TRACKING THE EVOLUTION OF STARE DECISIS

Sam C. Ehrlich & Ryan M. Rodenberg*

I.   INTRODUCTION

At the United States Supreme Court, what is old is new again. In a series of recent opinions,[1] the justices have repeatedly offered differing views on how stare decisis should be positioned when tasked with justifying or rejecting existing precedent.[2] Indeed, in three recent Supreme Court decisions the justices have wrestled with the effect of stare decisis on future decisions.[3] Reversing a decision, according to Justice Kagan, "demand[s] a 'special justification.'"[4] In contrast, Justice Thomas posited that "[w]hen faced with demonstrably erroneous precedent, my rule is simple: We should not follow it."[5] Chief Justice Roberts, in explaining his switch in direction from a prior dissent, concluded that "[t]he legal doctrine of stare decisis requires us, absent special circumstances, to treat like cases alike."[6] The result is that some precedent is retained while other precedent is discarded, which ensues the debate over whether the justices "practice what they preach."[7]

---

*   Sam C. Ehrlich, J.D./Ph.D., Assistant Professor of Legal Studies, Boise State University College of Business and Economics; Ryan M. Rodenberg, J.D./Ph.D., Associate Professor, Department of Sport Management, Florida State University.

[1]   Gamble v. United States, 139 S. Ct. 1960, 1984 (2019); Gundy v. United States, 139 S. Ct. 2116, 2148 (2019); Kisor v. Wilkie, 139 S. Ct. 2400, 2422 (2019); Allen v. Cooper, 140 S. Ct. 994, 1003 (2020); Ramos v. Louisiana, 140 S. Ct. 1390, 1405 (2020); June Med. Servs. L.L.C. v. Russo, 140 S. Ct. 2103, 2134 (2020).

[2]   Stare decisis is much-debated at the appellate level, too. In a recent Fifth Circuit case, a dissenting judge opined that his colleagues "suddenly discover[ed] that stare decisis is for suckers." Consumer Fin. Prot. Bureau v. All Am. Check Cashing, 953 F.3d. 591, 603 (5th Cir. 2020) (Smith, J., dissenting).

[3]   *Allen*, 140 S. Ct. at 994; *Gamble*, 139 S. Ct. at 1960; *June Med. Servs. L.L.C.*, 140 S. Ct. at 2103. Such a conversation will undoubtedly arise again in the Court's pending abortion case, *Dobbs v. Jackson Women's Health*, No. 19-1392 (2021), where the petitioners have explicitly called for the Court to set aside its previous decisions in *Roe v. Wade*, 410 U.S. 113 (1973), and *Planned Parenthood of Se. Pa. v. Casey*, 505 U.S. 833 (1992). *See* Brief for Pet., Dobbs v. Jackson Women's Health Org., No. 19-1392, 2021 WL 3145936, at *1 (2021) ("The stare decisis case for overruling *Roe* and *Casey* is overwhelming."). *See also* Richard Re, *Stare Decisis as Crying Wolf*, RE'S JUDICATA (Sep. 21, 2021, 7:48 AM), https://richardresjudicata.wordpress.com/2021/09/21/stare-decisis-as-crying-wolf/ [https://perma.cc/2FVQ-5ZHN[/

[4]   *Allen*, 140 S. Ct. at 1003 (citing Halliburton Co. v. Erica P. John Fund, Inc., 573 U.S. 258, 266 (2014)).

[5]   *Gamble*, 139 S. Ct. at 1984 (Thomas, J., concurring).

[6]   *June Med. Servs. L.L.C.*, 140 S. Ct. at 2134 (Roberts, C.J., concurring).

[7]   Adam Feldman, *Empirical SCOTUS: Precedent: Which Justices Practice What They Preach*, SCOTUSBLOG (July 7, 2020, 2:35 PM), https://www.scotusblog.com/2020/07/empirical-scotus-

The often inconsistent manner by which stare decisis is followed, distinguished, and (occasionally) overruled naturally leads to evolution in ever-growing networks of case law.[8] This evolution can effectively change law over time as different judges reinterpret and reapply precedent in ways that actually change the stare decisis effect of decisions throughout the course of history.[9]

This article aims to take a new approach to studying this phenomenon. Using a much-derided Supreme Court case from a century ago as an illustrative case study,[10] this article systematically captures and tracks the evolution of precedent and stare decisis—quantitatively and visually (*See*

Figure 1)—through a novel mixed methods legal citation network analysis that helps explain the lifespan of precedent at the Supreme Court level and below. This approach also lends itself to potentially revealing previously-unrecognized circuit splits in authority and, in turn, provides a metric for anticipating and recommending Supreme Court cert grants.[11]

---

precedent-which-justices-practice-what-they-preach/ [https://perma.cc/TF9D-T654].

[8] *See* Iain Carmichael et al., *Examining the Evolution of Legal Precedent through Citation Network Analysis*, 96 N.C. L. REV. 227, 254 (2017).

[9] *See id.*

[10] Fed. Baseball Club, Inc. v. Nat'l League of Prof'l Baseball Clubs, 259 U.S. 200 (1922).

[11] *See, e.g.*, Sam C. Ehrlich, *A Three-Tiered Circuit Split: Why the Supreme Court Needs to Hear Alston v. NCAA*, J. LEGAL ASPECTS SPORT (forthcoming 2022) (on file with author) (employing the mixed methods citation network analysis methodology described herein to find, define, and visualize a three-tiered circuit split among cases citing *NCAA v. Board of Regents*, 468 U.S. 85 (1984) for the purposes of applying antitrust law to NCAA amateurism rules). This article formed the basis of an amicus brief in *Alston* encouraging the Supreme Court to grant certiorari based on the identified three-tiered circuit split, which they did in December 2020. *See* Brief for Professor Sam C. Ehrlich as Amicus Curiae Supporting Petitioners, NCAA & Am. Athletic Conf. v. Alston, 141 S. Ct. 1231 (2020) (Nos. 20-512, 20-520); s*ee also* NCAA et al. v. Alston, 141 S. Ct. 1231 (2020) (*cert. granted*); NCAA v. Alston, 141 S. Ct. 2141 (2021) (affirming the decision of the Ninth Circuit while rejecting the other sides of the identified three-tiered circuit split).
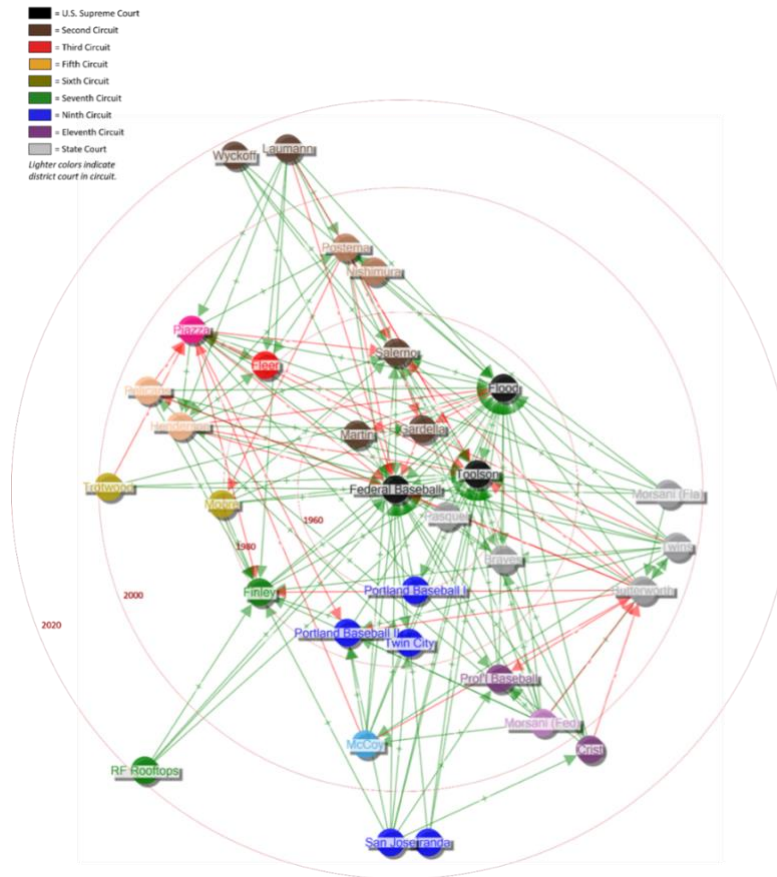
Figure 1: Complete Baseball Exemption Signed Network Graph Visualization.[12]

The component parts of the foundation for this article have a long history. As explained by Fowler and Jeon, "[e]ach judicial citation contained in an opinion is essentially a latent judgment about the case cited."[13] In this regard, law can "evolve" as judges constantly apply, evaluate, reevaluate, and distinguish prior precedent from the current case at hand, thereby expanding

---

[12] As will be explained in this article, this signed network graph represents the entirety of the *Federal Baseball v. National League* citation network with citations color-coded based on whether they followed the prior precedent (green lines with a "+" notation) or whether they rejected or distinguished the prior precedent (red lines with a "-" notation). *See infra* Part III(C)(i).

[13] James H. Fowler & Sangick Jeon, *The Authority of Supreme Court Precedent*, 30 Soc. Networks 16, 17 (2008).

or contracting the scope of different fields of law or fleshing out topic areas in particular applications.[14] As such, the legal literature can be seen as a vast tapestry, where the common law is created through connections between judicial decisions in the form of citations.[15] Each citation links a judicial opinion to a large group of other cases through the judgments made in and about that particular opinion within a field of law.

According to Post and Eisen, "[c]itation to precedent in judicial opinions is a seriously understudied phenomenon."[16] Some scholars within the literature have employed vertex centrality metrics through citation network analysis (CNA) in order to study that tapestry by "provid[ing] a way of quantifying the notion of importance of a case in a citation network."[17] However, legal scholars' previous attempts to use such methods to study judicial opinions have been limited to measurements of the prestige of particular judges,[18] analyses of the citation behavior of appellate courts,[19] and broad-scale efforts to develop a means to predict future case citations.[20] Others, as Carmichael noted, simply use vertex centrality measures "as a proxy for overall importance and, in turn, use them to rank cases."[21]

Along the same lines, Tiller and Cross observed that "legal researchers have extensively dealt with doctrine as a normative matter but have given little attention to the manner in which it actually functions," while "social scientists, who have done important descriptive work about how courts

---

[14]   *Id.*

[15]   *Id.* at 16.

[16]   David G. Post & Michael B. Eisen, *How Long is the Coastline of the Law? Thoughts on the Fractal Nature of Legal Systems*, 29 J. LEGAL STUD. 545, 545 (2000).

[17]   Carmichael et al., *supra* note 8, at 228; *see also* James H. Fowler et al., *Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court*, 15 POL. ANALYSIS 324, 325 (2007) ("[W]e show that network analysis is a viable way of measuring how central a case is to law at the Court[.]").

[18]   *See, e.g.*, Montgomery N. Kosma, *Measuring the Influence of Supreme Court Justices*, 27 J. LEGAL STUDIES 333 (1998); William M. Landes et al., *Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges*, 27 J. LEGAL STUDIES 271 (1998).

[19]   *See, e.g.*, Gregory A. Caldeira, *The Transmission of Legal Precedent: A Study of State Supreme Courts*, 79 AM. POL. SCIENCE REV. 178 (1985); Peter Harris, *Ecology and Culture in the Communication of Precedent Among State Supreme Courts, 1870–1970*, 19 LAW & SOC'Y REV. 449 (1985).

[20]   *See, e.g.*, Carmichael et al., *supra* note 8, at 232; Fowler et al., *supra* note 17, at 326.

[21]   Iain Carmichael, Probabilistic and Geometric Approaches to the Analysis of Non-Standard Data 70 (2019) (Ph.D. dissertation, University of North Carolina at Chapel Hill), https://idc9.github.io/assets/carmichael_dissertation.pdf [https://perma.cc/W92U-DKGF]. *See, e.g.*, Frank B. Cross & James F. Spriggs II, *The Most Important (and Best) Supreme Court Opinions and Justices*, 60 EMORY L.J. 407, 439-40 (2010) (using "legal relevance scores"—their application of centrality metrics within a legal context—to provide a measure of "opinion significance" based on "the number of citations received by an opinion and the significance of the citing cases as measured by the number of citations their cited cases receive.").

actually function, have largely ignored the significance of legal doctrine."[22] Moreover, most scholars employing CNA to study legal citations have done so using huge sample sizes of tens-of-thousands of cases that do not and cannot consider the particular nuances of individual case decisions, citations, and discrete applications of law.[23]

This article aims to fill that void. As network analysis scholars have noted, a focus on a "small, well-defined" network—rather than indiscriminate networks with large samples—can allow for the exploration of "features of network structure unobtainable in representative samples."[24] Applying this observation within the legal context, a focus on a "small, well-defined" network of case law can allow the researcher to give the legal doctrine created by each case its proper contextual due alongside quantitative measures of network influence. This is both important and novel within legal studies because—as Fowler and Jeon noted in a review of approaches to legal CNA papers in 2008—most authors within the legal CNA literature have failed "to consider the *quality* of judicial citations in their research."[25]

To fill this gap, this article offers a mixed methods legal CNA design to allow for both the quantification of legal precedent and the identification of the relative importance of particular cases within a certain field of law from both quantitative and doctrinal perspectives. This new framing to legal CNA studies specifically allows for the identification and study of discrete citation networks, which are defined for this study as citation networks specifically bound by a particular novel application of a field of law to a specific industry or factual context. This new application of legal CNA allows for the use of traditional quantitative CNA metrics to study which cases are the most centralized (i.e. important) in each network. But since these networks are necessarily bound by these unique applications of law, this CNA application also allows for the qualitative legal doctrinal study of these focused citation

---

[22] Emerson H. Tiller & Frank B. Cross, *What is Legal Doctrine?*, 100 N.W. L. REV. 517, 518 (2006).

[23] *See, e.g.*, Fowler et al., *supra* note 17, at 327 (noting a sample size of "26,681 signed or per curiam majority opinions" at the U.S. Supreme Court level); Fowler & Jeon, *supra* note 13, at 18 (noting a sample of "30,288 cases connected together by 220,500 citations"); Carmichael et al., *supra* note 8, at 243 (noting the two samples employed in the study: a Supreme Court network containing "27,885 cases and 235,881 citations" and a Federal Appellate network containing "959,985 cases and 6,649,916 citations.").

[24] Matthew T. Pietryka et al., *From Respondents to Networks: Bridging Between Individuals, Discussants, and the Network in the Study of Political Discussion*, 40 POL. BEHAV. 711, 711-12 (2018).

[25] Fowler & Jeon, *supra* note 13, at 17. *But see* Joseph Scott Miller, *Law's Semantic Self-Portrait: Discerning Doctrine with Co-Citation Networks and Keywords*, 81 U. PITT. L. REV. 1 (2019) (employing co-citation network analysis methods to look at the relative strength of the connections between cases by looking at the number of times cases are cited together while also employing keyword analysis methods to develop tables of frequently used keywords in the context of judicial citations).

networks to analyze the context and quality of each citation to give new insights into the evolution of doctrine in the studied field of law.

To demonstrate the usefulness of this CNA application, this article employs mixed methods CNA to study a discrete citation network with a century-worth of data points: *Federal Baseball v. National League*,[26] an ideally-situated decision that the Supreme Court has re-considered twice since its issuance in 1922.[27] Sports antitrust law is a field of law that is particularly suitable for a study of these discrete networks, as the differing nature of competition in sports relative to more traditional industries often leads to judicially-crafted exemptions that have rarely—if ever—been extended outside of the specific context of sports.[28] This disparate treatment led to the creation of the distinct exemption for baseball from antitrust law, which has created an insular and discrete network of precedent with little application to the larger field of law.[29] Thus, it provides a perfect example for the methodology discussed in this article.

Part II of this article provides the theoretical foundations of this new legal CNA approach by outlining the methodologies and findings of prior interdisciplinary legal studies, with a particular focus on those studies engaging in CNA to study the nature of legal precedent. Part III defines the specific methodological approach in this article by outlining the three-step mixed methods scheme used to define and analyze discrete citation networks of case law, using the baseball exemption as a case study for the process and benefits of such an approach.

II.   PRIOR APPROACHES TO INTERDISCIPLINARY LEGAL RESEARCH

Legal research is difficult to classify within traditional paradigmatic divides. On one hand, legal research is at its core a qualitative research methodology.[30] As Hutchinson and Duncan described, legal doctrinal research—the "research into the law and legal concepts"—is an inherently qualitative field since "[m]any aspects of the law are contingent on context, and need to be interpreted and analysed for meaning."[31] According to

---

[26] Fed. Baseball Club, Inc. v. Nat'l League of Prof'l Baseball Clubs, 259 U.S. 200 (1922).

[27] *See* Toolson v. New York Yankees, Inc. 346 U.S. 356 (1953); Flood v. Kuhn, 407 U.S. 258 (1972).

[28] *See* Claudia G. Catalano, *Application of Federal Antitrust Laws to Professional Sports*, 79 A.L.R. FED. 2d 1 (2003).

[29] *See* Christian L. Neufeldt, *Redeeming the Supreme Court: The Structure Behind the Baseball Trilogy and the Scope of the Baseball Antitrust Exemption*, 27 J. INTELL. PROP. L. 21 (2019).

[30] *See* Terry C. Hutchinson & Nigel Duncan, *Defining and Describing What We Do: Doctrinal Legal Research*, 17 DEAKIN L. REV. 83, 116–17 (2012).

[31] *Id.* at 85, 116.

Hutchinson and Duncan, a major aspect of legal doctrinal research involves applying the law to the facts wherein context is "a highly subjective process" in which "the outcome varies according to the expertise of the individual scholar and cannot be replicated exactly by another researcher" and "is totally dependent on the voice and experience of the individual."[32] As such, Hutchinson and Duncan maintained that legal doctrinal research "is necessarily a qualitative" methodology that requires "a specific language, extensive knowledge and a specific set of skills involving precise judgment, detailed description, depth of thought and accuracy."[33]

On the other hand, Hutchinson and Duncan also argued that legal doctrinal research shares many of the core qualities of quantitative paradigms, positing that "doctrinal research is underpinned by positivism and a view of the world where the law is objective, neutral and fixed."[34] Along these lines, there has been a rapid growth in recent years of the use of interdisciplinary scholarship—including the use of quantitative methods—to discuss legal issues.[35] For example, there is a large body of quantitative legal research that looks to track court opinions based on political ideology in an attempt to predict future cases decided by that judge.[36] Other related research has worked to: (1) determine the likelihood of success of a petition for *certiorari* (review by the Supreme Court) based on a number of factors, including frequent petitioners, the originating appellate court, and the attorney or law firm involved with the two parties to the case[37] and (2) theorize whether artificial intelligence can predict the likelihood of success of a particular case.[38]

According to Posner, the growth of interdisciplinary legal scholarship has been due in part to the fact that "a number of important legal doctrines turned out to be isomorphic with economic theory," requiring a mix of

---

[32] *Id.* at 116.

[33] *Id.*

[34] *Id.*

[35] *See* Richard A. Posner, *Legal Scholarship Today*, 115 HARV. L. REV. 1314, 1317 (2001).

[36] *See, e.g.*, Jeffrey A. Segal & Albert D. Cover, *Ideological Values and the Votes of U.S. Supreme Court Justices*, 83 AM. POL. SCI. REV. 557 (1989); Deborah H. Gruenfeld, *Status, Ideology, and Integrative Complexity on the U.S. Supreme Court: Rethinking the Politics of Political Decision Making*, 68 J. PERSONALITY & SOC. PSYCH. 5 (1995); Michael A. Bailey et al., *Signals from the Tenth Justice: The Political Role of the Solicitor General in Supreme Court Decision Making*, 49 AM. J. OF POL. SCI. 72 (2005); Adam Feldman, *Empiral SCOTUS: Is the Court Tracking Right or Roberts Left?*, SCOTUSBLOG (Mar. 19, 2019, 1:19 PM), https://www.scotusblog.com/2019/03/empirical-scotus-is-the-court-tracking-right-or-roberts-left/ [https://perma.cc/QS9V-LMTF].

[37] Adam Feldman & Alexander Kappner, *Finding Certainty in Cert: An Empirical Analysis of the Factors Involved in Supreme Court Certiorari Decisions from 2001–2015*, 61 VILL. L. REV. 795 (2017).

[38] Daniel Martin Katz, *Quantitative Legal Prediction—Or—How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909 (2013).

economic principles and legal doctrine to make law "more transparent and comprehensible."[39] As an example of this concept, Eisenberg argued that the use of quantitative historical legal data as part of legal scholarship is important to ensure that the policymakers who create law are well-informed and that the legal system is in tune with economic and health-care institutions.[40] However, Posner warned of the dangers of the opposite approach, in which outside disciplines are integrated with the law in so-called "breakthrough scholarship," whereby scholarship is aimed at a readership of primarily scholars rather than legal practitioners.[41]

Posner concluded that "interdisciplinary research is problematic unless subjected to the test of relevance, of practical impact."[42] This conclusion is especially poignant given that, about thirty years earlier, Posner worked with Landes to develop one of the first uses of quantitative research to explain the nature and quality of legal doctrine.[43] In this piece, Landes and Posner discussed how "the rule promulgated by a decision is not the court's express statement, if any, of a rule; rather, it is the court's *holding*, that is, the minimum rule . . . necessary to explain the outcome of the case."[44] Accordingly, Landes and Posner attempted to shift the discussion in legal research away from outside predictive measures and towards the examination of precedent to create legal rules, which they felt at the time had "received little theoretical or empirical analysis."[45]

Indeed, Landes and Posner showed through their research that the divide between quantitative and qualitative legal research does not matter, so long as the focus is on studying case law as influential precedent within a larger chain or network of law rather than as unique and separate elements.[46] That idea informs and guides the legal CNA application in this article.

A. The Role of Precedent in Legal Research

Based on his statements about the nature of legal research, Posner would likely agree that to fully understand legal doctrine—whether it be through

---

[39] Posner, *supra* note 35, at 1318.

[40] Theodore Eisenberg, *Use It or Pretenders Will Abuse It: The Importance of Archival Legal Information*, 75 UMKC L. REV. 1, 23 (2006).

[41] Posner, *supra* note 35, at 1320–21.

[42] *Id.* at 1326.

[43] William M. Landes & Richard A. Posner, *Legal Precedent: A Theoretical and Empirical Analysis*, 19 J. L. & ECON. 249 (1976).

[44] *Id.* at 249–50.

[45] *Id.* at 250.

[46] *Id.* at 250.

traditional legal doctrinal research or through interdisciplinary research—scholars must understand and be able to speak to the role of precedent within the law. Indeed, Hutchinson and Duncan note that within the legal doctrinal research methodology, the term "doctrinal" is "closely linked with the doctrine of precedent—legal rules take on the quality of being doctrinal because they are not just casual or convenient norms, but because they are meant to be rules which apply consistently and which evolve organically and slowly."[47]

In formulating a means to measure the force of legal precedent through empirical analysis, Landes and Posner noted that the true weight of a case decision does not begin to take shape until after it has been cited in other decisions.[48] As they explained, a judge-made ruling created by a single decision will "tend to be extremely narrow in scope," and "a broader rule will generally require a series of judicial decisions" in order to have a wider effect on the judicial landscape.[49] As such, Landes and Posner saw judicial precedent more as "inputs into the production of judge-made rules of law than as the rules themselves," where an earlier decision "provides a reason for deciding a subsequent similar case the same way, and a series of related precedents may crystallize a rule having almost the same force as a statutory rule."[50]

Landes and Posner's work represented a shift within the field of quantitative legal research from viewing cases as separate data points to studying the interconnectivity of case precedent.[51] To this end, they wrote that precedent can be explained in ordinary language as "something done in the past that is appealed to as a reason for doing the same thing again."[52] They felt that this ordinary definition was also true in law, wherein an earlier decision "provides a reason for deciding a subsequent similar case the same way," but more importantly also provides "a series of related precedents may crystallize a rule having almost the same force as a statutory rule."[53] Consequently, they described legal precedent as "inputs into the production of judge-made rules of law" rather than the rules themselves.[54] As such,

---

[47] Hutchinson & Duncan, *supra* note 30, at 84–85.
[48] Landes & Posner, *supra* note 43, at 250.
[49] *Id.*
[50] *Id.*
[51] *Id.*
[52] *Id.*
[53] *Id.*
[54] *Id.*

Landes and Posner posited that the true weight of a case within the larger body of law could not be discerned until after the case has been cited.[55]

Working along similar lines, Lindquist and Cross sought to formulate and test a theory to "measure the effect of precedent on judicial decisions quantitatively over different case areas and over time."[56] Lindquist and Cross's study was based in large part on Dworkin's comparison of the path of precedent to a chain novel, where a novel is passed from author to author with each subsequent novelist adding a new chapter or section to the work until the novel is deemed complete.[57] Similarly, Lindquist and Cross tested the idea that judicial decisions are affected in large part by prior precedent by using quantitative methods to test the correlation between judicial treatment of precedent and judges' ideological preferences through a comparison of cases of first impression to cases involving issues with a longer string of foundational precedent.[58] However, they found only limited support for the idea that precedent affords much constraining effect on judges' ideological tendencies.[59]

However, Balkin argued that the law should be seen as decidedly nonlinear, choosing instead to theorize the nature of the legal thought as "a descending series of rule choices of increasing factual complexity and specificity."[60] Post and Eisen would later add to this approach by theorizing that judicial doctrine is better described through fractal geometry, with a "branching doctrinal structure" that "has no natural stopping place but continues indefinitely downward" with individual rule choices as the "branching motif, endlessly repeated at finer and finer scale."[61] But Post and Eisen concluded that, unlike with fractal shapes, "the branching process [of law] does not continue indefinitely," as "[e]ach decided case represents a single instantiation of this process that has come to rest at some point" where "each opinion [is] a single tree in the forest of judicial opinions."[62] One tree with enough precedential force within a forest could then beset thousands of

---

[55] *Id.*

[56] Stefanie A. Lindquist & Frank B. Cross, *Empirically Testing Dworkin's Chain Novel Theory: Studying the Path of Precedent*, 80 N.Y.U. L. REV. 1156, 1157 (2005).

[57] *Id.* at 1156; RONALD DWORKIN, LAW'S EMPIRE 228–38 (1986). *See also* Cass Sunstein, *How Star Wars Explains Constitutional Law*, WASH. POST (Apr. 29, 2015), https://www.washingtonpost.com/news/wonk/wp/2015/04/29/how-star-wars-explains-constitutional-law/?utm_term=.74f6a4b0b153 [https://perma.cc/36DG-ESB4] (using Dworkin's chain novel theory as applied to the Star Wars universe to explain the authorship involved in the creation and evolution of Constitutional law jurisprudence).

[58] Lindquist & Cross, *supra* note 56, at 1158.

[59] *Id.* at 1205.

[60] J.M. Balkin, *The Crystalline Structure of Legal Thought*, 39 RUTGERS L. REV. 1, 8 (1986).

[61] Post & Eisen, *supra* note 16, at 553–54.

[62] *Id.* at 559.

other individual trees in that forest, each linked through the roots, or citations, and "the web of citations from one case to another" would then become "a critical component of the network of rules that comprise 'the law' in any area."[63]

In his seminal 1954 law review piece, "The Authority of Authority," Merryman theorized about the nature of precedent and legal authority.[64] For instance, Merryman wrote that "the choice of an authority by the court, to the extent that it raises the reasonable assumption that the court thought the authority applicable to the legal problem before it, becomes a factor in the decision of future cases."[65] The determination of which cases to cite in a decision, according to Merryman, "has a profound impact on the way the law grows and the shape legal doctrines take."[66] Indeed, Merryman reasoned that not only does "past judicial practice in determining applicability condition[] the judge's choice in the new case," but also the judge's "choice there influences the course of future decisions."[67]

According to Merryman, judicial citations have substantial power.[68] When a court cites a Restatement, for example, Merryman argues that a lawyer practicing in that jurisdiction "is justified in believing that [the] court . . . considers it persuasive, if not binding, authority" and is thus "influenced to use the Restatement himself, and perhaps to think of it as authoritative."[69] This process "becomes cumulative," as "the more frequently the Restatement is cited in judicial opinions the more frequently it will be cited in subsequent ones."[70] In this regard, "addition of prestige to a work by judicial citation has an unavoidable effect on future decisions" since "[a]s a work increases in stature it becomes more authoritative—more capable of influencing the actual consideration of cases by judges."[71]

While Merryman's work here primarily focused on secondary sources like the Restatements and legal encyclopedias, the same theories can obviously be applied to citations of case law. Just as a secondary source gains authority the more often it is cited,[72] the more often a case is cited the more authority it can possess. In fact, Merryman commented that in legal practice "it is quite possible for a secondary authority to be *more* authoritative than

---

63   *Id.* at 545.
64   John Henry Merryman, *The Authority of Authority*, 6 STAN. L. REV. 613, 613 (1954).
65   *Id.* at 614–15.
66   *Id.* at 615.
67   *Id.*
68   *Id.* at 618.
69   *Id.*
70   *Id.*
71   *Id.* at 619.
72   *Id.*

primary authorities" even though "primary authority is binding" and "secondary authority is at most 'persuasive[,]'" since a judge has "so much discretion in determining what is applicable to his case and how it is to be applied that he can, if he wishes to support a conclusion he has reached in the case, find that which supports it applicable and that which does not inapplicable."[73] By the same token, a judge can find certain cases—even, in some situations, out-of-circuit authority—more applicable than others based on the conclusion that he or she wishes to reach. That judge thereby gives that case more power within the discrete topic area than others by citing it and "influenc[ing] the course of future decisions" in that topic area by encouraging the continued citation of the chosen case.[74]

The concepts of the precedent-focused nature of law brought forth by Landes and Posner and Merryman provide the theoretical foundation for the specific CNA application in this article. If, as Merryman conjectured, judges' choices of cases to support specific rationale in legal opinions gives that cited precedent additional power over other similar case law, that power can be both measured using quantitative methods and closely analyzed using qualitative methods to determine both the statistical influence and qualitative interconnectivity of those cases within both broader and discrete networks of case law. This basic philosophy of the power of interconnected precedent on fields of law has also informed other more large-scale studies of precedent and has informed the specific employment of certain quantitative measures to study the nature of law.[75]

B. *Quantitative Approaches to Studying Legal Precedent*

In supposed response to Landes and Posner's mid-1970s statement about the "little theoretical and empirical analysis" existing that focused on the use of precedent to create legal rule,[76] much of the interdisciplinary quantitative legal research that has arisen since Landes and Posner's writing has been in the study of the influence and effect of legal precedent.[77] In light of Merryman's observations about the nature of citations and authority, much

---

[73] *Id.* at 620–21.

[74] *Id.* at 615.

[75] *See, e.g.*, Carmichael et al., s*upra* note 8, at 228 (developing an empirical perspective on precedent through citation networks); Fowler et al., s*upra* note 17, at 325 (studying precedent through network analysis).

[76] Landes & Posner, *supra* note 43.

[77] *See, e.g.*, Carmichael et al., s*upra* note 8, at 228 (developing an empirical perspective on precedent through citation networks); Fowler et al., s*upra* note 17, at 325 (studying precedent through network analysis).

of this research has been to take quantitative looks at the power of particular opinions by counting citations to those opinions.[78]

For example, Merryman himself later conducted a study looking at the different types of authority cited by the California Supreme Court in various decisions, finding that the courts tended to favor "more recent decisions" with a "predictable decline in the 'citation power' of a decision" as the opinion ages.[79] Merryman also found through quantitative analysis that the California Supreme Court had increased in how often it cited federal court decisions, though he did not come to a conclusion as to whether this was due to the increasing "federalization" of law, the "'inherent' quality or authority" of U.S. Supreme Court decisions over state court decisions, or simply due to increased access to federal decisions compared to the great expense of acquiring and maintaining access to all of the state reporters.[80]

In a more recent and focused example, Re conducted a citation count study to describe the "increasing regularity" of the use of the *Marks* rule—a rule determining which judicial opinion becomes the 'binding' precedent when there is a fragmented court and no single opinion receives a majority vote.[81] By counting and mapping the frequency of citations to the case that lent this rule its name—*Marks v. United States*[82]—Re was able to determine that the *Marks* rule has become an authoritative legal principle in both federal jurisdictions and even in state courts, thereby "offer[ing] a case study in precedential expansion."[83] This finding supports Merryman's contention regarding the increasing authority of a case the more that it is cited: "the more frequently [a particular case] is cited in judicial opinions the more frequently it will be cited in subsequent ones."[84]

This nonlinear theorization of legal precedent and legal thought as a "web of citations from one case to another" has led to the application of network theory to case precedent and case law as a means to visualize and

---

[78]   *See* Carmichael et al., s*upra* note 8, at 234–235.

[79]   John Henry Merryman, *Toward a Theory of Citations: An Empirical Study of the Citation Practice of the California Supreme Court in 1950, 1960, and 1970*, 50 S. CAL. L. REV. 381, 394–99 (1978).

[80]   *Id.* at 399–401.

[81]   Richard M. Re, *Beyond the Marks Rule*, 132 HARV. L. REV. 1942, 1944 (2019).

[82]   Marks v. United States, 430 U.S. 188, 193 (1977).

[83]   Re, *supra* note 81, at 1965. Of general interest but of no particular relevance to this study, Re also found that the number of times that *Marks* has been cited for the express purpose of the *Marks* rule has "come to dominate all *Marks* citations." *Id*. at 1955. Indeed, Re found that citations to the First Amendment and ex post facto aspects of *Marks* declined, showing that "most precedents decline in value over time, yielding progressively fewer citations." *Id*. This observation supports Merryman's finding about the "predictable decline in the 'citation power' of a decision" as the opinion ages. Merryman, *supra* note 79, at 394. *See also* Neal Devins & David Klein, *The Vanishing Common Law Judge?*, 165 U. PA. L. REV. 595, 614 (2017) (noting that "precedents typically depreciate in value over time.").

[84]   Merryman, *supra* note 64, at 618.

quantify the "branching doctrinal structure" described by Post and Eisen.[85] Network theory as used in the social sciences is defined as "the mechanisms and processes that interact with network structures to yield certain outcomes for individuals and groups."[86] Put more simply, network theory can be seen as focusing on "the consequences of network variables, such as having many ties or being centrally located."[87]

Generally speaking, analysts focusing on network relationships "focus on two concepts: 'nodes' (i.e., the people, information, etc., within a given setting) and 'links' (i.e., the relationships between nodes)" to both map and measure relationships between entities and "to develop quantitative indicators of interest regarding, for instance, the centrality or prestige of the nodes within a network."[88] Network analysis studies have become popular within the social sciences in recent years; in 1994, Wasserman and Faust noted that the methodology of social network analysis has "attracted considerable interest and curiosity in the social and behavior science community in recent decades," in large part due to "the appealing focus of social network analysis on relationships among social entities, and on the patterns and implications of these relationships."[89]

This "interest and curiosity" in network studies within the social sciences has led to novel applications of network theory in legal studies. This has led to an emerging trend of the use of CNA methodologies to study the nature of precedent and judicial citations. Legal CNA considers judicial opinions as "nodes in a legal network" where specific nodes are linked to other case notes through citation to existing precedent.[90] According to Fowler et al., while "[p]recedent plays a central role in the judiciary by providing information to judges and other decision makers about the relevance or weight of particular facts for a legal issue and by defining legal consequences or tests that pertain to those facts," not all judicial opinions are "equally positioned to serve as precedent for a given dispute."[91] This concept interconnects well with Merryman's theory that precedent is given power by judges finding certain cases more applicable than others based on the conclusion that they wish to reach; it can be said that judges' choices in which cases to cite and give power

---

[85] Post & Eisen, *supra* note 16, at 545, 553.

[86] Stephen P. Borgatti & Daniel S. Halgin, *On Network Theory*, 22 ORG. SCIENCE 1168, 1168 (2011).

[87] *Id.* (citing Daniel J. Brass, Social Networks in Organizations (2002) (unpublished manuscript), http://www.analytictech.com/Networks/Brass_Consequences.doc. [https://perma.cc/Z8P4-E89T]).

[88] Fowler et al., *supra* note 17, at 325. "Nodes" are sometimes also referred to as "vertices," while "links" are sometimes referred to as "edges." Carmichael et al., *supra* note 17, at 229.

[89] STANLEY WASSERMAN & KATHERINE FAUST, SOCIAL NETWORK ANALYSIS: METHODS AND APPLICATIONS 3 (1994).

[90] Fowler et al., *supra* note 17, at 325.

[91] *Id.*

to "influences the course of future decisions," thereby often sparking either evolution or divergent paths within a citation network.[92]

As such, the legal CNA methodology is positioned and crafted to determine the "legal relevance of a case," or "the degree to which the information in a given case remains germane for deciding contemporary legal disputes," which, according to Fowler et al., "lies at the heart of law and legal development."[93] Likewise, Fowler and Jeon noted that an important virtue of the legal CNA approach is the ability to "determine which rulings were thought to be most important and which were most carefully grounded in prior precedent *at any point in time*," granting researchers the unique ability to "test several hypotheses about the rise or fall of precedent."[94] In this regard, the legal CNA methodology allows the legal researcher the ability to observe the evolution of precedent over any given period of time, including any doctrinal shifts or divergences that may result.[95]

Fowler and Jeon employed the legal CNA methodology in a broad study of 30,288 Supreme Court majority opinions ranging from 1754 to 2002 in an effort to study the nature of judicial citations and look for patterns in the larger precedent network of Supreme Court doctrine.[96] Fowler and Jeon found that the doctrine of stare decisis—where judges defer to existing case law when crafting decisions—was not the driving force of judicial decision making in the earliest part of the country's history.[97] Furthermore, Fowler and Jeon were also able to use CNA methods to identify the reign of former Supreme Court Chief Justice Earl Warren as an activist court that revolutionized constitutional law by overruling more precedents than any other court.[98] Fowler and Jeon were able to come to this conclusion based on the fact that the Warren Court "cited fewer cases in their opinions and handed down a greater number of cases without any citations," which is indicative of less reliance on previous court opinions to make decisions.[99] On the other hand—and perhaps predictably—Fowler and Jeon noted fewer citations to Warren-era decisions than normal, as the future Burger and Rehnquist Courts "were unable to justify their [conservative-leaning] policy choices with liberal Warren Court precedents, forcing them to cite more conservative rules

---

[92] Merryman, *supra* note 64, at 615. *See supra* notes 64–74 and accompanying text.

[93] Fowler et al., *supra* note 17, at 325.

[94] Fowler & Jeon, *supra* note 13, at 17.

[95] *See generally* Carmichael et al., *supra* note 17, at 228 (using CNA generally to "understand the factors driving the evolution of law.").

[96] Fowler & Jeon, *supra* note 13, at 17.

[97] *Id.* at 18.

[98] *Id.* at 19.

[99] *Id.*

that were more consistent with their preferences, such as their own or pre-Warren precedents."[100]

More importantly, Fowler and Jeon used CNA metrics to determine that most Supreme Court authority has a tendency to exhibit a "rise and fall" in authoritative influence, with more important cases gaining much more influence over time and declining in influence much more slowly.[101] Fowler and Jeon hypothesized that this tendency was due to the idea of "precedential vitality," in which the Court is more likely to overturn precedents of higher authority.[102] Under this theory, legal decisions "that have not received much attention in the network of precedent are less likely to have an influence over future decisions and are less in need of revision," while the more important precedent would later be overturned to allow the court to clarify the created legal doctrine, thereby allowing these cases to be replaced by new case law.[103]

However, most legal CNA studies can be criticized through their focus on large samples of case law that inherently cannot "consider the *quality* of judicial citations in their research."[104] This limitation is exhibited by one particularly important use of legal CNA in legal studies: its foundation and application in computer sciences, specifically in its use in the development of search engine algorithms in legal databases like Westlaw and LexisNexis.[105] Both Fowler et al. and Fowler and Jeon drew from the computer science literature—particularly the work of Kleinberg[106]—to develop their importance metrics.[107] Conversely, Zhang and Koppaka—two researchers with LexisNexis—based the development of a new algorithm for LexisNexis's search engine in large part on the network nature of legal precedent.[108] However, Zhang and Koppaka went beyond the purely quantitative metrics used by most others studying legal citations to create what they called a "Semantics-Based Legal Citation Network," which was built to "dissect the general, multi-dimensional network into subnetworks

---

[100]   *Id.* at 19–20.

[101]   *Id.* at 25.

[102]   *Id.* (emphasis omitted).

[103]   *Id.* at 25–26.

[104]   *Id.* at 17.

[105]   Paul Zhang & Lavanya Koppaka, *Semantics-Based Legal Citation Network*, INT'L CONF. ON ARTIFICIAL INTELLIGENCE AND LAW 123 (2007).

[106]   *See* Jon M. Kleinberg, *Authoritative Sources in a Hyperlinked Environment*, A.C.M.-S.I.A.M. SYMP. ON DISCRETE ALGORITHMS 668 (1998); Jon M. Kleinberg, *Authoritative Sources in a Hyperlinked Environment*, 46 J. OF THE A.C.M. 604 (1999).

[107]   *See* Fowler et al., *supra* note 17, at 325; Fowler & Jeon, *supra* note 13, at 17.

[108]   Zhang & Koppaka, *supra* note 105, at 123.

each focusing on one specific legal issue."[109] To do so, Zhang and Koppaka wrote an algorithm that allows search-engine users to look not only at the quantitative citation metrics of a case when employing the LexisNexis software but also at the semantics and qualitative context behind a citation.[110]

Zhang and Koppaka observed that legal citations by nature are "semantically multi-dimensional," since "a case can cite several cases each supporting a different proposition; and, likewise, a case can be cited by other cases for different reasons."[111] For example, they noticed that one case, *People v. Green*,[112] had at the time been cited by hundreds of other cases, but the legal issues for which that case had been cited could be grouped into over ten different categories.[113] For that reason, Zhang and Koppaka actually argued that the semantic multi-dimensionality of citations has "made the use of legal citation network impractical" within the LexisNexis search engine function "because a general network traversing function would retrieve [an] indiscriminately huge number of cases and soon fill the network space, making the viewing and reading impossible."[114] A similar argument could be made regarding scholarly applications of legal citation networks.

Miller's research attempted to rectify the limitation posed by purely quantitative approaches to legal CNA by introducing the use of semantic links between cases that can be mapped into what he referred to as "semantic self-portraits" of clusters of similar cases.[115] Miller accomplished this goal through the combination of quantitative CNA and keyword analysis studies, which "compares two corpora to determine which words 'are significantly more frequent in one corpus than another.'"[116] Miller applied these methods to a group of Supreme Court cases bonded by "a common substantive focus and an unmistakably clear boundary" to find common keywords that are common between different groups of cases.[117]

In summing up his findings, Miller noted that the combined methodological framework that he created is merely intended as "an aid to, a further input for, interpretive judgment" and that there is much more to be interpreted based on his findings.[118] To this end, Miller called for further use

---

[109]   *Id.* at 125.

[110]   *Id.* at 124.

[111]   *Id.*

[112]   People v. Green, 609 P.2d 468 (Cal. 1980).

[113]   Zhang & Koppaka, *supra* note 105, at 124.

[114]   *Id.* at 125.

[115]   Miller, *supra* note 25, at 58.

[116]   *Id.* at 18–19 (quoting SUSAN HUNSTON, CORPORA IN APPLIED LINGUISTICS 1, 2 (Cambridge Univ. Press ed., 2002)).

[117]   *Id.* at 21.

[118]   *Id.* at 58.

of the CNA methodology to "deepen our understanding of any patch of the semantic fabric of our decisional law" and called for further research using network-analysis-driven methods to explore the tapestry that is legal doctrine.[119]

In the same vein as Fowler and Jeon's statement that previous studies employing network analysis approaches have often failed to "consider the *quality* of judicial citations in their research,"[120] Pietryka et al. argued in the political science network analysis context that a focus on a "small, well-defined" network allows for the exploration of "features of network structure unobtainable in representative samples."[121] As Miller showed when he focused his CNA and keyword analyses on a "group of Supreme Court cases with a common substantive focus and an unmistakably clear boundary," the same principle can be applied to CNA methodology.[122]

Along these lines, the mixed methods CNA application employed herein is based on the idea that the goal of looking more to the "quality of judicial citations" can be even better accomplished through the dual use of quantitative CNA methods, to outline the scope of discrete networks of case law, and qualitative legal doctrinal methods, to explore the relationships between each citation and gather information about the quality of each citation that is "unobtainable in representative samples" of case law.[123] In his study of the *Marks* rule,[124] for example, Re found that "the *Marks* rule itself generates intractable disagreement," making the rule both "important and uncertain" and thereby ripe for further scrutiny.[125] Such a finding could not be possible with a broader and pure-quantitative focus on wider fields of law, as seen in other legal CNA studies in which conclusions were limited to broad generalizations about the nature of precedent and of legal doctrine based on a sample of tens-of-thousands of cases. In this regard, a more limited and quality-oriented focus on discrete networks focused (as Re and Miller introduced) around a singular issue allows the researcher to better determine differences in judicial application of legal principles—including potential circuit splits.

Whereas Miller used quantitative co-citation network analysis methods to identify commonly used keywords used by judges when citing multiple cases together, the opposite approach—where the semantic keywords are

---

[119]   *Id.*

[120]   Fowler & Jeon, *supra* note 13, at 17. *But see* Miller, *supra* note 25, at 6.

[121]   Pietryka et al., *supra* note 24, at 711–12.

[122]   *See* Miller, *supra* note 25, at 21.

[123]   Pietryka et al., *supra* note 24, at 711–12. *See also* Fowler & Jeon, *supra* note 13, at 17.

[124]   *See supra* notes 81–84 and accompanying text.

[125]   Re, *supra* note 81, at 1965.

used to create and distill a discrete and focused network of case law around particular issues—could be an effective method to both identify and analyze how citations are employed to develop precedent.[126] By drawing both upon that wisdom and the findings of Zhang and Koppaka in developing a legal CNA application that reflects both quantitative importance and centrality and the qualitative semantic quality of citations, future studies could employ the mixed methods CNA application to powerful effect within a broad range of topics in legal studies.

III.   DEFINING THE MIXED METHODS LEGAL CITATION NETWORK ANALYSIS SCHEME

Employing this mixed methods legal CNA methodology involves employing two broad levels of analysis: one quantitative, one qualitative. This leads to a three-step methodological approach. First, the boundary of the discrete citation network must be specified and defined. Second, quantitative network analysis metrics are employed to determine the most centralized (i.e. important, given the specific context of the citation network studied) cases in each citation network. Third, the results of those quantitative network analyses are used in tandem with qualitative legal doctrinal analysis, where case citations are qualitatively identified as either positive or negative to produce signed network graphs that show the character of the network and can be used to demonstrate and further define circuit splits.

A.   *Step One: Boundary Definition and Data Collection*

An important initial consideration and challenge for any network analysis study is to define the network to be analyzed.[127] Unlike traditional quantitative social science research methods in which a dataset consists mostly of composition variables like gender or ethnicity of actors, network analysis data contain "at least one structural variable measured on a set of units," meaning that the focus of social network analysis is to find the "interrelatedness of social elements" rather than to determine several variable

---

[126]   Miller, *supra* note 25, at 34–58.

[127]   Sue Heath et al., *Chasing Shadows: Defining Network Boundaries in Qualitative Social Network Analysis*, 9 QUALITATIVE RES. 645, 650 (2009). *See also* David Knoke & Song Yang, *Social Network Analysis*, 154 QUANTITATIVE APPLICATIONS IN THE SOC. SCI. 15 (2008) ("In any empirical network research, investigators must initially attend to three important issues before beginning to collect data: boundary specification, network sampling, and measurement of relations.").

differences between individuals.[128] Because of this, determining a sample of actors within a study is determined by questions such as "which actors to include," "who are the relevant actors," and "which actors are in the population."[129]

The need for boundary specification is particularly important in network studies of legal case law, as many cases boast an enormous quantity of citations that can potentially be included in an all-encompassing CNA study. For instance, the case at the center of the sports antitrust exemption network chosen to demonstrate the CNA application—*Federal Baseball v. National League*—has been cited in 136 subsequent decisions across dozens of different courts as of January 1, 2020.[130] These citations often have varying qualitative rationales defining the citation that are not simply limited to applying professional baseball to antitrust law, and often do not even discuss *Federal Baseball*'s antitrust holding, let alone its holding specific to baseball.[131]

In accordance with other studies employing CNA to study networks of legal opinions, each node within the identified networks represents one legal opinion in which each edge represents the connection between those cases, or a judicial citation from one case to another. As citations cannot exist on both sides of a given edge (i.e., a past case cannot cite a future case), the citation network studies herein inherently are directional one-mode networks, with each edge linking from one past case to one future case. Each edge is consistently studied as one unweighted citation for the purposes of calculating centrality and other quantitative metrics.

Most CNA research studying legal citations has simply included all cases within broadly defined networks, and thus these studies include tens-of-thousands of cases within their purview.[132] As a result, these studies look at citations without regard to why these particular cases are being cited and

---

[128] Amy Chan Hyung Kim, Knowledge Structure in Sport Management: Bibliometric and Social Network Analyses at 10, 46-47 (2012), https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?acc ession=osu1338378142&disposition=inline [https://perma.cc/4NVC-D3QU] (Ph.D. dissertation, Ohio State University).

[129] *Id.*

[130] *Shepard's Citing Decisions Analysis:* Fed. Baseball Club, Inc. v. Nat'l League of Prof'l Baseball Clubs, LEXISNEXIS (last visited Aug. 15, 2021), https://plus.lexis.com/shepards/shepardspr eviewpod/?pdmfid=1530671&crid=3ed86148-da17-4c7c-b182-10e699a7cdb4&pdshepid=urn%3Ac ontentItem%3A7XWN-0BW1-2NSF-C3FW-00000-00&pdshepcat=initial&prid=de51bc6c-0a1d-4f5 e-b408-d43fbf5f826c&ecomp=8gktk#/citingref.

[131] *See, e.g.*, NLRB v. Hopwood Retinning Co., 98 F.2d 97, 100–01 (2d Cir. 1938) (examining the defendants' claim that they were not under the purview of the National Labor Relations Act (NLRA)—a federal statute that, like the Sherman Act, is limited to interstate commerce under the Commerce Clause—because "it does not engage in commerce but merely performs a service.").

[132] *See* Fowler et al., *supra* note 23 and accompanying text.

instead study, for example, different quantitative characteristics of citations, such as the changing likelihood of a case being cited over time[133] and the relative authoritative power of cases within that larger sample.[134] However, as Zhang and Koppaka noted, legal citations are "semantically multi-dimensional" in nature as "a case can cite several cases each supporting a different proposition; and, likewise, a case can be cited by other cases for different reasons."[135]

A foundational motivation in constructing the CNA application was to study discrete citation networks bound by specific qualitative citation characteristics with the goal of "consider[ing] the *quality* of judicial citations" as called for by Fowler and Jeon.[136] As such, the boundary of the studied networks should be defined by qualitative, doctrinal characteristics. For the case study used in this article, for example, case citations were included only if they discussed or otherwise advanced the threshold issue of whether antitrust law applies generally to professional baseball.

Knoke and Yang identified two boundary specification strategies for social network analysis studies: realist and nominalist strategies.[137] In the realist strategy, the actors within the network set the boundaries themselves, as "[a]ctors and their relations are included or excluded to the extent that the other actors judge them to be relevant."[138] In the nominalist strategy, however, the researcher "impos[es] an a priori conceptual framework that serves an analytic or theoretical purpose for a particular project," where the limitations are drawn on, for example, membership in particular clubs or organizations.[139] Instead of focusing on the actors within the network to set their own boundaries, the nominalist approach "focuses on the theoretical concerns of the researcher."[140]

While he did not explicitly define his boundary-creation approach as such, Miller's boundary creation strategy can be seen as an a priori nominalist framework in his legal CNA study as he set his boundaries by focusing on a data set comprised of "a set of citations within a discrete group of U.S. Supreme Court decisions" that had "a common substantive focus and an unmistakably clear boundary"—namely, a citation-set centered around "the [Supreme] Court's eight decisions about the scope and application of the

---

[133] *See, e.g.*, Carmichael et al., *supra* note 17 at 234–36.

[134] *See, e.g.*, Fowler & Jeon, *supra* note 13, at 31; Fowler et al., *supra* note 17, at 327.

[135] Zhang & Koppaka, *supra* note 108, at 124.

[136] Fowler & Jeon, *supra* note 13, at 17.

[137] Knoke & Yang, *supra* note 127, at 15.

[138] *Id.*

[139] *Id.*

[140] Kim, *supra* note 128, at 48.

Warsaw Convention."[141] Miller used these eight cases to "ma[ke] a two-column list of citing and cited cases" to create a list of connections between cases, which he then paired with common keywords in clusters created by frequent citations of the same cases in multiple subsequent opinions.[142] The use of these keywords—the qualitative data giving context behind each citation—gave Miller's research a mixed-methods flavor that represents a meaningful step toward the focus on "the *quality* of judicial citations" in CNA research as called for by Fowler and Jeon.[143]

The same was true with Re's study, even though it was centered not around several cases like Miller but around one case—the Supreme Court's decision in *Marks v. United States*[144]—and the procedural rule created by that case determining which written opinion governs when no opinion receives more than a plurality of votes from the sitting judges.[145] While Re did not explicitly employ CNA in his study, his use of basic descriptive statistics to study citation data possesses several similarities with legal CNA studies, including his general focus on interpreting citation data using quantitative metrics to study the historical nature of citations through the number of times *Marks* had been cited in both state and federal courts.[146] In order to contextually delineate the sample, Re used qualitative data to distinguish in his quantitative graphs and citation counts between "*Marks* Rule" citations and "Non-*Marks* Rule" citations.[147] This is similar to the work of Miller, who included case citations within his network "without respect to the stated reason, if any, for the citation" but later studied the different semantic applications within citations to "fully discern[] multiple layers of a legal doctrine's texture."[148]

This CNA application combines Miller's and Re's approaches to delineate the discrete citation network surrounding *Federal Baseball* with some key differences in furtherance of the purpose of the study. To that end, the methodology employs a two-column approach similar to the approach described by Miller to outline all of the connections between cases to be represented as network edges in the data visualization stage. Data was

---

[141] Miller, *supra* note 25, at 21–22.

[142] *Id*. at 22–34.

[143] *See* Fowler & Jeon, *supra* note 13, at 17.

[144] Marks v. United States, 430 U.S. 188 (1977).

[145] Re, *supra* note 81, at 1944–47. *See Marks*, 430 U.S. at 193 ("When a fragmented Court decides a case and no single rationale explaining the result enjoys the assent of five Justices, 'the holding of the Court may be viewed as that position taken by those Members who concurred in the judgments on the narrowest grounds.'").

[146] Re, *supra* note 81, at 1952–65.

[147] *Id.*

[148] Miller, *supra* note 25, at 23.

inputted into Microsoft Excel through the NodeXL plugin, which allows for network data collection and visualization.[149] Specifically, data was inputted into NodeXL's edges sheet (which automatically fills in the vertices sheet) with the citing case inputted into the Vertex 1 column and the cited case inputted into the Vertex 2 column, creating an organized matrix of each network edge (i.e., each case citation within the network). (*See* Figure 2).

---

[149] *See* Brian Britt, *Making Social Network Analysis Accessible: A Review of NodeXL*, THOUGHT ARK (Apr. 7, 2012), http://thoughtark.net/making-social-network-analysis-accessible-a-review-of-nodexl/.

Figure 2: NodeXL's setup within Microsoft Excel while visualizing the
baseball exemption citation network.

However, Miller included all citations to his included cases "without
respect to the stated reason, if any, for the citation" and only later used
qualitative semantic data to "fully discern[] multiple layers of a legal
doctrine's texture."[150] By contrast, the mixed-methods CNA uses semantic
data collected at the same time through legal doctrinal analysis and
qualitative coding to limit the boundaries of the collected data, as those
cases not citing the two centralized cases for the purposes of the legal
doctrine being studied—for example, the application of antitrust laws to
professional baseball—were not appropriate for inclusion in these
networks.[151] In this regard, while Miller included all citations within his

---

[150]   Miller, *supra* note 25, at 23.

[151]   Hypothetically, the discrete citation networks could be defined using community detection
algorithms, which allows for the identification of discrete 'communities' within large, complex
networks based on the density of links between nodes. Andrea Lancichinetti & Santo Fortunato,
Community Detection Algorithms: A Comparative Analysis, 80 PHYSICAL REV. E 1 (2008). As further
explained by Lancichinetti and Fortunato:

> Complex systems are usually organized in compartments, which have their own role and/or
> function. In the network representation, such compartments appear as sets of nodes with a
> high density of internal links, whereas links between compartments have a comparatively

study "without respect to the stated reason for the citation" but later examined those reasons as part of his study, the reasons for the citation in this CNA application act as a limiting factor for network inclusion, so cases not citing the focal case for the purpose of their unique approach to antitrust application should not be included in the discrete citation network studied using this methodology in order to keep in place "a common substantive focus and an unmistakably clear boundary" for the citation network.[152]

Similarly, this CNA application employs Re's strategy of focusing on one key case representing the foundation of a doctrinal rule at the center of the citation networks to create the two citation networks. Accordingly, the use of this methodology seeks to find similar results as Re, who argued

---

> lower density. These subgraphs are called communities, or modules, and occur in a wide variety of networked systems.

Id. Put more simply, a community within a network can be defined as "a group of nodes more densely interconnected, relatively to the rest of the network" where "it is clear the community structure conveys some very important information." Günce Keziban Orman et al., *Comparative Evaluation of Community Detection Algorithms: A Topological Approach*, J. STAT. MECH.: THEORY EXP. 1 (2012). However, the boundary detection strategy employed herein already captures the tenets of community detection by qualitatively defining the discrete networks in a way that makes community detection algorithms duplicative; relevant links between cases can be more easily and more accurately found by qualitatively examining each citation using online databases like LexisNexis and Google Scholar than by relying on algorithms, which may not accurately include cases that are at the outskirts of the identified community. *See generally* Rodrigo Aldecoa & Ignacio Marín, *Exploring the Limits of Community Detection Strategies in Complex Networks*, 3 SCI. REPORTS 1, 2 (2013) (noting that community detection algorithms generally "require the community structure to be known a priori.").

[152] Miller, *supra* note 25, at 21. Because of the care needed to determine the particular purpose for why a case cites another case, the identification of suitable and non-suitable cases was done manually by the researcher rather than relying on a computer algorithm as Miller did in his study. *See* Miller, *supra* note 25, at 25–26 (explaining the use of the corpus management software Sketch Engine to perform textual analysis of the collected data from Google Scholar). Miller's full list of citations included 1,648 rows representing 1,648 unique citations. Miller, *supra* note 25, at 27. For the baseball exemption case study, by contrast, the network contains only 149 unique citations, mostly due to the *a priori* exclusion of cases based on qualitative boundaries before forming the network, which contrasts with Miller's approach. For instance, this boundary definition necessitated in the baseball exemption network the threshold exclusion of cases that cite *Federal Baseball* topically for the court's general deference to baseball when that deference is not applied specifically to antitrust law. As such, this boundary definition excluded a case like *Cincinnati Reds v. Testa*, 122 N.E.3d 1178 (Ohio 2018), as while that case cited *Federal Baseball*'s finding that professional baseball did not constitute interstate commerce, the citation was for the purposes of interpreting state tax law, rather than antitrust law. *See Cincinnati Reds*, 122 N.E.3d at 1180–81 (finding that the giveaway of promotional items (i.e. bobbleheads) does not constitute "sales" under Ohio state tax law). The boundary definition for the *Federal Baseball* network study also excluded cases that cited *Federal Baseball* for antitrust law, but in application to another industry (or even another sport) rather than professional baseball. *See, e.g.*, Radovich v. National Football League, 352 U.S. 445 (1957) (distinguishing *Federal Baseball* and the baseball exemption from application in professional football as "the volume of interstate business involved in organized professional football places it within the provisions of the [Sherman Antitrust] Act.").

that his study of the *Marks* rule "offers a case study in precedential expansion" through his finding that citation to *Marks* for the identified *Marks* rule "largely lay dormant for years" but "accelerated in the early 2000s" and continues to accelerate to this day.[153] But while Re limited his study to cases directly citing *Marks*, this study expands on Re's approach to formulate a methodology crafted to determine—using the selected case study as an example—whether *Federal Baseball* offers a case study in both "precedential expansion" and precedential evolution.[154]

This goal is accomplished by going beyond cases that merely cite the central case(s) to create complete networks that "are based upon *all* of the links that exist between entities within a predefined and bounded population."[155] While Kim noted that "it is assumed that scholars can acquire information on all of the important actors in the actors set," it can be possible "to miss some actors unintentionally or for specific reasons."[156] The great depth of online legal databases and the thoroughness of citation dataset tools, including LexisNexis's "Shepardizing" service, lessens this concern for legal CNA research designs.[157] In accounting for the fact that these legal databases can vary regarding the number and relevance of returned cases after a given search, it is helpful to use multiple databases to ensure that the final data set accounts for the entire population of cases within the specified boundaries.[158] For the case study included herein, for example, the databases used to collect data included Google Scholar and LexisNexis, both of which include tools to perform searches for citations to a particular opinion, which allows for the efficient collection of network data by jumping from one case's citation data to another.

For the included case study, the network studied contains cases with citations to *Federal Baseball* for their unique application of antitrust law to professional baseball, citations to cases that cite *Federal Baseball* for

---

[153] Re, *supra* note 81, at 1965.

[154] *Id*. *See supra* notes 93–95 and accompanying text.

[155] Heath et al., *supra* note 127, at 648 (the concept of a "complete network" is in contrast to an "egocentric network," which within the social network analysis methodology consists of "(i) a starting point individual who is variously referred to as an *ego*, a *focal individual* or an *entry point* . . . and (ii) the individuals who are directly linked to the ego."). For instance, although Re did not explicitly utilize CNA methods, Re's research in essence was focused on a network centered around *Marks v. United States* and thereby represents an example of an egocentric network strategy within the legal CNA framework. *See generally* Re, *supra* note 81, at 1956–61.

[156] Kim, *supra* note 128, at 49.

[157] *See Shepard's Citation Service*, LEXISNEXIS, https://www.lexisnexis.com/en-us/products/lexis-advance/shepards.page [https://perma.cc/C7KV-55H7] (last visited May 18, 2019) (explaining LexisNexis's "Shepard's Citations Service," which allows researchers to look up all subsequent opinions that have cited a particular case and filter those citations in a variety of different ways).

[158] *See* Susan Nevelow Mart, *The Algorithm as a Human Artifact: Implications for Legal [Re]Search*, 109 LAW LIBR. J. 387 (2017); Laura K. Justiss, *A Survey of Electronic Research Alternatives to LexisNexis and Westlaw in Law Firms*, 103 LAW LIBR. J. 71 (2011).

those purposes, and so on without limitation of the number of citations between an included case and the nexus case. Put more simply, the networks include main and subsequent citations that directly or indirectly—by the basis of citation to any other case that is directly or indirectly connected—to *Federal Baseball* to the *n*th degree. For example, a case like *Portland Baseball Club v. Kuhn*[159]—which cited *Flood v. Kuhn* to stand for the baseball exemption but did not itself cite *Federal Baseball*—was included within the network by virtue of its indirect citation to *Federal Baseball* through *Flood v. Kuhn*.[160] Similarly, if any case cited *Portland Baseball* but not *Federal Baseball* or *Flood v. Kuhn*, that case would have still been included by virtue of its *third*-degree citation to *Federal Baseball*.[161]

At the same time, given the CNA application's focus on *citation* networks, any cases found with opinions that discuss the substantive legal doctrine at hand but do not cite any network cases to make that decision (or are themselves cited into the network) should be excluded. While this is necessary because any such cases were not part of the citation network, the exclusion of cases that may hypothetically have impacted the doctrinal field of law does create something of a systematic limitation on this study of these doctrinal networks.

Additional boundary specification is necessary to account for the fact that many cases often have several judicial opinions, often at different levels of the court system.[162] For instance, within a particular case, there

---

[159] *See* Portland Baseball Club, Inc. v. Kuhn, 491 F. 2d 1101 (1974).

[160] *Id.* at 1103.

[161] While not defining a quantitative limitation to the number of edges in each network pathway made it hypothetically possible that the defined network boundaries would deliver unlimited and unbound networks, the extremely narrow substantive focus of each network itself was correctly assumed to limit the number of cases in each network substantially. Indeed, the strong influence of *Federal Baseball* on its respective substantive doctrinal networks limited the number of included cases that are even one-step-removed from the nexus point to only a tiny amount of cases. In other words, for example, it was noted as possible but unlikely that a case discussing the application of antitrust law to professional baseball would fail to cite *Federal Baseball*, as *Federal Baseball* is the common law originator of the baseball exemption. To be sure that all cases that should be included within the defined boundaries of the network were included within the network, all cases included within the network were themselves "Shepardized" on LexisNexis to determine whether there were any cases citing those cases for the qualitative antitrust discussion that defines the boundaries that were not previously "vacuumed up" through searches of cases citing *Federal Baseball*. This process was repeated until the point where no new cases could be found. To this end, the assumption that the substantive boundary would itself limit and bind the network was proven correct; just four cases included within the baseball exemption network did not cite the central case of the network, and no cases were found with more than two degrees of separation between itself and the centralized case.

[162] *See, e.g.*, Postema v. Nat'l League of Prof'l Baseball Clubs, 799 F. Supp. 1475, 1487–88 (S.D.N.Y. 1992) (citing and quoting the district court, court of appeals, and Supreme Court decisions in *Flood v. Kuhn* to address the application of the baseball exemption to state antitrust law).

often are different opinions at the district court and at the court of appeals that courts cite in future cases for various reasons.[163] In these cases, a decision must be made whether to include all cited cases, even in situations in which the inclusion of certain cases was made unnecessary by superseding appellate history. For instance, if an appellate court overrules or affirms a district court opinion, a decision must be made as to what to include in the network: the Court of Appeals decision, the District Court opinion, or both.

Because of the superseding nature of precedent within the common law, however, the authors of this study deemed it appropriate to include only the highest-level court decisions that discuss antitrust application to the sports industry sectors discussed in the two identified citation networks, so long as that higher court case was appropriate for inclusion in the network based on these boundary definitions.[164] This occasionally resulted in the exclusion of otherwise worthy cases; for instance, the very detailed and thorough district court opinion in the baseball exemption case *Wyckoff v. Office of the Commissioner of Baseball* was excluded in favor of the much shorter Second Circuit decision that spent little time and cited very few cases to support affirmance of the judgment of the district court.[165] However, such exclusions were necessary to keep consistent boundaries in each citation network and to ensure the inclusion of only the most relevant precedent to determine the impact and influence of other cases within the networks.

In summary, 135 cases directly citing to *Federal Baseball* were collected from Google Scholar's "How Cited" and LexisNexis's

---

[163] *Id.*

[164] The inclusion of a limiting criterion where the case must focus on antitrust application was necessary for cases with multiple legal issues. In these circumstances, the highest court may only review one of those legal issues, and to that end, that higher court review may not discuss the antitrust issue at all. *Compare, e.g.*, *Postema*, 799 F. Supp. at *passim* (discussing in depth the plaintiff's employment law claims and claims under New York's human rights law before turning to whether her antitrust claim was barred under the baseball exemption), *with* Postema v. Nat'l League of Prof'l Baseball Clubs, 998 F.2d 60 (2d Cir. 1993) (reviewing only procedural matters rather than discussing any substantive legal issues).

[165] *Compare* Wyckoff v. Ofc. of the Comm'r of Baseball, 211 F. Supp. 3d 615 (S.D.N.Y. 2016), *with* Wyckoff v. Ofc. of the Comm'r of Baseball, 705 Fed. App'x 26 (2d Cir. 2017). Superseding higher court cases were not included over a preceding lower court case if the higher court case was not appropriate for inclusion into the network. An excellent example of when the exclusion of a higher court decision in favor of a lower court decision was discussed in *supra* note 164 and accompanying text. There, the Second Circuit Court of Appeals decision in *Postema* would be deemed excluded in favor of the district court decision since the Court of Appeals reviewed only procedural issues and did not discuss the substance of the antitrust claim. Because the Second Circuit decision—unlike the district court opinion it affirmed and unlike the Second Circuit opinion in *Wyckoff*—(naturally) does not either cite to *Federal Baseball* for the purposes of its unique application of antitrust law in professional baseball or even to any cases citing *Federal Baseball* for that purpose, it cannot be included within this study since, per the network boundary definitions outlined *supra* notes 159–61 and accompanying text, it cannot be considered part of the citation network without actually being linked to other network cases through citation. *See supra* note 159–61 and accompanying text.

"Shepardizing" tools.[166] However, this raw total included twenty-two cases superseded by stronger precedent that was more recent or at a higher court level, or both, and were therefore removed from the study.[167] Removing these cases left a sample of 113 top-level decisions directly citing *Federal Baseball* that next must be determined to either be within or outside of the boundary of the baseball exemption network.[168]

Of the 113 direct citations to *Federal Baseball* discussed in the prior section, 26 of the 113 cases were identified as part of the *Federal Baseball* citation network, as they involved direct in-degree citations to *Federal Baseball v. National League* for the explicit purpose of both: (A) applying antitrust law; (B) to legal issues involving professional baseball.[169] Of interest, it took some time for the baseball exemption to gain traction as a continuing legal issue (*See* Figure 3), as the first citation to *Federal Baseball* confirming, modifying, or discussing the baseball exemption was not until twenty-four years after *Federal Baseball* in *American League Baseball Club v. Pasquel*.[170]

---

[166]   *See supra* notes 155–58 and accompanying text.

[167]   "Superseding precedent" in this context refers to when a case also has a later decision at the same court and/or at an appellate court. For example, the Southern District of New York and Second Circuit Court of Appeals decisions and opinions in *Flood v. Kuhn* that preceded the Supreme Court's 1972 decision are included in the overall 135-case sample (since they directly cite *Federal Baseball*), but were discarded because they were superseded by the Supreme Court opinion. *See supra* note 164 and accompanying text. There was one case where a superseding Court of Appeals opinion was excluded in favor of a district court opinion: *Postema v. Nat'l League*, 799 F. Supp. 1475 (S.D.N.Y. 1992), *rev'd on other grounds*, 998 F.2d 60 (2d Cir. 1993). While the district court decision includes a robust discussion of the applicability and scope of the baseball exemption to the present facts (and found that Postema's claim was *not* within the scope of the exemption), the Second Circuit in a four paragraph *per curiam* opinion reversed wholly on a holding regarding the retroactivity of damages provision in the Civil Rights Act of 1991—the primary issue in the case relative to the more ancillary antitrust issue. *See Postema*, 998 F.2d at 61–62. Postema eventually voluntarily dismissed the case with prejudice after a confidential settlement. Grant Wahl, *Catching Up with Baseball Umpire Pam Postema*, SPORTS ILLUSTRATED (Apr. 28, 1997), https://www.si.com/vault/1997/04/28/226156/baseball-umpire-pam-postema-march-14-198 [https://perma.cc/HM8P-5GL7].

[168]   *See supra* notes 163–48 and accompanying text. Citations to *Federal Baseball* in non-Article III judicial bodies (e.g., Tax Courts, Bankruptcy Courts, and the NLRB) are not included in the sample, since they do not have jurisdiction over antitrust issues.

[169]   In two cases, arguably superseding opinions collected due to their citation of *Federal Baseball* were not included within the network in favor of prior opinions in those cases. For both cases, the second, arguably superseding opinion was a denial of a motion by MLB to appeal the prior decision to the relevant Court of Appeals. *See* Garber v. Ofc. of the Comm'r of Baseball, 120 F. Supp. 3d 334 (S.D.N.Y. 2014) (denying a motion to appeal Laumann v. NHL, 56 F. Supp. 3d 280 (S.D.N.Y. 2014)); Piazza v. MLB, 836 F. Supp. 269 (E.D. Pa. 1993) (denying a motion to appeal Piazza v. MLB, 831 F. Supp. 420 (E.D. Pa. 1993)). The decision was made to include the prior opinion rather than the opinion on the motion for interlocutory appeal because in each case: (A) the prior opinion contains much more of the judge's substantive thoughts as to the merits of the case; and (B) the judge incorporated by reference the prior opinion.

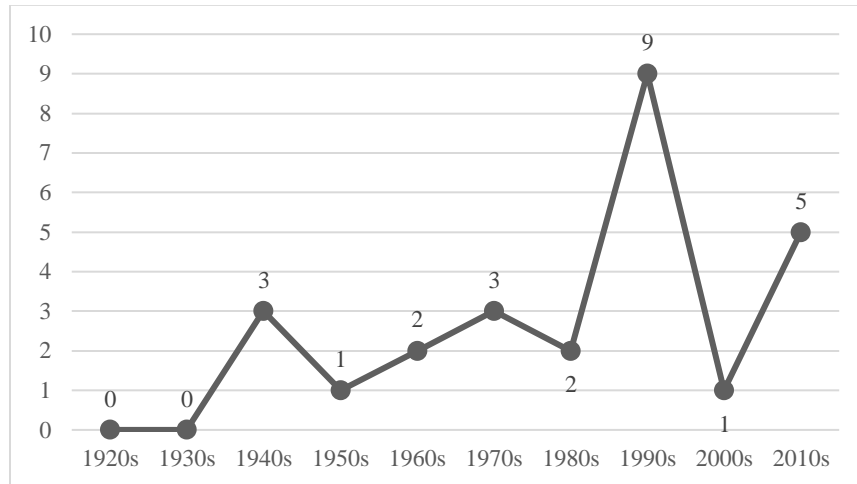[170]   Am. League Baseball Club v. Pasquel, 66 N.Y.S.2d 743 (N.Y. Sup. Ct. 1946).

Figure 3: Citations to Federal Baseball Related to the Baseball
Exemption.

Given the overall importance of *Federal Baseball* as the foundational
Supreme Court opinion establishing the baseball exemption, it stands to
reason that when the baseball exemption citation network was expanded
to include second and third-level citations to *Federal Baseball* and the
baseball exemption, few cases were found that did not cite *Federal
Baseball* directly. Indeed, the complete network includes just four cases
that do not directly cite *Federal Baseball* and instead link to the rest of the
network through second-level citation: *Portland Baseball Club v. Kuhn*,[171]
Twin City Sportservice v. Finley,[172] *Fleer v. Topps Chewing Gum*,[173] and
Nishimura v. Dolan.[174] Adding these cases expands the network to 32

---

[171] *Id.* The Ninth Circuit did not spend much time on the antitrust claim in this case; the court's full
discussion of the antitrust issue was limited to one sentence and a citation to *Flood. See id.* at 1103.

[172] Twin City Sportservice v. Finley, 512 F.2d 1264 (9th Cir. 1975).

[173] Fleer v. Topps Chewing Gum, 658 F.2d 139 (3d Cir. 1981).

[174] Nishimura v. Dolan, 599 F. Supp. 484 (E.D.N.Y. 1984). *Twin City*, *Fleer*, and *Nishimura* do
not themselves discuss or apply the baseball exemption at all. However, they are included within the
network because they are cited and discussed by other network cases discussing these cases' failure to
discuss the baseball exemption as exhibiting a limited scope to the application of the
exemption. *See* Henderson Broad. v. Houston Sports, 541 F. Supp. 263, 270–71 (S.D. Tex. 1982);
Postema v. Nat'l League, 799 F. Supp. 1475, 1489 (S.D.N.Y. 1992). Another case, MLB Properties
v. Salvino, 542 F.3d 290 (2d Cir. 2008) (dismissing antitrust counterclaim by the intellectual property
holding company of MLB and its clubs because the counterclaimant could not point to evidence of a
horizontal agreement to limit output or injury to competition), also considered an antitrust claim
against a baseball-related property without discussing the baseball exemption, but is omitted because
unlike *Twin City*, *Fleer*, and *Nishimura* it was not cited for reasons related to the baseball
exemption. *Laumann v. NHL*, 56 F. Supp. 3d 280, 290 n.54 (S.D.N.Y. 2014)—an in-network case—
cited *Salvino,* but only to discuss when the quick-look doctrine should be applied. Because of this
fact, *Salvino* can be distinguished from a case like *Baseball at Trotwood v. Dayton Prof. Baseball
Club*, 113 F. Supp. 2d 1164 (S.D. Ohio 1999), which was included because it cited *Federal*

cases (including *Federal Baseball* itself) with 142 total citations between them (*See* Table 1).[175]

Table 1 - List of Baseball Exemption In-Network Cases:

Full Name	Label for Charts/Tables	Citation
Federal Baseball v. National League	Federal Baseball	259 U.S. 200 (1922)
American League v. Pasquel	Pasquel	188 Misc. 102 (N.Y. Sup. Ct. 1946)
Gardella v. Chandler	Gardella	172 F.2d 402 (2d Cir. 1949)
Martin v. National League	Martin	174 F.2d 917 (2d Cir. 1949)
Toolson v. New York Yankees	Toolson	346 U.S. 356 (1953)

Baseball and other network cases to note that the baseball exemption was not raised as an issue, and *Twin Cities*, *Fleer*, and *Nishimura*, which were included since network cases cited them for the explicit reason that they do not raise the baseball exemption as an issue and thereby provide arguments for a more limited scope of the exemption.

[175]   The 32-case figure comes from:
  · Twenty-six cases citing *Federal Baseball* directly. (*see supra* 169 and accompanying text);
  · One case that cites *Flood v. Kuhn* but not *Federal Baseball* (*Kuhn*, 491 F.2d at 1101; *see supra* note 171 and accompanying text);
  · Three cases cited by other in-network cases without having cited themselves *Federal Baseball* or any other in-network cases (*Twin City*, *Fleer*, *Nishimura*; *see supra* notes 172–74 and accompanying text);
  · *Federal Baseball* itself; and
  · One additional case, *Moore v. Nat'l Assoc. of Prof'l Baseball Clubs* (*Nat'l League*, 259 U.S. at 200), was not found during the initial sweep of citations to *Federal Baseball* but was instead found by its citation by an in-network case, *Postema v. National League*.

Two cases were cited within other network cases but were not located on Google Scholar or LexisNexis: *Hale v. Brooklyn Baseball Club*, No. 1294 (N.D. Tex. 1958) and *Moore v. National Association of Professional Baseball Clubs*, No. C78-351 (N.D. Ohio 1976). After a search through court dockets and archives, the opinions for both cases were found as attachments to filings within the citing cases. After reviewing both cases, the decision was made to include *Moore* in the network but not *Hale*. While a written opinion and dismissal order represented *Moore* with clear citations to other cases, the only record of the ultimate decision in *Hale* was a transcript where the judge issued a bench ruling during oral argument with no concrete citations. While the judge in *Hale* did informally refer to other cases in issuing his ruling (*e.g. Toolson*, 346 U.S. at 356), the transcript cannot be read to cite other cases definitively. Furthermore, it is unknown whether this transcript reflects a dispositive ruling on the case; while the judge does respond "[y]es" when asked by one of the lawyers whether "everybody is dismissed as to the antitrust [issue]," it is impossible to know whether that did end the case without the case docket, which could not be located. *Moore,* No. C78-351, at 12.

Portland Baseball v. Baltimore Baseball	Portland Baseball I	282 F.2d 680 (9th Cir. 1960)
State v. Milwaukee Braves	Braves	31 Wis.2d 699 (Wis. 1966)
Salerno v. American League	Salerno	429 F.2d 1003 (2d Cir. 1970)
Flood v. Kuhn	Flood	407 U.S. 258 (1972)
Portland Baseball v. Kuhn	Portland Baseball II	491 F.2d 1101 (9th Cir. 1974)
Twin City Sportserv v. Finley	Twin City	512 F.2d 1264 (9th Cir. 1975)
Finley v. Kuhn	Finley	569 F.2d 527 (7th Cir. 1978)
Moore v. National League	Moore	No. C78-351 (N.D. Ohio 1978)
Fleer v. Topps Chewing Gum	Fleer	658 F.2d 139 (3d Cir. 1981)
Henderson Broadcasting v. Houston Sports	Henderson	541 F. Supp.263 (S.D.Tx. 1982)
Prof'l Baseball Schools & Clubs v. Kuhn	Prof'l Baseball	693 F.2d 1085 (11th Cir. 1982)
Nishimura v. Dolan	Nishimura	599 F. Supp. 484 (E.D.N.Y. 1984)
Postema v. National League	Postema	799 F. Supp. 1475 (S.D.N.Y. 1992)
Piazza v. MLB	Piazza	831 F. Supp. 420 (E.D. Pa. 1993)
New Orleans Pelicans v. National League	Pelicans	1994 U.S. Dist. LEXIS 21468 (E.D. La. 1994)
Butterworth v. National League	Butterworth	644 So.2d 1021 (Fla. 1994)
McCoy v. MLB	McCoy	911 F. Supp. 454 (W.D.Wash. 1995)
Morsani v. MLB	Morsani (Fla)	663 So.2d 653 (Fla. Ct. App. 1995)
Baseball at Trotwood v. Dayton Prof. Baseball Club	Trotwood	113 F. Supp. 2d 1164 (S.D. Ohio 1999)
Minnesota Twins P'Ship v. State	Twins	592 N.W.2d 847 (Minn. 1999)
Morsani v. MLB	Morsani (Fed)	79 F. Supp.2d 1331 (M.D. Fla. 1999)

MLB v. Crist	Crist	331 F.3d 1177 (11th Cir. 2003)
Laumann v. NHL	Laumann	56 F. Supp. 3d 280 (S.D.N.Y. 2014)
City of San Jose v. Office of the Commissioner	San Jose	776 F.3d 686 (9th Cir. 2015)
Miranda v. Selig	Miranda	860 F.3d 1237 (9th Cir. 2017)
Wyckoff v. Office of the Commissioner	Wyckoff	705 Fed. Appx. 26 (2d Cir. 2017)
Right Field Rooftops v. Chicago Cubs	RF Rooftops	870 F.3d 682 (7th Cir. 2017)

B. Step Two: Quantitative Legal Citation Network Analysis

1.   Network Visualization

After the list of citations in each of the two networks was collected using a two-column method for data collection, the citation data was inputted into network analysis software to create graphs used to visualize the scope and range of the collected citations. To accomplish this task in crafting the *Federal Baseball* case study, the software NodeXL was used due to its inherent compatibility with the spreadsheet software Microsoft Excel and, therefore, Miller's two-column approach. NodeXL is a well-established network analysis software maintained by the Social Media Research Foundation that has been extensively used in social network analysis, including in at least four legal CNA studies.[176] The software operates as a Microsoft Excel template that processes and analyzes an Excel spreadsheet that allows users to outline and analyze identified nodes (cases) and edges (citations) in individual Excel workbooks to create

---

[176] Britt, *supra* note 149. *See, e.g.*, Nidha Khanam & Rupali Sunil Wagh, *Application of Network Analysis for Finding Relatedness Among Legal Documents by Using Case Citation Data*, 6 J. INFO. TECH. 23 (2017); Ramalingam Jeyshankar & Elangovan Nishavathi, *Legal Citation Network Analysis: An Overview*, 59 PRODUCTIVITY 282 (2018); Feng Zhang & Guohua Jiang, *Patent Citations and Value: Through the Lens of a Social Network Approach*, 4 INT. J. MGMT. & NETWORK ECON. 115 (2018); Cecilia Hasner et al., *Technology Advances in Sugarcane Propagation: A Patent Citation Study*, 56 WORLD PATENT INFO. 9 (2019)).

network graphs.[177] As NodeXL's two-column setup (*See* Figure 2 above) is nearly identical to the methods described by Miller, the NodeXL software is well-suited for the data collection techniques outlined above.

NodeXL also allows for formatting and coloring of individual nodes and lines, which allowed for increased visualized categorization including, for example, color coding based on the court where the decision originated.[178] These network graphs were also organized chronologically using concentric rings that give a rough visual estimation of when a specific citation occurred by the date the citing case was decided with a larger radius from the center of the graph indicating that the case was decided more recently. (*See* Figure 4).[179]

---

[177] Britt, *supra* note 149. *See* Peter Aldhous, *NodeXL for Network Analysis*, NAT'L INST. FOR COMPUTER-ASSISTED REPORTING (2012), https://www.peteraldhous.com/CAR/NodeXL_CAR2012 .pdf [https://perma.cc/V999-E5J6] (explaining key features of NodeXL for network analysis).

[178] Aldhous, *supra* note 177.

[179] One potential limitation of the NodeXL software is that since it functions as a Microsoft Excel spreadsheet plugin, it is often limited in its ability to process larger datasets; indeed, the main collaborators working on the software note that its ability to create network analysis visualizations and analyze network analysis metrics to "small and medium size network with thousands to tens of thousands of nodes." Marc A. Smith et al., *Analyzing (Social Media) Networks with NodeXL*, 4 PROC. INT. CONF. ON COMM. & TECH. 255, 257 (2009), https://doi.org/10.1145/1556460.1556497 [https://perma.cc/PYP8-KSGX]. *See also* Britt, *supra* note 149 (noting that "Excel worksheets contain a maximum of 1,048,576 rows" which "strict limit on the number of nodes and edges that NodeXL can handle, as each node takes up one row on the 'Vertices' worksheet and each edge consumes a row on the 'Edges' worksheet" and "[o]nce the worksheet is out of rows, no more data can be entered."). However, this limitation poses no problem for the CNA application due to its focus on discrete and qualitatively-limited citation networks based on small and concentrated systems of case law and thus generally below one hundred nodes (cases). *See supra* notes 159–61 and accompanying text.
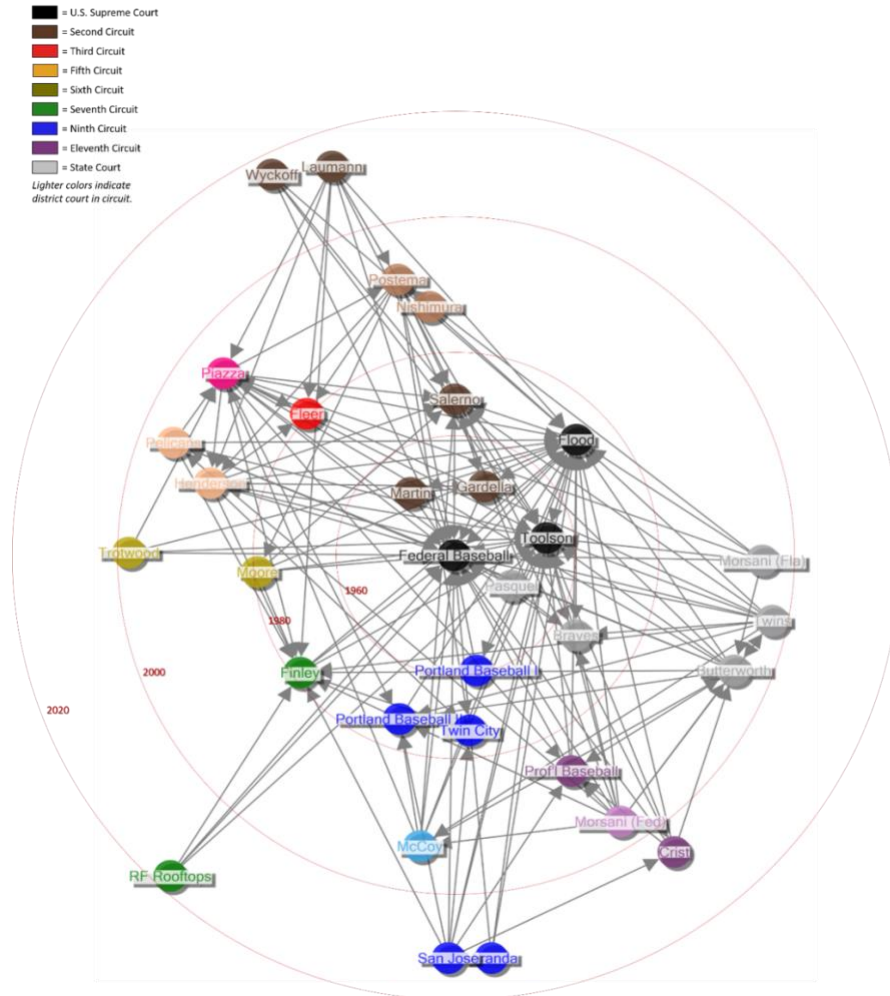
Figure 4: Complete Baseball Exemption Network Graph Visualization.[180]

---

[180] The graph is organized through a polar layout, with rotational angle and node color indicating jurisdictional circuit and polar radius indicating the date the case was decided. The network includes two similarly captioned cases referred to as "Portland Baseball I" and "Portland Baseball II." "Portland Baseball I" refers to *Portland Baseball Club v. Baltimore Baseball Club*, 282 F.2d 680 (9th Cir. 1960). "Portland Baseball II" refers to *Portland Baseball Club v. Kuhn*, 491 F.2d 1101 (9th Cir. 1974). *See* Table 1.

2. Traditional Quantitative Legal Citation Network Analysis

After collecting and visualizing the data in each discrete citation network, the next step is to measure the importance, or centrality, of each case within each of the two discrete citation networks chosen for this study. Measurement and analysis is accomplished first by formulating a count of in-degree and out-degree citations for each case in a network and detailing each link between cases within a network.[181]

Within the network analysis framework, in-degree citations (also called "inward citations" or "inlinks") are the number of nodes that link to the node in question.[182] Out-degree citations (also called "outward citations" or "outlinks") are the number of nodes that the node in question establishes a link.[183] For citation networks, the number of in-degree citations is "the count of citations a case has received" while the number of out-degree citations "is the count of cases cited in an opinion."[184] As Niinivaara notes, references within an article or case are out-degree citations while citations to an article or case are in-degree citations.[185]

Out-degree and in-degree counts are used to determine the centrality of a particular node within a larger network. As defined by Carmichael et al., centrality metrics "measure how *important* a [node] is in a network in different ways" and "provide a way of quantifying the notion of importance of a case in a citation network."[186] In a broader sense, centrality metrics are means to identify the actors "who are the most important or the most prominent" within the network, as graph theory in social network analysis holds that these actors "are usually located in strategic locations within the network."[187]

Determining centrality, according to Wasserman and Faust, is "one of the primary uses of graph theory in social network analysis" since the metrics to determine centrality and prestige "yield actor indices which attempt to quantify the prominence of an individual actor embedded in a network."[188] Prominence, in this case, means that "the ties of the actor

---

[181] Carmichael et al., *supra* note 8, at 230.

[182] *Id.*

[183] *Id.*

[184] *Id.*

[185] Olli Niinivaara, Qualified Information Networks 13 (Jun. 8, 2005) (unpublished manuscript), https://s3.amazonaws.com/academia.edu.documents/12441/niinivaara05qualified.pdf?AWSAccessKeyId=AKIAI.

[186] Carmichael et al., *supra* note 8, at 228.

[187] WASSERMAN & FAUST, *supra* note 8989, at 169.

[188] *Id.* Wasserman and Faust distinguish between "centrality" and "prestige," where both metrics are "examples of measures of the prominence or importance of the actors in a social network" but prestige, unlike centrality, can only be used in directed graphs since it incorporates measures of the "choices received by actors." *Id.* The difference, according to Wasserman and Faust, is that "[t]he prestige of an actor increases as the actor becomes the object of more ties but not necessarily when the actor itself initiates the ties"; or, in other words, prestige metrics "must look at ties directed *to* an actor to study that actor's prestige. *Id.* at 174. While

make the actor particularly visible to other actors in the network" and thereby "extensively involved in relations with other actors."[189] These concepts are naturally suited to the measurement of sociological and economic concepts such as brokerage of information since, as Wasserman and Faust point out, "the difference between the receiver is less important than just participating in many interactions."[190] The best *quality* of information is not necessarily important; "the actors with the most access or most control . . . will be the most central in the network" because they have the strongest relational links to others.[191] This concept also fits naturally within the legal CNA framework, as a case with the highest quality of legal analysis (if that can even be measured) will pale in power to a similar case that has been cited frequently, since the frequently cited case is much more visible within the network and accessible to judges looking for cases to cite, gaining "power" and influence according to Merryman's theory of cumulative judicial authority.[192]

Within that context, Carmichael et al. identified three different centrality metrics with various inputs and outputs with the goal of "develop[ing] a methodology to evaluate . . . centrality metrics in an evolving network based on how predictive a metric is of future citations."[193] The three metrics identified by Carmichael et al.—degree-based metrics, eigenvector-based metrics, and positional metrics—each "are driven by different structural properties of the network."[194] Degree-based metrics are based on simple in-degree and out-degree counts; the more a case cites and is cited, the more important it is deemed to be.[195] Eigenvector centrality metrics, which are "based on the idea that a case is important if it is cited by a lot of cases that are themselves important," gives a larger weight to out-degree citations from cases that themselves have a larger number of out-degree citations.[196] Positional metrics look to

_____

the metrics that Carmichael et al. call "centrality" metrics may be better termed by Wasserman and Faust as "prestige" metrics, this article refers to these metrics as "centrality" metrics since that is the term that Carmichael et al. use in their writing. It may also be the case that within the specific context of legal CNA centrality and prestige metrics are indistinguishable; unlike in traditional social network analysis wherein directed graphs, Actor A can both identify Actor B as a relation and have Actor B identify Actor A as a relation, a cited case cannot retroactively cite back to the case citing it. In this regard, both centrality and prestige metrics "look at ties directed *to* an actor" specifically since those are the only ties that can even exist in the first place. *Id.*

[189]   *Id.* at 172–73.
[190]   *Id.* at 173.
[191]   *Id.* at 173.
[192]   *See* Merryman, *supra* note 64, at 618–19. *See also supra* notes 64–74.
[193]   Carmichael et al., *supra* note 8, at 228.
[194]   *Id.* at 237.
[195]   *Id.*
[196]   *Id.* This definition is nearly identical to the "legal relevance score" metric used by Cross and Spriggs to provide a measure of "opinion significance" in comparison to other opinions. Cross &

determine how "close" a case is to other cases through the "distance" from one case to another measured by the number of citations that it would take to go from once case to another.[197]

Based on Merryman's observation that the process of judicial citations gaining power "becomes cumulative" since the more often cases are cited "the more frequently it will be cited in subsequent ones," initial studies using this CNA application (including the case study detailed later in this article) primarily employed eigenvector-based and degree-based metrics at the analysis stage, as the structural bases of those metrics are more in line with Merryman's theory and other root theoretical bases explored in this study.[198] Positional metrics—while useful in certain contexts—were not deemed to necessarily reflect the conceptualization of case "centrality" in this study, since finding the shortest path between nodes is not as important in this study as it is in studies determining the flow of information.[199]

In fact, Borgatti argued that betweenness centrality, a specific type of positional metric, is not appropriate for analyzing trail-based network processes like citation networks (where each node and edge cannot be revisited within the network flow) since betweenness "is conventionally thought to measure the volume of traffic moving from each node to every other node that would pass through a given node."[200] Similarly, Borgatti noted that the other positional metric, closeness centrality, is mainly for use in more temporal contexts like determining the shortest possible path.[201] Given the network flow of case citation networks, which like the gossip process discussed by Borgatti "does not necessarily follow shortest paths," it is not appropriate to consider measures that consider shortest paths since "the rank ordering of who receives information earliest on average will not correspond to the ordering provided by the closeness centrality measure."[202]

By contrast, degree-based metrics much more accurately describe case importance because, as Borgatti termed it, degree-based metrics can be considered "as a measure of immediate influence" that is appropriate for all parallel-duplication flow processes.[203] Similarly, Borgatti termed eigenvector centrality as a measure of "a long-term direct and indirect [influence]" in contrast to degree-based measure of "immediate

---

Spriggs, *supra* note 21, at 439–43. *See supra* note 21 and accompanying text.

[197] Carmichael et al., *supra* note 8, at 239.

[198] Merryman, *supra* note 64, at 618.

[199] *See generally* Stephen P. Borgatti, *Centrality and Network Flow*, 27 SOC. NETWORKS 55, 59–60 (2005).

[200] *Id.* at 60.

[201] *Id.* at 59–60.

[202] *Id.* at 60.

[203] *Id.* at 62.

[influence] only."[204] Since this form of CNA looks to measure both direct and indirect influence of prior precedent on the network flow of law, both the indirect influence measured by eigenvector centrality and the immediate influence measured by degree-based metrics were useful for discovering which cases have been the most influential within the two discrete case networks.

Even though Carmichael et al. found that simple out-degree metrics were the most predictive of future citations, studies employing this CNA application will look to accomplish a different task than what Carmichael et al. studied: namely, the idea that a few cases within a citation network will drive the qualitative evolution of law within a discrete network. Indeed, in recommending which vertex centrality metric to use in a given situation, Carmichael himself offered in his doctoral dissertation that the central assumption of his methodology, "that a good vertex centrality metric has the ability to predict future citations," is not always "the most appropriate starting assumption in picking a vertex centrality metric."[205] Instead, Carmichael suggested that "[a]wareness of starting assumptions and qualities of statistical tools help ground the scope of questions that can be asked, and in turn the situations in which a tool might be appropriately used."[206] Given this specific form of CNA's theoretical "starting assumption" based on Merryman's theory of the nature of precedent[207] and the "chain novel" theory posited by Dworkin,[208] it made more sense for the data analysis discussed in this CNA application to rely on eigenvector-based metrics for data analysis, as eigenvector centrality metrics give additional value to precedent that accumulates power over time in the manner that Merryman observed.[209]

Fortunately, NodeXL—the software used for data visualization as described above—also contains the ability to calculate several quantitative network analysis metrics.[210] Indeed, according to the original authors of the NodeXL software, "NodeXL includes a number of software routines for calculating statistics about individual nodes including in-degree, out-degree, clustering coefficient, and closeness, betweenness, and eigenvector centrality" (*See* Figure 5), with the ability to integrate

---

[204] *Id.*

[205] Carmichael, *supra* note 21, at 92–93.

[206] *Id.* at 93.

[207] Merryman, *supra* note 64, at 618. *See supra* notes 64–74 and accompanying text.

[208] DWORKIN, *supra* note 57, at 228–38. *See* Lindquist & Cross, *supra* note 56, at 1157–59; Sunstein, *supra* note 57. *See also supra* notes 56–59 and accompanying text.

[209] Merryman, *supra* note 64, at 618.

[210] Britt, *supra* note 176 and accompanying text.

additional analysis features.[211] After calculation, NodeXL places these metrics into the Excel spreadsheet where collected data is displayed, and allows for the creation of groups and other graph visualizations based on these results. (*See* Figure 6).[212]

---
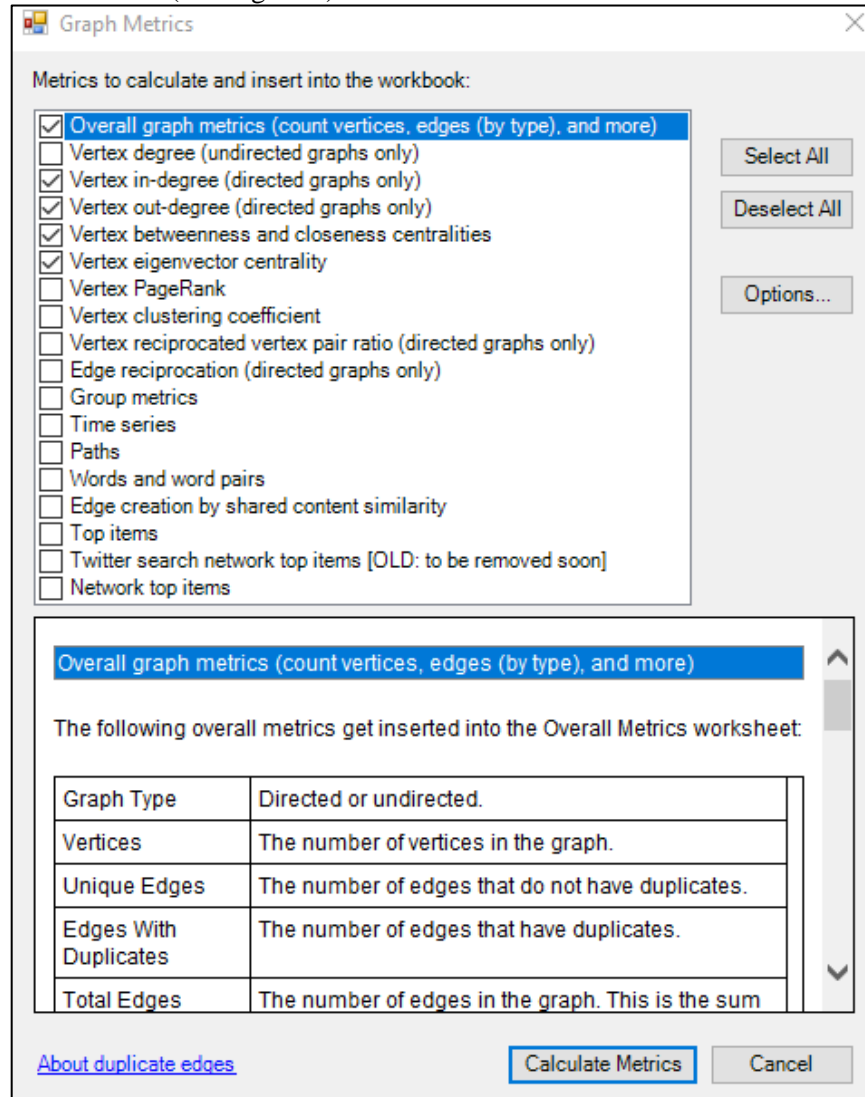
[211] Smith et al., *supra* note 179.

[212] *Id.* at 256–57; Britt, *supra* note 149 and accompanying text. *See also* Brian Britt (brianbritt87), *NodeXL Tutorial (part 2 of 3)*, YOUTUBE (Apr. 6, 2012), https://www.youtube.com/watch?v=owl9we4ldFI (demonstrating the use of analysis features in NodeXL in a video tutorial).

Figure 5: The NodeXL dialog box outlining graph metrics available for calculation.

	A	K	R	S	T	U	V	W
1			Graph Metrics					
2	Vertex	Tooltip	Degree	In-Degree	Out-Degree	Betweenness Centrality	Closeness Centrality	Eigenvector Centrality
3	Federal Baseball	259 U.S. 200 (1922)		23	0	210.625	0.029	0.078
4	Gardella	172 F.2d 402 (2d Cir. 1949)		3	1	3.774	0.018	0.016
5	Toolson	346 U.S. 356 (1953)		20	1	130.158	0.027	0.076
6	Corbett	202 F.2d 428 (6th Cir. 1953)		2	1	1.667	0.013	0.005
7	Kowalski	202 F.2d 413 (6th Cir. 1953)		2	2	10.790	0.017	0.010
8	Pappas	125 F.Supp. 343 (N.D.Ill. 1954)		0	2	0.000	0.016	0.014
9	Int'l Boxing	348 U.S. 236 (1955)		5	2	1.975	0.019	0.037
10	Radovich	352 U.S. 445 (1957)		8	3	7.752	0.020	0.053
11	Milwaukee Braves	31 Wis.2d 699 (Wis. 1966)		2	3	0.571	0.018	0.027
12	Salerno	429 F.2d 1003 (2d Cir. 1970)		5	4	15.284	0.020	0.039
13	Haywood	401 U.S. 1204 (1971)		3	0	1.200	0.014	0.011
14	Flood	407 U.S. 258 (1972)		14	4	80.739	0.025	0.070
15	Portland Baseball	491 F.2d 1101 (9th Cir. 1974)		4	1	1.775	0.016	0.022
16	Twin City	512 F.2d 1264 (9th Cir. 1975)		3	0	1.708	0.015	0.011
17	Finley	569 F.2d 527 (7th Cir. 1978)		6	4	10.183	0.020	0.049
18	Fleer	658 F.2d 139 (3d Cir. 1981)		3	0	0.000	0.015	0.013
19	Henderson	541 F.Supp.263 (S.D.Tx. 1982)		2	10	45.147	0.021	0.049

Figure 6: A test Federal Baseball network workbook with degree and centrality metrics calculated and inputted into the spreadsheet by the NodeXL software.

The traditional CNA metrics collected for the case study included:

- o In-degree and out-degree counts, which simply measure the number of times a case has been cited by other cases (in-degree) and the number of cases that a case cites internally (out-degree);[213] and
- o Eigenvector centrality metrics, which as explained above "judge a case to be more important if it is cited by many cases that are themselves cited by many other cases."[214]

Additionally, in-degree citations for the two central cases were calculated and analyzed in the same manner as done by Re, who tracked citations over time and compared results between *Marks* rule citations and

---

[213]   Carmichael et al., *supra* note 8, at 230.

[214]   *Id.* at 237. *See supra* notes 196, 198, and accompanying text.

non-*Marks* rule citations.[215] Use of these temporal measures allows for study of "precedential expansion" over time[216] and conforms with the observations of Carmichael et al., who noted that time must be considered when determining the value of a particular case as "[t]ime plays a large role in the evolution of the citation network."[217]

Interpretation of these centrality metrics is fairly straightforward; a case that has "a high centrality level, as measured by its degree, is 'where the action is' in the network."[218] In other words, the results of centrality metrics are proportional relative to each other and are not scaled based on external sources, including prior research.[219] The relativity of these metrics lends perfectly to the idea that a smaller discrete citation network still allows for the ability to fully and completely interpret the importance of certain cases within that network despite a smaller body of cases.

For the baseball exemption case study, it was unsurprising that the three most influential cases by both in-degree and eigenvector centrality metrics were the three Supreme Court cases shaping the baseball exemption: *Toolson v. New York Yankees*, *Flood v. Kuhn*, and *Federal Baseball* itself. (*See* Table 2). *Flood* somewhat unexpectedly has a slightly stronger eigenvector centrality number than *Federal Baseball*, but this can be explained by the fact that *Flood* has a much higher out-degree score, which makes sense given that *Federal Baseball* obviously could not itself have cited any in-network cases.

---

[215] Re, *supra* note 67, at 1954–65.

[216] *Id.* at 1965.

[217] Carmichael et al., *supra* note 8, at 260. *See also* Fowler et al., *supra* note 17, at 330 (discussing citation as "a time-dependent process" that creates inherent biases in time-agnostic measures of citation centrality). While Fowler et al.'s observations about the inherent biases of time-agnostic centrality metrics can also be applied to the other centrality metrics employed in this study, it is important to note that this specific form of CNA seeks to accomplish different goals than Fowler et al.'s study. While Fowler et al. (along with Carmichael et al.) looked to determine the predictive power of different types of centrality metrics in an attempt to position centrality as a way to predict future citations to cases, this study looks to the past—namely, to determine how certain cases influenced the evolution of precedent over time. Fowler et al., *supra* note 17, at 338. As such, Fowler et al.'s warnings about the inherent biases in standard centrality metrics, including eigenvector centrality, do not apply to the starting assumptions that provide the theoretical foundation for this CNA application. The reason for this is simple: a just-filed case that would therefore attain an importance score of 0 would, in this study, rightly be considered not important within the network since it has not (yet) contributed to the evolution of precedent over time in that network. *See* Carmichael, *supra* note 21, at 92–93 ("Awareness of starting assumptions and qualities of statistical tools help ground the scope of questions that can be asked, and in turn the situations in which a tool might be appropriately used."). *See also supra* notes 205–09 and accompanying text.

[218] Carmichael et al., *supra* note 8, at 179.

[219] Robert A. Hanneman & Mark Riddle, *Centrality and Power*, INTRODUCTION TO SOC. NETWORK   METHODS,   http://faculty.ucr.edu/~hanneman/nettext/C10_Centrality.html [https://perma.cc/7N5C-VX4L] (last visited May 18, 2019).

Table 2: Top ten scoring cases in the baseball exemption network, sorted by in-degree count:

Case	Citation (Year)	In-Degree	Out-Degree	Eigenvector Centrality
Federal Baseball	259 U.S. 200 (1922)	27	0	0.068
Toolson	346 U.S. 356 (1953)	23	1	0.066
Flood	407 U.S. 258 (1972)	20	7	0.070
Finley	569 F.2d 527 (7th Cir. 1978)	10	4	0.045
Salerno	429 F.2d 1003 (2d Cir. 1970)	10	3	0.044
Piazza	831 F. Supp. 420 (E.D. Pa. 1993)	7	11	0.054
Portland Baseball II	491 F.2d 1101 (9th Cir. 1974)	6	1	0.025
Butterworth	644 So.2d 1021 (Fla. 1994)	5	9	0.047
Prof'l Baseball	693 F.2d 1085 (11th Cir. 1982)	5	3	0.031
Braves	31 Wis.2d 699 (Wis. 1966)	5	2	0.030

Ignoring the three landmark Supreme Court cases, the two most influential cases within the network by in-degree count are *Finley v. Kuhn* and *Salerno v. American League*, which each have ten in-degree citations. These scores make sense; these cases can in some way be considered to be the most 'recent' cases at the courts of appeals of the Seventh Circuit (Finley) and Second Circuit (Salerno) that have had time to be cited by future cases. In this way, it can be correctly assumed that the recently-decided cases in each of those circuits—*Right Field Rooftops v. Chicago Cubs* at the Seventh Circuit and *Wyckoff v. Office of the Commissioner of Baseball* at the Second Circuit—each cite their circuit predecessors as outlining the particular contours of the baseball exemption within that particular circuit.[220] Two other high-scoring cases, the Ninth Circuit's

---

[220] Right Field Rooftops v. Chicago Cubs, 870 F.3d 682, 689 (7th Cir. 2017) ("In [*Finley*], we found that '[the] exemption does not apply wholesale to all cases which may have some attenuated relation to the business of baseball' . . . [b]ut we do not view the Cubs' conduct as attenuated to the business of baseball"); Wyckoff

Portland Baseball v. Kuhn ("*Portland Baseball II*") and the Eleventh Circuit's *Professional Baseball Clubs and Schools v. Kuhn*, were similarly cited by later cases at the same circuit, *City of San Jose v. Office of the Commissioner of Baseball* and *MLB v. Crist*.[221]

A surprisingly high scoring case in both in-degree count and especially eigenvector centrality count was *Piazza v. MLB*. *Piazza*—a district court case—tallied seven in-degree citations, placing it as the network decision with the sixth most relational links to the rest of the network. A district court decision ranking in relative importance alongside court of appeals decisions is unusual given the much narrower scope of such decisions' precedential effect.[222]

The reason for *Piazza*'s high scores is related to its unique holding, however; *Piazza*'s reasoning supporting its holding to limit the baseball exemption solely to issues involving the reserve clause was frequently cited by plaintiffs looking to argue around other courts' much wider definitions of the exemption, and was thereby addressed by the courts hearing that case in making its decision.[223] Similarly, another surprisingly high scoring case using in-degree metrics was *Butterworth v. National League*: a Florida Supreme Court case that largely followed the reasoning in *Piazza* to hold the Florida attorney general's investigation outside the scope of the exemption. While other states and federal courts are generally not bound by one state's authority, the large number of Florida baseball exemption cases that followed *Butterworth*—either by virtue of federal court remands in the same case[224] or other plaintiffs arguing similar theories within Florida[225]—would obviously look to *Butterworth* to

---

v. Ofc. of the Comm'r of Baseball, 705 Fed. Appx. 26, 29 (2d Cir. 2017) (citing *Salerno* to note that "[o]ur Court has applied [the *Federal Baseball*, *Toolson*, and *Flood*] precedent to exempt from antitrust regulation certain claims brought by professional baseball umpires against the American League.").

[221] City of San Jose v. Ofc. of the Comm'r of Baseball, 776 F.3d 686, 689 (9th Cir. 2015) (citing *Portland Baseball II* to note that "[u]nder the baseball exemption, we have rejected an antitrust claim that was wholly unrelated to the reserve clause"); Major League Baseball v. Crist, 331 F.3d 1177, 1183 (11th Cir. 2003) (citing *Prof'l Baseball* alongside *Federal Baseball*, *Toolson*, and *Flood* to collectively hold that "[t]he 'business of baseball' is exempt from the federal antitrust laws."). *But see* Miranda v. Selig, 860 F.3d 1237, 1242 (9th Cir. 2017) (noting that "[i]n [*City of San Jose*] . . . we held that restrictions on franchise relocation fall squarely within the 'business of baseball' and are therefore exempt from federal antitrust laws under Flood" while failing to cite *Portland Baseball II*). *Miranda*'s failure to cite *Portland Baseball II* can be easily explained by the fact that *City of San Jose* has likely replaced *Portland Baseball II* as the landmark Ninth Circuit baseball exemption case but is still too recent to have gained relative importance using network centrality metrics.

[222] Joseph W. Mead, *Stare Decisis in the Inferior Courts of the United States,* 12 NEV. L.J. 787 (2012).

[223] *See, e.g.*, McCoy v. Major League Baseball, 911 F. Supp. 454, 457 (W.D. Wash. 1995); Minnesota Twins P'ship v. State, 592 N.W.2d 847, 855–56 (Minn. 1999).

[224] *See Crist*, 331 F.3d at 1181 (noting that "the Supreme Court of Florida held in an earlier decision that the antitrust exemption established by federal law extends only to the reserve system rather than broadly exempting the 'business of baseball.'").

[225] Morsani v. Major League Baseball, 663 So.2d 653, 657 (Fla. Ct. App. 1995) ("*Morsani (Fla)*") (overruling a district court's grant of dismissal in favor of MLB on the grounds that "[t]he trial judge below did not have the benefit of the Florida Supreme Court's decision which reversed the appellate court and held . . . that the antitrust exemption for baseball extended only to the reserve system"); Morsani v. Major League

delineate the current stance of the Florida courts on the baseball exemption's application and scope. Whether these courts actually *followed Piazza* and *Butterworth* is a different story, however, and one that will be explored in the qualitative portion of the study.[226]

A few key differences were observed when exploring the highest scoring cases within the network by out-degree scores and eigenvector centrality. Fowler and Jeon referred to out-degree citations as "hub scores" and noted that this metric indicates the degree to which a case is well-grounded in previous important rulings.[227] As discussed when discussing the applicability of traditional network analysis metrics in the legal CNA context, Carmichael et al. found that out-degree scores beats in-degree centrality in predicting the future citation of a case and "that the number of citations [within an opinion] is strongly associated with future legal relevance."[228]

While this study and use of the employed methodology—given the limited sample size and specific contextual network boundaries inherent in a discrete citation network—is not intended nor set up to provide either support or opposition to Carmichael et al.'s findings, it is interesting to see that in the context of the baseball exemption there seems to be little correlation between out-degree citation and future in-degree citation. In other words, in baseball exemption cases, there is little association between the number of citations within each case and its future legal relevance within the network. (*See* Table 3).

Table 3: Top ten scoring cases in the baseball exemption network, sorted by out-degree count:

Short Name	Citation (Year)	In-Degree	Out-Degree	Eigen-vector Centr-ality
Morsani (Fed)	79 F. Supp. 2d 1331 (M.D. Fla. 1999)	0	12	0.041
Piazza	831 F. Supp. 420 (E.D. Pa. 1993)	7	11	0.054
Twins	592 N.W.2d 847 (Minn. 1999)	1	11	0.044

Baseball, 79 F. Supp. 2d 1331, 1334 n.10 (M.D. Fla. 1999) ("*Morsani (Fed)*") (calling the *Butterworth* decision "[u]tterly foreign to the unquestionable weight of governing federal authority.").

[226] Piazza v. Major League Baseball, 831 F.Supp. 420 (E.D. Pa. 1993); Butterworth v. Nat'l League, 644 So.2d 1021 (Fla. 1994).

[227] Fowler & Jeon, *supra* note 13, at 17.

[228] Carmichael et al., *supra* note 8, at 252–53.

Postema	799 F. Supp. 1475 (S.D.N.Y. 1992)	3	10	0.037
Butterworth	644 So.2d 1021 (Fla. 1994)	5	9	0.047
Laumann	56 F. Supp. 3d 280 (S.D.N.Y. 2014)	0	9	0.032
Henderson	541 F. Supp.263 (S.D. Tex. 1982)	4	8	0.039
San Jose	776 F.3d 686 (9th Cir. 2015)	1	8	0.028
Flood	407 U.S. 258 (1972)	20	7	0.070
McCoy	911 F. Supp. 454 (W.D.Wash. 1995)	2	7	0.036

Indeed, the top scoring case here—the federal court iteration of *Morsani v. MLB*—has zero in-degree citation; *Morsani* was not even cited by the later *MLB v. Crist*, a case within the same circuit that came to largely the same conclusions as *Morsani*.[229] Conversely, when looking back at the top-scoring cases by in-degree count, it is observed that many of the top scoring cases by in-degree have few in-network out-degree citations. In *Toolson v. New York Yankees*, for example, the Supreme Court cited only *Federal Baseball*, completely ignoring the decisions between *Federal Baseball* and *Toolson* that called into question the continued relevance of *Federal Baseball* after the expansion of the Commerce Clause.[230] In *Portland Baseball II,* the Ninth Circuit cited only *Flood v. Kuhn*,[231] even though just one year later the Ninth Circuit would hear a case applying antitrust law to the baseball industry without even considering whether the baseball exemption should apply.[232]

The one major exception here is *Piazza v. MLB*. *Piazza*, which as observed above scored a surprisingly high in-degree count for a district court case, was also a top scorer by out-degree count, with 11 in-network out-degree citations.[233] Here, the reason for *Piazza*'s high out-degree count

---

[229] *Crist*, 331 F.3d at 1177. Of course, *Crist* was an Eleventh Circuit Court of Appeals decision while *Morsani* was merely at the U.S. District Court for the Middle District of Florida. Indeed, *Crist* makes it fairly clear why *Morsani* was not cited: even though *Morsani* was mainly notable for its finding of a broad scope to the exemption following the passage of the CFA (*See Morsani*, 79 F. Supp. 2d at 1335 n.12), the Eleventh Circuit in *Crist* noted in a footnote that the defendant (the Florida Attorney General) had dropped the argument that the baseball exemption should be confined to the reserve clause after that notion was, according to the Eleventh Circuit, "forcefully destroyed" by the district court. *Crist*, 331 F.3d at 1181 n.10. The district court did cite *Morsani*, but only as an example of what they deemed an "unbroken line" of precedent shaping the baseball exemption. *See* Major League Baseball v. Butterworth, 181 F. Supp. 2d 1316, 1322 n.4 (N.D. Fla. 2001).

[230] Toolson v. New York Yankees, Inc. 346 U.S. 356, 356–57 (1953). *See, e.g.*, Gardella v. Chandler, 172 F.2d 402 (2d Cir. 1949). Even the dissenting opinion written by Justice Burton in *Toolson* ignored *Gardella*'s findings in respect to the continued viability of *Federal Baseball*; while Justice Burton did cite *Gardella*, he did so in a footnote while explaining the function and application of the reserve clause. *Toolson,* 346 U.S. at 362 n.10 (Burton, J., dissenting).

[231] Portland Baseball Club, Inc. v. Kuhn, 491 F.2d 1101, 1103 (9th Cir. 1974).

[232] *See* Twin City Sportservice v. Finley, 512 F.2d 1264 (9th Cir. 1975).

[233] Piazza v. Major League Baseball, 831 F. Supp. 420 (E.D. Pa. 1993).

is related to its painstaking analysis of the continued expansiveness of the baseball exemption, which both cited and heavily discussed not only the Supreme Court trio of *Federal Baseball*, *Toolson*, and *Flood,* but also other cases that either explicitly or implicitly limited the exemption's scope.[234] For example, the court in *Piazza* examined *Finley v. Kuhn*'s discussion of the scope of the exemption post-*Kuhn*,[235] cited *Henderson Broadcasting*'s finding that the exemption does not apply to broadcasting rights,[236] and pointed to several other cases including *Postema* and *Henderson Broadcasting* to support its contention that the scope of the exemption is not unlimited.[237] The *Piazza* court clearly felt that its heretical finding that *Flood* confined the baseball exemption only to the limited purpose of the baseball exemption must be strongly backed by the law, and thus necessitated substantial citation to a variety of different authority (including, notably, the Third Circuit's decision in the seminal abortion case *Planned Parenthood of Southeastern Pa. v. Casey*) to justify its sharp misdirection from other decisions.[238] In accordance with Carmichael et al.'s observations, that strong backing would later be noticed by parties attacking the baseball exemption and by other courts, the latter of whom often pointed to and discussed *Piazza*'s painstaking analysis in deciding whether to agree or disagree with *Piazza*'s ultimate holding.[239]

---

[234] *Id.* at 433–41.

[235] *Id.* at 436–37.

[236] *Id.* at 439.

[237] *Id.* at 440.

[238] *See id.* at 437–38. The citation to *Planned Parenthood v. Casey* was for the purpose of differentiating between two types of *stare decisis*: "rule stare decisis," which requires a particular *legal rule* to be followed and broadly maintained; and "result stare decisis," which only requires the result of the decision to be applied to particularly similar fact patterns. *Id. See* Planned Parenthood of Southeastern Pa. v. Casey, 947 F.2d 682, 691–92 (3d Cir. 1991), *aff'd in part and rev'd in part on other grounds*, 505 U.S. 833 (1992). The *Piazza* court thus reasoned that *Flood*—through its actions overturning the legal rationale supporting *Federal Baseball* and *Toolson* (that professional baseball was not interstate commerce)—only bound lower courts to apply the baseball exemption in fact patterns involving similar fact patterns to *Flood*: namely, players' challenges to the reserve clause. *Piazza*, 831 F. Supp. at 438.

[239] *Compare* Butterworth v. Nat'l League, 644 So.2d 1021, 1025 (Fla. 1994) (agreeing with *Piazza*'s analysis, finding that while "[t]here is no question that *Piazza* is against the great weight of federal cases regarding the scope of the exemption . . . none of the other cases have engaged in such a comprehensive analysis of *Flood* and its implications"), *with* Minnesota Twins P'Ship v. State, 592 N.W.2d 847, 855–56 (Minn. 1999) (disagreeing with *Piazza*'s analysis, writing that while "[t]he *Piazza* opinion is a skillful attempt to make sense of the Supreme Court's refusal to overrule *Federal Baseball . . . Piazza* ignores what is clear about *Flood* — that the Supreme Court had no intention of overruling *Federal Baseball* or *Toolson* despite acknowledging that professional baseball involves interstate commerce.").

Analyzing eigenvector centrality in discrete citation networks presents some of the same challenges as out-degree count. Eigenvector centrality looks at the indirect value of both in-degree and out-degree citations to and within a case, thereby providing a measure of "a long-term direct and indirect [influence]" that is "based on the idea that a case is important if it is cited by a lot of cases that are themselves important."[240] As Carmichael et al. noted, since eigenvector centrality is "based on the idea that a case is important if it is cited by a lot of cases that are themselves important," it gives a larger weight to out-degree citations from cases that themselves have a larger number of out-degree citations.[241]

Here, eigenvector centrality can be used to present a clear picture of the most important cases within the baseball exemption citation network (*see* Table 4), including the three top scorers: the Supreme Court trio. But what is interesting about the eigenvector centrality results here are the three cases that come after: *Piazza v. MLB*, *Butterworth v. National League*, and *Minnesota Twins Partnership v. State*. None of these three cases are prominent Court of Appeals decisions; in fact only *Piazza*, a district court case, is even a federal case.[242]

But what these three cases—along with the next highest scoring, *Finley v. Kuhn*—have in common are their collective discussion of the scope of the baseball exemption.[243] In fact, each of the three cases are highly interconnected; *Butterworth* relies heavily on *Piazza* in coming to the same conclusion as *Piazza* while *Twins* discusses *Piazza* and cites *Butterworth* in ultimately rejecting *Piazza*'s holding.[244] So while the use of eigenvector centrality perhaps suffers from some of the limitations as out-degree count in overvaluing citations by a case to other in-network cases in a context where that is not dispositive of the case's influence, the results of the eigenvector centrality analysis point strongly to one key finding: the discussion of the scope of the baseball exemption in the 1990s was important to the baseball exemption's development to date.

---

[240] Borgatti, *supra* note 199, at 62; Carmichael et al., *supra* note 8, at 237–38.

[241] Carmichael et al., *supra* note 8, at 237–38.

[242] *See, e.g.*, *Piazza*, 831 F. Supp. at 420; *Butterworth*, 644 So.2d at 1021; *Minnesota Twins P'ship*, 592 N.W.2d at 847.

[243] *See, e.g.*, *Piazza*, 831 F. Supp. at 420; *Butterworth*, 644 So.2d at 1021; Finley v. Kuhn, 238 So.2d at 527; *Minnesota Twins P'ship*, 592 N.W.2d at 847.

[244] *See, e.g.*, *Piazza*, 831 F. Supp. at 420; *Butterworth*, 644 So.2d at 1021; *Minnesota Twins P'ship*, 592 N.W.2d at 847.

Table 4: Top ten scoring cases in the baseball exemption network, sorted by Eigenvector centrality:

Case	Citation (Year)	In-Degree	Out-Degree	Eigenvector Centrality
Flood	407 U.S. 258 (1972)	20	7	0.070
Federal Baseball	259 U.S. 200 (1922)	27	0	0.068
Toolson	346 U.S. 356 (1953)	23	1	0.066
Piazza	831 F. Supp. 420 (E.D. Pa. 1993)	7	11	0.054
Butterworth	644 So.2d 1021 (Fla. 1994)	5	9	0.047
Finley	569 F.2d 527 (7th Cir. 1978)	10	4	0.045
Salerno	429 F.2d 1003 (2d Cir. 1970)	10	3	0.044
Twins	592 N.W.2d 847 (Minn. 1999)	1	11	0.044
Morsani (Fed)	79 F. Supp. 2d 1331 (M.D. Fla. 1999)	0	12	0.041
Henderson	541 F. Supp.263 (S.D. Tex. 1982)	4	8	0.039

3.   Measuring Case 'Stickiness' Over Time

The relatively smaller size of these networks also allows for more creative use of quantitative metrics that allow for a fuller determination of the network influence of particular case opinions throughout the perceived 'evolution' of legal precedent. This can be accomplished by measuring the continuing influence of precedent over time. As Fowler et al. noted, citation to case law is "a time-dependent process" in which time-agnostic centrality metrics "inherently bias[] downward the legal relevance of recent cases."[245] After all, a hypothetical case that has only been cited within the network would not be deemed by network centrality metrics as important, but the case may have entirely shifted the trajectory of network doctrine through new application of old precedent in a novel way. While the case studies included herein employed qualitative network analysis

---

[245]   Fowler et al., *supra* note 13, at 330.

and legal doctrinal methods at the second methodological step of this CNA application in a way that allowed for identification of any such recent shifts, it was helpful for that evaluation to also use newer, more nontraditional quantitative metrics to determine which cases have 'stuck' throughout the course of the citation network and therefore have become particularly relevant within recent case law.

In a 2019 study, Bennardo and Chew created the metric of "citation stickiness," a few simple formulae developed to track which cases cited by case parties have "stuck" and therefore been cited by the judges themselves. (*See* Figure 7).[246] According to these authors, determining citation stickiness "provides a window into judicial decision making" because it can "help uncover to what extent judges are conducting independent legal research" rather than merely relying on the parties' writings to provide the legal background of relevant precedent.[247] Generally, Bennardo and Chew hypothesized that citations in a party's brief can either be "*sticky*," in that "it appears in one of the parties' briefs and then again in the court's opinion," or "*unsticky*," where it does not appear in the court's opinion.[248] In the same regard, citations in a court's opinion either originate from a party's brief, which they again call being "*sticky*," or appear for the first time in the opinion, which gives them an "*endogenous*" nature.[249]

Stickiness Equations

$$\text{Sticky cites in opinion} = \frac{cases\ cited\ in\ both\ opinion\ and\ at\ least\ one\ brief}{cases\ cited\ in\ opinion}$$

$$\text{Sticky cites in brief} = \frac{cases\ cited\ in\ brief\ and\ opinion}{cases\ cited\ in\ brief}$$

$$\text{Super-sticky cites in opinion} = \frac{cases\ cited\ in\ opinion\ and\ both\ parties'\ briefs}{cases\ cited\ in\ opinion}$$

Figure 7: Bennardo and Chew's "citation stickiness" formulae.[250]

To determine whether the party brief citations in the cases they studied were "sticky" or "endogenous," Bennardo and Chew studied a sample of 325 court decisions and briefs from each of the thirteen federal courts of appeals while excluding cases where there were supplemental briefs or

[246] Kevin Bennardo & Alexa Z. Chew, *Citation Stickiness*, 20 J. APP. PRAC. & PROCESS 61, 64 (2019).

[247] *Id.* at 67.

[248] *Id.* at 64.

[249] *Id.*

[250] *Id.*

amicus briefs that may also provide influence on judicial decision-making and collected citation data by scanning the brief's table of authorities and inputting the results into a Microsoft Excel spreadsheet.[251] Bennardo and Chew then compared the 7,552 unique case citations collected from the judicial opinions and the 23,479 citations from the briefs and found, using basic statistical analyses, that overall, 49 percent of cases cited in the courts' opinions were also cited in at least one party's brief, meaning that 51 percent of the cases "originated from somewhere else, most likely the courts' own research.[252] Conversely, only 38 percent of cases cited by both parties and only 16 percent of the cases cited by one of the two parties in their briefs were later cited by the courts in their opinions, meaning that there was "an 84% likelihood that the court would not mention the case" that was cited by both parties in support of their positions.[253] As such, Bennardo and Chew reasoned that while they could not determine "*why* so many cases cited in opinions were endogenous," the fact that so many cases did originate from the courts' own research was novel and worth further exploration.[254]

While Bennardo and Chew did not make any direct statements about any causal natures of the citation stickiness theory—including whether a "sticky" citation appeared in the opinion as a direct result of the party's citation of a case—their formulae proved extremely adaptive to the network analysis context. Indeed, their conception of citation stickiness also proved useful in determining not only which cases have stuck from a party brief to a judicial opinion, but also in establishing 'stickiness' of cases from one judicial opinion to another along a network path.

The adaption of Bennardo and Chew's citation stickiness formulae from examining party briefs to examining subsequent legal opinions took two forms. (*See* Figure 8). The first new formula was created by adapting Bennardo and Chew's "sticky cites in opinion" formula, which divides the number of cases cited in both the studied opinion and at least one brief by cases cited in the opinion, into a formula that now divides the number of cases cited in both the studied opinion and a previous opinion by the number of cases cited in the current opinion. This adaption allows for evaluation as to how reliant a given opinion was on prior network precedent, and whether the same network citations relied upon in the previous case are being recycled in the new opinion. It also identifies whether certain citations remained in specific qualitatively defined groups

---

[251] *Id.* at 81 nn.78, 82.
[252] *Id.* at 82–84.
[253] *Id.* at 84.
[254] *Id.* at 112.

of cases—in which one case in the same court or circuit largely builds off of the reasoning in a prior case—by analyzing how that case stuck within particular groups of cases, including in individual precedential silos like appellate circuits. In accordance with Carmichael et al.'s determination that *out*-degree centrality beats in-degree centrality in terms of predicting future citation, this first adaption allowed for a determination of current influence based on Carmichael et al.'s finding "that the number of citations [within an opinion] is strongly associated with future legal relevance."[255]

Stickiness Equations (adapted for case-to-case network evaluation)

$$\text{Sticky cites in opinion} = \frac{cases\ cited\ in\ both\ past\ and\ current\ opinion}{cases\ cited\ in\ current\ opinion}$$

$$\text{Enduring stickiness of a case} = \frac{number\ of\ times\ a\ case\ is\ cited\ throughout\ subsequent\ cases\ in\ network}{number\ of\ subsequent\ cases\ in\ network}$$

Figure 8: Citation stickiness formulae adapted for case-to-case network evaluation.

More critically, however, this first adapted metric was also used to identify whether certain citations remained in specific qualitatively defined groups of cases, in which one case in the same court or circuit largely builds off of the reasoning in a prior case. The results of this metric are used to set up further qualitative doctrinal analysis and allow for contextual conclusions in the qualitative portion of studies employing this specific form of CNA.

The second adapted citation stickiness formula developed for use in this CNA application, called the "enduring stickiness formula," was adapted by combining Bennardo and Chew's citation stickiness concept with the methods employed by Re in tracking the historical trend of citations to *Marks v. United States*,[256] thereby creating a metric used to capture the enduring stickiness of a case throughout the entire network path. To do so, this formula divides the number of times a case is cited throughout subsequent cases in the network by the number of citation opportunities, which is simply an expression of the number of subsequent cases in the network that had the opportunity to cite the case.

In some instances, a case that is originally relevant may be replaced by a later case that overturns, amends, or otherwise supplants the original precedent. For example, one can argue that in the *Federal Baseball* network the Supreme Court in its 1972 decision in *Flood v. Kuhn*

---

[255] Carmichael et al., *supra* note 8, at 252–53.
[256] *See generally* Re, *supra* note 81, at 1948.

supplanted the original *Federal Baseball* decision.[257] The Supreme Court did this by reversing *Federal Baseball*'s finding that professional baseball was not interstate commerce while still upholding the baseball exemption on the basis that Congress had not acted legislatively, and thus, "as yet has had no intention to subject baseball's reserve system to the reach of the antitrust statutes."[258] As such, a subsequent opinion could conceivably cite *Flood v. Kuhn* to stand for the case precedent representing the doctrinal basis of the baseball exemption without citing *Federal Baseball*.[259] Similarly—and even for non-central cases within a network—a case may conceivably fall out of favor with judges writing later opinions and thereby lose its influential value within the network.

As case citation is "a time-dependent process,"[260] it is also informative to track the enduring stickiness of a case on a temporal basis, tracking how "sticky" a case remains or becomes over time. To accomplish this goal in the case studies, the enduring stickiness metric was calculated both for the overall network and on a decade-by-decade basis, where, for example, a case like *Flood v. Kuhn* decided in 1972 had its enduring stickiness tracked for all subsequent cases decided in the 1970s (aside from any cases decided before *Flood*'s filing), the 1980s, the 1990s, the 2000s, and the 2010s.[261] Essentially, this metric was measured by dividing a case's in-degree count by the number of in-degree opportunities, as represented by the number of subsequent cases that hypothetically could cite the target case, which measures the raw percentage of later network cases in which a target case is cited. If, for instance, 15 of 18 subsequent cases cite *Flood v. Kuhn*, *Flood*'s enduring stickiness value will be 0.8333, meaning that *Flood* was cited in 83.33% of subsequent network cases.

---

[257] Flood v. Kuhn, 407 U.S. 258 (1972).

[258] *Id.* at 282–83.

[259] *See, e.g.*, Portland Baseball Club v. Kuhn, 491 F.2d 1101, 1103 (9th Cir. 1974) (dismissing the plaintiff's antitrust claim based on *Flood v. Kuhn* while not citing *Federal Baseball*).

[260] Fowler et al., *supra* note 13, at 330.

[261] Given the limited denominators (i.e., number of subsequent cases and thus citation opportunities) in each decade, the results of this metric per decade can often be presented as fractions, as opposed to percentages. (*See* Figure 9.)
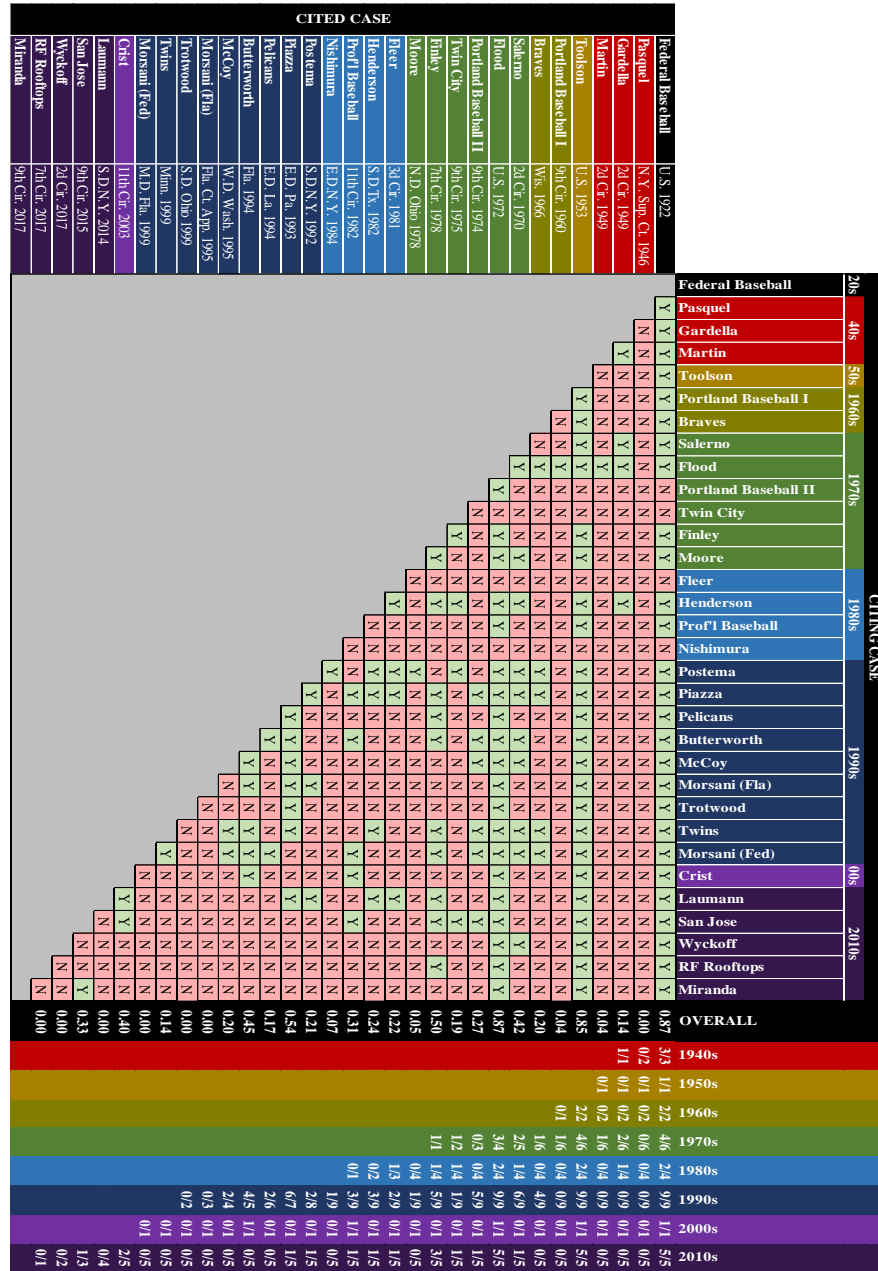
Figure 9: A table used to track and visualize the results of the adapted enduring stickiness formula, as applied here to the test-run Federal Baseball network.

These findings are presented in tabular form (as well as in bar graph form as warranted by the results) in the same manner as Re's calculating of citations to the Marks rule in each decade so that the full visual effect of temporal stickiness can be seen over time.[262] For the first metric, the sticky citations in the opinion metric, the data are presented largely in-text through tables showing the stickiness values for each network opinion and selected bar graphs for cases identified as important within the narrative. For the second metric, which measures the enduring stickiness of a case, the data are presented through a table that gives a binary result as to whether a citation relationship exists between two cases (e.g., a 'Y' for yes or 'N' for no). In this fashion, the results of this data analysis are presented in a table with the cited/prior case on the *y*-axis and the citing/subsequent case on the *x*-axis. (*See* Figure 9.) Cases that remain sticky over time—and therefore stand in contrast to Merryman's observation of the "predictable decline in the 'citation power' of a decision" as the opinion ages[263]—were seen as quantitatively influential within the citation network.

As discussed, when analyzing the more traditional quantitative network centrality metrics,[264] three cases stood out as particularly noteworthy for their importance within the network: *Piazza v. Major League Baseball*, *Butterworth v. National League*, and *Finley v. Kuhn*. The interconnectivity of these three cases is evident by using the adapted citation stickiness metrics tailored for network use.[265] When analyzing *Butterworth* as directly following *Piazza* to determine which citations stuck, it is found that 87.5% (7/8) of the in-network citations in *Butterworth* stuck from *Piazza*.[266] Similarly, 80% (8/10) of the in-network citations in *Twins* were carried over from *Piazza* and 70% (7/10) were carried over from *Butterworth*. (*See* Figure 9).[267] When *Finley* is added into the mix, however, it is clear that *Piazza* and

---

[262] *See, e.g.*, Re, *supra* note 81, at 1954 (presenting the number of federal court citations to *Marks* for each five-year period after it was decided up until the date the study was concluded). While Re presented his graphs with "Non-*Marks* Rule Cites" superimposed against "*Marks* Rule Cites," this CNA application presents either pure in-network citations to a case as the sole presented data within the graph or citations to multiple cases within one graph as warranted by the results. *See* Re, *supra* note 81, at 1954. Because of the number of cases within each network and the fact that many of the network cases were relatively unimportant to the network, only selected stickiness results representing novel results and/or the stickiness of particularly important cases are presented in bar graph form.

[263] Merryman, *supra* note 79, at 394.

[264] *See supra* Section III.B.2.

[265] *See supra* Section III.B.3.

[266] The count of out-degree citations for the following case used as the denominator in the "sticky cites in opinion" formula does not include the past case, for obvious reasons. In other words, the count of in-network citations in *Butterworth* does not include *Piazza*—even though *Piazza* was cited by *Butterworth*— since that citation obviously cannot be sticky: *Piazza* cannot cite itself.

[267] *See* Minnesota Twins P'ship v. State, 592 N.W.2d 847 (Minn. 1999); Piazza v. Major League

its successors drew from different sources. Each of *Piazza* (30%; 3/10), *Butterworth* (37.5%; 3/8), and *Twins* (30%; 3/10) carried over just three citations from *Finley*.[268] Unsurprisingly, these three cases in each case were the three Supreme Court cases.[269]

This finding suggests that *Piazza*, *Butterworth*, and *Twins* each drew their knowledge from different sources than *Finley*. This finding makes sense for two reasons. First, *Finley* only cited four in-network cases: the three Supreme Court cases to illustrate the history of the exemption and *Twin City Sportservice v. Finley* as an example of a case in which the baseball exemption was not cited.[270]

Second, *Piazza* heavily relied on two cases decided after *Finley* that limited the scope of the baseball exemption in their own right: *Henderson Broadcasting v. Houston Sports* and *Postema v. National League*. This shows that the precedential landscape changed between *Finley* and *Piazza*; while *Finley* considered and rejected the idea that *Flood* limited the baseball exemption to the reserve clause, *Piazza* accepted that conclusion in large part due to this newly decided precedent. Indeed, *Piazza* shared a striking amount of cited case law with both *Henderson* and *Postema*, carrying over 80% (8/10) of its citations from *Henderson* and 90% (9/10) from *Postema*.[271] This shows that while *Piazza* certainly relied heavily on *Finley*'s analysis to reach its conclusion, it relied even more strongly on *Postema* and *Henderson* to support its rationale.

The results of the sticky cites in opinion formula also demonstrate how important certain events were in changing the course of the baseball exemption. For example, when looking at the pair of *Portland Baseball*-led cases at the Ninth Circuit, *Portland Baseball II* carried over none of its citations from *Portland Baseball I*, preferring instead to rely exclusively on *Flood* to show that the baseball exemption was still in full force and effect. Similarly, the passage and adoption of the CFA severely hurt the quantitative importance of *Piazza*'s holding. Of the two cases initiated after the passage of the CFA[272] that debated the scope of the baseball exemption—*Laumann v. NHL* and *City of San Jose v. Office of the Commissioner of Baseball*—only

---

Baseball, 831 F. Supp. 420 (E.D. Pa. 1993); Butterworth v. Nat'l League, 644 So.2d 1021, 1025 (Fla. 1994).

[268]   *See* Charles O. Finley & Co. v. Kuhn, 569 F.2d 527 (7th Cir. 1978).

[269]   *See* Fed. Baseball Club, Inc. v. Nat'l League of Prof'l Baseball Clubs, 259 U.S. 200 (1922); Toolson v. New York Yankees, Inc., 346 U.S. 356 (1953); Flood v. Kuhn, 407 U.S. 258 (1972).

[270]   Twin City Sportservice, Inc. v. Charles O. Finley & Co., 512 F.2d 1264, 1268 (9th Cir. 1975).

[271]   Predictably, the case that *Piazza* shared with *Postema* but not *Henderson* was, in fact, *Henderson*, which stands to reason since *Henderson* obviously could not have cited itself.

[272]   While *Minnesota Twins P'Ship v. State*, 592 N.W.2d 847 (Minn. 1999), was decided after the passage of the CFA, its holding was not affected by the CFA since the principles of *ex post facto* applied.

Laumann (87.5%; 7/8) carried over the majority of its citation from *Piazza* while also citing *Piazza*.[273]

Unsurprisingly, *Laumann* is also the case that continued to limit the baseball exemption in refusing to apply it to issues related to baseball broadcasting rights.[274] While *City of San Jose*, on the other hand, also carried over 87.5% (7/8) from *Piazza*, it did not cite *Piazza* nor any of the exemption-limiting cases that *Piazza* relied on to justify its holding (i.e., *Henderson Broadcasting v. Houston Sports* and *Postema v. National League*).[275] Indeed, *City of San Jose*'s discussion in response to the plaintiffs' arguments that the exemption is limited to the reserve clause was dismissed in large part by not only citing relevant case law but by also citing the CFA, which according to the *San Jose* court "withdrew baseball's antitrust exemption with respect to the reserve clause and other labor issues, but explicitly maintained it for franchise relocation."[276]

The final quantitative formula employed in this study, the enduring stickiness formula,[277] also shows that *Piazza* and similar cases lost persuasive force after the passage of the CFA. As with the traditional centrality metrics, the top-scoring cases by the enduring stickiness formula are *Federal Baseball*, *Flood*, and *Toolson*, which are expectedly cited frequently to outline the history of the baseball exemption.[278] The results of this metric rejected one of the possible outcomes discussed when outlining this new applied network statistic, in which it was discussed that some cases may cite a later Supreme Court case within a network as the new landmark case illustrating the baseball exemption.[279] But *Flood*'s enduring stickiness closely mirrors the enduring stickiness of *Federal Baseball*, and only one case (*Portland Baseball II*[280]) cites *Flood* without citing *Federal Baseball* as well. (*See* Table 5).

---

[273] *See* Laumann v. NHL, 56 F. Supp. 3d 280 (S.D.N.Y. 2014).

[274] *Id.* at 297.

[275] *See* City of San Jose v. Ofc. of the Comm'r of Baseball, 776 F.3d 686 (9th Cir. 2015).

[276] *Id.* at 690–91.

[277] *See supra* notes 256–62 and accompanying text.

[278] *See* Table 5.

[279] *See supra* note 261 and accompanying text.

[280] Portland Baseball Club, Inc. v. Kuhn, 491 F.2d 1101, 1103 (9th Cir. 1974).

Table 5: Top twenty scoring cases in the baseball exemption network by enduring stickiness:[281]

CASE	Year	Overall	1960s	1970s	1980s	1990s	2000s	2010s
Federal Baseball	1922	0.87	2/2	4/6	2/4	9/9	1/1	5/5
Flood	1972	0.87		3/4	2/4	9/9	1/1	5/5
Toolson	1953	0.85	2/2	4/6	2/4	9/9	1/1	5/5
Piazza	1993	0.54				6/7	0/1	1/5
Finley	1978	0.50		1/1	1/4	5/9	0/1	3/5
Butterworth	1994	0.45				4/5	1/1	0/5
Salerno	1970	0.42		2/5	1/4	6/9	0/1	1/5
Crist	2003	0.40						2/5
San Jose	2015	0.33						1/3
Prof'l Baseball	1982	0.31			0/1	3/9	1/1	1/5
Portland Baseball II	1974	0.27		0/3	0/4	5/9	0/1	1/5
Henderson	1982	0.24			0/2	3/9	0/1	1/5
Fleer	1981	0.22			1/3	2/9	0/1	1/5
Postema	1992	0.21				2/8	0/1	1/5
Braves	1966	0.20		1/6	0/4	4/9	0/1	0/5
McCoy	1995	0.20				2/4	0/1	0/5
Twin City	1975	0.19		1/2	1/4	1/9	0/1	1/5
Pelicans	1994	0.17				2/6	0/1	0/5
Twins	1999	0.14					0/1	0/5
Gardella	1949	0.14	0/2	2/6	1/4	0/9	0/1	0/5

Just as it was with the traditional centrality metrics, *Piazza v. MLB* was a surprisingly high scorer using the enduring stickiness metric, which shows that *Piazza* achieved importance beyond its base precedential status as a district court decision. But when examining *Piazza*'s enduring stickiness over time, it was interesting to discover that while *Piazza* was extensively cited by the cases that followed it in the 1990s, it was cited just once in the

---

[281] *See supra* notes 256–63 and accompanying text.

six cases that followed in the 2000s and 2010s.[282] This finding—just as with the sticky citations in opinion formula—shows that *Piazza*'s unique holding limiting the baseball exemption to the reserve clause lost prominence after the passage of the CFA, which future cases have interpreted to have explicitly defined the scope of the exemption as near-boundless.[283] Indeed, the one case from the 1990s that did not cite *Piazza* indirectly rebutted its reasoning on the grounds of the CFA's passage, writing that "Congress' preservation of the broadest aspects of the antitrust exemption" in the CFA "casts in sharp relief" the Florida Supreme Court's holding in *Butterworth v. National League* following *Piazza*.[284]

Collectively, the results obtained through quantitative analysis lead to one particularly noteworthy conclusion: the importance of *Piazza v. MLB* in questioning the expansive scope of the baseball exemption leading up to the passage of the CFA. The importance of *Piazza* considering the contextual doctrine created will be further explored in the qualitative analysis to follow.

C. Step Three: Qualitative Legal Citation Network Analysis

1.   Signed Network Graph Visualization

The purpose of the qualitative step of this mixed methods CNA application is to provide context to the determined links and importance levels by adding qualitative data into the created network visualizations. To accomplish this goal, legal doctrinal methods are employed to analyze the nature and quality of each judicial citation in the citation networks collected in the quantitative portion of the study, with a focus and emphasis on the links between cases identified as statistically relevant, important, and central to the discrete citation network. The use of legal doctrinal methods to provide qualitative context to the network graphs involves two substeps: (1) data collection and signed graph visualization through the infusion of qualitative legal doctrinal data into network graphs; and (2) the doctrinal analysis of the context of network links identified in the signed graphs to provide a full explanation of evolutions in precedent over time.

The infusion of qualitative data into the earlier-created legal citation network graphs makes it possible to visually capture the essence of the qualitative shifts in each discrete citation network. Broadly speaking, this infusion allows for the use of network graphs to investigate how judicial

---

[282] *See* Major League Baseball v. Butterworth, 181 F. Supp. 2d 1316, 1323 (N.D. Fla. 2001).

[283] *See* Morsani v. Major League Baseball, 79 F. Supp. 2d 1331, 1335 (M.D. Fla. 1999).

[284] *Id.* at 1335 n.12.

language has shifted and evolved over time in connection to future interpretation and application by later courts in a manner similar to the "chain novel theory" conceptualization of the evolution of legal precedent proposed by Dworkin.[285]

The use of qualitative data to provide additional context to citation networks is not a new strategy within network analysis scholarship.[286] Heath et al. wrote that "it is widely acknowledged that [social network analysis] has at least some of its roots in the qualitative tradition" and noted that the British Sociological Association's network analysis study group has mentioned that "in some of the best examples [of social network analysis research] quantitative data is integrated with a rich contextual understanding and analysis derived from archival and ethnographic sources."[287] While Heath et al. observed that modern-day social network analysis studies "which are positioned by their investigators as examples of qualitative social network analysis . . . appear to be rare," they argued that the qualitative data "can make a powerful contribution to the ongoing debates within the field of social network analysis concerning the dynamic and shifting nature of social networks."[288] Additionally, Zhang and Koppaka's work for the legal database LexisNexis creating what they term "semantics-based legal citation networks" within the computer science literature allowed for the implementation of smarter database searches, which let researchers using the LexisNexis database filter citations of a given case based on specific legal issues that are the basis for the citation.[289]

The inclusion of qualitative data in network analysis is arguably essential for the true determination of relationships in network analysis.[290] As Wasserman and Faust note, "the term 'prestige' is perhaps not the best label" for the concept of determining influence and centrality within networks as in some situations actors can be considered "prestigious" according to quantitative metrics but in practice "are not held in very high regard by their peers."[291] These actors "are certainly renowned, but it is for negative feelings rather than positive."[292] In this regard, Wasserman and Faust opine that "the

---

[285] DWORKIN, *supra* note 57, at 228–38. *See* Lindquist & Cross, *supra* note 56, at 1157–59; Sunstein, *supra* note 57. *See also supra* notes 56–59 and accompanying text.

[286] *See* Heath et al., *supra* note 127, at 646.

[287] *Id.*

[288] *Id.* at 646, 658.

[289] Zhang & Koppaka, *supra* note 108, at 129.

[290] Dominik E. Froehlich, *Linking Quantitative and Qualitative Network Approaches: A Review of Mixed Methods Social Network Analysis in Education Research*, REV. OF RES. IN EDUC. 1 (April 2020), https://journals.sagepub.com/doi/pdf/10.3102/0091732X20903311.

[291] WASSERMAN & FAUST, *supra* note 89, at 175.

[292] *Id.*

substantive nature of the measured relation is quite important when interpreting the property" of that measured relation.[293]

Along these lines, some network analysis scholars have argued for the use of certain network analysis strategies as specifically allowing for the beginnings of qualitative interpretation of data within CNA.[294] Niinivaara theorized the concept of "qualified information networks," which he calls networks that "have qualitative properties formalized into explicit quantitative properties" by "attaching qualifying data to nodes and/or links."[295] Similarly, within the network analysis methodological literature, the use of signed graphs—which are graphs that contain a positive ("+") or negative ("-") sign to assign positive or negative connotations to the link in a network graph—can be useful to interpret relations between nodes "as being either positive or negative in affect, evaluation, or meaning."[296]

A signed directed graph, according to Iacobucci, extends that idea to directed graphs.[297] For example, a social network researcher can identify both friends and enemies of a particular person; as shown below (*See* Figure 10), a person's nomination of a friend would be represented by a "+" while their nomination of an enemy would be represented by a "-."[298]

---

[293] *Id.*

[294] *See* Niivaivaara, *supra* note 185, at 2.

[295] Niivaivaara, *supra* note 185, at 2.

[296] Dawn Iacobucci, *Graphs and Metrics*, *in* SOCIAL NETWORK ANALYSIS: METHODS AND APPLICATIONS 92, 136 (Stanley Wasserman & Katherine Faust eds., 1994).

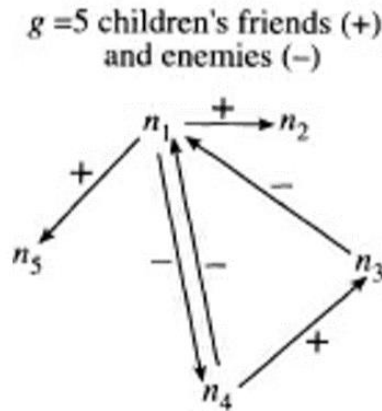[297] *Id.* at 139.

[298] *Id.*

Figure 10: Signed directed graph example.[299]

The signed directed graph concept can naturally be applied within the legal CNA design scheme to provide qualitative context to judicial citations. As Dorelan and Mrvar argue, "there is no sensible basis for treating any *overturning* 'citation' tie as a positive citation to the overturned decision,"[300] since that citation is completely different in purpose and form from a positive following citation. But along these lines, citations by judges to non-binding judicial opinions in other courts truly only have two true potential outcomes—judges can analogize the case being cited to draw a favorable comparison between the facts or rule outlined in the cited case and the judge's opinion in the current case to establish background or follow the same line of reasoning; or they can distinguish the case either by reasoning that the circumstances of the cited case are different enough that the rule and precedent should not be followed in the present case or by simply disagreeing with the reasoning of the cited non-binding precedent.[301] When binding

---

[299] Reproduced from *id.* at 139.

[300] Patrick Dorelan & Andrej Mrvar, *Signed Networks for the US Supreme Court Overturning its Prior Decisions*, 39 CONNECTIONS 1, 2 (2019).

[301] For an example of a clear and obvious negative citation in the case study network, *see, e.g.*, McCoy v. Major League Baseball, 911 F. Supp. 454, 457 (W.D. Wash. 1995) (citing Piazza v. Major League Baseball, 831 F. Supp. 420 (E.D. Pa. 1993) to represent a recent case that "recognized that the scope of the antitrust exemption was limited by *Flood*" but explicitly rejecting *Piazza*'s alternative holding, preferring to side with "the great weight of authority [which] recognizes that the scope of the antitrust exemption covers the business of baseball."). For more on how cases are either analogized or distinguished in legal writing, *see generally* Eugene Volokh, *Analogizing and Distinguishing Cases*, VOLOKH CONSPIRACY (Aug. 10, 2009, 2:26 PM), http://volokh.com/posts/1249928819.shtml [https://perma.cc/7JJR-SNHA]; *How to Craft an Effective Case Comparison*, WRITING CTR. AT GEORGETOWN LAW (2017), https://www.law.georgetown.edu/wp-content/uploads/2018/07/How-to-

precedent is cited by the same or a higher court, the difference is even simpler: the judicial opinion can either follow and uphold the prior opinion *or* it can overturn the prior precedent in favor of new law.[302]

At the same time, the degree to which a court can decide to use or dismiss cited precedent can obviously vary; for example, a court dismissing a cited case as against "the great weight of authority"[303] has a much more negative quality to the citation link than a court stating that the cited case has precedential effect but the facts surrounding the two cases differ enough that the case does not need to be followed. Further, citations could hypothetically be interpreted as neutral, where, for example, a judge simply states that a case exists without noting its effect on the court's decision.[304] But when judicial citations are distilled to their bare essence, all judicial citations of cases must be either positive or negative; even the aforementioned 'neutral' example is still a positive citation, since the judge is clearly using the case to give the necessary legal background for his or her holding and not explaining why that case is different from the present case or incorrectly decided.[305]

For the purposes of studies employing the CNA application—where the positive or negative values are merely used to set the stage for later qualitative legal doctrinal research on the links between cases—coding takes a strictly binary form; a "+" and a green-colored line is assigned to citations deemed as positive while a "-" and a red line is assigned to citations deemed as negative. This is despite the fact that the perceived degree of quality of those citations (i.e., how comparatively positive or how comparatively negative a citation may be) can absolutely vary. This coding process is analogous to

---

Craft-an-Effective-Case-Comparison.pdf [https://perma.cc/Y8KA-5WJV].

[302] *See generally Stare Decisis*, CORNELL LAW SCHOOL LEGAL INFORMATION INSTITUTE (last visited Sept. 5, 2021), https://www.law.cornell.edu/wex/stare_decisis [https://perma.cc/Q8PN-W36E].

[303] McCoy v. Major League Baseball, 911 F. Supp. 454, 456 (W.D. Wash. 1995).

[304] *Id.*

[305] A future research project could work to assign quantitative values to the degree in quality that a case is either analogizing or distinguishing a case, where a citation that more strongly refutes a case could have a more negative number than a citation that simply distinguishes the case based on the facts. *See* Iacobucci, *supra* note 296, at 139–42 (explaining valued directed graphics, where "the strength and frequency of each tie is recorded" using numerical values). For example, the link between two cases where the latter judicial opinion criticizes but ultimately follows the prior opinion could be coded as a positive 1 to indicate the relative weakness of that citation compared to the link between two cases where the latter judicial opinion that completely follows the rule established in a case, which could be coded with a higher number. *See, e.g.*, Salerno v. American League, 429 F.2d 1003, 1005 (2d Cir. 1970) (noting the court's "belief that *Federal Baseball* was not one of Mr. Justice Holmes' happiest days, that the rationale of *Toolson* is extremely dubious and that, to use the Supreme Court's own adjectives, the distinction between baseball and other professional sports is 'unrealistic,' 'inconsistent' and 'illogical'" but ultimately following *Federal Baseball* because "we continue to believe that the Supreme Court should retain the exclusive privilege of overruling its own decisions."). Such a decision could—and perhaps should—be quantitatively distinguished between a decision that followed the cited opinion more favorably, but the methodology is not designed for such an endeavor.

what quantitative researchers term dummy variables, in which data are forced into one of two distinct and mutually exclusive categories.[306] This coding is completed using a directed content analysis approach, where each case citation is analyzed and coded as either positive (creating a "+" sign) or negative (creating a "-" sign) as related to the case being cited.[307]

Notes on why a case is being coded as positive or negative should be coded into the NodeXL workbook, saved in appendices, and explained further in the doctrinal analysis portion of the study. To do so, three columns can simply be added to the NodeXL spreadsheet used for the prior quantitative analysis: (1) a column with a "+" or "-" which will allow for signed graph visualization in NodeXL; (2) a text column giving the specific page numbers in the citing case where the citation occurs; and (3) a text column briefly explaining the reasoning for the code, including whether the citing case either followed, distinguished, simply acknowledged, etc., the cited case.

After these codes were collected, the directed network graphs created earlier using the NodeXL software were edited to include signs on each network edge. The chronological organization scheme used to create the concentric rings on the original graphs was kept in place for the signed graph visualizations to give further context behind each citation beyond simple plusses and minuses in order to show any substantive evolution in the application of the law to the specific sports industry sectors being analyzed in each discrete citation network. Further, to increase readability, each edge was colored as either green (indicating a positive citation) or red (indicating a negative citation) in the network graph visualizations. In this regard, the creation of signed directed citation graphs in the context of legal CNA proves valuable to visually capture the essence of the qualitative shifts in each discrete citation network.

Using these codes in the baseball exemption study, a signed network graph was created to show the entire baseball exemption while illustrating how each case within the network was cited by future cases. (*See* Figure 11). For the sake of clarity, each edge in the network graph was color coded as either green (signifying a positive citation) or red (signifying a negative citation) along with the "+" or "-" designation.

---

[306] DIMITRIOS ASTERIOU & STEPHEN G. HALL, APPLIED ECONOMETRICS 209–10 (3d ed. 2015).

[307] Hsiu-Fang Hsieh & Sarah E. Shannon, *Three Approaches to Qualitative Content Analysis*, 15 QUALITATIVE HEALTH RES. 1277, 1281–83 (2005) (describing the directed content analysis approach as "a deductive use of theory" where the goal of the researcher is "to validate or extend conceptually a theoretical framework or theory" using a structured process that uses "existing theory or prior research" to "identify[] key concepts or variables as initial coding categories.").

As seen in Figure 11, there were a surprising amount of negative citations to *Federal Baseball* signifying a fair amount of disagreement as to the exact scope of the baseball exemption within the network. Immediately following the Seventh Circuit's questioning of the continued broad scope of the exemption in *Finley v. Kuhn* in cases with only "some attenuated relation to the business of baseball,"[308] a growing number of cases distinguished *Federal Baseball* and the baseball exemption generally as shown by a growing number of negative network links originating from cases decided in the 1980s and 1990s.[309] This general finding conforms with the findings in the quantitative portion of the study, where it was observed that the 1990s contained a number of surprisingly important cases (i.e., *Piazza* and *Butterworth*) given these cases' status as district court and state Supreme Court cases rather than federal appellate court cases.[310]

---

[308] Charles O. Finley & Co. v. Kuhn, 569 F.2d 527, 541 n.51 (7th Cir. 1978) ("We recognize that this exemption does not apply wholesale to all cases which may have some attenuated relation to the business of baseball") (citing Twin City Sportservice v. Finley, 365 F. Supp. 235 (N.D. Cal. 1972), *rev'd on other grounds*, 512 F.2d 1264 (9th Cir. 1975)).

[309] *See supra* note 180; *See* Figure 4.

[310] *See* Piazza v. Major League Baseball, 836 F. Supp. 269 (E.D. Pa. 1993); Major League Baseball v. Butterworth, 181 F. Supp. 2d 1316 (N.D. Fla. 2001).

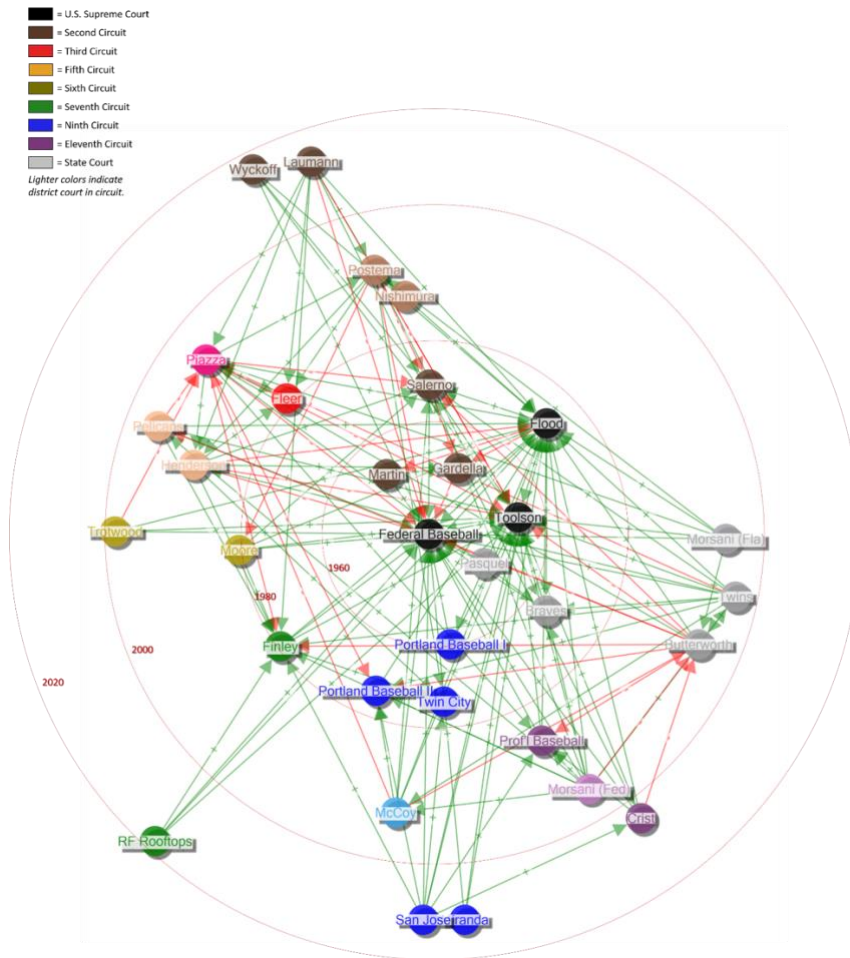Figure 11: Complete Federal Baseball Signed Network Graph
Visualization[311]

The use of the term "robustness" here does not meet the traditional use of the term as normally employed in quantitative methods, which is generally defined as the use of a statistical test to remove the effects of underlying assumptions from statistical models.[312] In this context, however, the infusion of qualitative legal doctrine into the analysis of the created network graphs

---

[311]   *See supra* note 180; *see* Figure 4.
[312]   PETER J. HUBER, ROBUST STATISTICS (1st ed. 2004). *See also* WASSERMAN & FAUST, *supra* note 89, at 114–15 (noting that tests of the connectivity of a graph is the usual method for measuring robustness in quantitative social network analysis).

does add robustness to this study when analyzing the importance of cases by citations, as despite one citation's treatment by the quantitative metrics as neutral, a qualitatively-defined negative citation (especially in the case of a rejection of the studied judicial reasoning) of a studied case would logically not be as important within the citation network information flow as a positive citation. The use of mixed methods in this regard thus allows the study to speak to the "quality of judicial citations" beyond their simple quantity.[313]

2. A Note of Caution: The Need for Supplemental Legal Analysis

Despite the demonstrated value of signed network graph visualizations in legal CNA, it is acknowledged that these graphs require further explanation to fully discern the nuances of judicial citation. Standard quantitative network analysis graphs cannot parsimoniously account for the difference between citing a case to follow it and citing a case to diverge from it.[314] Accordingly, this makes the infusion of qualitative legal doctrinal data through signed directed citation graphs especially important within the legal CNA methodology.

But these signed graphs only do half of the necessary work. Taking two citations to within the baseball exemption, for example, the cases *McCoy v. MLB*[315] and *Piazza v. MLB*[316] each cite one of the benchmark Supreme Court baseball exemption cases, *Flood v. Kuhn*, in a positive fashion, following *Flood*'s holding regarding the baseball exemption (as district courts are want to do when citing Supreme Court cases). However, they do so in radically different ways. *McCoy* took *Flood* to preserve the baseball exemption in its broadest forms until such a time that Congress were to act to remove or limit the exemption.[317] On the other hand, *Piazza* had interpreted *Flood* as having "stripped from *Federal Baseball* and *Toolson* any precedential value those cases may have had beyond the particular facts there involved, *i.e.,* the

---

[313] *See generally* Fowler & Jeon, *supra* note 13, at 17. The use of mixed methods in this regard is not a new concept even when looking at the sport management literature. For example, a sport marketing study employing qualitative methodologies noted from the reverse perspective that "since qualitative work depends on the author's unique perspective and interpretation of results, follow-up mixed methods or quantitative studies would help ensure the robustness of results." Ashley Stadler Blank, Kristi Sweeney, & Rhema D. Fuller, *Room for Growth in Professional Sport: An Examination of the Factors Affecting African-American Attendance*, 23 SPORT MARKET. Q. 225, 238 (2014).

[314] *See* Carmichael et al., *supra* note 8, at 228.

[315] McCoy v. Major League Baseball, 911 F. Supp. 454 (W.D. Wash. 1995).

[316] Piazza v. Major League Baseball, 831 F. Supp. 420 (E.D. Penn. 1993).

[317] *McCoy*, 911 F. Supp. at 456–57.

reserve clause," and thus found that the baseball exemption did not apply to the franchise relocation facts at issue in that case.[318]

Despite the stark difference in how those two cases cited and followed *Flood*, both of those citations are shown in an identical fashion on a signed network graph.[319] As

Figure 12 shows, a signed network graph may give some clues as to those differences, as the later-decided *McCoy* did cite and reject *Piazza*'s reasoning—a fact that would be (and is) reflected on the full signed network graph.[320]
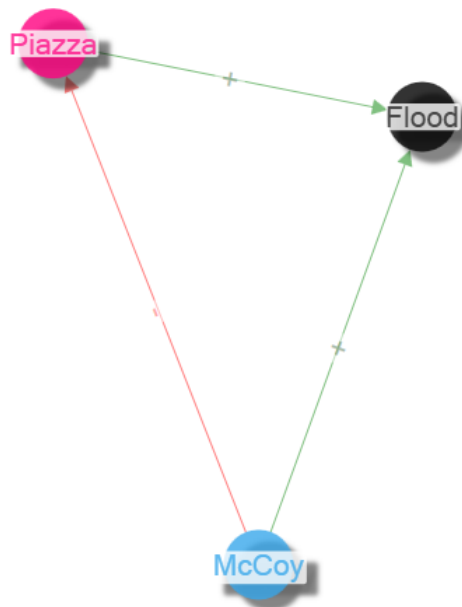

Figure 12: An example of a signed directed graph to show the citation relationships between Flood v. Kuhn, Piazza v. MLB, and McCoy v. MLB.

As such, it is necessary to not simply input qualitative data into the quantitative signed graphs, but to use the signed graphs as a vehicle to show

---

[318]   *Piazza*, 831 F. Supp. at 435–36.

[319]   *See* Figure 12.

[320]   *See McCoy*, 911 F. Supp. at 457 (rejecting *Piazza* on the basis that it failed to consider the full weight of *Flood*'s language, specifically its closing paragraph, which the court interpreted to explicitly affirm *Federal Baseball* and *Toolson*).

further qualitative data through legal doctrinal analysis. The signed network graphs should not be seen as wholly demonstrative of any legal conclusion on their own, but should instead be used to supplement and drive qualitative doctrinal legal analysis. This invites a two-substep process for this qualitative step of the mixed methods CNA: first, the collection of qualitative data and the creation of the signed directed graphs; and second, the explanation and analysis of each of the network edges using more in-depth legal doctrinal research and analysis.

IV.   CONCLUSION

The mixed methods approach to legal CNA proposed in this article can be a powerful tool to take a deep dive into the network flow that defines the history of discrete areas of law. In addition, the quantitative methods can potentially help reveal latent circuit splits that have remained opaque through the decades.[321]

The CNA application discussed herein has substantial value to any area of law bound by "a common substantive focus and an unmistakably clear boundary" in the same way that this article applied it to the case law surrounding the baseball exemption, despite its 100 years of existence.[322] Any judicially-made exemption, specific application of a statute, or even a unique manner by which law has been applied to a particular industry could be studied using the methodology, which can lead to valuable new insights into the nature of law and how precedent changes over years of judicial review and development.

The use of this unique form of legal CNA does present some inherent limitations, however. One major limitation exists with the manner by which case citations were coded in the qualitative network analysis study conducted on each citation network. As specifically discussed in this article,[323] each case citation must be either positive or negative because the ultimate application or articulation of a rule in each cited case can either be followed, resulting in a positive citation; or distinguished/rejected, resulting in a negative citation.[324] Even a neutral citation of a case is inherently a positive citation, as the court citing that case is acknowledging its authoritative influence on the instant issues.

---

[321]   *See, e.g.*, Ehrlich, *supra* note 11.

[322]   Miller, *supra* note 25, at 21.

[323]   *See supra* note 305 and accompanying text.

[324]   *Id.*

This did not present a problem in the baseball exemption network, where the clarity of the legal doctrine forged through three Supreme Court cases ensured that citations—particularly to those cases—were very clear as to their positive or negative nature. In practice, however, cases may not be so easily coded in a binary fashion as either 'positive' or 'negative'; these codes depend heavily on explanation of the context of the citation and—perhaps more importantly—*which* particular citation in a case was picked for coding and analysis. As such, there is some limitation as to the wider use value of the signed network graph visualizations and final quantitative conclusions in this study, as taken out of context they may create some confusion and/or disagreement as to the nature of each studied citation in the network.

But even with those limitations in mind, the mixed methods CNA application in this article has proven useful to quantitatively identify centralized cases within discrete networks, distinguish between positive and negative citations, and visualize case networks and circuit splits, and can have similar important usages to explore other discrete citation networks of legal doctrine. As such, the methods herein look to add to the growing field of quantitative and mixed-methods legal studies by developing a new way to look at citation networks and the manner by which legal doctrine changes and evolves in response to continued addition to the "chain novel" of law through judicial decision-making.[325] While the use of this new application proved useful to studying the unique circumstances of the baseball exemption, continued use of the methodology can seamlessly extend to other discrete networks as well.

---

[325] DWORKIN, *supra* note 57, at 228–38. *See* Lindquist & Cross, *supra* note 56, at 1157–59; Sunstein, *supra* note 57. *See also supra* notes 56–59 and accompanying text.