



Scientific consensus, the law, and same sex parenting outcomes



jimi adams^{a,*}, Ryan Light^b

^a Department of Health & Behavioral Sciences, University of Colorado Denver, Campus Box 188, PO Box 173364, Denver, CO 80217-3364, United States

^b Department of Sociology, University of Oregon, 1291 University of Oregon, Eugene, OR 97403, United States

ARTICLE INFO

Article history:

Received 17 November 2014

Revised 4 May 2015

Accepted 9 June 2015

Available online 10 June 2015

Keywords:

Same-sex parents
Scientific consensus
Citation networks
Sociology of law
Same-sex unions

ABSTRACT

While the US Supreme Court was considering two related cases involving the constitutionality of same-sex marriage, one major question informing that decision was whether scientific research had achieved consensus regarding how children of same-sex couples fare. Determining the extent of consensus has become a key aspect of how social science evidence and testimony is accepted by the courts. Here, we show how a method of analyzing temporal patterns in citation networks can be used to assess the state of social scientific literature as a means to inform just such a question. Patterns of clustering within these citation networks reveal whether and when consensus arises within a scientific field. We find that the literature on outcomes for children of same-sex parents is marked by scientific consensus that they experience “no differences” compared to children from other parental configurations.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

A central component of the legal debate concerning recognition of same-sex unions addressed whether and how outcomes for children raised by same-sex parents differ in comparison to children raised in other family configurations. Decisions in two cases by the US Supreme Court hinged, in part, on the scientific consensus regarding these questions (*United States v. Windsor*, 2013; *Hollingsworth v. Perry*, 2013). Reflecting the centrality of children's outcomes in this decision, during oral arguments, Justice Scalia suggested that “there’s considerable disagreement [a]mong sociologists as to what the consequences of raising a child in a [s]ingle-sex family, whether that is harmful to the child or not” (*Hollingsworth v. Perry*, 2013: 19). Accordingly, scientific consensus became a central question for scientific papers and amici curiae briefs addressing this case (APA, 2013a, 2013b; ASA, 2013; Pafford, 2013; Siegel et al., 2013). For example, a brief submitted by the American Sociological Association (ASA) concludes, “Whether a child is raised by same-sex or opposite-sex parents has no bearing on a child’s wellbeing” (ASA, 2013: 3); and Siegel and colleagues recommend that decisions about parenting competency should be made “without regard to their parents’ gender or sexual orientation” (Siegel et al., 2013: 829). In this recent case and others, such as *Wal-Mart v. Dukes* (2011), the relationship between social science and the law has been brought to the fore.

Indeed, a significant strain of legal thought relies, in part, on social science for guidance in arguing and adjudicating legal cases (Monahan and Walker, 2009). For example, expert witnesses trained in a social science may attest to the likelihood that a defendant will re-offend during the sentencing phase of a criminal trial (Uggen and Inderbitzin, 2010). Alternatively, more general evidence of the persistence of racial disparities may influence a judge’s decision about a potential violation of the

* Corresponding author.

E-mail addresses: jimi.adams@ucdenver.edu (j. adams), light@uoregon.edu (R. Light).

Voting Rights act (for a recent example, see references to sociologist Vincent J. Roscigno's expert testimony in *NAACP v. Husted*, 2014). The success of these empirically-based arguments is likely to depend in part on where scientific consensus lies on legal issues confronting the courts (Acker, 1990).

Social science has become widely incorporated into the American legal system. Indeed, every Supreme Court Justice at least since 1986 has cited social science in at least one opinion or joined an opinion that has (Acker, 1990; Monahan and Walker, 1986: 477).¹ With this in mind, lawyers have, for some time, tried to wrangle social science to their favor collecting amicus curiae briefs and hiring social scientists as expert witnesses. Social science becomes a key element of many legal cases and the balance of these cases often hinges on arguments about the trust placed in empirical research (Collins and Evans, 2002). Further, this trust is often explicitly linked to whether a particular social scientific issue has reached a state of consensus. Although a good deal of research addresses the formation (e.g. Shwed and Bearman, 2010) and perception of scientific consensus within society generally (e.g. Oreskes (2004) on climate change), less attention has connected the tools of sociology of science to key issues before the courts. Lawyers, juries, and judges are left to determine scientific consensus on their own (Acker, 1990).

Systematic approaches from the sociology of science provide strategies for evaluating scientific consensus and can assist in the legal determination of the value of social science research. These approaches are particularly needed for social science evidence that aims to address issues beyond narrow determinations about issues specific to a case to, as Acker states, "general, empirical propositions about social events or relationships that may be instrumental to legal rule-making" (1990: 26). This evidence most often enters the legal record via amici curiae briefs, which may—or may not—necessarily directly engage scientific consensus on the question(s) before them.

Shwed and Bearman (2010) demonstrate one method for identifying when scientific consensus emerges. This method employs citation networks among the relevant scientific literature to provide a structural overview of consensus within a scientific community as a means of capturing content shifts. Like Shwed and Bearman (2010), we conceive of consensus as a type of closure. In other words, consensus forms when boundaries around a contentious issue are resolved or enclose around one correct answer (Shwed and Bearman, 2010: 820; Collins and Evans, 2002). This type of consensus forms a key dimension of how and whether social scientific research is accepted within the courts (Daubert v. Merrell, 1993). Here, we build upon their method to examine whether the research on outcomes for children of same-sex parents is marked by scientific consensus, and estimate when the attained level of consensus was achieved. These techniques can inform court decision-making, but are also worth answering independent of whether "no difference" should be the basis for adjudicating a legal decision about union legalization (Yurkewicz, 2012).

We adapt Shwed and Bearman's approach to the case of same-sex parenting outcomes. Our adaptation is consistent with their theoretical and methodological bases in two ways. First, a purely search-term based means for identifying the salient literature has some limitations, which requires the identification of strategies for estimating the robustness of identified patterns to potential noise in the data. Second, we note that in absence of expert "consensus statements," akin to the one made by the Intergovernmental Panel on Climate Change about anthropomorphic causes of climate change, the method also requires scientific consensus to be interpreted as a process, not solely an end-state. To preview, our analyses reveal that consensus exists on outcomes for children of same-sex parents, and arose early in the twenty first century.

2. Background

The relationship between scientific consensus and policy is complicated. On one hand, policies can respond to scientific consensus, putting into practice recommendations that stem from empirical evidence. Alternatively, policies may be enacted prior to the identification of scientific consensus, or even agnostic to the empirical basis for such policy recommendations. Either way, if policy is to be linked to empirical scientific evidence—whether in the former situation as the basis of enacting policies or in the latter for evaluating their impacts—we require systematic means for identifying whether scientific consensus exists, and when it arises. Additionally, it is important to understand this as a dynamic unfolding process, i.e., scientific consensus itself has a life course reflecting various stages of (un-)certainty, which must be identifiable to appropriately assess its association with policy. The dynamic characteristic of consensus highlights how closure around a particular set of ideas differs from "correctness" in any metaphysical sense as the state of consensus is subject to change. Nonetheless, the question of consensus remains a key issue affecting how social policy is adjudicated in the courts.

We are specifically interested in how the legal community incorporates science, particularly social science. To that end, we provide a means to evaluate whether consensus has formed around a particular topic.² The following overview intends to summarize pertinent legal scholarship on the relationship between social science, consensus, and the law. This overview will

¹ As discussed in greater detail below, social science plays a role and is taken seriously by Justice Thomas, in his concurring opinion, and Justice Breyer in his dissenting opinion on *Parents Involved in Community Schools v. Seattle School District No. 1* (2007), a case on school racial integration policies. Justice Scalia, joined by Justices Kennedy, Ginsburg, Sotomayor and Kagan, engages in a serious discussion of causality within social science research on the effects of video game violence in his *Brown v. Entertainment Merchants Association* (2011) ruling opinion. Justice Alito's concurring opinion disagrees with this aspect of Scalia's argument and reinterprets the social science on video game violence. Chief Justice Roberts's majority opinion, *Shelby County v. Holder* (2013) deeming parts of the Voting Rights Act of 1965 unconstitutional, uses some social science as evidence for decline in racial differences in political participation, although Kimball (2013: 324) argues that it is "jarring" how little of the social science evidence is introduced by Roberts.

² While somewhat related, here we focus less on legal frameworks or philosophies that are informed by social scientific disciplines, like rational choice and the law or critical legal studies.

also illustrate that determining the state of social science research is a pressing one and reaches the highest levels of the US courts. We demonstrate one way that sociologists can help the legal community evaluate the state of consensus within a (social) scientific field.

2.1. Social science and the law

Social science has entered legal discussions with mixed success since the 1908 *Muller v. Oregon* case widely seen as the first to systematically engage social science research. Numerous social scientists and legal scholars recognize the inconsistent reception of social science as a serious problem (Acker, 1990; Grunwald, 2012; Monahan and Walker, 1986). At the level of the US Supreme Court, several key decisions, most notably *Brown v. Board of Education*, have drawn heavily on social science. However, the value of social science for legal arguments has not led to its systematic use. In fact, in several key decisions Justices in the majority explicitly identify their own disagreement with summarizations of social scientific findings made by social scientists. The majority opinion in *Lockhart v. McCree* serves as one example (see Bersoff, 1987 for an insider's view). This case centered upon the issue of whether the exclusion of anti-death penalty jurors in capital criminal cases, or cases that could result in a sentence of death, could affect the guilt-phase of a trial. The Supreme Court had previously established that jurors could be excluded for holding anti-death penalty views as these potential jurors were essentially telling the judge that they would not follow the law during the sentencing phase (*Witherspoon v. Illinois*, 1968). As most death penalty states have a sole jury determining both guilt and sentence, jurors with anti-death penalty beliefs were excluded from both the guilt and sentencing phases.

As Bersoff (1987: 54) describes, *Witherspoon* is one of the few cases where the Court issues an "invitation to social scientists to develop data to help it in resolving a crucial point of law." This crucial point is whether those willing to impose a death sentence are more likely to convict a defendant than those anti-death penalty jurors who are excluded from all phases of a capital trial. Social psychologists took the Court's challenge seriously and built a small, but robust literature consistently describing how the exclusion of anti-death penalty jurors indeed increases the likelihood of convictions. The American Psychological Association submitted an amicus brief to *Lockhart v. McCree* (1986), a case challenging the *Witherspoon* decision, summarizing these findings. Yet, the five Justices signing the majority opinion written by Justice Rehnquist fundamentally disagreed with this summary arguing that there were "several serious flaws" in the presented social science evidence (see *Lockhart v. McCree*, 1986; Bersoff, 1987).

More recently, in his dissenting opinion on the school desegregation case *Parents Involved in Community Schools v. Seattle School District No. 1* (2007), Justice Thomas writes that serious disputes within social science on the positive effects of diversity programs call to question the educational benefits of "coerced racial mixing." He concludes that we should not leave our "equal-protection jurisprudence at the mercy of elected government officials evaluating the evanescent views of a handful of social scientists" (2007:766).³ Needless to say, the reception of social scientific data has been neither universal nor unanimous, but is subject to interpretation and debate. While some of the disagreement regarding social scientific data may be based on credibility gaps between the social and natural sciences (Wechsler et al., forthcoming), numerous scholars argue that it is often the variation within standards of practice placed on social scientific data by both social scientists and jurors that are most in question (e.g. Sorenson and Sharkey, 2011; Acker, 1990; Monahan and Walker, 1986). What properties of social science lead to its inconsistent incorporation by the legal community?

One major sticking point for wider incorporation of social science is the diverse tasks asked of social scientific research. Social scientists are asked to both contribute to establishing facts specific to a case, or "adjudicative facts," and to address issues that speak to facts more generally, or "legislative facts." To the latter point, Monahan and Walker (1986) argue that social scientific evidence should not solely be treated as facts in a legal sense, but as "social authority" more akin to precedent. In this and subsequent work on "social frameworks," Monahan and Walker (1986) outline a systematic approach to incorporating social science into normative determinations based on the principle of generalizability. As they write:

Research used in the creation of a rule of law – for example, studies on the effects of school segregation on self-esteem, the effects of exposure to pornography on anti-social behavior, or the deterrent value of the death penalty – has the same kind of future-oriented generality that case precedent possesses (1986: 41).

Social science research should not be limited to adjudicating facts – psychological analysis of a particular person or case – but can pertain to more fundamental and less transitory aspects of the law. Indeed, much of the social science research advanced within the courts, especially the higher courts, seeks to do just that.

To more deeply engage and gain the trust of the legal community, Monahan and Walker (1986) argue that social science research must meet several standards that relate both to the content and the structure of research. These standards parallel familiar principles of sound methodology, such as validity and generalizability. Moreover, these issues as they pertain to expert scientific witnesses were more formally resolved by the Court in its *Daubert v. Merrell Dow Pharmaceuticals* (1993). This opinion outlines guidance for how judges are to interpret scientific testimony. The characteristics include whether (1) the expert's theory can and has been tested, (2) it has been peer reviewed, (3) it identifies and measures error, (4) this

³ Breyer's noteworthy dissent on this case offers an alternative view of what standard of consensus should be applied: "If we are to insist upon unanimity in the social science literature before finding a compelling interest, we might never find one."

error is controlled, and (5) the social scientific research in question has “attracted widespread acceptance within a relevant scientific community” (1993: 12).

Identifying the heavy-burden placed on the courts to determine the value of a particular body of social science research, Acker (1990) advocates an organizational approach that incorporates greater participation by professional associations in submitting amici curiae briefs. Acker (1990) argues that Justices appreciate social science so much that they expend a great deal of time trying to figure it out. Like Monahan and Walker (1986), Acker further argues that professional organizations “can make an important contribution to the transmission of social science information to the justices, and that social scientists need not serve only as after-the-fact critics of the Court’s opinions” (1990: 40). In this context, professional organizations, like the ASA or APA, provide a key sorting mechanism for summarizing relevant literature and identifying points of consensus, a necessary function when there is a glut of information as is the case in many scientific subfields. This encourages professional organizations to play a direct role in guiding public interpretations of its findings.

The state of consensus can be and often is identified cognitively as experts writing briefs or participating as scientific witnesses use their training to identify the state of their field (Shwed and Bearman, 2010). However, this lack of formalism permits easy disagreement as one scholar’s immersive inspection may considerably differ from another’s. This difference can result from the state of information glut itself: The parsing of a literature is far more complicated than it once was because many subfields have grown dramatically. How and whether expert witnesses provide a fair summary of the literature is one point of contention identified in the debate over the testimony in *Wal-Mart v. Dukes* (2011). The respondents in this case alleged that their employer, Wal-Mart, discriminated against them because of their sex. They sought relief for themselves and for 1.5 million other female Wal-Mart employees. A sociologist, William Bielby, participated as an expert witness in this case concluding based on his research that Wal-Mart’s organizational practices contributed to “disparities between men and women in their compensation and career trajectories at the company” (Bielby, 2003: 408). Wal-Mart petitioned the Court to deny the class action. The ASA sided with Bielby and filed an amicus curiae brief on behalf of the respondents. But, the Court ultimately sided with Wal-Mart and singled out the expert testimony as being particularly faulty stating that it “does nothing to advance the respondents’ case” (*Wal-Mart v. Dukes*, 2011).

Following *Wal-Mart v. Dukes*, a debate developed about the role of sociology and the law (Mitchell et al., 2011; Nielsen et al., 2011; Sørensen and Sharkey, 2011; Tomaskovic-Devey, 2011). This debate illustrates the multidimensional standards required for successful dialogue between social science and the legal system as it both addresses standards of evidence, which can be read as an issue of content, and arguments regarding the state of a field, which can be read as a question of both content and structure. For example, the majority opinion challenges whether Bielby’s evidence provides “significant proof” of discrimination (*Wal-Mart v. Dukes*, 2011). Nielsen et al. (2011) defend the ASA brief and the standard of evidence provided by Bielby arguing that his analysis is replicable and systematic. On the other hand, Sørensen and Sharkey (2011) argue that the ASA brief falsely claims that a consensus has formed around the effects of personnel practices. Nielsen et al. (2011) respond that this literature is more resolved than Sørensen and Sharkey allow. This case illustrates the importance of accounting for both standards of evidence and systematically addressing the description of research fields. Formal techniques for evaluating consensus provide one way out of this interpretive conundrum as, at minimum, it explicates a systematic strategy used to operationalize consensus.

2.2. In search of consensus

Research in the sociology of science and information sciences provides promising directions for analyzing the trajectory of research fields. For example, Shwed and Bearman (2010) propose a network-based strategy for evaluating the temporal evolution of scientific consensus within a literature. Their approach relies on identifying the potentially salient corpus of literature on a subject, then constructing and analyzing a citation network among the papers in that corpus. A citation network is a directed network that captures the potential ordering of influence from cited papers to citing papers. That is, the ties in a citation network represent that each reference in a citing paper reflects the potential influence of each cited paper in the development of the ideas in the focal paper. A network of these citations can then be analyzed to identify common citation practices within the corpus.

Scientific consensus in these citation networks is best reflected by a concept known as community detection, which can evolve over time (Shwed and Bearman, 2010). Conceptually, communities are tightly knit segments of a population. In turn, the analytic idea behind identifying network communities relies on finding subsections of the network where most ties (in our case, cites) occur within the group as compared to across groups (Fortunato, 2010; Porter et al., 2009). The overall level of segmentation into different communities can be used to identify how many communities there are in a given network, and how well those communities account for the general pattern of ties (citations) within the network (of papers). Shwed and Bearman (2010) demonstrate that consensus in a literature is consistent with identifying when modularity—a summary statistic of the overall segmentation among identified communities—attains a sustained low equilibrium level or resolves to a consistently low score.

While citation networks provide a means for estimating if and when scientific consensus arises, the approach proposed by Shwed and Bearman (2010, hereafter SB) requires two assumptions that we extend here. First, relying primarily on search

terms, typically keyword-based, to identify the salient corpus can introduce potential noise into the evaluated corpus. In most of the examples provided by SB, there is a clear unidirectional potential relationship to evaluate. For example, when examining the carcinogenicity of UV radiation one can reasonably use a string of search terms combining solar radiation and cancer.⁴ In that case, there is little chance of finding work also claiming a “reverse” causal pathway of cancer causing UV radiation. As such, the identified corpus is relatively well bounded by the search terms used.

The case of outcomes for children of same-sex parents is not similarly well-bounded. Here, the search terms used (described in the Data section below) returns a corpus which includes: (1) literature that is salient to the question at hand, along with (2) research concerned with children’s developmental concerns generally, *regardless of parental sexual orientation*. These appeared in the set via three primary routes – (i) research on outcomes for children who are themselves LGBTQ, (ii) genetic comparisons that explicitly limited their comparisons to same-sex (sibling) pairs (e.g., twin studies), and (iii) comparisons of intrafamilial outcomes across generations, which often limited comparisons between children and their same-sex parent. We therefore develop a strategy for adapting the SB method to evaluate its robustness to noise in the search-term identified corpus. Our findings remain largely consistent with this adjustment, so we present initial results in a manner consistent with the SB approach, followed by analyses that estimate the robustness of these findings to the noise in the corpus.

A second extension of their method requires interpretation of content in cases that either empirically appear to be still contested, or are claimed to be by researchers. Here, we lean on two separate approaches. First, SB provide several different ideal-typic pathways to consensus; the “cyclical” form that they describe best fits the narrative discourse around the case examined here, while the “spiral” form seems more consistent with the citation network evidence we provide. In the cases SB examine, “expert” consensus claims (e.g., the Surgeon General’s report on smoking) serve both to adjudicate and at times accelerate the consensus processes in contested cases (Evans, 2007). In the case we examine, there are no similarly prominent expert consensus statements. As such, second, we rely on reading and interpreting included content from important articles identified within the examined corpus as a means to interpret the content of consensus on outcomes for children of same-sex parents.

3. Data

We begin by identifying a corpus of literature on same-sex parenting, as represented within ISI Web of Science. From this corpus, we extract an analytic sample of 19,430 publications. We identified this literature corpus from a search for *all* sources that included the following terms in their topic: (same*sex OR homosexual* OR gay* OR lesbian*) AND parent*, which did not restrict the dates of publications returned by this search, and thus started from a list of 21,369 publications between 1900 and 2013 (as of March 2, 2013).⁵ Given that our analyses are based on citations, we drop papers that include no cited references *and* are cited by no other papers in the set ($N = 1770$). From this corpus, we also excluded publications from 1900 to 1965 from analyses, because there were too few publications within this period to suggest the presence of a substantial scholarly literature on the topic as a whole, or within any given window.

Citation networks consist of scientific papers connected to one another via the citations each paper makes. Traditionally in citation networks, a tie is recorded (as “sent”) from paper i to paper j when paper j includes paper i among its cited references (Yan and Ding, 2012). The direction of such ties is intended to reflect the “direction of influence” in that j ’s citation of i indicates that the ideas in paper i somehow influence the ideas in paper j , and is therefore coded as a tie from i to j . SB then examine the evolutionary patterns of community structure within these citation networks. We similarly construct a series of temporal moving-window citation networks, the width of which is determined by computing the time-varying median citation age across the examined corpus.⁶ Citation age provides a concrete means of identifying the salient “epistemic period” of literature from which new research develops its arguments (Shwed, 2015). All data pre-processing was completed with Sci2 (Sci2 Team, 2009), and analyses were conducted in igraph 0.7 for R 3.1.1 (Csardi and Nepusz, 2006).

Analytic results below are further limited to citation networks with windows ending from 1977 to 2013 (papers published beginning in 1970). This exclusion was due to the relative scarcity of publications within the early period of an identifiable literature. This allowed small numbers of papers to produce wide fluctuations in the metrics examined, suggesting patterns containing more noise than meaningful variation. In a sense, this is the most basic form of “no consensus.” There is so little literature that everyone is virtually working independently on a new problem, lacking any basis from which to build or contest consensus. Fig. S2 presents supplementary analyses, which extend the analyses presented in Fig. 1 to also cover this excluded period; doing so largely reproduces the primary findings presented here.

⁴ The precise combination they used was: “(Sun OR solar OR photo OR “UV radiation” OR tanning) AND (cancer OR carci* OR melanoma OR sarcoma OR basal) (Shwed and Bearman, 2010: 826).

⁵ This cutoff reflects the level of consensus in the literature at the time that competing claims regarding scientific consensus were being made before the Court. While we could also extend this analysis to examine how consensus has evolved since then, this would be in service of a different question. The aim here is to assess what level of consensus could reasonably be claimed when competing claims about that consensus were being made before the United States Supreme Court.

⁶ Citation age is computed for each cited reference as: the publication year of the citing paper minus the publication year of the cited paper. See Appendix Fig. S1 for a plot of the moving-window widths.

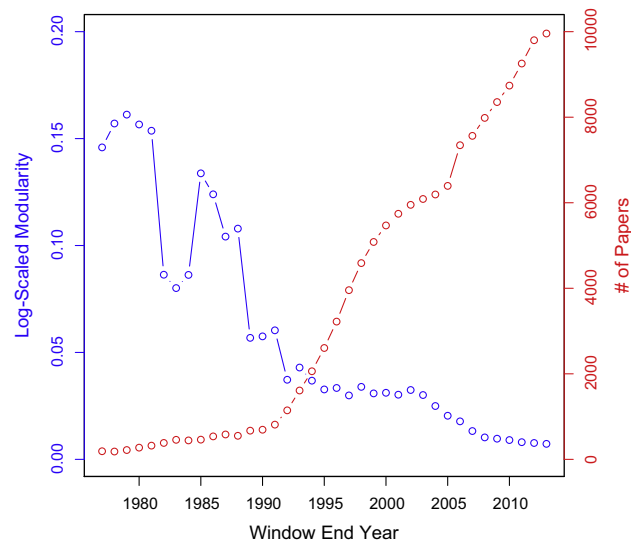


Fig. 1. Modularity and Corpus Size Changes 1977–2013. NOTE: The x-axis denotes the end-year of the moving window for each epistemic period in the corpus. The y-axis presents the changes across those windows in: modularity score (blue/solid line, left axis), and growth in number of publications (red/hatched line, right axis). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Methods

SB estimate the evolution of scientific consensus based on the number and extent of divisions between scientific “communities” in the field. Communities identify subsections of the network where ties are predominantly within group, and limited across groups (Fortunato, 2010; Porter et al., 2009). This community structure can be summarized with a **modularity score** (Brandes et al., 2007; Newman, 2006). We compute this modularity score for each of the temporal slices, using the edge-betweenness method (Newman and Girvan, 2004), then examine how it changes across the observed period.⁷ Modularity is network-size dependent (Shwed and Bearman, 2010). To allow comparability among reported scores—accounting for the increasing size of the corpus across the observed period (see Fig. 1) – we scale modularity scores by the logged number of papers in the network for each observed window (Navon and Shwed, 2012; Shwed and Bearman, 2010).

As described above, the search-term based identification of the corpus examined appears to have introduced more noise into the data than was the case for examples in SB. As such, in addition to plotting these complete corpus modularity trends, we also devise an approach to examine how robust the patterns identified are to this noise. Given that the amici curiae briefs and the consensus statement published in *Pediatrics* take our central question here as their focus, we take the citations from this “core set” as a starting point, assuming that their citations are substantially less subject to this data noise limitation. From this core set, we then extract three additional citation networks from the complete corpus and reproduce the same analyses on each. These restricted networks follow the (out) citations from within the core set, to 1-step, 2-steps and 3-steps out from the core set. Respectively, these extracted sub-networks include 523, 1667 and 2912 publications of the initial corpus’s 19,761. We then examine how the modularity trends in Fig. 1 compare to computations on each of these subsets.

Examining the temporal shifts in modularity allows us offer a structural basis for identifying whether and when consensus arises within the field. However, given consensus, it does not allow us to directly determine what that consensus is. In other words, following the two dimensions of consensus, the structural approach answers the question of whether consensus has developed in a particular field - a key question asked of social science by the courts -, but leaves open the question of what conclusions the field has drawn. To estimate the content of consensus, we sample the top 10% of papers based on citations received in a selection of distinct stages in the identified modularity curves: the period of high modularity, suggestive of “epistemic rivalry” (network slice ending in 1985), declining modularity, suggestive of an arising consensus (network slice ending in 1990) and two different points from the period of sustained low modularity, suggestive of achieved scientific consensus (windows ending in 2000 and 2010). Degree centrality is frequently used as an indicator of perceived importance (Borgatti and Everett, 2006). We examine the content of these “important” articles to assess the primary claim(s) being made in each representative period about outcomes for children of same-sex parents.

⁷ We obtained similar results using the Louvain method as an alternative algorithm (Blondel et al., 2008), as implemented in Pajek 3.01 (Batagelj and Mrvar, 2003).

5. Findings

5.1. Consensus trajectory

Fig. 1 presents the modularity trends for moving windows ending from 1977 to 2013. The pattern in our case is consistent with two examples in SB: research on anthropogenic sources of climate change and carcinogenicity of ultraviolet radiation. In each of these cases including our own, the trend is marked by an early period of high modularity (here occurring in periods that end prior to the mid 1980s), or strong community segmentation, indicative of early epistemic rivalries in the field. This is followed by a sharp modularity decline (occurring in the mid 1980s through mid 1990s), representing the consolidation of those previously separated communities, a pattern that is indicative of an arising consensus. Finally, each of these cases resolve into a sustained period of low modularity representing consolidated community structure, which SB claim is indicative of achieved scientific consensus (here sometime after the mid-1990s). This is consistent with what SB label a “spiral” pattern of consensus formation, wherein “some initial epistemic rivalry is quickly resolved, and scholars then move to secondary questions. This leads to increasing numbers of papers linked to a common core, keeping modularity low” (Shwed and Bearman, 2010: 830).

Next, to assess how robust this pattern is to the noise in the data described above, we examine the modularity trajectory for three alternative samples of the literature, drawn from the citation networks emanating from the “core set” of citations in the three amicus curiae briefs and the *Pediatrics* consensus statement. Fig. 2 presents these results, with a line for each of the examined networks. The general pattern of a spiral consensus process is reproduced regardless of which sample is used, suggesting that the temporal trend observed above is largely robust to the substantial noise in the data. These ancillary analyses differ from those presented in Fig. 1 in two ways. The closer we move to the “core set” of citations drawn on, the later it appears that consensus was attained (see how the downward sloping part of the curve shifts to the right as you move to the more restrictive samples). However, the pattern among all of these samples remains virtually indistinguishable since the late 1990s. A second difference is that these more restrictive samples appear to draw on citations that are slightly older than those identified in our search-term based corpus (see Fig. S2). This would be consistent with researchers drawing on work they are aware of, which necessarily has some temporal limitations.

For fields with achieved scientific consensus, SB overlay the sorts of modularity trends described above with the timing of “consensus statements” (e.g., reports by the IPCC, or The Surgeon General’s reports on the carcinogenicity of smoking). These overlays allow comparisons between when modularity indicates consensus is achieved and when scientists make declarative statements about that consensus. In the case of outcomes for children of same-sex parents, there have been no similarly authoritative “consensus statements”, though the amicus curiae briefs and *Pediatrics* statement each makes claims about this consensus. It is worth noting, however, that they do so well after consensus was apparently achieved; among the examples in SB, only the timing of the IPCC’s 2007 report on climate change is similarly delayed (Shwed and Bearman, 2010: 829).

Given the delay in consensus statements in our case, the trends in Fig. 1 can also be compared to other landmark moments *indirectly* connected to this consensus. The American Bar Association issued a statement in 1989 encouraging sexual orientation as a class for non-discrimination protection (ABA, 2006), before modularity equilibrated (i.e., prior to consensus formation), which is similar in timing to the Surgeon General’s first report on the carcinogenicity of smoking (Shwed and Bearman, 2010: 829). DOMA, the Defense of Marriage Act passed in 1996 (US Congress, 1996), during the period of modularity decline (suggesting developing consensus). The “Don’t Ask, Don’t Tell” policy passed in 1993 (US Congress, 2010), just after the period of sharp modularity decline (suggesting approaching consensus), a pattern consistent with the International Agency for Research on Cancer’s 1997 report on UV-radiation’s carcinogenicity, and the IPCC’s 1995 report on climate change (Shwed and Bearman, 2010: 830). The period of sustained low modularity (suggesting high consensus) coincided with the repeal of DADT (2010).

5.2. Consensus content

Using citation degree centrality, we then sample the top 10% of articles in a set of representative periods from the curves illustrated in Fig. 1, to examine their content as a means of understanding what the literature reveals about the state of knowledge on outcomes for children of same-sex parents at each of these respective points in consensus development. Before describing the temporal differences in this content, it is worth noting two common patterns that did not vary across time within the sampled important articles. First, the literature focuses more heavily on lesbian mothers (single or partnered), rather than on same-sex parents generally. Second, studies disproportionately focus on children’s sexual orientation over other outcomes.

How content *changes* align with representative points in the modularity curves allows interpretation of the evolving consensus. Articles from the early period (slice ending in 1985) of epistemic rivalry show a literature addressing a new question and finding mixed results, not strong evidence of advantages/disadvantages by parental configuration. This pattern is consistent with the high modularity evidence suggestive of epistemic rivalry (see Fig. 1). Moreover, studies from this period largely drew on interview data within small samples. Some authors appeared surprised to find similarities of outcomes, while those finding differences found that children of same-sex parents experienced higher likelihood of dealing with parental divorce and shifting parental/child roles.

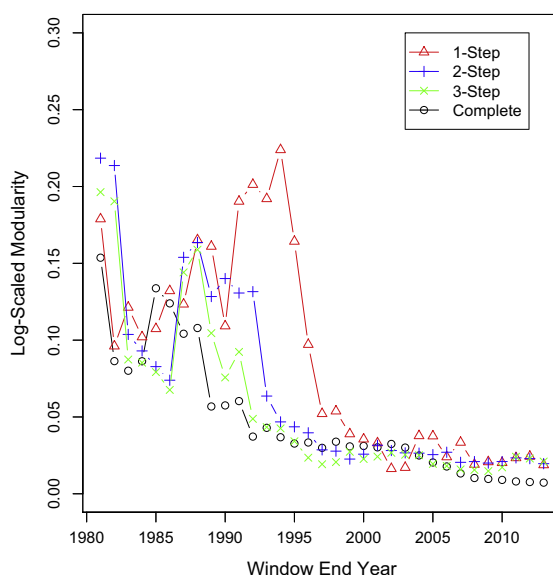


Fig. 2. Modularity comparisons for complete corpus and extracted network subsets. NOTE: The x-axis denotes the end-year of the moving window for each epistemic period in the corpus. The y-axis presents the changes modularity score across those windows. The black line (circles) reproduces the complete corpus line from Fig. 1. The other lines refer to the sub-networks selected from sent citations among the “core set” of cited papers (see text for description), separately for the subset as defined by the number of steps away from that core set.

During the period of declining modularity (slice ending in 1990), key articles reflected the suggested *developing* consensus. Research in this period increasingly focused on comparisons between single parents, to specifically isolate the effect of parental same-sex preference from that of family instability more generally (Perrin et al., 2013). In this period, articles disproportionately focused on same-sex parenthood that occurred following dissolved heterosexual partnerships; i.e., not as much research involving same-sex parents who adopted or used IVF. Thus, research increasingly attempted to separate effects of parents’ partnership instability from parental sexual identity. With these “controls” in place, such studies began pointing to finding non-differences. During this period, several researchers also specifically noted lesbian mothers’ socialization efforts toward instilling gender-*normative* behaviors in their children.

Finally, highly cited articles from two recent periods (slices ending in 2000 and 2010), also supported the recent period’s low modularity suggestion of high consensus. Additionally, the overwhelming consensus view among these articles is consistent with the *Pediatrics* statement and ASA/APA briefs claims of “no difference” (APA, 2013a, 2013b; ASA, 2013; Siegel et al., 2013). In fact, the few studies noting any differences within this period identified *advantages* for children of same-sex parents (e.g., increased parental bonding). Recent research also extended to include samples of *adults* raised in same-sex relationships (whereas earlier research focused on outcomes during childhood), and children *born into* same-sex partnerships.

6. Discussion

Consensus is not something that should be considered to simply exist or not. Instead, as the temporal evolution approach used here demonstrates, its state at any given point in time should be seen as a cross-sectional snapshot of a process that unfolds over time. For the questions posed here – whether there is consensus on outcomes for children of same-sex parents, when and how it arose, and what it is composed of – this dynamic perspective has three implications for interpreting the answers our results provide. First is to characterize the nature of that unfolding process rather than simply identifying the state at any given point in time. Second various constituents can draw upon those patterns to make various prescriptive claims (e.g., for how the Supreme Court should go about making a decision) that can either rely upon or contest that existing state of knowledge. Third recognizes that any current state that arose from a temporal process can be altered in the future.

The pattern evidenced in Fig. 1 is characteristic of a “spiral” consensus formation pattern, which coincides with a rapid cascade towards consensus, after which most work continues to link to that common core of ideas (Shwed and Bearman, 2010). In other words, as many have claimed—including opponents of same sex marriage (e.g., Regnerus, 2012)—the field is one that is largely marked by an achieved consensus. Moreover, the supplementary interpretation of important articles from the latest periods indicates that the achieved consensus is consistent with the perspective of “no differences” on outcomes for children of same-sex parents.

Various constituents may subsequently draw upon the existing scholarship in ways that are consistent with the empirical results above or raise questions about it. It is important to note that any prior attempts to *systematically review* the literature, have supported the no differences conclusion (Siegel et al., 2013). If opponents to this consensus engage with the literature comprising it, they must either ignore that literature, question the basis of the claims within it, or draw selectively from it

to support their position. Among those who have made the most recent and critical claims about a “no differences” consensus are Regnerus (2012) and Marks (2012). Each of these authors questions the evidence base underlying the consensus, by arguing that because the literature does not rely primarily on population-based samples, the claims of consensus are premature without such evidence. There are two things worth noting about this approach. First, it actually concedes that there is consensus in the scientific literature—in other words it would be in agreement with the primary findings presented above. Second, in assessing the evidence-base questions these authors raise, recent re-analyses show that applying the additional standard of evidence requiring reliance on population-based samples—via Regnerus’s own New Family Structures Survey data, when conducted appropriately—would only provide additional support to the “no differences” consensus (Cheng and Powell, 2015).

Beyond ignoring or questioning the evidence about consensus, those making counter claims could rely on findings peripheral to the literature comprising that consensus. That is, perhaps they are more likely to draw on less authoritative literature to make their claims. Regnerus’s article (2012) and the amicus brief he coauthored (Pafford, 2013) each leans heavily on this strategy. For example, the first few pages of the brief draw heavily on literatures that have no direct bearing on same-sex parenting, but are instead concerned with topics such as gender norms of communication patterns and advantages for children to married couples compared to single or cohabiting parents—i.e., they rely on indirect evidence to develop their argument, rather than relying on the work that directly engages the question they seek to answer. Moreover, this strategy appears to have inflated the importance of Regnerus’s own claims-making toward this literature, by championing their opposition sufficiently to have forced several sources in the supportive literature to cite them, solely for the purposes of debunking the work they represent (Cheng and Powell, 2015; Perrin et al., 2013), including each of the briefs that claim the no differences consensus (APA, 2013a, 2013b; ASA, 2013).

Questioning and reanalyzing fields marked by consensus is a pattern identified by Shwed and Bearman (2010), one they refer to as the “cyclical” form. In this pattern, “similar questions are revisited without stable closure” (2010: 818). They provide the example of how this marked research on the carcinogenicity of smoking for a large portion of the twentieth century. Shwed (2015) later showed how the tobacco industry’s involvement in delayed the production and dissemination of knowledge about these questions. Considering the application of a similar course to the case here, we could in the future see such contestation re-open the question to such a degree that the field would become less settled than it is currently. However, projecting such future shifts is beyond the scope of our empirical question of whether consensus currently exists, which our evidence suggests is the case, and few seem to be questioning—even those opposing the implications of that consensus.

7. Conclusion

Previously the method described by Shwed and Bearman (2010) was primarily used to calibrate when known consensus was achieved (Navon and Shwed, 2012; Shwed and Bearman, 2010), to identify cases where it has not yet arrived (Shwed and Bearman, 2010), or make sense of anomalous patterns in such trajectories (Shwed, 2015). Here, we have shown how the approach can also be used as a *systematic* means for adjudicating between competing claims when consensus remains contested. The scientific community examining outcomes for children of same-sex parents has achieved consensus, and the consensus is that children of same-sex parents do not experience comparative disadvantages on important outcomes from children in other parental configurations. This provides a template for social scientific research to more directly engage with legal questions, and can provide a basis for expert opinions that rely more directly on the state of evidence in the field than upon the indirect interpretations of single expert witnesses. Furthermore, following the suggestion by Yurkewicz (2012), it is additionally reasonable to conclude that even if the opposition position was not refuted by current scientific consensus, it should not bear on decisions whether or not to recognize same-sex unions. Continuing to allow for barring them would reinforce conditions known to be associated with negative outcomes for children from *all* parental backgrounds, e.g., unstable parental partnerships and identity-based discrimination (Hatzenbuehler et al., 2010; Herek, 2006; Manning et al., 2004).

These methods extend to how social scientists engage the courts more generally. Despite recent visible efforts towards more public forms of social science, the role that social science plays within the legal system remains under-examined: This is a public sphere worthy of more attention. The courts have established criteria for social science testimony and evidence, which includes establishing the extent of consensus. The method demonstrated here provides one means of answering this question. In addition to this structural model, future research should continue to explore how this can be combined with computational and qualitative means for systematically and intensively examining the content of literatures on pertinent social science issues before the courts.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ssresearch.2015.06.008>.

References

- ABA, American Bar Association, 2006. Resolution 122b.
 Acker, James R., 1990. Social science in supreme court criminal cases and briefs: the actual and potential contribution of social scientists as Amici Curiae. *Law Hum Behav.* 14 (1), 25–42.

- APA, American Psychological Association, 2013a. American Medical Association, American Academy of Pediatrics, California Medical Association, American Psychiatric Association, American Psychoanalytic Association, American Association for Marriage and Family Therapy, National Association of Social Workers and its California Chapter and California Psychological Association. Brief of Amici Curiae on the Merits in Support of Affirmance; Hollingsworth V. Perry, Article 12-144.
- APA, American Psychological Association, 2013b. The American Academy Of Pediatrics, The American Medical Association, The American Psychiatric Association, The American Psychoanalytic Association, The California Medical Association, The National Association of Social Workers and its New York City and State Chapters and The New York State Psychological Association. Brief of Amici Curiae on the Merits in Support of Affirmance; Us V. Windsor, Article 12-307.
- ASA, American Sociological Association, 2013. Brief of Amicus Curiae in Support of Respondent Kristin M. Perry and Respondent Edith Schlain Windsor. Hollingsworth V. Perry & Us V. Windsor, Article 12-144, 12-307.
- Batagelj, Vladimir, Mrvar, Andrej, 2003. *Pajek: analysis and visualization of large networks*. In: Junger, M., Mutzel, P. (Eds.), *Graph Drawing Software, Mathematics and Visualization*. Springer.
- Bersoff, Donald N., 1987. Social science data and the supreme court: lockhart as a case in point. *Am. Psychol.* 42 (1), 52–58.
- Bielby, William T., 2003. Expert Report of William T. Bielby, Ph.D., Betty Dukes, et al. v. Wal-Mart Stores. <<http://walmartclass.com/staticdata/reports/r3.html>> (retrieved 20.04.15).
- Blondel, Vincent D., Guillaume, Jean-Loup, Lambiotte, Renaud, Lefebvre, Etienne, 2008. Fast unfolding of communities in large networks. *J. Stat. Mech.* 10, 10008–10020. <http://dx.doi.org/10.1088/1742-5468/2008/10/P10008>.
- Borgatti, Stephen P., Everett, Martin G., 2006. A graph-theoretic perspective on centrality. *Soc. Networks* 28, 466–484.
- Brandes, Ulrik, Delling, Daniel, Gaertler, Marco, Gorke, Robert, Hoefler, Martin, Nikoloski, Zoran, Wagner, Dorothea, 2007. On finding graph clusterings with maximum modularity. *Lect. Notes Comput. Sci.* 4769, 121–132.
- Brown v. Entertainment Merchants Association, 2011. U.S. Supreme Court, 564 U.S.
- Cheng, Simon, Powell, Brian, 2015. Measurement, methods, and divergent patterns: reassessing the effects of same-sex parents. *Soc. Sci. Res.* (0). <http://dx.doi.org/10.1016/j.ssresearch.2015.04.005>.
- Collins, H.M., Evans, Robert., 2002. The third wave of science studies: studies of expertise and experience. *Soc. Stud. Sci.* 32 (2), 235–296. <http://dx.doi.org/10.1177/0306312702032002003>.
- Csárdi, Gábor, Nepusz, Tamás, 2006. The Igraph software package for complex network research. *Int. J. Complex Syst.*, 1695
- Daubert v. Merrell Dow Pharmaceuticals Inc., 1993. U.S. Supreme Court. 509 U.S. 579.
- Evans, John H., 2007. Consensus and knowledge production in an academic field. *Poetics* 35 (1), 1–21. <http://dx.doi.org/10.1016/j.poetic.2007.01.001>.
- Fortunato, Santo, 2010. Community detection in graphs. *Phys. Rep.* 486, 75–174. <http://dx.doi.org/10.1016/j.physrep.2009.11.002>.
- Grunwald, Ben K., 2012. Suboptimal social science and judicial precedent. *Univ. Pennsylv. Law Rev.* 161, 1409–1443.
- Hatzenbuehler, Mark L., McLaughlin, K.A., Keyes, K.M., Hasin, D.S., 2010. The impact of institutional discrimination on psychiatric disorders in lesbian, gay and bisexual populations: a prospective study. *Am. J. Public Health* 100, 452–459.
- Herek, G.M., 2006. Legal recognition of same-sex relationships in the United States: a social science perspective. *Am. Psychol.* 61, 607–621.
- Hollingsworth v. Perry, 2013. U.S. Supreme Court 570 U.S.
- Kimball, David C., 2013. Judges are not social scientists (yet). *Elect. Law J.* 12, 324–326. <http://dx.doi.org/10.1089/elj.2013.12.38>.
- Lockhart v. McCree, 1986. U.S. Supreme Court, 476 U.S. 162.
- Manning, Wendy D., Smock, Pamela J., Majumdar, D., 2004. The relative stability of cohabiting and marital unions for children. *Popul. Res. Policy Rev.* 23, 135–159.
- Marks, Loren, 2012. Same-sex parenting and children's outcomes: a closer examination of the American Psychological Association's Brief on Lesbian and Gay Parenting. *Soc. Sci. Res.* 41 (4), 735–751. <http://dx.doi.org/10.1016/j.ssresearch.2012.03.006>.
- Mitchell, Gregory, Monahan, John, Walker, Laurens, 2011. The Asa's missed opportunity to promote sound science in court. *Sociol. Methods Res.* 40 (4), 605–620. <http://dx.doi.org/10.1177/0049124111412714>.
- Monahan, John, Walker, Laurens, 1986. Social authority: obtaining, evaluating, and establishing social science in law. *Univ. Pennsylv. Law Rev.* 134 (3), 477–517.
- Monahan, John, Walker, Laurens, 2009. *Social Science in Law: Cases and Materials*, vol. 7th ed. Foundation Press.
- Muller v. Oregon, 1908. U.S. Supreme Court 208 U.S. 412.
- NAACP v. Husted, 2014. U.S. Court of Appeals, 6th Circuit 14–3756.
- Navon, Daniel, Shwed, Uri, 2012. The chromosome 22q11.2 deletion: from the unification of biomedical fields to a new kind of genetic condition. *Soc. Sci. Med.* 75 (9), 1633–1641. <http://dx.doi.org/10.1016/j.socscimed.2012.06.023>.
- Newman, Mark E., Girvan, M., 2004. Finding and Evaluating Community Structure in Networks. *Physical Review E.* 69: no.026113.
- Newman, Mark E.J., 2006. Modularity and community structure in networks. *Proc. Natl. Acad. Sci.* 103 (23), 8577–8582.
- Nielsen, Laura Beth, Myrick, Amy, Weinberg, Jill, 2011. Siding with science: in Defense of ASA's Dukes Vs. Wal-Mart Amicus Brief. *Sociol. Methods Res.* 40 (4), 646–667. <http://dx.doi.org/10.1177/0049124111424128>.
- Oreskes, Naomi, 2004. The scientific consensus on climate change. *Science* 306 (5702), 1686. <http://dx.doi.org/10.1126/science.1103618>.
- Pafford, Abram J., et al., 2013. Amici Curiae Brief of Social Science Professors in Support of Hollingsworth and Blag. Hollingsworth V. Perry & Us V. Windsor, Article 12-144, 12-307.
- Parents Involved in Community Schools V. Seattle School District No. 1 et al., 2007. U.S. Supreme Court. 551 U.S. 701.
- Perrin, Andrew J., Cohen, Philip N., Caren, Neal, 2013. Are children of parents who had same-sex relationships disadvantaged? A scientific evaluation of the no-differences hypothesis. *J. Gay Lesbian Mental Health* 17 (3), 327–336. <http://dx.doi.org/10.1080/19359705.2013.772553>.
- Porter, Mason A., Onnela, Jukka-Pekka, Mucha, Peter J., 2009. Communities in Networks. *Notices of the AMS* 56 (9), 1082–1097, 164–166.
- Regnerus, Mark, 2012. How different are the adult children of parents who have same-sex relationships? Findings from the new family structures study. *Soc. Sci. Res.* 41 (4), 752–770. <http://dx.doi.org/10.1016/j.ssresearch.2012.03.009>.
- Sci2 Team, 2009. *Science of Science (Sci2) Tool*. Bloomington, IN: Indiana University and SciTech Strategies.
- Shelby County V. Holder, 2013. U.S. Supreme Court. 570 U.S.
- Shwed, Uri., 2015. Robust science: passive smoking and scientific collaboration with the tobacco industry in the 1970s. *Sociol. Sci.* 2, 158–185.
- Shwed, Uri, Bearman, Peter S., 2010. The temporal structure of scientific consensus formation. *Am. Sociol. Rev.* 75 (6), 817–840.
- Siegel, Benjamin S., Perrin, Ellen C., et alCommittee on psychosocial aspects of child and family health, 2013. Promoting the well-being of children whose parents are Gay or Lesbian. *Pediatrics* 131 (4), 827–830. <http://dx.doi.org/10.1542/peds.2013-0376>.
- Sørensen, Jesper B., Sharkey, Amanda J., 2011. The perils of false certainty: a comment on the Asa Amicus Brief in Dukes Vs. Wal-Mart. *Sociol. Methods Res.* 40 (4), 635–645. <http://dx.doi.org/10.1177/0049124111424126>.
- Tomaskovic-Devey, Donald, 2011. The politics and practice of sociology in the courts. *Sociol. Methods Res.* 40 (4), 621–634. <http://dx.doi.org/10.1177/0049124111422797>.
- Uggen, Christopher, Inderbitzin, Michelle, 2010. Public criminologies. *Criminol. Public Policy* 9, 725–750.
- United States V. Windsor, 2013. U.S. Supreme Court 570 U.S.
- US Congress, 1996. Defense of Marriage Act, Article HR 3396.
- US Congress, 2010. Don't Ask, Don't Tell Repeal Act, Article HR 2965 & S.4023.
- Wal-Mart V. Dukes, 2011. U.S. Supreme Court 564 U.S.
- Wechsler, Hayley J., Kehn, Andre, Wise, Richard A., Cramer, Robert J., 2015. Attorney beliefs concerning scientific evidence and expert witness credibility. *Int. J. Law Psychiatry, forthcoming*. <http://dx.doi.org/10.1016/j.ijlp.2015.03.008>.

Witherspoon v. Illinois, 1968. U.S. Supreme Court. 391 U.S. 510.

Yan, Erija, Ding, Ying, 2012. Scholarly network similarities: how bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks and cword networks relate to each other. *J. Am. Soc. Inform. Sci. Technol.* 63 (7), 1313–1326.

Yurkewicz, Ilana, 2012. Why Mark Regnerus' Study Shouldn't Matter, Even If It Were the Most Scientifically Robust Study in the World. *Scientific American*, June 16.

Supplement for “Scientific Consensus, the Law, and Same Sex Parenting Outcomes”

Figure S1. Median Citation Age-based Moving Window Widths

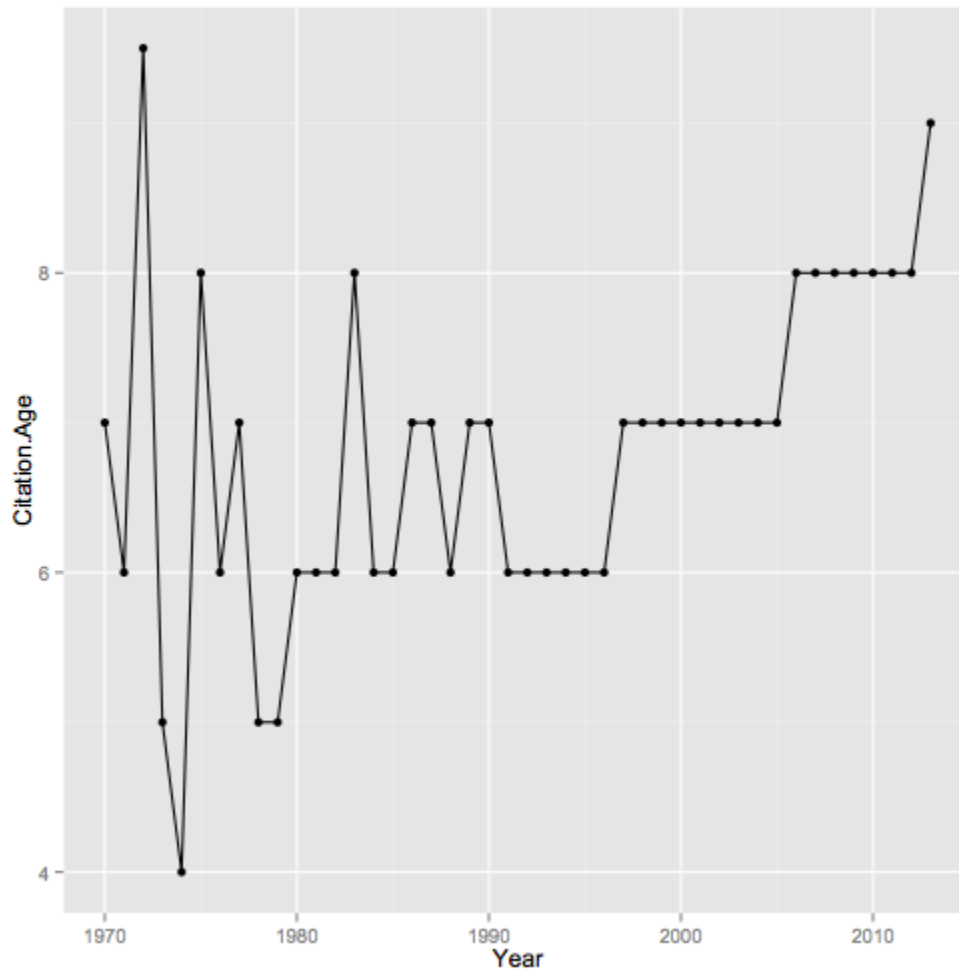
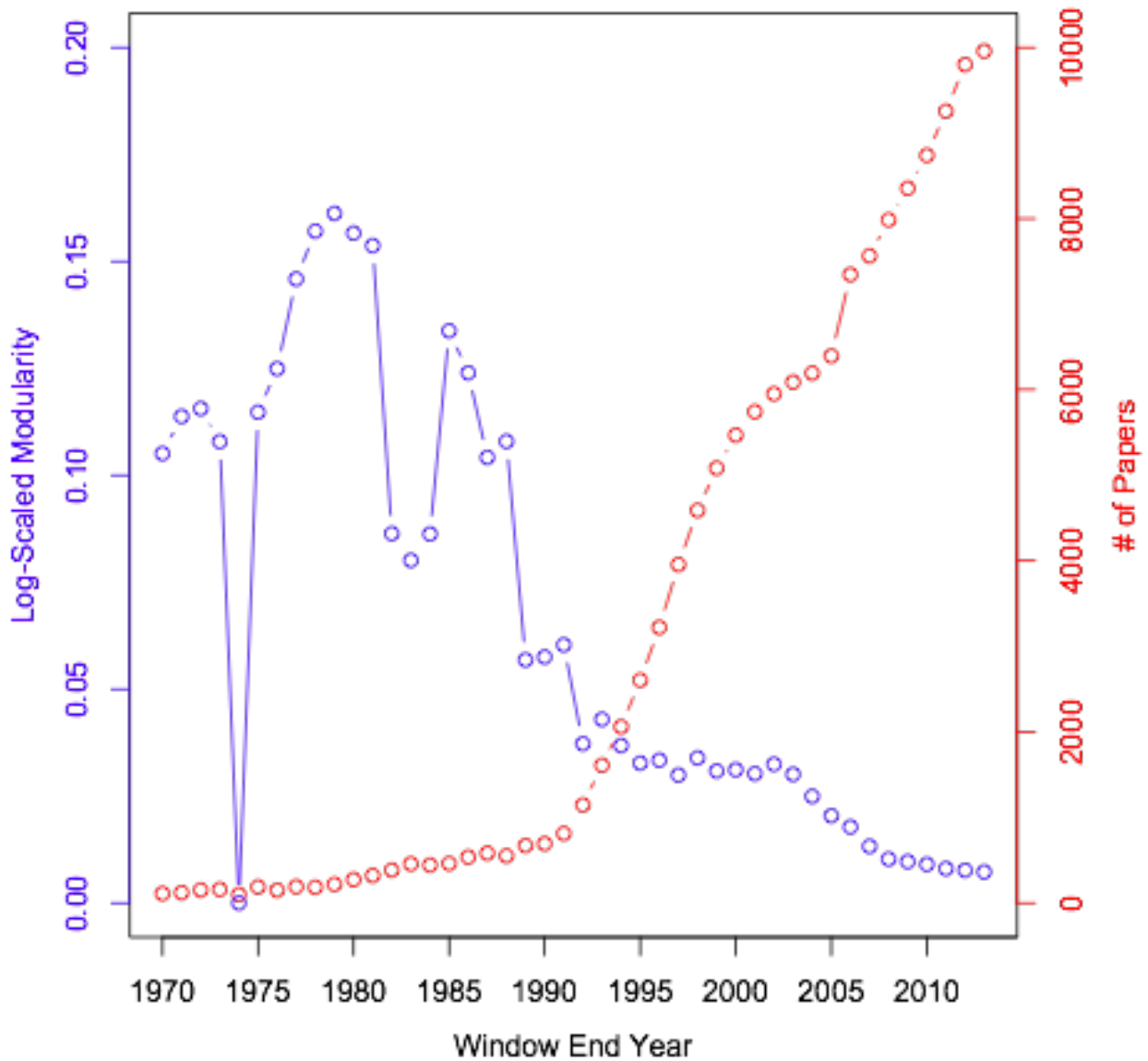
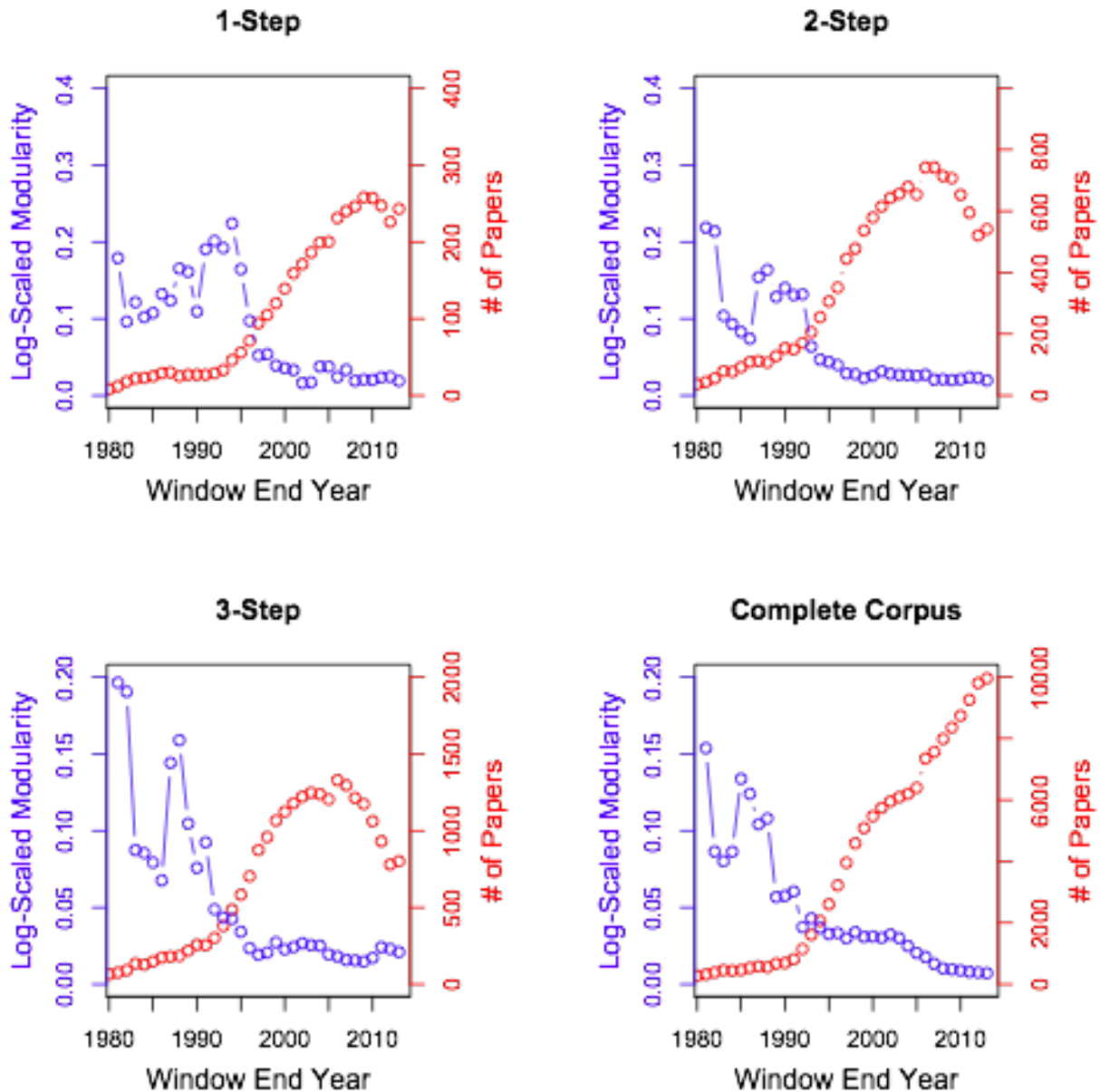


Fig S2. Complete Modularity and Corpus Size Changes 1970-2013.



NOTE: The x-axis denotes the end-year of the moving window for each epistemic period in the corpus. The y-axis presents the changes across those windows in: modularity score (blue/solid line, left axis), and growth in number of publications (red/hatched line, right axis).

Fig S3. Comparing Modularity Corpus Size: Complete Corpus & Extracted Network Subsets



NOTE: Within each panel, the x-axis denotes the end-year of the moving window for each epistemic period in the corpus. The y-axes presents the changes across those windows in: modularity score (blue/solid line, left axis), and growth in number of publications (red/hatched line, right axis). Panel D reproduces Figure 1 for the paper, while Panels A-C respectively provide the comparable information for the subset networks that are defined by distance in sent citations from the core set.