the district courts. A motion to hear a discretionary second-level appeal is usually decided by a single justice. If review is granted, the panel may be expanded to include at least three justices (see Eisenberg et al. 2010, who explore the different decision-making patterns of appeals originating from the Israeli Court's mandatory and discretionary jurisdiction).
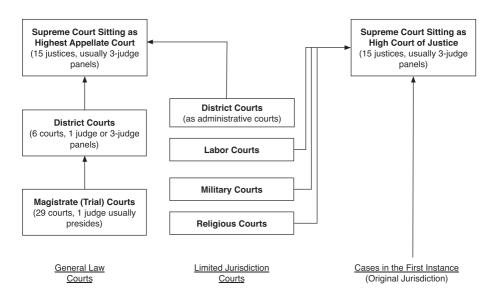
Figure 1:   The Israeli legal system—routes to the Supreme Court.

Note: The Israeli legal system, with emphasis on routes to the Supreme Court. Limited jurisdiction courts exclude the Antitrust Tribunal, Court of Admiralty, and the Standard Contracts Court because very few of their decisions go to the Supreme Court (a total of 21 between 2010 and 2018). Also, decisions of magistrate courts can be appealed to the Supreme Court in a limited set of circumstances (see note 7).

sits as the HCJ, it also hears appeals from some specialized courts,[8] as Figure 1 shows. However, the vast majority of HCJ cases (nearly 75 percent) arrive as they do in many European constitutional courts: in the first instance (i.e., on original jurisdiction) as administrative or constitutional petitions against actions taken by the government.

Whether sitting as the highest court of appeals or the High Court of Justice, panels of three, five, seven, and nine justices are assigned to hear "main" cases,[9] though three-justice panels resolve 99 percent of the disputes. Note, too, that "resolve" need not be a decision on the merits. In fact, for HCJ cases, "On the Merits" occurs in just under 50 percent of the cases, as Figure 2 shows.

This is hardly a comprehensive description of the Israel judicial system,[10] but it should suffice to understand the structure and contents of the Database, to which we now turn.

---

[8]Eighty-seven percent are appeals from Military Courts.

[9]In addition to the responsibility described in note 7, a single justice may also hear motions for injunctions, temporary restraining orders, detentions, and other interim rulings. The Israeli Court refers to cases decided by a single justice as "Minor Proceedings" and panel cases as "Main Cases." We follow suit.

[10]For readers interested in more information, see Mautner (2011), Meydani (2011), Friedmann (2016), and Galnoor and Blander (2018).
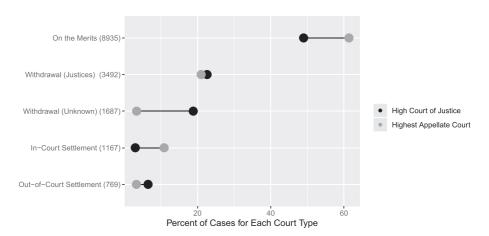
Figure 2:   Case disposition by the Israeli Supreme Court capacity.

Note: Disposition of petitions and appeals in the Israeli Supreme Court by whether the Court is sitting as the highest court of appeals or the High Court of Justice (HCJ), 2010–2018. The dark horizontal lines help visualize the difference between the HCJ and highest appellate court on the categories of disposition. Number of total cases = 8,661 for highest appellate court and 7,389 for the High Court of Justice (totals exclude 59 cases disposed of by other means). *N*s for each disposition are in parentheses. "Withdrawal (Justices)" indicates that the justices recommended withdrawal. Calculated from the Israeli Supreme Court Database.

III. The Israeli Supreme Court Database

Most datasets for the analysis of judges and their courts are designed to answer specific research questions and are hand coded in method. Sometimes the data are released in full; sometimes only the variables necessary to replicate the study's findings are archived; and sometimes the data never see the light of day (see Section IV.C).

The Israeli Supreme Court Database is different. It joins a handful of public multi-user datasets focused on courts and judges (see note 3). The idea behind these datasets is that they are so rich in content that multiple users, even those with distinct projects, can draw on them. Researchers can also add new content to suit their interests (see Section IV.E).

The Israeli Database and others of its type are not tailored to a particular project; that is a cost. But it may be well outweighed by the benefits if the dataset's developers understand local conditions and embed their product in tried-and-true principles governing high-quality infrastructure. Section IV describes these principles, and how they guided construction of the Israeli Database. In what directly follows, we preview the Database's structure and content.

III.A. Structure and Case Inclusion

The Database comes in two flavors: case- and justice-centered (both available at https://iscd. huji.ac.il/data). They are identical in format (cases in the rows and variables in the columns) but differ in the unit of analysis. In the case-centered version, each case is the unit; in the justice version, each case-justice is the unit (i.e., the case receives a row for each participating justice).

The rule for case inclusion is simple. The Database encodes information about the population of final decisions in cases opened between 2010 and 2018 and issued by a panel of three to nine justices in the Court's two capacities (see Figure 1): when presiding as the highest appellate court, including all criminal, civil, and administrative appeals; and when sitting as the High Court of Justice, including all constitutional and administrative petitions submitted to the Court in the first and final instance.

Note that under this rule the Database houses all petitions and appeals regardless of how the justices resolved them. This is a noteworthy difference from other apex court datasets (including the U.S. Supreme Court Database), which focus on merits decisions even though almost all courts have developed other methods for disposing of cases (see, e.g., Figure 2). Including non-merits decisions mitigates selection-effect concerns that arise in studies of only adjudicated cases (Kastellec & Lax 2008; but see Klerman & Lee 2014).

III.B. Case-Level Variables

The case-centered version of the Database contains 61 variables (columns of data) that come in four sets: identification, background, substantive, and outcome and voting variables.

As the name suggests, identification variables provide information about the full text of the Court's decision, such as the identification number assigned by the Israeli Supreme Court, the case name, and a link to the decision. Also included is a variable indicating the number of petitions or appeals the Court consolidated under a single decision. As in the U.S. Supreme Court,[11] consolidation is relatively rare, occurring in 15 percent of the cases. Nonetheless, to the extent that some studies (e.g., of the parties or lawyers) require including all consolidated cases, this variable enables researchers to make appropriate choices.

Background variables focus on the parties, their lawyers, the courts whose decisions the Israeli Supreme Court reviewed, and key dates in the case's history. The top panel of Figure 3 visualizes one of these variables, the types of parties, by whether the party was the petitioner or the respondent.[12] Note that in the petitioner category individuals dominated, bringing over 75 percent of the cases, but were far less often respondents (about 17 percent of the time). For the government, the picture is reversed: it was the most frequent respondent (50 percent of the cases) and one of the least frequent petitioners (5 percent). Business is roughly evenly divided between appealing and defending decisions, as it is in the U.S. Supreme Court (Epstein et al. 2017).

Substantive variables encode information about the procedure the Court used to take the case (e.g., criminal appeal, civil appeal) and the controversy's subject matter. The subject-matter variable consists of 64 possible categories; the middle panel of Figure 3 collapses them into 10, and then by whether the Court heard the case in High Court or appellate mode. By design, the two dockets are quite different: civil and criminal law

---

[11]The U.S. Supreme Court consolidated petitions in 9 percent of its merits cases between the 2010–2018 terms. Calculated from the U.S. Supreme Court Database.

[12]The Database includes the types of petitioners and respondents for the first three in each category. The figure uses only the first listed: *respondentType1* and *petitionerType1*.
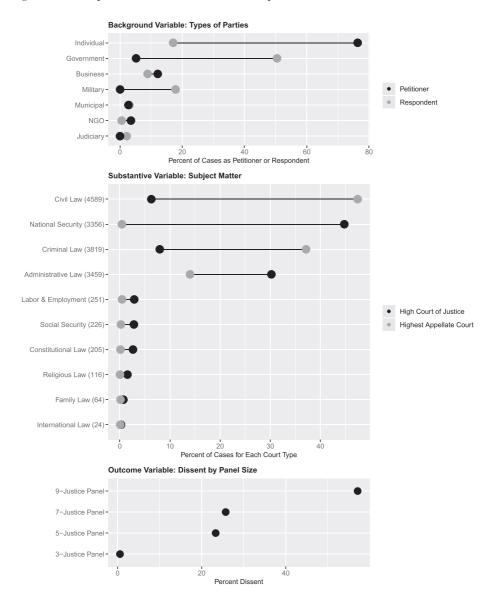
Figure 3:   Examples of variables in the Israeli Supreme Court Database.

predominate when the Court is in highest appellate mode and national security matters are most frequent when it sits as the High Court of Justice.

Turning to outcome and voting variables, one set draws attention to the Court's ruling, including *Disposition*, the manner in which the Court closed the case

(see Figure 2); *Outcome*, the Court's treatment of the decision below; and *WinnerCourt*, whether the petitioner or respondent won the case. Another set focuses on the justices: the number on the panel, whether they were unanimous, and the identity of the opinion writer, among others. The bottom panel of Figure 3 provides a simple example: split decisions (one or more dissent) by panel size. Note that the very low rate of dissensus on three-judge panels resembles the U.S. courts of appeals (Epstein et al. 2011); at the other extreme, dissent is as likely in the Israeli Supreme Court as it is in the U.S. Supreme Court when nine justices sit—a rare event occurring only in cases that raise the most pressing legal issues, much as in the U.S. high court.

III.C. Justice-Level Variables

For analyses focused on the Israeli Court and its decisions the four variable groupings just denoted are likely to prove most useful. Justice-centered studies will require some of these variables too, but also information on the voting, opinions, and characteristics of the individual justices.

   The Database supplies this information for the 25 justices serving since 2010. Many of the variables are standard; for example, whether the justice dissented, held for the petitioner, voted to dismiss the case, and so on. Others are more novel because they encode the justices' biographies. Figure 4 depicts two: gender and religiosity. Note that over time the percentage of votes cast by female justices shows a small decline, while the percentage by religious Jewish justices has increased significantly.

   Even for close followers of Israeli politics, these data may come as a surprise. Roughly half the lawyers in Israel are female but their vote share on the Court is only 25 percent. As for religiosity, in response to charges that the justices were overly "activist"—too eager to invalidate government acts—right-wing politicians tried to pack the court with religious (Jewish) justices on the assumptions that (1) conservative justices are less inclined to intervene in the government's policies and (2) religiously observant people tend to be more conservative (or at least share the ruling regime's preferences) (Friedmann 2016). To the extent that the percentage of religious Jewish justices has risen from around 10 percent in the 1990s to 27 percent in 2018—over twice the population (Jasper 2019)—and their vote share has increased from under 25 percent to over 40 percent, the conservative politicians succeeded. (Then again, as readers will soon see, the data lend little support to charges of activism or to assumptions about the self-restraint of religious Jewish justices.)

IV. Principles for the Creation of High-Quality Data Infrastructure

Developing these variables and other content is about making choices—from which cases to include to how to label the data, and everything in between. Regardless of the court or judges under study, arriving at these decisions requires knowledge of local conditions and also of the five basic principles that distinguish high-quality infrastructure: capable of
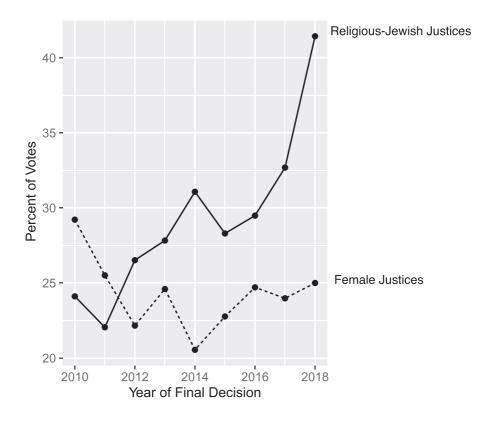
Figure 4: Percentage of votes by females and religious justices, 2010–2018.

Note: Two examples of biographical variables in the Israeli Supreme Court Database: percentage votes cast by female justices and religious (Jewish) justices, 2010–2018. The Database labels justices as Jewish religious, Jewish secular, and non-Jewish. Because one justice's religiosity is unknown, his votes are omitted from calculation of the percentage.

addressing real-world problems, accessible, reproducible and reliable, sustainable and updatable, and foundational.

IV.A. *Capacity to Address Real-World Problems*

By definition, data infrastructure should promote innovation, inventions, and insights. Although no product can guarantee these ends, infrastructure aimed at solving (or developing implications for) real-world problems increases the odds of success. Happily, this principle is easy to meet in the context of judicial behavior. Data and studies drawing on the U.S. Supreme Court Database have made their way into congressional debates,

lawyers' briefs, judges' writings, and journalists' stories.[13] Then there is Grendstad et al.'s (2015) dataset on the Norwegian Supreme Court. Not only have the developers used it to offer interesting insights into a court that had received almost no systematic attention; the data also have informed contemporary debates about the relationship between the Court's political, geographical, and experiential composition and its decisions.

Taking lessons from these and other existing tools, while tailoring variables to policymakers, scholars, and citizens interested in Israel, the Database also allows for evidence-based assessments of the Court and its work. Two examples suffice to make the point.

The first centers on the Court's use of it judicial review powers (Hirschl 2001; Dotan 2002, 2013). As suggested earlier, politicians on the conservative side of the political map have long accused the Court of being too activist, interfering unnecessarily and unwisely with the regime's policies. This criticism has mostly related to the Court's final rulings when deciding petitions against the government as the High Court of Justice (HCJ).

The data, however, seem to tell different story. Contrary to public and political discourse, the HCJ overtly rules in favor of a petitioner and against the government's policies in only 3 percent of the petitions, as Figure 5 shows. This is a significantly lower rate than for all other forms of appeals.

Focusing on "covert" rulings against the government does not change the picture much.[14] Although several studies have suggested that HCJ tends to pressure the state to change its policy before it makes a final ruling (Hofnung & Weinshall-Margel 2010; Dotan 2013), this happens in less than 15 percent of the cases; and it is not unique to HCJ petitions. All and all, far from intervening in the executive's and parliament's policies, the Court seems to defer to the other branches—especially when sitting as the HCJ.[15]

A second example of how the Database could contribute to public discourse focuses on the plan to pack the Court with religious Jewish justices in response to the Court's alleged activism. On this theory, to reiterate, religious Jewish justices should be less likely to rule against the government. But, again, the data seem to provide little support. Consider Figure 6, which shows the percentage of votes, by religiosity of the justice, to invalidate actions taken by the other branches in constitutional and non-constitutional disputes adjudicated by the HCJ and resolved on their merits.

In non-constitutional cases, neither religious nor secular Jewish justices are extreme activists—though, if anything, religious Jewish justices tend to invalidate acts by the other branches at a slightly higher (but not significantly different) rate. Turning to constitutional petitions, which are usually submitted directly against the parliament, once again religious Jewish justices invalidate government policies more often than secular Jewish

---

[13]For one recent example, see Liptak and Parlapiano (2019).

[14]Overt rulings (for or against the government) are decisions on the merits of the dispute; covert rulings (again, for or against the government) are those withdrawn or settled.

[15]Our emphasis on "seems" is no mistake. Just as in the United States and other contexts, careful multivariate work is needed before reaching more definitive conclusions (Weiden 2011; Epstein & Landes 2012). With some additional covariates, the Database could accommodate that work.
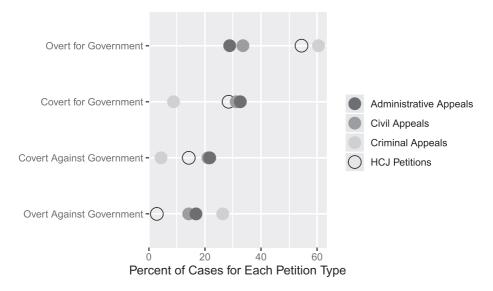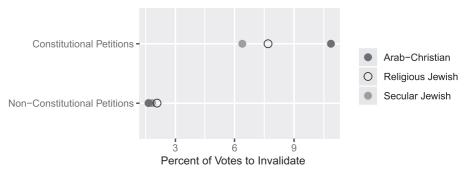
Figure 5:   Government win rates, in the Israeli Supreme Court by type of legal procedure, 2010–2018.

Note: HCJ = High Court of Justice. For definitions of "covert" and "overt," see note 14. The number of observations for Administrative Appeals = 1,082, Civil Appeals = 1,252, Criminal Appeals = 2716, and HCJ = 6,268.

justices and the two Arab-Christian justices more often than the 23 Jewish justices in the study. (These differences are statistically significant at $p < 0.05$.)

The point in presenting these data is not to reach any strong conclusions about the circumstances under which the Court or its members do or do not hold in favor of the

Figure 6:   Overt votes against the government in High Court of Justice cases, by justice's religion, 2010–2018.

Note: An overt vote is one cast against the government on the case's merits.

government; that would require much more analysis (see note 15). It is instead to show the potential to engage the Database to develop real-world implications and contribute to public and academic discourse on pressing legal-political issues.

IV.B. Accessibility

Another basic principle in the creation of high-quality infrastructure is that members of the community should be able to access it with no barriers to entry or use. That break-throughs cannot be had (not even for the researcher who developed the infrastructure) from "secret" data is so obvious that it would seem unnecessary to rehearse the point. Apparently, though, many scholars still impose long embargoes on their data, or perhaps never release them at all (Houtkoop et al. 2018). A recent study showed that only 20 per-cent of papers published in a journal requiring submissions to include a "data availability statement" actually deposited the data in a repository (Federer et al. 2018). Perceived obstacles include a belief that data sharing is not common because it requires extra work or special skills (Houtkoop et al. 2018). Then there is a fear of being "scooped" by another research term even when safeguards are in place.

This thinking is deeply flawed.[16] In the first place, articles that use publicly avail-able data are more likely to be cited regardless of the journal's impact factor, the study's year of publication, or the author's country of origin (Piwowar et al. 2007). Sec-ond, no embargoed infrastructure can be labeled "high quality," if only because the community cannot possibly make that assessment: Accessibility is essential to reproduce and verify the data and results, not to mention advance knowledge by enabling others to ask new questions. For these reasons many disciplines now mandate transparency in methods and data. In political science, for example, 27 of the top journals signed a joint statement "requiring authors to ensure that cited data are available at the time of publication through a trusted digital repository."[17] Finally, it is the duty of developers to share their data as a public good, as most national science foundations imply in their

---

[16]There are some exceptions. In the extreme, in France it is now a "criminal offence to 'evaluate, analyse, compare or predict' the behaviour of individual judges" (law translated in Langford & Madsen 2019). Were researchers to release datasets revealing the judges' names, they could face a maximum sentence of five years in prison. For other possible exceptions, see note 17.

[17]The statement further notes:

> If cited data are restricted (e.g., classified, require confidentiality protections, were obtained under a non-disclosure agreement, or have inherent logistical constraints), authors must notify the editor at the time of submission. The editor shall have full discretion to follow their journal's policy on restricted data, including declining to review the manuscript or granting an exemption with or without conditions. The editor shall inform the author of that decision prior to review.

Worth noting, though, is that the discipline of law and legal studies seems to lag behind in data-sharing norms, which may reflect the relative novelty of empirical methods in the field of law or a conception of databases as intel-lectual property owned by their creators (Heise 2011). By designating valuable space to databases, the *Journal of Empirical Legal Studies* has become a force for change.

requirements that researchers release the data from publicly funded research (Borgman 2012).[18]

Following the near-bedrock principle of accessibility, the Israeli Supreme Court Database is freely and publicly available. Anyone can download the data (from https://iscd.huji.ac.il/data); credentials are not needed. In addition, the website houses an analysis tool that permits users to access variables without having to download the dataset (see https://iscd.huji.ac.il/analysis/). Under the premise that output from research funded by the Israel Science Foundation belongs to the public, the tool was constructed to ensure access to citizens, decisionmakers, students, and journalists who might otherwise have difficulty analyzing or even viewing the data.

IV.C. Reliability and Reproducibility

This principle is simple: high-quality infrastructure should be reproducible, and the encoded data, reliable. Reproducibility means that users and developers alike must understand how to duplicate the data housed in the infrastructure. Reliability is related: it is the extent to which encoded data can be replicated, producing the same value using the same standard for the same subject at the same time, regardless of who or what is doing the replicating (Epstein & Martin 2014). Because these standards are about consistency, not accuracy (that is, validity[19]), establishing high-quality infrastructure does not require consensus over how to measure each and every concept of interest. All the infrastructure must provide are reliable data essential for building (or restructuring) variables to suit favored conceptual understandings that scholars can later evaluate and debate (see Section IV.E).

Enhancing reliability and reproducibility requires, first, that database developers and users treat all facts—observations and variables—as if they were part of a chain of evidence, asking (or knowing) how they were generated. If the infrastructure houses court cases, researchers can reproduce it only if they know how and on what basis cases were selected for inclusion. In the Israeli Supreme Court Database, all resolved cases assigned to a panel of judges between 2010 and 2018 were harvested from the Israeli Supreme Court's open website. As a means of enhancing credibility, as well as facilitating add-on variables and text analysis, the Database includes a link to the Court's final decision in each case.

A second step toward ensuring reproducibility and reliability is to provide all information required to replicate the dataset. The idea is that anyone with sufficient skill should be able to understand, evaluate, and reproduce the tool's content without any additional information from the creators (Epstein & King 2002). To meet this standard, the Israeli Database's website includes documentation that provides precise definitions

---

[18]For example, the U.S. National Science Foundation's policy: "Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing." Available at https://www.nsf.gov/bfa/ dias/policy/dmp.jsp.

[19]More specifically, validity is the extent to which a reliable measure reflects the underlying concept being measured.

and coding notes for each variable (see https://iscd.huji.ac.il/documentation). In addition to advancing the principles of reliability and reproducibility, detailed information helps promote a database's credibility and legitimacy (i.e., compliance with accepted standards). By the same token, documentation of coding protocols is essential for scientific assessment of research yielded from the database if only because coding decisions can influence the reported outcomes (as researchers know all too well!).

Finally, reliability in coding is best achieved by removing or at least minimizing human judgment. This could entail automating input (e.g., through data scraping), or writing clear and precise instructions, training coders to follow them, and assessing consistency among coders. Consensus between independent and blinded coders is critical to the objectivity sought by systematic coding (Hall & Wright 2008). If coders disagree at a high rate, asking them to confer or using a tie-breaker is inappropriate; that would produce uniform data points, but the coded data would not necessarily be replicable by others. Instead, disagreements should be reduced by refining the coding protocol.

We sought to ensure reliability of the Database's coding through a combination of computer scripts and humans. After harvesting the cases, an automated (data-scraping) technique was employed to allow efficient and accurate data entry for most variables. In the process of writing the code, we performed computerized and human reliability tests, including dual coding. The values for most variables were consistently coded in over 99 percent of the cases (sometimes after a process of rewriting or fixing the code).

Unfortunately, the lack of a structured paragraph or standardized phrasing of votes and outcomes did not permit for consistent scraping of these variables, and the same problem emerged for subject matter. To hand-code these variables, we employed 10 advanced law students, and to facilitate their task, we created a custom-designed back-end web application that made the process easier, faster, and more reliable. The app identifies the court document to be coded and pastes a form on it with the variables to be collected. It also randomly assigns cases to coders and transfers the coded data to the larger database without human intervention. At the end of the process, a second set of law students recoded a random sample of 20 percent of the cases; coding was consistent for 90 percent.

IV.D. Sustainability

Sustainability, standing the test of time, is a fourth principle crucial to the construction of high-quality infrastructure. Because of the difficulty in making predictions about the endurance of any tool, this principle may seem more aspirational than practical. Then again, just as Elkins et al. (2009) demonstrate for constitutions, certain features of infrastructure will increase the odds of longevity—chiefly, whether it can and will be maintained and updated. This is a legitimate concern of many grant-making panels: after all, why invest thousands, perhaps millions of dollars in databases and other products that will become outdated or otherwise incapable of producing breakthroughs down the road? They would not, and yet the lack of sustainability has been the downfall of many fine datasets (see note 3).

Seen in this way, staving off obsolescence is a challenge all data tools must meet—and there are at least four strategies for so doing. One is to repel (irrational) data exuberance. When planning infrastructure, regardless of the form, it is tempting to include every

characteristic of, say, judges, courts, and decisions that anyone has ever mentioned, and then to develop detailed coding protocols for each that must be implemented by humans. Resisting this temptation is difficult (as we know all too well), but giving in to it is the surest way to create a product that will die a slow death because it will be too difficult, time con-suming, and expensive to maintain and update. Overexuberance also increases the odds of an abandoned project, not to mention unreliable data (see Benoit et al. 2016).

Second, to the extent possible, developers should minimize human involvement in the data-generation process. Once again, we accomplished this by writing scripts to scrape data (facts) from the Court decisions. Not only does this form of automation enhance reliability, it also reduces the effort and costs of updating information—though it does not eliminate them altogether. The code will need to be periodically adjusted to account for changes in opinion format or Court procedures. Additionally, the extracted text almost always requires some housework.

A third strategy to advance sustainability, which may be feasible for future versions of the Database, is to exploit methods for automating content analysis: not scraping facts (e.g., case names) but developing algorithms to help organize the texts—Court decisions—into categories of interest (i.e., classification) (for a nontechnical introduction, see Grim-mer & Stewart 2013). Researchers can use these methods regardless of whether they know the categories in advance and simply wish to train computers to replicate (or extend) their hand coding or they hope to learn about/discover new categories via unsupervised or computer-assisted methods (Grimmer & King 2011; Livermore & Rockmore 2019).

Finally, researchers might consider outsourcing coding to non-experts. True, this approach seems to maximize, not minimize, human contact with the data. And true, it may seem heretical, even crass, because many of us (like to) think that our jobs require specialized knowledge. But retaining experts to code all non-automated variables may be infeasible on a periodic basis, and, more to the point, it is unnecessary. Social scientists have now shown beyond any reasonable doubt that non-expert online workers can com-petently code even very technical data (Benoit et al. 2016; Carlson & Montgomery 2017).

IV.E. Foundational

This last principle is related to sustainability: data infrastructure should serve as a founda-tion upon which researchers can build by adding content, backdating, updating, or oth-erwise adapting it to their own needs; it should not be the be-all, end-all. If there is no way to adapt the infrastructure to future or even present purposes, it will sentence researchers to a life of using the same materials to do the same work over and over again. This has the benefits of consistency and promoting shared norms in the field, but it is not sustainable if only because new generations of scholars looking to innovate will inevi-tably turn to other sources or develop their own "one-off" products.

The U.S. Supreme Court Database is an example of a foundational tool. On the surface it would seem comprehensive, consisting of over 200 pieces of information on every case decided with an opinion since 1791. However, almost no study published over the last decade uses the U.S. database "as is"; authors have created new classifications,

collapsed existing values, and, best of all, added scores of variables, such as characteristics of the justices, the identity of the lawyers, the public's "mood," case facts, and many more.

And so it should be for most data infrastructure—including the Israeli Supreme Court Database. Although it will hopefully prove useful in its own right, it was designed as a foundational tool capable of accommodating present and future research needs. One happy consequence of hewing to this principle was that it furthers the Database's sustainability, as we were less tempted to include all content of interest, trusting that others could and would develop their own add-ons.

V. Concluding Remarks

After decades of developing datasets and other tools to answer very specific research questions, the community of law and courts scholars is now recognizing the value of high-quality data infrastructure aimed at a range of research needs. Not only will such projects advance knowledge and drive discovery, but they also have the potential to bring together diverse scholars worldwide—goals on which we all can agree.

We are proud to take part in this growing movement, accelerated by the *Journal of Empirical Legal Studies*' establishment of a section devoted to building and sharing large legal datasets, and encourage other scholars, regardless of their subject or approach, to contribute to this exciting exchange.

References

Alarie, Benjamin, & Andrew J. Green (2017) *Commitment and Cooperation on High Courts.* New York: Oxford Univ. Press.

Anidjar, Leon, Ori Katz, & Eyal Zamir (forthcoming) "Enforced Performance in Common-Law Versus Civil Law Systems: An Empirical Study of a Legal Transformation," *American J. of Comparative Law.*

Arguelhes, Diego Werneck, & Ivar A. Hartmann (2017) "Timing Control Without Docket Control: How Individual Justices Shape the Brazilian Supreme Court's Agenda," 5(1) *J. of Law & Courts* 105.

Benoit, Kenneth, Drew Conwy, Benjamin E. Lauderdale, & Michael Laver (2016) "Crowd-Sourced Text Analysis: Reproducible and Agile Production of Political Data," 110(2) *American Political Science Rev.* 278.

Borgman, Christine L. (2012) "The Conundrum of Sharing Research Data," 63(6) *J. of the American Society for Information Science & Technology* 1059.

Carlson, David, & Jacob Montgomery (2017) "A Pairwise Comparison Framework for Fast, Flexible, and Reliable Human Coding of Political Texts," 111(4) *American Political Science Rev.* 835.

Chen, Kong-Pin, Kuo-Chang Huang, & Chang-Ching Lin (2015) "Party Capability Versus Court Preference: Why Do the 'Haves' Come Out Ahead?—An Empirical Lesson from the Taiwan Supreme Court," 31(1) *J. of Law, Economics, & Organization* 93.

Cichowski, Rachel, & Elizabeth Chrun (2014) *European Court of Human Rights Database.* Available at http://depts.washington.edu/echrdb/.

Dotan, Yoav (1999) "Judicial Rhetoric, Government Lawyers, and Human Rights: The Case of the Israeli High Court of Justice During the Intifada," 33(2) *Law & Society Rev.* 319.

——— (2002) "Judicial Accountability in Israel: The High Court of Justice and the Phenomenon of Judicial Hyperactivism," 8(4) *Israel Affairs* 87.

——— (2013) *Lawyering for the Rule of Law: Government Lawyers and the Rise of Judicial Power in Israel.* Cambridge, UK: Cambridge Univ. Press.

Eisenberg, Theodore, Talia Fisher, & Issi Rosen-Zvi (2010) "Israel's Supreme Court Appellate Jurisdiction: An Empirical Study," 96(4) *Cornell Law Rev.* 693.

——— (2012) "Does the Judge Matter? Exploiting Random Assignment on a Court of Last Resort to Assess Judge and Case Selection Effects," 9(2) *J. of Empirical Legal Studies* 246.

——— (2013) "Group Decision Making on Appellate Panels: Presiding Justice and Opinion Justice Influence in the Israel Supreme Court," 19(3) *Psychology, Public Policy, & Law* 282.

Elkins, Zachary, Tom Ginsburg, & James Melton (2009) *The Endurance of National Constitutions.* New York: Cambridge Univ. Press.

——— (2020) *Comparative Constitutions Project.* Available at http://comparativeconstitution sproject.org.

Epstein, Lee, James L. Gibson, & Andrew D. Martin (forthcoming) "Using Databases to Study Constitutional Law," in M. Langford & D. Law, eds., *Handbook of Research Methods in Constitutional Law.* Edward Elgar Publishing.

Epstein, Lee, & Gary King (2002) "The Rules of Inference," 69(1) *Univ. of Chicago Law Rev.* 1.

Epstein, Lee, & William M. Landes (2012) "Was There Ever Such a Thing as Judicial Self-Restraint?" 100(3) *California Law Rev.* 557.

Epstein, Lee, William M. Landes, & Richard A. Posner (2011) "Why (and When) Judges Dissent," 3 *J. of Legal Analysis* 101.

——— (2013) *The Behavior of Federal Judges.* Cambridge, MA: Harvard Univ. Press.

——— (2017) "When it Comes to Business, the Right and Left Sides of the Court Agree," 54(1) *Washington Univ. J. of Law & Policy* 33.

Epstein, Lee, & Andrew D. Martin (2014) *An Introduction to Empirical Legal Research.* Oxford: Oxford Univ. Press.

Epstein, Lee, Thomas G. Walker, Nancy Staudt, Scott Hendrickson, & Jason Roberts (2020) *U.S. Supreme Court Justices Database.* Available at http://epstein.wustl.edu/research/justicesdata. html.

Federal Judicial Center (2020) *Biographical Directory of Article III Federal Judges.* Available at https:// www.fjc.gov/history/judges/ biographical-directory-article-iii-federal-judges-export.

Federer, Lisa M., Christopher W. Belter, Douglas J. Joubert, Alicia Livinski, Ya-Ling Lu, Lissa N. Snyders, & Holly Thompson (2018) "Data Sharing in PLOS ONE: An Analysis of Data Availability Statements," 13(5) *PLOS ONE* 1.

Friedmann, Daniel (2016) *The Purse and the Sword: The Trials of Israel's Legal Revolution.* New York: Oxford Univ. Press.

Galnoor, Itzhak, & Dana Blander (2018) *The Handbook of Israel's Political System.* Cambridge, UK: Cambridge Univ. Press.

Gliksberg, David (2014) "Does the Law Matter? Win Rates and Law Reforms," 11(2) *J. of Empirical Legal Studies* 378.

Grendstad, Gunnar, William R. Shaffer, & Eric Waltenburg (2015) *Policy Making in an Independent Judiciary: The Norwegian Supreme Court.* Colchester, UK: ECPR Press.

Grimmer, Justin, & Gary King (2011) "General Purpose Computer-Assisted Clustering and Conceptualization," 108(7) *Proceedings of the National Academy of Sciences* 2643.

Grimmer, Justin, & Brandon M. Stewart (2013) "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts," 21(3) *Political Analysis* 267.

Hall, Mark A., & Ronald F. Wright (2008) "Systematic Content Analysis of Judicial Opinions," 96(1) *California Law Rev.* 63.

Hamann, Hanjo (2019) "The German Federal Courts Dataset 1950–2019: From Paper Archives to Linked Open Data," 16(3) *J. of Empirical Legal Studies* 671.

Hanretty, Chris (forthcoming) *A Court of Specialists: Judicial Behaviour in the UK Supreme Court.* Oxford, UK: Oxford Univ. Press.

Haynie, Stacia L., Reginald S. Sheehan, Donald R. Songer, & C. Neal Tate (2003) *National High Courts Database.* Available at http://artsandsciences.sc.edu/poli/juri/highcts.htm.

Heise, Michael R. (2011) "An Empirical Analysis of Empirical Legal Scholarship Production, 1990–2009," 2011(5) *Univ. of Illinois Law Rev.* 1739.

Hirschl, Ran (2001) "The Political Origins of Judicial Empowerment Through Constitutionalization: Lessons from Israel's Constitutional Revolution" 33(3) *Comparative Politics* 315.

Hofnung, Menachem, & Keren Weinshall-Margel (2010) "Judicial Setbacks, Material Gains: Terror Litigation at the Israeli High Court of Justice," 7(4) *J. of Empirical Legal Studies* 664.

Honnige, Christoph, & Thomas Gschwend (2010) "The Bundesverfassungsgericht in the German Political System—The Great Unknown?" 51(3) *Politische Vierteljahresschrift* 507.

Houtkoop, Bobby Lee, Chris Chambers, Malcolm Macleod, Dorothy V. M. Bishop, Thomas E. Nichols, & Eric-Jan Wagenmakers (2018) "Data Sharing in Psychology: A Survey on Barriers and Preconditions," 1(1) *Advances in Methods & Practices in Psychological Science* 70.

Jasper, Alon (2019) "On the Social and Professional Composition of the Judiciary," paper presented at the Empirical Studies of Public Law and Human Rights Workshop at Hebrew University.

Kapiszewski, Diana, & Matthew C. Ingram, eds. (forthcoming) *Concepts, Data, and Methods in Comparative Law and Politics.* Cambridge University Press.

Kastellec, Jonathan P., & Jeffrey R. Lax (2008) "Case Selection and the Study of Judicial Politics," 5 (3) *J. of Empirical Legal Studies* 407.

Klerman, Daniel, & Yoon-Ho Alex Lee (2014) "Inferences from Litigated Cases," 43(2) *J. of Legal Studies* 209.

Kuersten, Ashlyn K., & Susan B. Haire (2002) *Update to the Appeals Courts Database.* Available at http://artsandsciences.sc.edu/poli/juri/appct.htm.

Langford, Malcolm, & Mikael Rask Madsen (2019) *France Criminalises Research on Judges.* Available at https://verfassungsblog.de/france-criminalises-research-on-judges/.

Liptak, Adam, & Alicia Parlapiano (2019) "A Supreme Court Term Marked by Shifting Alliances and Surprise Votes," June 29 *New York Times.*

Livermore, Michael A., & Daniel N. Rockmore, eds. (2019) *Law as Data: Computation, Text, and the Future of Legal Analysis.* Santa Fe, NM: Santa Fe Institute Press.

Mautner, Menachem (2011) *Law & Culture in Israel.* Oxford, UK: Oxford Univ. Press.

Meernik, James (2014) *The International Criminal Tribunals.* Available at http://www.psci.unt.edu/~meernik/International%20Criminal%20Tribunals%20Website.htm.

Meydani, Assaf (2011) *The Israeli Supreme Court and the Human Rights Revolution: Courts as Agenda Setters.* Cambridge, UK: Cambridge Univ. Press.

Muro, Sergio, Alejandro Chehtman, Jorge Silva M'endez, & Nelson Amaya Dur'an (2018) "Testing Representational Advantage in the Argentine Supreme Court," 6(1) *J. of Law & Courts* 1.

Patty, John (2015) *Responding to a Petition to Nobody (or Everybody).* Available at http://www.mathofpolitics.com/2015/11/06/responding-to-a-petition-to-nobody-or-everybody/.

Piwowar, Heather A., Roger S. Day, & Douglas B. Fridsma (2007) "Sharing Detailed Research Data Is Associated with Increased Citation Rate," 2(3) *PLOS ONE* 1.

Posner, Richard A. (2008) *How Judges Think.* Cambridge, MA: Harvard Univ. Press.

Rosenthal, Maor, Gad Barzilai, & Assaf Meydani (2018) "The Personalization of Judicial Review: The Cohesiveness of Judicial Nominations and Constitutional Courts," paper presented at the at the Comparative High Court Decision Making Workshop, at Hebrew University.

Shamir, Ronen (1990) "'Landmark Cases' and the Reproduction of Legitimacy: The Case of Israel's High Court of Justice," 24(3) *Law & Society Rev.* 781.

Sommer, Udi (2009) "Crusades Against Corruption and Institutionally-induced Strategies in the Israeli Supreme Court," 15(3) *Israel Affairs* 279.

Songer, Donald R. (1996) *The Original U.S. Appeals Courts Database 1925–1996.* Available at http://artsandsciences.sc.edu/poli/juri/appct.htm.

Spaeth, Harold J., Lee Epstein, Ted Ruger, Jeffrey A. Segal, Andrew D. Martin, & Sara Benesh (2020) *U.S. Supreme Court Database.* Available at http://supremecourtdatabase.org.

Weiden, David L. (2011) "Judicial Politicization, Ideology, and Activism at the High Courts of the United States, Canada, and Australia," 64(2) *Political Research Q.* 335.

Weinshall-Margel, Keren (2011) "Attitudinal and Neo-Institutional Models of Supreme Court Decision Making: An Empirical and Comparative Perspective from Israel," 8(3) *J. of Empirical Legal Studies* 556.

———— (2016) *Law and Ideology in Supreme Court Decision-Making: A Comparative and Quantitative Analysis.* Jerusalem: Hebrew Univ. Sacher Press.

Wistrich, Andrew J., Jeffrey J. Rachlinski, & Chris Guthrie (2015) "Heart Versus Head: Do Judges Follow the Law or Follow Their Feelings," 93(4) *Texas Law Rev.* 855.