

Analyzing Contracts: State of the Field, Mixed-Methods Guiding Steps, and an Illustrative Example

Karra Greenberg¹, Evgenia Jane Kitaevich, Siddharth Chaudhari and Anna Kirkland²

Contracts are the underpinning of functional business, governmental, organizational, and inter-personal relations across the developed world. It is a methodological advance for contractual research scholars to transform qualitative contractual information into quantitative data to analyze if, how, and why the law is reflected and reproduced in contracts—potentially explaining the continuation or amelioration of social injustices—and/or if contractual content informs human behavior or vice versa. However, efforts to “quantitize” contracts and statistically study them are scattered across subfields without much methodological guidance. The linguistic complexity of many legal contacts, paired with few repositories, makes their comparative quantitative analysis challenging. Prior attempts to quantitatively analyze contracts often lack shared systematic methods for measuring specified textual content, transforming this information into a quantitative format, or statistically analyzing the produced data set. This article presents the power and promise of transforming legal contracts and other similar documents into quantitative data for analysis, reviews the research to date that does so in different ways, and then provides concrete guiding steps and an illustrative example of how to transform contractual content into quantitative data that are valid, reliable, and reproducible, increasing the quality of the data produced and supporting the important socio-legal conclusions that can be derived from their analysis.

INTRODUCTION: WHY ANALYZE CONTRACTS?

Individuals’ and organizations’ responsibilities, protections, entitlements, and remedies in personal, social, and business matters are often contained in a variety of formal, contractual agreements. By the word “contracts,” we mean the world of mostly nongovernmental written agreements that set out obligations between parties, such as the health benefits that an employer provides in exchange for work, the homeowner’s association contract that spells out rules, benefits, and fees for residents, or even prenuptial agreements that regulate two people’s entering into and dissolving their marriage.

Karra Greenberg is Research Affiliate, Institute for Research on Women and Gender, University of Michigan Population Studies Center, University of Michigan, Ann Arbor, MI, United States. Email: karrag@umich.edu

Evgenia Jane Kitaevich is Doctoral Candidate, Department of Political Science, University of Michigan, Ann Arbor, MI, United States. Email: evgenia@umich.edu

Siddharth Chaudhari is Doctoral Candidate, Department of Political Science, University of Michigan, Ann Arbor, MI, United States. Email: singhsid@umich.edu

Anna Kirkland is Kim Lane Scheppelle Collegiate Professor of Women’s and Gender Studies and Arthur F. Thurnau Professor, University of Michigan, Ann Arbor, MI, United States. Email: akirklan@umich.edu

While some of these are governmental (such as police labor union contracts with a city), we generally exclude documents such as international treaties and constitutions, which are contracts in some sense but are already well studied and do not fit the research challenges we outline here. We use the term “contract” more broadly than attorneys might (for example, we call summary plan descriptions of health benefit offerings by that term, although they are actually required disclosures of the benefit terms). We adopt a schematic definition for contracts that researchers and ordinary people confront: a dense written document that sets out important rights and duties that is important but not easily analyzed for its parts and meanings when there are dozens or even hundreds in the relevant set.

Analyzing the content of contracts and explaining the reasons for, and implications of, their variation is a critical research agenda for legal and socio-legal scholarship. Contracts contain important information for research in governance, financing, reproduction, health, civil rights, incarceration, punishment, employment, marriage, housing, and commercial transactions. We may want to know, for example, if police union contracts have changed following Black Lives Matter protests, if the terms and provisions for nurses’ access to personal protective equipment (PPE) varies across hospital types or states, or if and how commercial and residential leasing contracts have changed amidst the COVID-19 pandemic. Contracts are text-based documents, so a qualitative approach is clearly fitting. “Quantitizing” contracts for statistical analyses (Tashakkori and Teddlie 1998)—through transforming contractual text into distilled, finite categories or numbers—empowers scholars and policy makers to uncover associations that otherwise may remain hidden. The challenge is to approach contracts with methods that recognize the importance of their textual meaning while also scaling up to be able to describe patterns in large numbers of complex documents.

In his 2002 article encouraging contractual law scholars to quantize and statistically analyze contractual data, Russell Korobkin (2002, 1045) notes that statistical analyses of contracts powerfully demonstrate correlations between contractual provisions and contextual variables, providing “a deeper understanding of the law than any doctrine can provide.” Systematic statistical analyses of contracts can demonstrate variation in provisions or ideas across contracts and whether the observed variation is associated with other contractual provisions or contextual variables such as location, policy climate, or time period. Without the systematic comparison of sizable numbers of contracts and tests for associations between contractual elements and their greater societal context, explanations for variation across contracts is otherwise unknown in the legal contractual world (Korobkin 2002). Almost twenty years after Korobkin’s (2002) exhortation to contractual legal scholars, however, little contractual research uses mixed qualitative and quantitative methods, and, of the research that does, very little uses uniform statistical practices to ensure the quality of the data and their analysis.

It is time for socio-legal analysis of contracts to cohere as a research strategy apart from similarly complex analyses of other types of legal texts such as judicial opinions. In fact, some of the most pressing, current socio-legal research questions are best answered through a mixed methods approach that combines topic-specific knowledge of the law, policy, and interests that contracts help to govern with qualitative analysis of their language and quantitative analysis of the relationships between contractual elements and their contexts. Quantitizing contractual data for statistical analysis

enormously eases comparative analysis while centering the important content from the contracts that has been qualitatively derived. In this article, we discuss how scholars have made their way through this challenging research area, explaining the obstacles as well as scholars' different approaches. We bring together work on contracts published by law professors, economists, sociologists, political scientists, and policy scholars, all with different disciplinary training and writing for different audiences. Our approach is based on our analysis of hundreds of health insurance summary plan descriptions (SPDs), during which we developed methodologically informed steps for analyzing these complex contracts.¹ We argue that there are methodological steps that contractual socio-legal scholars should take to study contracts that can impact legislation and policy. The mixed-methods approach behind quantizing contracts that we advocate here is rarely realized in socio-legal studies, however, for a complex set of reasons. We introduce these complexities by first reflecting on how hard it can be to assemble the data one needs and to assess how complete it is.

Contracts are hard to find and collect. It is hard to know how many there are in a given field or industry, and their text is often lengthy, dense, and subject specific. The socio-legal field of contracts is itself vast and highly diverse, touching many areas of divergent content expertise. These realities account for many of the methodological challenges that scholars have faced in studying what contracts say, how they change over time, and how their terms relate to actual practices. A researcher may be interested in a specific type of contract or a social field in which certain contracts are nonpublic (such as upper management hiring contracts at a private company or personally held contracts like prenuptials), and, thus, the difficulty of gathering these contracts is already given in the research topic.

Sampling is greatly assisted in contract areas in which a census of contracts exists—in other words, the full universe of contracts in a given field or industry is known. For example, contracts could be collected using the mandatory reporting requirements for the Securities and Exchange Commission or Standard & Poor's 1500. For a census to exist, there must be some law or requirement that produces the census itself and a reporting system to support it or another research team that has exhaustively collected every contract and made their database public. There are very few contractual areas in which a census exists, however. There is no centralized, annual reporting requirement for all health insurance contracts sold in the United States, for example, where any plan offered must be publicly recorded. Instead, there is a scattered collection and reporting effort by fifty different state insurance regulators and the Department of Labor. When there are records of private health plans publicly available, they are very hard to assemble (because the actual plan does not have to be filed, just a form vaguely

1. By health insurance "contract," we mean the summary plan description, which is the roughly one hundred- to two hundred-page disclosure that an employer offering employee health insurance benefits is required by law to provide to his or her employees. Included in the requirements for the content of this document is information about the benefits to which they are entitled (which also includes any services or benefits that are explicitly excluded from coverage) and the deductible, maximum out of pocket, and cost-sharing requirements. Our team analyzed summary plan descriptions and also the Summary of Benefits and Coverage. These are distinct from other documents that may also be provided regarding coverage such as the member handbook, certificate of coverage, subscribers contract, medical policies, or drug formulary.

describing the plan) and impossible to verify for completeness (because a private employer may have incompletely reported their plans).

Much contract research probably exists in a middle realm, where it is possible, but complicated, to estimate the population of contracts that is represented by one's sample of contracts. Contracts by government entities may be retrievable by Freedom of Information Act (FOIA) requests, and one could make a reliable count of how many entities (such as police forces and sheriff's departments) hold these contracts and how many contracts are held in a given department, thus extrapolating the potential total number of contracts or the universe. Responsiveness to FOIA requests will shape the final count and representativeness of the collected contracts because responsiveness may or may not correlate to characteristics of the contracts. For example, are contracts from nonresponsive police forces more authoritarian, with greater police protections in their contracts, or are these contracts similar to those collected from other departments yet housed in disorganized departments? Stated differently, do individuals and organizations that select to share their contracts with researchers differ from those that do not, and may this difference relate to what is in the contracts themselves? A researcher in this situation will need to assemble secondary data about the targeted contracts and assess explanations for nonresponse and missing contracts in order to understand the representativeness of the collected sample.

Drawing a sample can be difficult. Contracts may be entirely private and protected from disclosure by legal privileges, as in prenuptial agreements. The most likely way to gather them is to ask the people who made them to voluntarily release them for research. Health services researchers have called insurance companies to ask what their health plans contain (Ngaage et al. 2020), and perhaps sampling people for interviews rather than documents would help. However, this method does not solve the sampling problem, and people may not reliably remember or accurately describe what is in the contracts. We presume here that the actual document language is critical data. Moreover, the sheer volume of total contracts in a socio-legal field, such as most consumer contracts and even many nonnegotiable employment contracts, would make sampling difficult because it is drinking from the fire hose (Luker 2010). Geographical and time boundaries would help further define the parameters of what the sample aims to represent. For example, a researcher might study a sample of rental or employment contracts from one year in two cities, gaining a better idea of what this sample represents. Contracts can also be hard to read. Health insurance contracts are hundreds of pages long and contain many definitions and provisions that require cross-referencing in order to decipher. While the Affordable Care Act (ACA) standardized some features of some health plans, for example, many features and plans are not subject to ACA regulations and thus lack imposed standardization.

Contracts are special as research objects, and they deserve their own approach. While none of the methodological combinations we describe below are in themselves new, it is novel to pull together existing scholarship on the topic of mixed-methods contractual research, describe the particular challenges of studying contracts, and propose a multi-method, interdisciplinary path forward. This article has the following structure. First, we present a range of important socio-legal research questions that can only be addressed by quantizing contracts and analyzing them comparatively and statistically. Second, we review the current state of mixed-method, quantitized

contractual research, noting that scholarship is on a spectrum of quantization: some researchers characterize qualitative content in a near-quantitative fashion, without expectation that it can or should be further quantized and statistically analyzed, while others quantize contracts and create complex variables for statistical analysis.

The extent to which contractual data is and should be quantized is highly dependent on the research question. We point the way for research to produce data sets and analyses that achieve desirable reliability or accuracy and consistency of the techniques used to create the data set; validity or grounds for accepting conclusions drawn because the concepts intended to be captured are captured; and reproducibility or the ability for other scholars to reach consistent results using the same contracts and methodology as the original study. We compare quantitative contractual research to quantitative research on judicial opinions, the most commonly analyzed other legal text, discuss mixed-methods approaches utilized in contractual research in law and economics, and discuss the use of machine learning.

Third, we empirically demonstrate the process of quantizing and analyzing contractual texts by detailing the eight steps we have developed for health insurance plan analysis. Contractual research requires particularistic approaches when applying mixed-methods methodology—in procuring a sample of contracts, navigating complexities in understanding its representativity, assembling a research team that may include subject-specific expertise, and, especially, developing content analysis and data transformation rules that account for the structural, linguistic, and semantic uniqueness of contracts. These particular features of quantizing contracts correspond to our Guiding Steps 1–3 and 5, which are discussed below. Our proposed guiding steps methodology presents best practices that can empower scholars to explore research questions from systematic description through full quantization of contractual data. If these practices are used consistently, they permit scholars to quantize and analyze data from prior studies and similarly produce new data that can be further used by others. There is currently no scholarship that outlines how to apply general quantizing practices to contracts and no unified conversation on the topic despite demonstrated interest across disparate subfields—hence, our outlining of eight steps. We conclude with a discussion of the expected challenges and intractable problems that scholars may face, suggesting helpful adaptations.

IMPORTANT SOCIO-LEGAL RESEARCH QUESTIONS ONLY ANSWERABLE BY QUANTITIZING CONTRACTS

Despite foundational research in law and society scholarship that reveals businessmen do not pay much attention to what is written in formal contracts (Macaulay 1963), scholars still care about whether and to what extent contracts reflect and reproduce the law in different sites, transactions, and contexts. Contract details can matter because they are the outcomes of policy and can thus be evaluated for whether the policy is working as intended. Contracts can be the formal legal embodiment of private actors' conceptions of their own duties and entitlements, showing in black and white how formal legal requirements make it into agreements. Contracts come in many varieties and are formed by different types of parties with varying levels of

power. Some contractual features may be simply sticky bits that no one understands but busy lawyers replicate for decades (Gulati and Scott 2011). All of this accumulates to create innumerable instances in which private actors' duties and entitlements, embodied in contracts, may or may not reflect and reproduce the law, potentially propagating or ameliorating social injustices. Yet contractual scholarship cannot effectively know the extent to which contracts in a variety of areas reflect and reproduce the law because currently little effort has been made to track the extent to which this is occurring or to provide systematic explanations for it.

Research questions about small numbers of particular types of contracts and their meaning are amply answered by qualitative contractual research. However, research questions that aim to understand how and potentially why variation in contractual characteristics is observed across contracts are best served by quantitative analyses of the qualitatively derived portions of contracts. Business and management scholars have advanced well in this area given the central place of business deals in their field, and there are examples of the type of empirical work on contract terms that we suggest should migrate to legal studies (Ryall and Sampson 2009). Socio-legal scholars likely focus less on business terms but wonder about whether contracts reflect policy aims, especially when regulations are designed to empower vulnerable contracting parties, but regulatory agencies may be underfunded and hobbled. How a private entity contracts with members of the public may be a critical policy outcome for a wide range of areas such as the environment, consumer protection, financial services, and rental housing. As Shauhin Talesh (2012) points out, insurance regulations are public interest law, and scholars should approach them as such. Legal protections for the public or less powerful groups and individuals may be widely ignored in contracts with extreme power differentials between the contracting parties. For example, many consumer contracts are entirely one-sided, with consumers likely completely unaware of what is in the terms and unable to do anything other than refuse the good or service. Health insurance plans are typically contracts of adhesion, and most individuals who have insurance through their employer or through a state exchange have no choice but to accept the terms offered for coverage of designated medical conditions. Without scholars' systematic documentation of concerning contractual terms across large numbers of contracts, no challenge may arise from the vulnerable contracting parties themselves because the power differential between the contracting parties is too great and the burdens of fighting are too onerous for the individual. Providing systematic explanations for how, where, and when contracts may incorrectly reflect policy aims helps scholars to understand why policies may be failing.

Contracts are also important routes for "studying up" (Nader 1974) the actions of "repeat players" (Galanter 1974) or powerful actors that use their greater knowledge and resources to secure contractual advantages. Labor union contracts are examples of contracts in which highly salient public issues such as police protections in lethal force cases or nurses' rights to adequate PPE in a pandemic may be bargained for in detail over many months and with relatively powerful parties on both sides. These topics and others have become particularly salient in contemporary times with public unrest regarding police behavior in lethal force cases and non-federally coordinated organization of the US medical supply response for health-care workers to battle COVID-19. Important research questions abound: in police contract provisions, are there correlations between

the incidence of non-legitimated police lethal response and internal disciplinary provisions in police contracts, and do these correlations vary by whether the precinct permits/requires police unionization or other precinct characteristics; over time, will police contracts demonstrate the removal of some of the process protections for officers who shoot or injure someone while on duty, such as delay periods and opportunities to view recorded footage before giving a statement; what are the terms and provisions for nurses' access to PPE, and do these meaningfully vary across hospitals, zip codes, states, or types of hospitals? We can successfully learn the answers to these types of questions with the methods we synthesize and mold to contractual research here.

MIXED APPROACHES THROUGHOUT PREVIOUS CONTRACTUAL MIXED-METHODS RESEARCH

Legal and socio-legal scholars have explored contracts using mixed methods on a variety of topics. For example, Hillary Berk (2020) has studied surrogacy agreements, asking to what extent and how are legal protections differently applied to different types of individual contracting parties in private contracts and is this variation associated with state legal institutions? Berk demonstrates that there is much to learn about how valuation of the rights of future parents, fetuses, and surrogates is captured in contracts and why variation across jurisdictions may not be observed. How lenient are the investigative and disciplinary procedures in cases of misconduct by police officers, and is this leniency associated with state labor laws and, more specifically, with the union contract provisions that some police organizations provide their officers (Rushin 2017)? How common are non-enforceable, oppressive, and misleading terms in residential leases, and are they associated with the characteristics of leaseholds or tenants (Furth-Matzkin 2017; Hoffman and Strezhnev 2022)? Have particular corporate contractual legal provisions for employees, such as noncompete clauses, become more prevalent and more powerful over time, thus increasingly restricting employee freedom (Bishara, Martin, and Thomas 2015)? President Joe Biden's recent executive orders against the widespread use of noncompete clauses in hiring contracts is a powerful example of change in contracting that has resulted in high-level policy change (Spiggle 2021). Are contractual terms for end-user software licenses biased against consumers, what software company characteristics inform this bias, and has this changed over the 2000s (Marotta-Wurgler 2007, 2008; Marotta-Wurgler and Taylor 2013)? Additionally, are inter-corporate legal provisions between unequally powerful parties sensitive to network effects, thus facilitating costly and poor contracting choices by the less powerful party (Kahan and Klausner 1997)?

Of course, many areas of social science use mixed methods to analyze complex documents (Allee and Peinhardt 2010; Koremenos 2013; Allee, Elsig, and Lugg 2017; Nyarko 2019). In political science, mixed methods have been extensively utilized in the study of judicial opinions in the United States (as in the US Supreme Court database) (Epstein et al. 2015; Rachlinski and Wistrich 2017; Frankenreiter 2018) and in courts around the world (Medvedeva, Vols, and Wieling 2020). However, while many socio-legal research questions require mixed-methods analysis of contract documents for answers, contracts between individuals and organizations and between

individuals are comparatively understudied and do not fit easily into a subfield of established scholarship. Socio-legal research in law and economics often focuses on contracts, yet the large majority of this research focuses on problems in contract law doctrine (Ben-Shahar 2011) or takes a formal modeling approach with hypothetical examples that does not depend on actual contractual language and mixed method empirical analyses of it (Hart 2017). Even so, Omri Ben-Shahar and James White (2006) used a sample of eighteen boilerplate contracts between the big auto manufacturers and their top suppliers and interviewed the lawyers who drafted them, asking how these contracts varied and conveyed transactional advantages. Across fields, finding and analyzing what contracts actually say or fail to say drives innovative scholarship.

The use of rigorous methods strengthens validity—that the concepts intended to be captured from the text and the population of contracts intended to be represented by the sample are correctly measured—and reliability—that is, accuracy and consistency in the techniques used and the measurements made. However, using rigorous methods in analyzing contracts presents challenges at three moments in the research process that each inform the reliability, validity, and replicability of the data produced and analyzed: (1) collating and/or accessing a sizable sample of contracts with substantive variation and clearly defined representativity; (2) customizing the content analysis for contractual texts that are subject specific and idiosyncratic in content, language, and structure; and (3) executing data collection/transformation on lengthy and dense text and its subsequent statistical analysis. In the sections below, we explore how scholars have approached these three moments in the research process in a variety of ways.

The Sample of Contracts

The first major hurdle to clear for socio-legal contractual scholars is obtaining an optimal sample of contracts. Access may be particularly difficult given the proprietary practices of many contracting parties and the relevant gatekeepers, as noted above. Moreover, it may be difficult for scholars to reach a particular threshold number. Observed results in small samples may not be reliable; they may be due to chance or random errors characteristic of small samples versus systematic characteristics found in a given population. Then, to successfully assess explanations for meaningful variation across contracts, the sample must contain theoretically important variation in the variable of interest—characteristics that may be unknowable before a sample has been obtained. Obtaining contracts that do not vary on the explanatory variable—often contextual, like geography, time period, or contracting party characteristics—makes it difficult to draw conclusions about the role of the explanatory variable on the contractual outcome. Contracts that do not vary on the outcome variable—often a contractual element—prevent the falsifiability of the results and thus challenge their validity. Effective sampling requires scholars to have an *a priori* sense of what the contracts will contain and where the most variation is likely to reside before they pursue their sample collection or access.

Lastly, it is difficult for researchers to have a sense of what population of contracts their sample, and, thus, their results, represent. It is particularly challenging for contract research that the full population of particular contract types is most often not known,

so there is not a reliable sense of how far off the sample, and its analyzed relationships, are from representing the entire population. This situation challenges the validity of the sample. It often requires extra research on the part of scholars to ascertain whether the contracts' inclusion in the sample is independent or not independent of the contracts' characteristics. Moreover, are contracts missing from the sample due to non-systematic reasons (missing completely at random), systematic characteristics unrelated to the contracts (missing at random), or systematic characteristics related to the contracts (missing not at random)? The best approach is to be honest in writing up one's results about the challenges in gathering the sample. Any one of these six possible combinations regarding independence and what is missing and why presents a useful sample that likely represents a specific subpopulation, and scholars can thus use conclusory language that matches any of the limitations.

Given these challenges of obtaining a sample representative of an intended population of contracts, legal and socio-legal contractual scholars have approached acquiring their samples in various ways. Some scholars have drawn their samples from the (relatively rare) equivalent of public organization censuses, such as the Securities and Exchange Commission (Chen and Bharadwaj 2009; Coyle 2016), the Canadian Security Administrator (Coyle 2019), or Standard & Poor's 1500 public corporations (Bishara, Martin, and Thomas 2015), where companies' contracts are included. Similarly, others have produced their sample of contracts by first obtaining a sample of contracting parties in a particular industry and location and then utilizing the contracts of these contracting parties, necessarily included in the human sample, to produce their sample of contracts (Allen and Lueck 1992). Researchers have also used existent repositories of contracts such as health insurance plans of employees at self-insured US corporations and residential leases attached to eviction hearings, collected by for-profit and nonprofit entities, respectively (Kirkland, Talesh, and Perone 2021; Hoffman and Strezhnev 2022). To study the contract terms in sovereign bonds between 1960 and 2011, Stephen Choi, Mitu Gulati, and Eric Posner (2012) drew from a preexisting data set of bonds (Thomson One Banker) but had to supplement with older bonds found in libraries and archives because the data set begins in 1990 (see also the description of similar data set construction with hybrid methods in Weidemaier and Gulati 2013). The bond contract data set was built from sales prospectus details about the contracts, not the actual bond contracts (similar to our situation below with the health benefit summary plan descriptions).

When existing repositories of contracts or samples of contracting parties have not been available, others have collated their sample of publicly available contracts through public record requests, municipal government websites, and web searches (Rushin 2017), requests of state regulators to procure contracts from regulated entities like homeowner insurance groups (Schwarcz 2011), or utilized accessible data from public or private organizations that note the pertinent contractual elements for all contracts of a particular kind (Kahan and Klausner 1997; Marotta-Wurgler 2007, 2008, 2009; Marotta Wurgler and Taylor 2013). Amy Monahan and Daniel Schwarcz (2022, 24) have studied medical necessity rules in health insurance contracts by searching state regulatory filings over five years for the top three insurers in a selection of states. After the initial study of this sample of health plans, they searched again to construct a final sample of 170 policies from every state that had filings publicly available, selecting only

the most recent plans by the top-three carriers (25). Samples of private contracts between individuals have been obtained by scholars directly contacting the drafting lawyers (Berk 2020) or contracting parties (Furth-Matzkin 2017).

With few exceptions (Marotta-Wurgler 2007, 2008, 2009; Choi, Gulati, and Posner 2012; Marotta-Wurgler and Taylor 2013), there is generally little uniform discussion in the literature of what contractual populations are represented by these samples, such as how they were drawn from organizational censuses to guarantee representativity of an entire population or how the consideration of independence and missing observations applied when collected by other means that likely produced representativity of a particular subpopulation. Most mixed-methods contractual scholars have procured samples (or curated assemblages, to use language from another methodological viewpoint) that vary in both the explanatory and outcome variables. Mixed-methods contractual samples vary in size from thirty (Berk 2020) to 170,000 (Hoffman and Strehnev 2022), with most analyses containing approximately one hundred to a few thousand contracts (Kahan and Klausner 1997; Marotta-Wurgler 2007, 2008, 2009; Chen and Bharadwaj 2009; Marotta-Wurgler and Taylor 2013; Bishara, Martin, and Thomas 2015; Coyle 2016, 2019; Rushin 2017).

Contractual Content Analysis and Variable Creation

The content analysis of contractual texts and their preparation for quantization via data entry presents the second area of difficulty for mixed-methods contractual scholars. Transforming text-based, industry-specific contractual data into quantitative data for statistical analysis is an interdisciplinary, multi-methods effort. Doing so in a manner that optimizes efficiency, reliability, and validity of data production is challenging because different subject-specific knowledge, tied to a particular logic, is required for each of the numerous steps involved: determining what substantive content to extract, analyzing the content, and creating the quantitative variables. In quantizing contractual data, legal and socio-legal scholars necessarily aim to reduce data from large quantities of text and concepts to defined, close-ended concepts tied to numbers in the form of variables. Decisions about what data to extract and what quantitative variables to create from this extracted data can be challenging for scholars because of the subject-specific nature of contracts, the large quantity of information they often contain, and the contingency or interrelatedness of the numerous concepts or contractual elements.

Moreover, in transforming qualitative, textual data into a quantitative format, comparative contractual scholars are particularly challenged to ensure validity and reliability. Contractual texts are complex in unique ways that vary from the complex texts of judicial opinions, international treaties, and so on. Contracts are often long, multifaceted, linguistically dense, hierarchically and cross-referentially organized, industry specific, and terminologically inconsistent when comparing across different contracting contexts or parties. They are simply not meant to be read publicly. (Even judicial opinions are written with some public-facing explanation and follow a predictable structure with reason giving included.) This structure makes comparative contractual research particularly vulnerable to measurement error or not capturing the

concepts that are intended to be measured and/or not measuring consistently due to the misinterpretation of the texts or their non-systematic review.

Contractual scholars have responded in a number of ways to the interdisciplinary, particularistic characteristics of contractual content analysis and the preparation of texts for quantitative transformation. To garner industry-specific knowledge for different contracting contexts, some investigators utilize their own expertise as area-specific legal scholars (Marotta-Wurgler 2007, 2008, 2009; Marotta-Wurgler and Taylor 2013; Bishara, Martin, and Thomas 2015) and non-law scholars in the specified subject area (Kahan and Kalusner 1997; Chen and Bharadwaj 2009) and their dual training in law and social sciences (Kirkland, Talesh, and Perone 2021), or they have sought insight from industry-specific practicing lawyers and contracting parties (Berk 2020). Other scholars have used media coverage, public opinion, and/or advocacy organizations to first understand the most important topic-specific content to glean from analyzed contracts (Rushin 2017; Hoffman and Strezhnev 2022) or have been guided by state regulatory codes or international conventions that explicitly list problematic or ideal contract terms, respectively (Coyle 2016, 2019; Furth-Matzkin 2017) or standard industry contracts from underwriting organizations (Schwarcz 2011). Some scholarly teams incorporate researchers with statistical expertise (Allen and Leuck 1992; Chen and Bharadwaj 2009).

In deciding the details of what qualitative content to extract and how to best transform it into quantitative variables, scholars take different approaches. Using inductive methods, some examine contractual texts for the most common themes, terms, provisions, or monetary values and then use this information to inform the variables to be created for quantitative analysis (Berk 2020). Some scholarship has further attempted to validate their inductively derived classification schemas of contract provisions for information technology outsourcing through review by an expert panel of law professors and practicing lawyers (Chen and Bharadwaj 2009). Others use deductive approaches, establishing *a priori* specific concepts or legal provisions that are the focus of research and then examining contracts for the presence, absence, or nature of this content, such as disciplinary provisions related to police misconduct investigations (Rushin 2017), the usual types of corporate take-over events considered in risk covenants (Kahan and Klausner 1997), rights and risks of software end-user purchasers (Marotta-Wurgler 2007, 2008, 2009; Marotta-Wurgler and Taylor 2013), oppressive, non-enforceable, and/or misleading lease terms (Furth-Matzkin 2017; Hoffman and Strezhnev 2022), stipulations of the Convention of International Sale of Goods (Coyle 2016, 2019),² homeowner insurance coverage provisions (Schwarcz 2011), or farmers' and landowners' choice of crop-share versus cash-rent farmland contracts (Allen and Leuck 1992).

Some research explicitly notes an iterative process of using inductive and deductive reasoning to arrive at their focal variables, such as three types of employment noncompetition covenants for chief executive officers (Bishara, Martin, and Thomas 2015). Berk (2020) describes her descriptive coding, process coding, and *in vivo* coding. However, scholarly descriptions of this part of the research process vary widely, and

2. Convention on the Law Applicable to Contracts for the International Sale of Goods 1985, 24 ILM 1575 (1985).

sometimes there is little detail about how complex contractual texts have been systematically navigated and what interpretive or analytic decisions have been made and why. This additional methodological detail is important when contracts do not have clearly or consistently defined terms and conditions, have multiple associated documents or hierarchically/interrelated provisions, or use industry-specific, idiosyncratic legal terms.

Data Entry, Management, and Statistical Analysis

Quantitative data entry, data management, and statistical analysis of the data present the final, broad challenge to contractual scholars. Comparative contractual scholars strive for reliability in their data entry. However, ensuring the reliability of data classification efforts with contractual text is challenging given the length and denseness of many contracts. Moreover, administrative organization can be daunting as storing and calling upon numerous, lengthy contracts, data sets, and statistical files for data analysis requires careful coordination when a research team is involved. Lastly, empirical legal and socio-legal research is theoretically driven and aims to understand (1) the content, structure, and meaning of contracts; (2) how contractual elements vary, and (3) how and under what circumstances contractual elements are associated with characteristics of the contracting parties, organizations, macro-context, costs, and so on.

Contractual scholars have addressed quantizing practices, data management, and meaningful data analysis in a variety of ways. Some scholars have aimed to increase data entry reliability by having a single person enter data from the same contracts at two separate times (Rushin 2017) or using statistical best practices of at least two individuals entering data separately on every contract in the analysis (Chen and Bharadwaj 2009). Others have relied on a close supervision process with regular corrections of research assistants' work coupled with spot checks throughout the data set (Kirkland, Talesh, and Perone 2021). Some research with dual coders has added a metric to assess—on a scale from 0 to 100 percent—the extent to which data entry was reliable, or consistently matching, between coders (Chen and Bharadwaj 2009). Some contractual research discusses administrative practices that enable the process of quantizing qualitative contractual information, such as the software packages used to perform and store the qualitative coding of contract content and interpretation, like *Atlas.it* computer-aided qualitative data analysis software (Berk 2020).

Research that empirically analyzes quantitative contractual data does so in a variety of ways. Some scholars present lists of findings without numeric analyses, such as displaying which of seven total disciplinary measures are mentioned in police union contracts for different US cities (Rushin 2017) or listing the major terms and provisions found in surrogacy contracts across different states (Berk 2020) or the findings of no mentions of the specified terms (Coyle 2019). These descriptive analyses may fully answer a research question on their own, such as whether a specific term has been incorporated into contracts or what terms are the most common, or they may lend themselves to theory development. Anna Kirkland, Shauhin Talesh, and Angela Perone (2021) similarly present the exact language of the transgender coverage exclusions they counted in health insurance contracts, creating a scale that indicates the clarity and generosity of coverage and giving the numbers and percentages of different types of

exclusions. Variable construction of an index in some research has become standardized, such as the Marotta-Wurgler End User License Agreement Bias Index that, using the Universal Commercial Code as a basis of comparison, assigns a positive, negative, or no point to each term of purchasers' rights and risks in software licenses and then sums these values for an overall index value (Marotta-Wurgler 2007, 2008, 2009; Marotta-Wurgler and Taylor 2013).

Going further to numerically analyze data, having created indices or scales by comparing contract terms to industry or legal standards, some contractual researchers provide a summary of these indices and scales or incidence of contractual terms through descriptive statistics like averages and proportions and further demonstrate associations between contractual terms and contracting context by calculating descriptive statistics across contexts (Schwarcz 2011; Coyle 2016; Furth-Matzkin 2017). Norman Bishara, Kenneth Martin, and Randall Thomas (2015), for example, tested for population-level differences across subgroups with chi-square tests, finding that covenants not to compete (CNCs) had become more common and more expansive over time and were comparatively less common in California, a state that does not permit CNC enforcement.

Other quantitative contractual research demonstrates correlations between contractual elements and contracting context, contracting parties, and/or opting in to contracts by using regression techniques. Findings include counterintuitive observations such as that non-enforceable (for example, pro-landlord) lease terms are associated with more expensive leaseholds in whiter, richer areas of an urban center (Hoffman and Strezhnev 2022). Some scholars have included numerous variables in their regression modeling to account for competing explanations or mechanisms to the focal relationship (Allen and Leuck 1992; Furth-Matzkin 2017). Using numerous indicator and control variables from the contractual texts, research on information technology outsourcing finds that the interdependence of processes increases contract extensiveness (Chen and Bharadwaj 2009). Similarly, accounting for numerous software company characteristics, research on consumer end-user software licenses finds pro-seller bias (Marotta-Wurgler 2007) that has increased over time (Marotta-Wurgler and Taylor 2013), yet the bias is not greater for companies with greater market power (Marotta-Wurgler 2008) and does not vary by whether contract terms are disclosed before or after the purchase (Marotta-Wurgler 2009). Empirical contractual scholarship has even ventured into quasi-experimental methods, fabricating contracts with onerous terms that vary in structure and mode of presentation (on a fabricated company website and so on) to test if this variation has influenced selected consumers' beliefs that these terms were legally enforceable and morally defensible (Wilkinson-Ryan 2017).

Machine Learning versus Human Power

In recent years, legal scholars have adopted machine learning-enabled text analysis. Using best practices to build and leverage custom-made machine-learning models on large corpuses of text (Lucas et al. 2015), court and judicial opinion researchers reveal relationships between gender and/or the political background of judges and their decision making, predict the actions of the US Supreme Court (Katz, Bommarito, and Blackman 2017), the French Court of Cassation (Sulea et al. 2017), and the European

Court of Human Rights (ECtHR) (Aletras et al. 2016) based on previous judgments, and predict outcomes in the ECtHR (Medvedeva, Vols, and Wieling 2020). These advancements were made possible by considerable research investment in public documents that have a somewhat predictable structure. However, despite the powerful methodological aid provided by machine learning for socio-legal research, contractual research may not benefit from machine learning for many years. Many contracts are simply more difficult for machines to learn to decipher, due to the inter-connected, contingent structure of contractual texts, their non-standardization even within the same topic and industry, and the resultant need for an extensive machine-learning effort that is costly in terms of time and money to which most scholars do not have access.

To arrive at a final conclusion for a specified contractual topic, many contractual texts require a reader to internally cross-reference numerous sections of the contract following a contingency logic that may also be hierarchical. Many contracts also have textual addenda that bear on the conclusions drawn. Machine-learning algorithms would need substantial time to learn how to extract just a handful of variables from such interconnected text for each contract of a particular structure and verbal style. Moreover, this time line would be necessarily multiplied for comparative contractual research because private contracts and those of many industries are not standardized in terms of structure, formatting, or even the particular vocabularies utilized to denote the same concepts. This variation translates into a very customized and lengthy machine-processing effort. Time and cost calculations are best informed by the number of contracts to be analyzed and scholars' access to financial resources. For scholars with very large samples, numerous variables to be created and extracted, significant research funding, and a time horizon of many years, machine learning may be the appropriate mechanism.

However, meaningful conclusions can be drawn from smaller samples such as two hundred, which is well within the temporal and financial range for human effort on many comparative contractual investigations. In comparison to a machine-learning approach, we enlisted two paid graduate students to complete our non-machine-learning approach, and they received extensive project-specific training in developing and using the established content analysis procedures and in helping to further define the variables required, totaling one-and-a-half months. Once the training was completed, the rate of data entry for twenty-four variables extracted from every contract was five contracts per hour per research assistant (with each research assistant quantizing the same contract for the purposes of double data entry, discussed below). For 435 contracts, data entry took approximately four months with each research assistant dedicating ten hours per week and a postdoctoral supervisor dedicating three to five hours per week.

ILLUSTRATION OF STEPS TO CREATING AND ANALYZING A QUANTITATIVE CONTRACTUAL DATABASE

In the following sections, we describe the steps we took in health insurance contract analysis and how we confronted the unique challenges of analyzing contracts. We explain the work in great detail to show exactly how these research techniques,

widely used across the social sciences, apply in this particular research area, where they do not seem to have permeated quite so well. Our research questions were: what is the incidence of coverage for outpatient mental health counseling (that is, talk therapy); to what extent is coverage provided, medically and financially, including out-of-pocket costs; and is the type of insurance plan or the type of employer sponsor important for the generosity of a plan? Our application of mixed-methods steps, utilized in the sample creation, content analysis, and quantizing of contractual information, assisted us in reducing two types of statistical error that can compromise data integrity: representation error, which is tied to how a sample of contracts is drawn, and measurement error, which is tied to how contractual information is extracted or measured. Careful attention to these factors allowed us to ensure the data set's validity, reliability, and reproducibility, thus allowing potential falsification. We explain below how we implemented these principles. Admittedly, many of these steps are common to qualitative and quantitative research projects of many kinds. And, yet, much of the current socio-legal research on what is in contracts does not seem to take these steps. We take care to point out along the way how our methodological pathway is particularly tailored to analyzing contracts specifically.

Step 1: Create or Obtain a Sufficiently Large Repository of Contracts That Contains Variation in Structurally Important Ways

A major goal of inferential socio-legal contract analysis is to assess meaningful variation across contracts in their content and the potential explanations for this variation. This comparison is enabled through obtaining a sample or census of contracts, which is potentially a great challenge for contract researchers. We accessed a resource called AXIACI from Leverage Global Consulting, a proprietary database that contains insurance plan offerings and coverage from private and public insurance market segments. This database includes plans of the Employee Retirement Income Security Act (ERISA),³ which governs companies who self-insure their health benefits using third-party administrators and are subject to filing at the Department of Labor using Form 5500.⁴ Our use of the proprietary database for the purposes of public policy research and analysis in health insurance is governed by an agreement with Leverage. We decided to use a repository created by others to overcome the significant research barrier of having to pull and organize all the contracts ourselves. We would not have had the time and resources to create our own repository. We extracted from the AXIACI database 435 health insurance contracts from 2019 for forty ERISA-governed self-insured corporations. There were no human subjects in this research and no personally identifying information involved

3. The Employee Retirement Income Security Act is applicable to private sector companies that offer pension plans to employees, including health benefit plans. Employee Retirement Income Security Act, 1974, 88 Stat. 829 (codified as amended at 29 U.S.C. §§ 1001-1461 and in scattered sections of 5, 18, and 26 of the U.S.C.).

4. There is another public filing system for insurance documents regulated at the state level (National Association of Insurance Commissioners 2021), though this public database is also difficult to navigate and incomplete.

because these are only the contract documents, not the records of anyone's health insurance claims or medical information.

Step 2: Understand What Population of Contracts Our Sample Represents

Legal and socio-legal contract researchers aim to know the population of contracts their sample represents in order to accurately relate for which population of contracts the results hold. This can pose a particular challenge for contract scholars. Our sample of 435 health plans from forty ERISA-governed companies includes all the available contracts of that type, as of July 2019, in the repository. Leverage Global Consulting collected plan documents using the public filings information (US Department of Labor 2021), web searches, and other internal strategies. Our example shows that weak regulation can produce an uncertain census of what contracts exist. Evidence suggests many companies file late or fail to file the reporting form completely (Leone 2006), and penalties are frequently reduced and can be avoided if the omission can be framed as accidental. The Department of Labor's filing does not include the actual plan documents, only the brief information about the plan on the form. We were left with a fairly large sample of contracts, but it was unclear as to whether the sample was independent of Leverage's own business needs or of the variable ease of locating documents. Moreover, there was no way to verify if the plans for each reporting company were the complete set offered that year or if particular plans from a company were missing and if they were missing due to their characteristics. Our sample is thus representative of insurance plans that would be reported by ERISA-governed corporations to the Department of Labor and publicly available for the diligent and informed searcher.

Step 3: Assemble an Interdisciplinary Research Team with the Right Blend of Methodological and Substantive Knowledge

Our aim was to transform text-based, industry-specific contractual data into quantitative data with efficiency, reliability, and validity. This generally requires subject-specific knowledge tied to a particular logic for each step of data production and analysis. This can be particularly challenging when researching contracts. We assembled a multidisciplinary research team with members versed in the different substantive and methodological logics required for the project: a faculty member with expertise in legal research, qualitative methods, and health insurance policy, who was also able to negotiate the repository access; a postdoctoral fellow with expertise in statistical methods, the sociology of mental health, and public health; a graduate student with expertise in content analysis, sampling, and data set construction; and another graduate student with expertise in statistical methods. We also relied upon information technology assistance and content expertise from the industry partners.

Substantive topic-specific knowledge of mental health illnesses and socially stratified access to their treatment informed the development of our research questions within a broader research area. The consulting firm shared their health insurance industry-specific perspective, pointing us to where meaningful variation in health

insurance coverage was likely to be found. Expertise in legal content analysis enabled us to navigate externally and internally referential language and complex structures within health insurance contracts. Our team's expertise in legal content analysis and familiarity with health insurance industry particularities improved our data's validity by ensuring that the qualitative concepts presented in the contracts were correctly located and interpreted before being transformed into predefined, quantitative variables. Our knowledge of statistics helped ensure that we understood the subpopulation of ERISA-governed insurance contracts that our sample represents with its limitations; quantitative variables were appropriately defined to capture the qualitative text, data entry errors were tracked and reduced, and the data were statistically analyzed—all improving the validity, reliability, and replicability of the quantitative data's creation and analysis.

Step 4: Determine the Content to Extract and the Variables to Create

To determine the data to extract and the quantitative variables to create from the contractual data, our scholarly team took a multi-step approach common to qualitative analysis projects generally. We first broadly defined the substantive area of legal interest—namely, health insurance coverage of outpatient mental health counseling. Before the analysis of the contracts began, extensive research using information from medical health and advocacy organizations informed the specific mental health conditions considered treatable by outpatient counseling. Research using these resources also informed what is considered ideal medical treatment and ideal health insurance coverage for these conditions. It is important to define a baseline or ideal that should be covered (for example, what a major medical association recommends) in order to understand the differences from it. These resources, taken together, also helped inform what common exclusions from coverage might be observed in the contracts. This preliminary research informed our research questions (noted above), and the research questions informed what substantive information needed to be extracted from the health insurance contracts.

Next, we assessed a subsample of contracts to ascertain the extent to which predefined, desired information was present and the level of detail with which coverage was discussed. This informed the first effort to create variables. Empirically understanding the structure of health insurance contracts for outpatient mental health treatment was particularly important because its coverage, like many medical conditions, was fully explained in a cross-referential pattern across multiple contract sections. For example, coverage may be stated if treatment is provided by “an approved healthcare professional,” yet the types of approved providers, based on their professional degree, would be listed in another document or a far-off section. Thus, the simple question “is this treatment covered or not?” for example, quickly grew complicated in ways not typically seen in qualitative interview coding. Not only was the qualitative coding of language and phrases challenged by the spread-out and cross-referential structure of the contracts, but the production of variables as numerical values also required reducing them, as Kristin Luker (2010) puts it, to a drop-down option after following a unique string of questions in a spreadsheet. We had to adjust predefined variables and create new

variables, specifically following the multi-step structures detailed below. We established an iterative process that allowed variables to be updated until data entry began, with the goal of having the variable values exhaustively cover all possible options found within the contracts.

Step 5: Develop Content Analysis Rules, Rule Application, and Rule Validation for Complex Contractual Texts, Avoiding Systematic Measurement Error

Contractual scholars are challenged to reduce measurement error and ensure validity and reliability in the content analysis of contractual texts that are long, cross-referentially and hierarchically connected, linguistically inconsistent, and idiosyncratic in a number of ways. These content analysis challenges are particular to contracts. For example, each of the health insurance contracts we analyzed was approximately sixty to three hundred plus pages in length, and there were two separate components in the plans—the Summary of Benefits and Coverage (SBC) and the main section—that needed to be analyzed for every case. To address concerns about measurement error, our project used an ordered process of rule development, rule application, and rule validation when utilizing content analysis to transform the contractual data into a quantitative format.

Rule Development

Our processes for rule development of the content analysis and quantitative variable creation were tied to each other and were necessarily iterative. After initially defining our variables of interest based on prior research and initial reviews of some contracts, we specified the procedure for navigating the contractual texts. There were three umbrella concepts particular to analyzing the health insurance contracts, and each concept required the use of health insurance industry-specific knowledge to create rules for coders to follow: (1) identifying possible substitutions of various terms; (2) creating a roadmap to navigate contingently and hierarchically related portions of the contracts; and (3) defining what non-mentions of coverage would mean.

Prior to working with the contracts, we aimed to identify potential variation in health-care terms related to particular phenomena. Yet it was only upon engaging with the text that the full range of potential terms used became clear, such as the fact that “telehealth” is similarly referred to as “virtual health,” “telemedicine,” “teladoc,” “online health services,” and so on. We worked to guarantee that the inclusion of these alternative terms was sufficiently broad so as not to risk inadvertently excluding a service that in actuality is distance health care yet also that the alternative terms were also substantively meaningful and not referring to a distinct phenomenon.

We developed a protocol to navigate the contractual texts, identifying the different segmented sections that needed to be reviewed, whether there were contingent or cross-referential relationships between the sections and, thus, the order in which to review each. The main contractual sections we identified for the health insurance

project were “Covered Benefits,” “Exclusions,” “Glossary of Terms,” “Cost Tables,” and “the SBC.” Within these sections were subsections that we similarly identified. These sections and subsections themselves often used different terminology in their headings, requiring us to make iterative use of the alternative terms lists we had constructed. We searched for each section by both scrolling through the full contract to get a sense of its structure and formatting and then by doing specific key word searches to confirm that we had not missed any relevant sections or information.

We first searched for a main section entitled by variants of “Covered Benefits” and then a subsection headed by variants of “Mental Disorders.” If such a section existed, we discerned whether mental health conditions were recognized and covered independently from health conditions related to substance abuse or behavioral health (the latter term usually used for indications of mental health conditions in children). In the event that the contract contained ambiguous language, we examined other pre-identified sections entitled with variants of “behavioral health” or “mental health” to clarify whether the plan guaranteed coverage for more than just substance abuse.

Once it was established that mental health coverage was provided, we engaged in a multi-step process to determine the details of coverage: whether it was for outpatient counseling therapy, whether all or only a subgroup of mental disorders were covered, whether coverage included “telehealth,” and what were the associated out-of-pocket costs. We first searched for any mention of the term “outpatient” in the Mental Disorder subsection. If mentioned, we examined this term in the “Glossary” to confirm that an outpatient was an individual who would be treated outside of a hospital or medical clinic. We then assessed coverage for verbal counseling therapy. To do this, we examined variant mentions of the term “medical provider” or “specialist” in the Mental Disorder section and potentially in connection with the mention of “outpatient” treatment in the Mental Disorder section. If a “medical provider” was mentioned, we examined the particular term in the Glossary. This was done to evaluate whether the type of provider covered included “psychologists,” “therapists,” “counselors,” “social workers,” and other mental health medical personnel not limited to psychiatrists. In the Mental Disorder section, we similarly examined the term “treatment,” either in association with, or independent from, clauses mentioning “outpatient.”

Carefully noting that non-counseling forms of outpatient mental health treatment exist, such as electroconvulsive therapy, we searched for language that stated coverage beyond this particular form of treatment and similarly searched for the definition of “treatment” in the Glossary. Together, these efforts demonstrated whether outpatient counseling treatment was covered with practitioners that were non-board-certified doctors of medicine or psychiatry—for example, individuals that facilitate counseling therapy. To validate or further elucidate our understanding based on undetailed or ambiguous language, we examined the Mental Disorder subsection of the Exclusions section where non-coverage is explicitly stated. We examined whether “outpatient” treatment for mental health conditions was explicitly excluded, if particular types of “treatments” were excluded, and if treatment provided by particular types of “medical providers,” such as non-board-certified doctors, was excluded—all further clarifying if indeed outpatient counseling therapy was covered.

We then evaluated the breadth of mental health conditions covered in the contract by examining “mental disorder” (or the particular term used in the “covered

benefits”) in the Glossary. These definitions shed light on which conditions were covered—either by defining mental disorder as any condition listed in the *Diagnostic and Statistical Manual of Mental Disorders (DSM)*—the broadest possible list of conditions—or by listing specific conditions or classes of conditions, thus indicating that conditions falling outside of the contract’s definitions may not be covered (American Psychiatric Association 2013). Similarly, to gather more complete information and further elucidate any potential ambiguity, we then examined the Exclusions section, searched for “mental disorder” (or its variants), and noted explicitly excluded conditions or classes of conditions, such as the commonly excluded “impulse disorders.” Next, we examined if coverage existed for “telemedicine” and whether outpatient mental health treatment was clearly stated as eligible.

Lastly, we looked at costs associated with receiving outpatient mental health counseling treatment and needed to utilize industry-specific knowledge about the hierarchical relationship between different contract elements in a health plan. The SBC is uniformly regulated and more regularly updated by insurance companies, and so we judged it to be the most reliable document for costs and general coverage. As such, in comparing hierarchically related documents, we compared the “effective date” of each plan’s SBC and main contract. If the date listed in those documents matched, we derived all cost information from the SBC. However, if a discrepancy existed, we relied on the main contract as the source of all cost information, ensuring internal validity in reference to the other coverage mentioned in the main body of the contract and knowing that some changes in coverage may have been instituted for that plan at some later point in the year but that the exact changes could not be verified. Whether using the SBC or the main contract, we captured costs noted for the “overall deductible” (taking note of whether it had to be met to avail outpatient mental health treatment), “out-of-pocket limit” for individuals, and cost in dollars or as a percentage that needed to be paid by the beneficiary to secure in-network and out-of-network outpatient services related to mental health. For all of the rules developed for the content analysis, we similarly stipulated under what circumstances an exception to the rule would be allowed.

The last portion of rule development for the content analysis involved industry-specific knowledge that particular medical conditions and components of treatment may be covered despite not being mentioned in any component of the health plan contract. Non-mentions posed a challenge to our analysis because it is often the case that a “non-mention” of a medical condition or particular treatment truly means that the plan offers no coverage for it, but it is not clearly excluded by not being mentioned. As a result, we worded our variables and substantive conclusions of our analyses to state that “coverage is not stated or known” instead of saying that “coverage is not provided.” We were careful to not erroneously conclude that, because “mental disorders” were not mentioned, they were not contractually determined or covered in other legal documents held by the health insurance company.

Rule Application

After rule development, we embarked on the stage of rule application, testing the rules developed in the preliminary stage of research to gauge how well they fit the

existing data (that is, the contractual text from different contracts). As we extracted information from the text using the protocol and earlier predefined terms and their alternatives, we kept a complete and exhaustive list of successful rule applications as well as instances where the rules did not yield desired results. We found it very useful to record explanations for any divergences in interpretation.

Rule Validation

Once our stage of rule application ended, we entered a rule validation stage. We updated existing rules using the feedback obtained and recorded during our rule application stage, with a specific emphasis on the list of divergences in interpretation/variable value and their respective explanations. Open to the possibility of creating new variables when it appeared necessary, we went through a process of aggregating and disaggregating information while testing the superiority and functionality of new rules by applying them to other test cases, such as a randomly selected set of contracts.

Simultaneously and iteratively related to the process of rule development, application, and validation for the content analysis, we developed our quantitative variables and outlined the process for recording the captured data in quantitative variables. This allowed us to validate the information based on references to different contractual sections and terms and then to disaggregate the information into more defined or new variables based on whether more contractual information was available. For example, we first defined our coverage variable broadly, such as “Yes/No: the contract contains a section that mentions coverage for mental health.” Then, as it became clearer that different contractual sections and terms needed to be cross-referenced and that there was more detailed information on coverage, more detailed variables were created. It was only upon iteratively seeing the coverage provided (or not) by several contracts that we could then better define what the appropriate values/categories/ranges should be for the variables. This process ultimately allowed us to define one of our main variables: “Yes/No: outpatient mental health counseling for the DSM’s major nine conditions is covered.”

Step 6: Our Contractual Text Data Extraction, Classification, and Validation—Avoid Measurement Error

To help ensure reliability and reduce measurement error during data entry for lengthy and linguistically dense contracts, we systematically structured our data entry effort in five ways: using double data entry by two coders for each contract, following protocols for data classification and its validation, utilizing a training phase for data entry before the official data entry phase, calculating the extent of inconsistencies between the double data entry efforts, and then having a third coder correct any coding inconsistencies. Independently of each other, two individuals entered data for each contract in our sample, allowing us to test the robustness of the initial rules developed for classification and better detect inconsistencies that would otherwise remain unnoticed with one coder. However, relying on more than one coder is not a foolproof

method: coders can make identical errors or explicitly develop an informal rule that will result in incorrect classification that is not immediately evident given the lack of discrepancy. Therefore, we additionally developed a protocol for data classification and its validation across training and non-training phases of data entry by the coders.

To achieve uniformity in coding, our goals for the training phase were threefold: (1) to introduce coders to the content analysis rules and preemptively troubleshoot any misunderstandings; (2) to gauge whether our initial content analysis rules performed well; and (3) to modify our existing content analysis rules, if needed, given the information from the text of the contract. Naturally, this training phase for data entry was the backbone to the iterative process of rule development, application, and validation for the content analysis and the iterative process of variable development. We used a subset of one hundred contracts out of the total 435 for this phase. The two coders' work was regularly (twice weekly) and systematically assessed to address coding inconsistencies across the two data sets that they were simultaneously constructing. With nonmatching variable values highlighted for our group meetings, with no information as to which coder provided which answer, the principal investigator (PI) adjudicated the difference without introducing bias.

We then discussed whether the inconsistencies were due to misunderstandings of rules or limitations of existing rules when addressing a novel case in the contract. If it was the latter, a new rule would be developed, the coders would return to the contracts to update those instances given the new rule, and then they would apply it to a new subset of the one hundred contracts. This process was repeated until we neared our specified threshold for accuracy. The training phase took the lion share of time and effort from the team, and its duration was informed by the coders' familiarity with the contract types, the complexity of the contractual text, whether inconsistencies were random or systematic, and the speed of their resolution. In the second phase, coders worked independently and without feedback to enter data for the remainder of the contracts in our sample. We set weekly goals for the number of contracts to be quantized and highlighted the discrepancies in batches as data entry proceeded.

When all of the data entry was complete, we evaluated the extent of divergence between the two data sets with a statistic of inter-coder reliability. Broadly speaking, we computed the percentage of non-divergences or matching spreadsheet cells, from the total number of spreadsheet cells. The total number of variables was twenty-five, and, for 435 plans, this produced 10,875 spreadsheet cells. Our inter-coder reliability was 0.75 (rounded), meaning that 75 percent of the data entered matched between the two coders. Lastly, because our produced data set was relatively small, and we had low tolerance of risk for having inaccurate data, the PI—that is, someone who was not an initial coder—addressed every mismatch and adjudicated the correct answer. The resulting data set was then exported into STATA statistical package for the analysis.

Step 7: Our Data Administration Practices

Data administration for mixed-methods contractual research can be complex as it involves storing and regularly accessing numerous, lengthy contracts, spreadsheets for

storing this content in quantitative form, statistical files that utilize these data for analysis, and numerous team members that can access each of these files and the software. (We should point out that this step is not a distinct phase but should be planned out in advance.) For our investigation, the contracts were stored as PDF files in the online database tool created and run by the consulting firm. Access to the contracts was granted to team members who entered their granted credentials for database entry. The contracts in this database had already been organized by the consulting firm in numerous ways, such as by year, state, sector, and so on, and the database software allowed contracts to be searched and grouped using these specified criteria. For data entry by our team, each coder was provided with an identical Excel spreadsheet entitled with their unique identifier and the unpopulated variables of interest, each with drop-down options for potential variable values. After each session of data entry, copies of the spreadsheets were saved on the coders' hard drives and uploaded to our cloud-based storage service (Dropbox) where they were visible to other team members. Based on these two spreadsheets, the person calculating the inter-coder reliability created a separate spreadsheet with populated variable values that indicated mismatching cells with highlighting. A new spreadsheet, with reconciled data performed by the third party, was then created from this spreadsheet and served as the master data set in Excel. This reconciling occurred only after data entry had been completed in order to not introduce bias or accidentally alter any coder's data. Data from this Excel spreadsheet was then transferred into a STATA readable data structure using Stat/Transfer and was saved on our cloud-based storage system, as were statistical files for the analysis.

Step 8: Analyzing Meaningful Relationships in Our Data

We took several steps to demonstrate the incidence of coverage and correlations between contextual variables and health insurance contractual elements in our substantive analysis of outpatient mental health counseling coverage. We first documented the percentage of plans that provided coverage for outpatient mental health counseling, the number that provided coverage for any mental health condition listed in the *DSM*, and the average patient responsibility costs associated with outpatient mental health counseling treatment if coverage was provided, and so on. Heretofore, this information was largely unknown, and prior research on these topics generally relied on individuals' perceptions of their mental health coverage. We then moved to more explanatory research, examining if there were associations between coverage (or lack thereof) for outpatient mental health counseling and ERISA-governed corporation size and industry type, also assessing competing explanations such as a health maintenance organization versus a preferred provider organization and carrier (such as an insurance company; in these cases, a third party administrator for a self-insured corporation). If outpatient mental health counseling was covered, we further examined if there were associations between the breadth of mental health diagnoses covered or associated costs for coverage and company size and industry type, also examining if costs for care were associated with exclusions or breadth of coverage. Testing for statistical significance, we could conclude whether the associations observed in the sample data reflected true associations in the population of ERISA-governed health insurance contracts. We were further able to

create composite scores of generosity of outpatient mental health coverage and rank ERISA corporations and carriers. We are preparing these results for independent publication and so do not reproduce the whole analysis here nor many of its findings. However, our statistical approach enabled us to assess the incidence of outpatient mental health coverage, its out-of-pocket costs, the relative rank in generosity of coverage, and potential explanations for variation in coverage for ERISA-governed corporations, substantially advancing prior research efforts on insurance coverage of mental health.

We share here some findings for a subset of 105 health insurance contracts from financial services corporations: some form of coverage for outpatient mental health talk therapy is stated by 97 percent of plans, and, of these, 99 percent do not explicitly state exclusion of treatment for any mental illness, while 1 percent restricts treatment to nine major mental disorders outlined in the *DSM* (National Collaborating Centre for Mental Health et al. 2011; National Alliance on Mental Illness 2016). Only 33 percent of plans that offer coverage do not require that a deductible be met for in-network treatment, while 31 percent of plans have an individual deductible amount that is greater than \$1,500 for in-network treatment. Although we further analyze and elaborate in a separate article being prepared for publication, we note here that these plans are relatively generous in their coverage; many plans from other types of corporate industries state limitations to the specific mental disorders covered for treatment. This finding raises questions as to why coverage for outpatient mental health talk therapy may be more generous in certain industries than in others.

CONCLUSION: THE POWER OF MIXED-METHODS CONTRACT RESEARCH AND MANAGING DIFFICULTIES

Contracts encapsulate individuals' and organizations' responsibilities, protections, entitlements, and remedies in numerous sectors of our society. In this article, we make the case for legal and socio-legal research on contracts to cohere as a research strategy that utilizes mixed qualitative and quantitative methods to assess social and legal explanations for if, how, and why the law is reflected and reproduced in contracts—potentially explaining the continuation or amelioration of social injustices—and/or if contractual content informs human behavior or vice versa. For contract research to be more impactful, researchers must produce valid, reliable, and reproducible contractual data and analyses that allow scholars to build on each other's work. This may also enable contractual scholarship to use causal inference—going beyond the documentation of associations and into the realm of delineating causal explanations. For this to be realized, a protocol is needed. We present here rigorous, methodological steps that explain how contractual data on any substantive topic can be collected, quantitized, and analyzed, with a focus on understanding the population of contracts represented by a sample and structuring variable development, data transformation, and analyses that can answer specified research questions, while reducing measurement error at each step in the process—all leading to data and analyses that are valid, reliable, and reproducible.

In the real world of research, however, there are often significant limitations to overcome. Sampling problems and labor supply are major obstacles to a fully realized mixed methods analysis of contract documents. Some problems are unsolvable, such as when there is no realistic way to gain access to purely private contracts. We focus our concluding thoughts on adaptations that researchers in different positions might need to make. For example, it is likely that a researcher could create a database of contracts but would not be able to say with confidence what population of contracts the sample represents. Authors should describe their samples as fully as possible, drawing on secondary data to understand if there are systematic characteristics of the contracts that explain their inclusion in the sample and if there are characteristics that explain why, within that sample, particular (types of) contracts are missing from the sample, using careful language to acknowledge the limitations of the sample while describing the likely, more narrow subpopulation of contracts that is represented by the sample.

We concede that much of what we outline here takes considerable resources and time. Ideas for how scholars can tackle resource challenges include (1) collaborating with colleagues that have shared interests in fields that regularly receive grant money, such as demography, public health, and medicine; (2) analyzing contracts that are short and straightforward and/or terms that are readily findable, clear, and not linked to numerous sections in a contract; and (3) analyzing a very specified subpopulation of contracts that requires a smaller quantity of contracts to guarantee representativity, such as contracts in a specific subfield, in a single city during a single year, and so on. If costs for double coding are insurmountable, some degree of reliability could be assessed by recoding a subsample with a new coder or having a single coder recode the entire sample after taking time away from coding. Scholars can also incorporate students to do double coding as part of course credit—perhaps in a research methodology course—making sure to double-check their work. Whether best practices or resource-restricted practices are undertaken for coding, it is important to spell out what has been done.

Regarding analyses, while quantitized contractual data holds enormous promise for socio-legal research to go beyond descriptions of the types of content found in contracts, most carefully structured contractual analyses will only be able to demonstrate correlations or that variation in one variable is matched by variation in another variable when accounting for competing explanations and mechanisms, indicated by additional variables in an analysis. These latter variables may be difficult to obtain in contractual research, leaving scholarship with demonstrated associations between focal variables, yet short of explanatory mechanisms. Demonstrating a causal relationship with contractual data will be most challenging—as is similarly the case in the social sciences—because it requires data and analyses demonstrating that a change in one variable is correlated with a change in another variable while accounting for all competing explanations. As a natural consequence, “why” research questions, with accompanying answers, may remain rare in the field. Again, scholars should use careful language to clearly relate what conclusions can be drawn from their data and analyses. However, as we have outlined here with research steps and an empirical example, it is possible to go beyond treating contracts as texts and language in qualitative software and describing what one finds in them. Contractual scholars can and should add quantitative analysis to explore socio-legal phenomena. Moreover, recent empirical research in fields like shadow governance (Nili and Hwang 2020) demonstrate that the time is ripe

for the guiding steps outlined herein to be applied to research on documents similar to contracts, such as charters, departmental or internal policies documents, state statutes, and municipal ordinances.

REFERENCES

- Aletras, Nikolaos, Dimitrios Tsarapatsanis, Daniel Preoțiu-Pietro, and Vasileios Lampos. 2016. "Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective." *Peer Journal of Computer Science* 2 (October): e93. <https://doi.org/10.7717/peerj-cs.93>.
- Allee, Todd, Manfred Elsig, and Andrew Lugg. 2017. "Is the European Union Trade Deal with Canada New or Recycled? A Text-as-data Approach." *Global Policy* 8, no. 2: 246–52. <https://doi.org/10.1111/1758-5899.12420>.
- Allee, Todd, and Clint Peinhardt. March 2010. "Delegating Differences: Bilateral Investment Treaties and Bargaining over Dispute Resolution Provisions." *International Studies Quarterly* 54, no. 1: 1–26. <https://doi.org/10.1111/j.1468-2478.2009.00575.x>.
- Allen, Douglas, and Dean Lueck. 1992. "Contract Choice in Modern Agriculture: Cash Rent versus Cropshare." *Journal of Law and Economics* 35, no. 2: 397–426. <https://doi.org/10.1086/467260>.
- American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Arlington, VA: American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890425596>.
- Ben-Shahar, Omri. 2011. "Fixing Unfair Contracts." *Stanford Law Review* 63, no. 4: 869–906.
- Ben-Shahar, Omri, and James J. White. 2006. "Boilerplate and Economic Power in Auto Manufacturing Contracts." *Michigan Law Review* 104, no. 5: 953–82.
- Berk, Hillary L. 2020. "Savvy Surrogates and Rock Star Parents: Compensation Provisions, Contracting Practices, and the Value of Womb Work." *Law & Social Inquiry* 45, no. 2: 398–431. <https://doi.org/10.1017/lsi.2019.57>.
- Bishara, Norman D., Kenneth J. Martin, and Randall S. Thomas. 2015. "An Empirical Analysis of Noncompetition Clauses and Other Restrictive Postemployment Covenants." *Vanderbilt Law Review* 68, no. 1: 1–52.
- Chen, Yuanyuan, and Anandhi Bharadwaj. 2009. "An Empirical Analysis of Contract Structures in IT Outsourcing." *Information Systems Research* 20, no. 4: 484–506. <https://doi.org/10.1287/isre.1070.0166>.
- Choi, S. J., M. Gulati, and E. A. Posner. 2012. "The Evolution of Contractual Terms in Sovereign Bonds." *Journal of Legal Analysis* 4, no. 1: 131–79. <https://doi.org/10.1093/jla/las004>.
- Coyle, John F. 2016. "The Role of the CISG in US Contract Practice: An Empirical Study." *University of Pennsylvania Journal of International Law* 38: 195–240.
- . 2019. "The Role of the CISG in Canadian Contract Practice: An Empirical Study." *Journal of Law & Commerce* 38: 65–82.
- Epstein, Lee, Jeffrey Allan Segal, Harold J. Spaeth, and Thomas G. Walker. 2015. *The Supreme Court Compendium: Data, Decisions and Developments*. 6th ed. Los Angeles: Sage Publications.
- Frankenreiter, Jens. 2018. "Are Advocates General Political? An Empirical Analysis of the Voting Behavior of the Advocates General at the European Court of Justice." *Review of Law & Economics* 14, no. 1: 20160037. <https://doi.org/10.1515/rle-2016-0037>.
- Furth-Matzkin, Meirav. 2017. "On the Unexpected Use of Unenforceable Contract Terms: Evidence from the Residential Rental Market." *Journal of Legal Analysis* 9, no. 1: 1–49. <https://doi.org/10.1093/jla/lax002>.
- Galanter, Marc. 1974. "Why the 'Haves' Come Out Ahead: Speculations on the Limits of Legal Change." *Law & Social Inquiry* 9: 95–160.
- Gulati, Mitu, and Robert E. Scott. 2011. "Introduction: The Three and a Half Minute Transaction: Boilerplate and the Limits of Contract Design." *Social Sciences Research Network Electronic Journal* 40, no. 1: 1–10. <https://doi.org/10.2139/ssrn.1945988>.

- Hart, Oliver. 2017. "Incomplete Contracts and Control." *American Economic Review* 107, no. 7: 1731–52. <https://doi.org/10.1257/aer.107.7.1731>.
- Hoffman, David A., and Anton Strezhnev. 2022. "Leases as Forms." *Journal of Empirical Legal Studies* 19, no. 1: 90–134. <https://doi.org/10.1111/jels.12309>.
- Kahan, Marcel, and Michael Klausner. 1997. "Standardization and Innovation in Corporate Contracting (or 'The Economics of Boilerplate')." *Virginia Law Review* 83, no. 4: 713–70. <https://doi.org/10.2307/1073747>.
- Katz, Daniel Martin, Michael J. Bommarito, and Josh Blackman. 2017. "A General Approach for Predicting the Behavior of the Supreme Court of the United States." *PLOS One* 12, no. 4: e0174698. <https://doi.org/10.1371/journal.pone.0174698>.
- Kirkland, Anna, Shaubin Talesh, and Angela K. Perone. 2021. "Transition Coverage and Clarity in Self-Insured Corporate Health Insurance Benefit Plans." *Transgender Health* 6, no. 4: 207–16. <https://doi.org/10.1089/trgh.2020.0067>.
- Koremegos, Barbara. 2013. "The Continent of International Law." *Journal of Conflict Resolution* 57, no. 4: 653–81. <https://doi.org/10.1177/0022002712448904>.
- Korobkin, Russell. 2002. "Empirical Scholarship in Contract Law: Possibilities and Pitfalls." *University of Illinois Law Review* 2002, no. 4: 1033–66.
- Leone, Marie. 2006. "Late ERISA Filers May Number 35,000." CFO: *Risk and Compliance*, June 12. <https://www.cfo.com/risk-compliance/2006/06/late-erisa-filers-may-number-35000/>.
- Lucas, Christopher G., Thomas L. Griffiths, Joseph J. Williams, and Michael L. Kalish. 2015. "A Rational Model of Function Learning." *Psychonomic Bulletin & Review* 22, no. 55: 1193–1215.
- Luker, Kristin. 2010. *Salsa Dancing into the Social Sciences: Research in an Age of Info-Glut*. Cambridge, MA: Harvard University Press.
- Macaulay, Stewart. 1963. "Non-Contractual Relations in Business: A Preliminary Study." *American Sociological Review* 28, no. 1: 55–67.
- Marotta-Wurgler, Florencia. 2007. "What's in a Standard Form Contract? An Empirical Analysis of Software License Agreements: Empirical Analysis of Software License Agreements." *Journal of Empirical Legal Studies* 4, no. 4: 677–713. <https://doi.org/10.1111/j.1740-1461.2007.00104.x>.
- . 2008. "Competition and the Quality of Standard Form Contracts: The Case of Software License Agreements." *Journal of Empirical Legal Studies* 5, no. 3: 447–75. <https://doi.org/10.1111/j.1740-1461.2008.00130.x>.
- . 2009. "Are 'Pay Now, Terms Later' Contracts Worse for Buyers? Evidence from Software License Agreements." *Journal of Legal Studies* 38, no. 2: 309–43. <https://doi.org/10.1086/596040>.
- Marotta-Wurgler, Florencia, and Robert Taylor. 2013. "Set in Stone: Change and Innovation in Consumer Standard-Form Contracts." *New York University Law Review* 88: 240–85.
- Medvedeva, Masha, Michel Vols, and Martijn Wieling. 2020. "Using Machine Learning to Predict Decisions of the European Court of Human Rights." *Artificial Intelligence and Law* 28, no. 2: 237–66. <https://doi.org/10.1007/s10506-019-09255-y>.
- Monahan, Amy B., and Daniel Schwarcz. 2022. "Rules of Medical Necessity." *Iowa Law Review* 107: 423–94.
- Nader, Laura. 1974. "Up the Anthropologist: Perspectives Gained From Studying Up." In *Reinventing Anthropology*, edited by Dell Hymes, 284–311. New York: Vintage Books.
- National Alliance on Mental Illness. 2016. "Public Policy Platform." <https://www.nami.org/Advocacy/Policy-Platform>.
- National Association of Insurance Commissioners. 2021. "System for Electronic Rates and Forms Filing." *SERFF Filing Access*. Accessed August 2022. https://www.serff.com/serff_filing_access.htm.
- National Collaborating Centre for Mental Health, National Institute for Health, Clinical Excellence, British Psychological Society, and Royal College of Psychiatrists. 2011. *Common Mental Health Disorders: The NICE Guideline on Identification and Pathways to Care*. London: Stanley L. Hunt Printers.
- Ngaage, Ledibabari M., Brooks J. Knighton, Caroline A. Benzal, Katie L. McGlone, Erin M. Rada, Devin Coon, Rachel Bluebond-Langner, and Yvonne M. Rasko. 2020. "A Review of Insurance

- Coverage of Gender-Affirming Genital Surgery.” *Plastic & Reconstructive Surgery* 145, no. 3: 803–12. <https://doi.org/10.1097/PRS.00000000000006591>.
- Nili, Yaron, and Cathy Hwang. 2020. “Shadow Governance.” *California Law Review* 108: 1097–1146.
- Nyarko, Julian. 2019. “We’ll See You in . . . Court! The Lack of Arbitration Clauses in International Commercial Contracts.” *International Review of Law and Economics* 58: 6–24. <https://doi.org/10.1016/j.irle.2018.12.004>.
- Rachlinski, Jeffrey J., and Andrew J. Wistrich. 2017. “Judging the Judiciary by the Numbers: Empirical Research on Judges.” *Annual Review of Law and Social Science* 13, no. 1: 203–29. <https://doi.org/10.1146/annurev-lawsocsci-110615-085032>.
- Rushin, Stephen. 2017. “Police Union Contracts.” *Duke Law Journal* 66, no. 6: 1191–1266.
- Ryall, Michael D., and Rachelle C. Sampson. 2009. “Formal Contracts in the Presence of Relational Enforcement Mechanisms: Evidence from Technology Development Projects.” *Management Science* 55, no. 6: 906–25. <https://doi.org/10.1287/mnsc.1090.0995>.
- Schwarcz, Daniel. 2011. “Reevaluating Standardized Insurance Policies.” *University of Chicago Law Review* 78, no. 4: 1263–1348.
- Spiggle, Tom. 2021. “President Biden’s Recent Executive Order Takes Aim At Non-Competes.” *Forbes*, July 16. <https://www.forbes.com/sites/tomspiggle/2021/07/16/president-bidens-recent-executive-order-takes-aim-at-non-competes/>.
- Sulea, Octavia-Maria, Marcos Zampieri, Mihaela Vela, and Josef van Genabith. 2017. “Predicting the Law Area and Decisions of French Supreme Court Cases.” <https://doi.org/10.48550/ARXIV.1708.01681>.
- Talesh, Shauhin. 2012. “Insurance Law as Public Interest Law.” *University of California Irvine Law Review* 2: 985–1009.
- Tashakkori, Abbas, and Charles Teddlie. 1998. *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Applied Social Research Methods Series, vol. 46. Thousand Oaks, CA: Sage Publications.
- US Department of Labor. 2021. “Form 5500 Datasets.” *Employee Benefits Security Administration*. <https://www.dol.gov/agencies/ebsa/about-ebsa/our-activities/public-disclosure/foia/form-5500-datasets>.
- Weidemaier, W. Mark C., and Mitu Gulati. 2013. “A People’s History of Collective Action Clauses.” *Virginia Journal of International Law* 54, no. 1: 51–96.
- Wilkinson-Ryan, Tess. 2017. “The Perverse Consequences of Disclosing Standard Terms.” *Cornell Law Review* 103, no. 1: 117–76.