

## On Social Network Position in Employment Law: Conjectures for Charlie

*Sachin S. Pandya\**

*This paper, part of a Festschrift for Charles A. Sullivan, shows how arguments from two of Sullivan’s papers on employment law would fare in a world in which employers can easily see a worker’s or job applicant’s relative position within a social or professional network. The paper then uses Sullivan’s corpus of legal scholarship to illustrate some challenges to using social network evidence in employment law.*

I. INTRODUCTION .....	1427
II. BACKGROUND .....	1428
A. Social Networks and Hiring.....	1428
B. Measures of Centrality .....	1430
III. DISCUSSION.....	1432
A. Hiring Algorithms.....	1432
B. Partial Enforcement and Non-Competes.....	1435
C. Estimating Charlie’s Network Position.....	1436
IV. CONCLUSION.....	1444
V. COLOPHON .....	1444

### I. INTRODUCTION

What if employers could see any worker or job applicant’s relative position within a professional network? In this Festschrift essay, I consider how legal arguments in two papers by Charles A. Sullivan would fare in the counterfactual world in which anyone’s professional network position is visible to employers and thus available as evidence in employment law litigation. Then, I use Sullivan’s corpus of legal scholarship to illustrate some challenges to collecting and using such evidence of network position.

---

\* Professor of Law, University of Connecticut ( <http://orcid.org/0000-0001-7387-1307>).  
© 2020 Sachin S. Pandya. This work is available after publication under a CC-BY-NC-ND 4.0 license. For comments on a prior draft, thanks to Charles A. Sullivan.

## II. BACKGROUND

### A. *Social Networks and Hiring*

Since the mid-20th century, researchers have studied how and how much a worker's odds of getting hired or promoted or of quitting depend on their relative position within a social network—the professional, friendship, familial, or other ties between individuals or other actors in some population of interest. In hiring, for example, employers tend to place more trust in applicants referred to them by their current employees or others they already know. Many employers also hire someone in part based on what they think an applicant's connections can do for the firm—more business, a better reputation, or better information flow about industry innovations.<sup>1</sup>

One worry: Because people tend to connect with other people like themselves (homophily),<sup>2</sup> when employers rely on social networks to hire, that tends to benefit those who are demographically or otherwise similar to those doing the hiring, and that may increase segregation or limit mobility within a firm, occupation, or industry based on race, gender, or other worker characteristics.<sup>3</sup> In the U.S., courts and commentators have discussed homophily in the context of employment discrimination lawsuits against employers for using word-of-mouth recruiting.<sup>4</sup>

In recent years, with the rise of online social media platforms (*e.g.*, Facebook, LinkedIn, Twitter), employers have used social media presence not only to vet job applicants,<sup>5</sup> but also to judge worker competence based

---

<sup>1</sup> See generally Lori Beaman, *Social Networks and the Labor Market*, in THE OXFORD HANDBOOK OF THE ECONOMICS OF NETWORKS (Yann Bramoullé, Andrea Galeotti, & Brian Rogers eds., 2016); Emilio J. Castilla, George J. Lan & Ben A. Rissing, *Social Networks and Employment: Outcomes (Part 2)*, 7 SOC. COMPASS 1013 (2013); Matthew O. Jackson, Brian W. Rogers & Yves Zenou, *The Economic Consequences of Social-Network Structure*, 55 J. ECON. LIT. 49 (2017); Jinhwan Jo & Jill E. Ellingson, *Social Relationships and Turnover: A Multidisciplinary Review and Integration*, 44 GROUP & ORG. MGMT. 247 (2019).

<sup>2</sup> Miller McPherson, Lynn Smith-Lovin & James M. Cook, *Birds of a Feather: Homophily in Social Networks*, 27 ANN. REV. SOC. 415 (2001); Jackson, Rogers, & Zenou, *supra* note 1, at 56–58.

<sup>3</sup> See Fabiana Silva, *The Strength of Whites' Ties: How Employers Reward the Referrals of Black and White Jobseekers*, 97 SOC. FORCES 741 (2018); David B. Bills, Valentina Di Stasio & Klarita Gërkhani, *The Demand Side of Hiring: Employers in the Labor Market*, 43 ANN. REV. SOC. 291, 298–99 (2017); Aaron Clauset, Samuel Arbesman & Daniel B. Larremore, *Systematic Inequality and Hierarchy in Faculty Hiring Networks*, 1 SCI. ADVANCES 1 (2015); Natalie Wreyford, *Birds of a Feather: Informal Recruitment Practices and Gendered Outcomes for Screenwriting Work in the UK Film Industry*, 63 SOC. REV. 84 (2015); Roberto M. Fernandez & Isabel Fernandez-Mateo, *Networks, Race, and Hiring*, 71 AM. SOC. REV. 42 (2006).

<sup>4</sup> See BARBARA LINDEMANN, PAUL GROSSMAN & GEOFF WEIRICH, EMPLOYMENT DISCRIMINATION LAW § 16.II.C (6th ed. 2020) (collecting cases).

<sup>5</sup> See Matthieu Manant, Serge Pajak & Nicolas Soulié, *Can Social Media Lead to Labor*

on their online network itself. Here is how a Swedish recruiter put it in a recent study:

Well, in some positions it may be very important to have a network in order to not get stuck in a rut. Because it also shows, if you have many business contacts or . . . . Do you have a large network on LinkedIn? Because you *must* be there [in LinkedIn] today, if you're working in the positions that I am recruiting for. Are you? 'Yes' Ok, 'check'. Then it's, do you have a large network showing that you're actually networking? Because I'm always asking, 'What networks do you have?' Because if you're an engineer without being active in networks, and don't really have any, how do you develop?<sup>6</sup>

Here, the recruiter focuses on the *size* of the applicant's network as it appears on the LinkedIn platform and infers from it how likely the job applicant will keep up with recent developments in the field.

While employers may glimpse the size or other features of the job applicant's *own* professional network, it is harder to see the applicant's overall network position within a profession, industry, or other population of interest. Platforms such as LinkedIn or Twitter allow users to export their own data<sup>7</sup> and sometimes notify those users about connections in common. But these online platforms do not provide users a way to see the network structure of everyone on the platform. They have offered researchers more extensive data on platform users, but only in anonymized form.<sup>8</sup> Moreover, online platform participation among users in the U.S. varies considerably by platform and is not representative of the overall population.<sup>9</sup> Nonetheless, some have studied social media activity on these platforms to approximate social networks within a profession or industry.<sup>10</sup>

---

*Market Discrimination? Evidence from a Field Experiment*, 28 J. ECON. & MGMT. STRATEGY 225 (2019); Yoram Bachrach, *Human Judgments in Hiring Decisions Based on Online Social Network Profiles*, in 2015 IEEE INTERNATIONAL CONFERENCE ON DATA SCIENCE AND ADVANCED ANALYTICS (DSAA) 1 (2015); Brenda L. Berkelaar & Patrice M. Buzzanell, *Online Employment Screening and Digital Career Capital: Exploring Employers' Use of Online Information for Personnel Selection*, 29 MGMT. COMM. Q. 84 (2015).

<sup>6</sup> Anna Hedenus, Christel Backman & Peter Håkansson, *Whom Do You Know? Recruiters' Motives for Assessing Jobseekers' Online Networks*, INT'L J. HUM. RESOURCE MGMT. 1, 15 (2019).

<sup>7</sup> E.g., *Exporting Connections from LinkedIn*, LINKEDIN, <https://www.linkedin.com/help/linkedin/answer/66844> (last visited, April 1, 2020).

<sup>8</sup> Jeremy Kahn, *LinkedIn Will Allow Economics Researchers to Mine Its Data*, BLOOMBERG NEWS (Aug. 20, 2018), <https://www.bloomberg.com/news/articles/2018-08-20/linkedin-will-allow-economics-researchers-to-mine-its-data>.

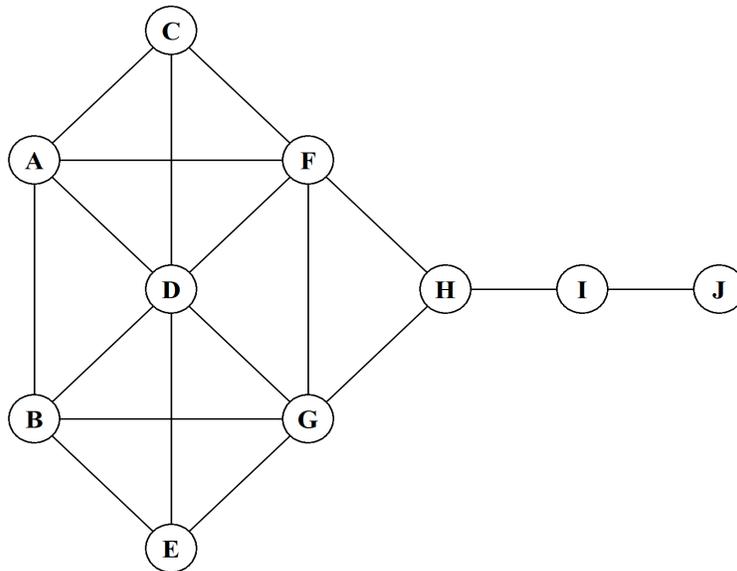
<sup>9</sup> Eszter Hargittai, *Potential Biases in Big Data: Omitted Voices on Social Media*, 38 SOC. SCI. COMPUTER REV. 10, 16 (2018).

<sup>10</sup> E.g., Folker Hanusch & Daniel Nölleke, *Journalistic Homophily on Social Media*, 7

### B. Measures of Centrality

If employers could fully see a job applicant's ties to others within a population of interest, *i.e.* their professional network, the next step would be for them to choose how to measure that worker's importance (their relative position) in that network. Here, we briefly discuss the intuitions behind three standard network *centrality* measures (among many) and how employers might use them to assess network position.<sup>11</sup> The point here is simple: just as "importance" depends on context, choosing which measure of network position to use depends on why the employer wants to measure network position in the first place.

Figure 1: Hypothetical Social Network



To illustrate, Figure 1 depicts the fictional social network of ten individuals presented in Krackhardt (1990).<sup>12</sup> Each circle ("node") represents an individual (A through J), and each line between nodes represents a social connection between them—here whether or not they regularly interacted. The lack of a line between nodes denotes no social connection between them.

---

DIGITAL JOURNALISM 22 (2019).

<sup>11</sup> For details, see, for example, Stephen P. Borgatti & Daniel J. Brass, *Centrality: Concepts and Measures*, in *SOCIAL NETWORKS AT WORK* (Daniel J. Brass & Stephen P. Borgatti eds., 2019).

<sup>12</sup> David Krackhardt, *Assessing the Political Landscape: Structure, Cognition, and Power in Organizations*, 35 *ADMIN. SCI. Q.* 342, 351 (1990).

2020]

## CONJECTURES FOR CHARLIE

1431

Any individual's centrality (importance) in this network depends on the measure. Here, we consider three centrality measures: degree, closeness, and betweenness.

Table 1: Centrality Scores for Figure 1

id	degree	closeness	betweenness
A	4	0.0588235	0.8333333
B	4	0.0588235	0.8333333
C	3	0.0555556	0.0000000
D	6	0.0666667	3.6666667
E	3	0.0555556	0.0000000
F	5	0.0714286	8.3333333
G	5	0.0714286	8.3333333
H	3	0.0666667	14.0000000
I	2	0.0476190	8.0000000
J	1	0.0344828	0.0000000

*Degree centrality* is a tally of the number of connections for each node. In Figure 1, individual *D* has the highest degree (6) and individual *J* has the lowest (1). This measure matters if, for example, the employer believes that an employee with the most connections with other employees at a firm is also the least likely of them to quit.

*Closeness centrality* is the reciprocal of the average length of the shortest paths to and from all the other nodes in the network. It captures the ability of any one person to reach anyone else in the network in the shortest number of connections. By this measure, *F* and *G* are the most important individuals in this network, because it takes the fewest connections on average to get from *F* or *G* to any other node in this network.

To show how this works, here is how to calculate closeness for *F*. Start by counting the number of connections in the shortest path from *F* to each of the other nine nodes in the network. For *F*'s shortest paths, it takes one connection to get from *F* to each of five nodes ( $= \{A, C, D, G, H\}$ ), two connections to each of three other nodes ( $= \{B, E, I\}$ ), and three connections to get from *F* to *J*. Finally, take the sum of these connections (14), divide by the total number of nodes (10) in the network, and (because we want higher scores to reflect the shorter distances) take its reciprocal. The result: *F*'s closeness score ( $= 10/14 \cong 0.7142857$ ). If we did this for all the nodes, we would get closeness scores for each (Table 1).

For employers, this measure of network position might be useful if the network depicts connections within an industry and the employer wants to

assess how likely the worker is to receive some information (*e.g.*, an industry innovation). Here, closeness captures who is most likely to receive such information, on average, if first sent from any other person in that network.

*Betweenness centrality* refers to the number of shortest paths passing through any particular node that connects it to otherwise disconnected subnetworks. Individuals with high betweenness might be attractive to employers because they can function as bridges or brokers between subgroups of people who otherwise are relatively isolated from each other and may themselves be more likely to have better ideas as a result.<sup>13</sup> Based on this measure, *H* is the most important actor because *H* connects otherwise isolated *I* and *J* to everyone else in the network.

### III. DISCUSSION

This section considers whether the legal arguments in two of Charlie's papers, if otherwise sound, would remain sound in the counterfactual world in which a worker's relative social network position (*i.e.*, that worker's centrality) is completely visible to any and all employers and thus available as evidence in employment law litigation. Then, we use Charlie's corpus of scholarship to illustrate the main current hurdle to employer use of worker network position: missing data.

#### A. Hiring Algorithms

In his 2018 paper on hiring algorithms, Charlie argued that the traditional Title VII "disparate treatment" and "disparate impact" legal frames fail when applied to hiring algorithms because, in the hard case, those frames let employers escape Title VII liability for practices that Title VII would prohibit if undertaken by humans alone.<sup>14</sup>

In general, machine-learning algorithms for hiring predict—based on a set of observable applicant attributes—the values for some target variable of interest that no one yet knows (*e.g.*, whether that applicant, if hired, will meet a quarterly sales quota). To do this, the algorithm first takes historical data on past employees. That data not only contain information on their attributes (*e.g.*, gender, race, education, experience), but also the values for those employees on the target variable of interest. The algorithm then "trains" on that data, *i.e.*, it calculates a function of the information provided (a model) that best predicts the values of the target variable in the historical data. Then, the algorithm uses that model on new job applicants to estimate the probability for each value in the target variable. Above a certain fixed probability threshold, the algorithm recommends hiring; below it, it

---

<sup>13</sup> See Ronald S. Burt, *Structural Holes and Good Ideas*, 110 AM. J. SOC. 349 (2004).

<sup>14</sup> Charles A. Sullivan, *Employing AI*, 63 VILL. L. REV. 395 (2018).

recommends rejecting the applicant.

In his paper, Charlie focused on the conceptual hard case for Title VII doctrine: The employer or its human agents always do what the algorithm recommends; they do not know in advance (of any litigation) the model that the algorithm will adopt; they train the algorithm on historical data that is itself unbiased and representative; and computational power is unlimited. Accordingly, set aside cases of employers who *intend* to use the algorithm to discriminate.<sup>15</sup> Also, set aside whether an employer's decision to add fairness criteria into the algorithm itself exposes the employer to Title VII liability if such criteria expressly accounts for race, gender, or other protected characteristics.<sup>16</sup>

Now, suppose the hard case in which the algorithm, once deployed, adopts a model that always recommends, for example, rejecting female job applicants of childbearing age. What result?

First, the disparate-treatment frame does not apply, Charlie argued. Even if we learn, upon inspection, that the algorithm assigned non-zero weight to job applicants' protected characteristics (*e.g.*, applicant "sex"), the algorithm itself cannot have the requisite illegal intent or motives, and neither can the employer or its human agents, because they could not know the algorithm's model in advance.<sup>17</sup> Thus, the plaintiff's lawyer is better off abandoning the disparate-treatment frame and arguing instead that the employer, by using the algorithm, violated § 703(a)(2) of Title VII for "classify[ing]" her in a "way which would deprive or tend to deprive any individual of employment opportunities . . . because of such individual's . . . sex."<sup>18</sup>

Second, the disparate-impact frame may not apply if the algorithm's model counts as facially discriminatory. Some courts today still take disparate-impact liability to require the challenged employment practice to be facially-neutral,<sup>19</sup> even though § 703(k), which Congress added in 1991, does not expressly require that.<sup>20</sup> That matters for Charlie because, while using the algorithm does not count as "intentional discrimination" within the

---

<sup>15</sup> See Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 692–93 (2016).

<sup>16</sup> See generally Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857 (2017); Barocas & Selbst, *supra* note 15; Sam Corbett-Davies & Sharad Goel, *The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning* (2018), <https://arxiv.org/pdf/1808.00023>.

<sup>17</sup> Sullivan, *supra* note 14, at 405–06; see also Barocas & Selbst, *supra* note 15, at 699–700.

<sup>18</sup> 42 U.S.C. § 2000e-2(a)(2) (2018). See Sullivan, *supra* note 14, at 407–08; see also Kim, *supra* note 16, at 910–16.

<sup>19</sup> See LINDEMANN & GROSSMAN, *supra* note 4, at § 3.1 n.18–19 (collecting cases).

<sup>20</sup> Sullivan, *supra* note 14, at 410–11 n.56.

meaning of § 703(k),<sup>21</sup> it arguably counts as a facially-discriminatory practice once the algorithm assigns non-zero weight to protected applicant attributes.

Third, suppose the algorithm assigns zero weight to any protected characteristics and *still* rejects all female job applicants of childbearing age, because it weights other factors that strongly co-vary with protected characteristics. If so, employers can escape liability by prevailing on its defense to disparate-impact liability that relying on the hiring algorithm is “job related for the position in question and consistent with business necessity.”<sup>22</sup> On this point, Charlie argued that, assuming the target variable is itself “job-related,” the employer will likely prevail on the business necessity defense if the disparate impact is based on actual differences in the target variable (*e.g.*, productivity, job tenure) between races or genders.<sup>23</sup>

Fourth, because hiring algorithms are complex and only as effective as the historical data on which they train, in the hard case, plaintiffs will have a hard time defeating a business-necessity defense by showing “an alternative employment practice” that produces less disparity but that the employer refuses to adopt.<sup>24</sup> Charlie also asserted that the requisite employer refusal-to-adopt in effect limits this argument to cases where the plaintiff “learns enough about the disputed practice to serve something very much like a demand letter on the employer, which then fails to adopt the proffered alternative.”<sup>25</sup>

If otherwise sound, do these arguments hold in the counterfactual world in which hiring algorithms can train on, among other things, data about worker network position? Answer: yes.

First, in the hard case, the algorithm still lacks intent or motive, and the disparate-treatment frame fails, because even if the algorithm assigns weight to applicant network position, it is a facially neutral practice in any case.

Second, if the algorithm still rejects all female job applicants of childbearing age, in the hard case, employers can still escape Title VII disparate impact liability by showing that the model that assigns weight to network position is “job-related” and “consistent with business necessity.” This depends on how well network position (alone or in combination with other worker attributes) predicts the target variable (*e.g.*, a salesperson’s degree-centrality as a predictor of annual sales). Indeed, because hiring algorithms are focused on prediction, not identifying causal mechanisms, we

---

<sup>21</sup> See 42 U.S.C. § 2000e-2(k)(2).

<sup>22</sup> 42 U.S.C. § 2000e-2(k)(1)(A)(i).

<sup>23</sup> Sullivan, *supra* note 14, at 411, 423–427; see also Barocas & Selbst, *supra* note 15, at 706–12; Kim, *supra* note 16, at 866–67, 905–09.

<sup>24</sup> 42 U.S.C. § 2000e-2(k)(1)(A)(ii).

<sup>25</sup> Sullivan, *supra* note 14, at 428 (footnote omitted).

can avoid debating which network centrality measure is most appropriate and instead have the algorithm train on all of them.

### B. *Partial Enforcement and Non-Competes*

In a 2009 paper, Charlie criticized the law that lets judges enforce partially invalid non-compete and other contract clauses.<sup>26</sup> When faced with partly-invalid contracts, some courts edit out or interpret away the clause's objectionable parts and enforce what remains in a way to preserve what is left of the private bargain. As a result, judges can modify contracts in this way in some states for employee non-compete agreements.<sup>27</sup>

The problem, Charlie argues, is that this judicial practice makes employers more likely to use, or keep on using, partly invalid non-compete clauses. In any case, the employer benefits if a worker erroneously believes that an invalid non-compete clause is enforceable and acts accordingly. If, however, a worker challenges the clause in court and wins, then because of partial enforcement, the employer has less to lose than it otherwise would have, because the court will still enforce the valid parts of the clause. This gives employers an incentive to continue using partly-invalid non-compete clauses. Charlie's argument implies that, all else equal, we should expect fewer partly-invalid non-competes in states that bar partial enforcement than in states that permit partial enforcement.<sup>28</sup> Canvassing the competing arguments, Charlie concludes that, given a partially-invalid non-compete, the courts should refuse to enforce the entire non-compete, especially if the employer uses that clause for other workers, unless the invalid part is minor and unintentional.<sup>29</sup>

If otherwise sound, does this conclusion remain as sound if we could submit evidence of a worker's network position when challenging a non-compete? It depends on how much less likely employers would use partly invalid non-competes, given the availability of such evidence.

To see why, consider a rival possible cause of partly invalid non-competes: legal uncertainty about what counts as an invalid non-complete in the first place. State common-law typically requires that non-competes impose "reasonable" restrictions in relation to an employer's "protectable"

---

<sup>26</sup> Charles A. Sullivan, *The Puzzling Persistence of Unenforceable Contract Terms*, 70 OHIO ST. L. J. 1127 (2009).

<sup>27</sup> DONALD J. ASPELUND & JOAN E. BECKNER, EMPLOYEE NONCOMPETITION LAW § 8.1 (2019); Ferdinand S. Tinio, Annotation, *Enforceability, Insofar as Restrictions Would Be Reasonable, of Contract Containing Unreasonable Restrictions on Competition*, 61 A.L.R.3d 397, §§ 16[a]–21[d], 30[a]–34[b], 37–41 (1975 & Supp. 2020).

<sup>28</sup> Compare WIS. STAT. § 103.465 (2019), with TEX. CODE BUS. & COM. CODE ANN. § 15.51(c) (West 2019).

<sup>29</sup> Sullivan, *supra* note 26, at 1175–76.

interests.<sup>30</sup> What counts as a “reasonable” restriction? In published court opinions, judges discuss many factors, typically treat none as dispositive, and write opinions that are often opaque as to how they weighed those factors.<sup>31</sup> If this makes it hard to predict what a court would find invalid in the first place, it may also thereby make it more likely that employers use all but obviously invalid non-competes as boilerplate.

Suppose we supplemented the “reasonable” restriction standard by providing that for anyone with a network position below a preset closeness-centrality score, the employer bears the burden under a more stringent standard of proof (*e.g.*, “clear and convincing”) that the challenged non-compete restriction is a “reasonable” one. In the fictional network in Figure 1, given a pre-set closeness-centrality threshold of 0.05, individuals *I* and *J* would benefit from the proposed rule. Here, the premise is that closeness centrality captures how likely each node receives information flowing through that network, such as information about new jobs within that profession. Thus, the challenged restriction is *prima facie* unreasonable, because it sufficiently reduces the odds that individuals below the threshold will get a suitable new job.

If we adopted this proposed change, Charlie’s argument against partial enforcement remains sound depending on how much partial enforcement would affect the incidence of partly invalid non-competes, *given* that change. For example, if the proposed change itself reduced the incidence of partly-invalid non-competes to virtually zero, partial enforcement may be far more tolerable to leave in place for the (now rare) case where a non-compete, though declared invalid, was likely valid based on what the parties knew or could have known *ex ante*. Conversely, if the proposal has zero effect on the incidence of invalid non-competes, Charlie’s argument remains as sound as it ever was.

### C. *Estimating Charlie’s Network Position*

For now, we are far from the counterfactual world where employers can easily see worker network position. This section identifies the main current hurdle to employer use of worker network position: missing data.<sup>32</sup> To illustrate, I show what happened when I attempted on my own to discover Charlie’s network position.

Any measure of social network position depends on how we define the

---

<sup>30</sup> See generally COVENANTS NOT TO COMPETE: A STATE-BY-STATE SURVEY (Brian Malsberger ed., 12th ed. 2018).

<sup>31</sup> ASPELUND & BECKNER, *supra* note 27, at §§ 6.8–6.45 (collecting cases).

<sup>32</sup> See generally M. E. J. Newman, *Estimating Network Structure from Unreliable Measurements*, 98 PHYSICAL REV. E 1 (2018); Gueorgi Kossinets, *Effects of Missing Data in Social Networks*, 28 SOC. NETWORKS 247 (2006).

population of interest, as well as how much we know about the actors in that population and how they are connected, if at all, with each other. The population of interest is always a subset of everyone who is or has been on planet Earth. For example, we have assumed that employers care about a worker’s network position solely with respect to a “professional” network—other people in the same occupation or industry. Accordingly, in measuring network position, one must exclude some nodes and the connections to and from them, either in defining the population of interest (boundary-specification),<sup>33</sup> or because one does not have any data about them. The worry is that any connections to and from excluded nodes non-negligibly affect a measure of network position. If, for example, we excluded *I* and *J* from the network in Figure 1, because they belonged to a different “profession,” we would not see how *H* is important on a betweenness-centrality measure, and thus well-positioned to serve as a broker or bridge between professions.

To illustrate, consider my attempt to discover Charlie’s network position on my own and unbeknownst to him, much as an employer might do the same prior to recruiting. I started first by defining the population of interest as “law professors in the U.S.” In academic year 2018, law schools in the United States ( $n = 203$ ) employed 9,663 full-time faculty members.<sup>34</sup> To figure out Charlie’s network position within this population, I started first online on the premise that electronic traces of Charlie’s connections would help roughly approximate his professional connections. Table 2 reports four online identifiers I found for Charlie.

Table 2: Online Identifiers for Charles A. Sullivan, 2020

Platform	ID
work email	charles.sullivan@shu.edu
LinkedIn	charles-sullivan-4003464
Facebook	charles.a.sullivan.75
Twitter	cas1234567

Each of these identifiers, if traceable, might have provided me with some data of his ties to other U.S. law professors. For example, his incoming and outgoing email would indicate who he emails, who emails him, and how often they communicate, much in the way researchers have approximated

<sup>33</sup> Kossinets, *supra* note 32; Peter V. Marsden, *Network Data and Measurement*, 16 ANN. REV. SOC. 435, 439–40 (1990).

<sup>34</sup> *DataSet Faculty (Academic Year 2018)*, ACCESSLEX, <http://analytix.accesslex.org/DataSet> (last visited Apr. 9, 2020).

social networks with data on emails between people within a firm.<sup>35</sup> His current employer (Seton Hall University) has this information on its servers, albeit only for emails sent or received via a Seton Hall email account. In contrast, online social media platforms can better capture cross-firm social networks, but this depends on how often law professors use these platforms to communicate to each other.

None of this data, however, is publicly available, and my best guess was that Seton Hall University, Facebook, LinkedIn, and Twitter were not going to hand them over. Moreover, even if they had let me see only Charlie's communications on these platforms, that would be an incomplete picture of the complete network of law professors, and thus his true network position among law professors. For example, it would completely exclude subnetworks of law professors with whom Charlie had no connection.

To be sure, some researchers have exploited industry and occupational norms to estimate professional networks independent of platform. For example, researchers have used co-authorship in the corpus of published scientific papers to estimate the networks between researchers within and across different scientific fields.<sup>36</sup> Despite similar studies of law professors,<sup>37</sup> co-authorship is rare in legal academia.<sup>38</sup>

Still, inspired by another of Charlie's papers,<sup>39</sup> I resorted to approximating a subset of Charlie's professional network by tracing incoming and outgoing expressions of gratitude that appeared in the typical acknowledgements footnote in his and other law professors' published legal scholarship.<sup>40</sup> To do this, we (myself and student research assistants) first obtained all of Charlie's published scholarly papers up through 2018 from the Hein Online database. For each paper, we identified all the professors he thanked, if any, in the traditional acknowledgements footnote. Where Charlie had a co-author, we assumed co-authors thanked each other, and we treated paper and book co-authors alike. Then, we searched the usual law databases for papers in which the author(s) thanked Charlie. For each of

---

<sup>35</sup> E.g., Amir Goldberg et al., *Fitting in or Standing out? The Tradeoffs of Structural and Cultural Embeddedness*, 81 AM. SOC. REV. 1190 (2016).

<sup>36</sup> E.g., M.E.J. Newman, *Coauthorship Networks and Patterns of Scientific Collaboration*, 101 PROC. NAT'L ACAD. SCI. 5200 (2004).

<sup>37</sup> E.g., Paul H. Edelman & Tracey E. George, *Six Degrees of Cass Sunstein: Collaboration Networks in Legal Scholarship*, 11 GREEN BAG 2D (2007).

<sup>38</sup> Michael I. Meyerson, *Law School Culture and the Lost Art of Collaboration: Why Don't Law Professors Play Well with Others?*, 93 NEB. L. REV. 547 (2015).

<sup>39</sup> Charles A. Sullivan, *The Under-Theorized Asterisk Footnote*, 93 GEO. L.J. 1093 (2005).

<sup>40</sup> See Jonathan I. Tietz & W. Nicholson Price II, *Acknowledgements as a Window into Legal Academia*, 98 WASH. U. L. REV. (forthcoming 2020), <https://ssrn.com/abstract=3552673> (reporting results of automated textual analysis of acknowledgement footnotes in 29,024 articles published from 2008 to 2017 in 183 law reviews in the U.S.).

these papers, we also identified the other law professors they thanked.

This method, however, approximates not the network of all law professors in the U.S., but only the subset of them connected to Charlie through at least one other law professor. To use the same method to approximate Charlie's network position among *all* law professors in the U.S., we would have had to collect and code all the (incoming and outgoing) thanks in all published papers by all law professors in the U.S. during the time of Charlie's career thus far, including those professors that neither thanked Charlie nor received thanks from Charlie.

This was infeasible, largely because we chose to code by hand to deal with name disambiguation. Typically, bibliometric researchers face the problem of author-name disambiguation in cases where different name strings—arising from typical variations in spelling (*e.g.*, “Charles A. Sullivan,” “Charles Sullivan,” “Charlie Sullivan”), nicknames, and data-entry errors—all refer to the same person, as well in cases as where the same string refers to two or more different people. These disambiguation problems also apply to the names of people thanked in the acknowledgements footnote. We also had to figure out whether the persons thanked were professors or not (*e.g.*, students, support staff) when the footnote text did not so indicate. To be sure, unlike most professions, law professors in the U.S. are specifically identified by name and other biographical information in the American Association of Law Schools' *Directory of Law Teachers*, an annual listing of law professors. Unfortunately, *Directory* volumes are not currently publicly available in a machine-readable format. Thus, we resorted to coding by hand and using both the source material, Internet search engines, and the *Directory* in an ad hoc manner to resolve ambiguities.

The result is a dataset that approximates the network of law professors connected to Charlie (either directly or indirectly through one other law professor) any some point during Charlie's academic career. For ease of exposition, we depict this network as four separate networks in four time periods: 1975–1995, 1996–2005, 2006–2010, 2011–2018 (Figures 2, 3, 4, 5, respectively). Each arrow from a node indicates an outgoing expression of gratitude, and each arrow to a node indicates an incoming expression of gratitude.



2020]

CONJECTURES FOR CHARLIE

1441

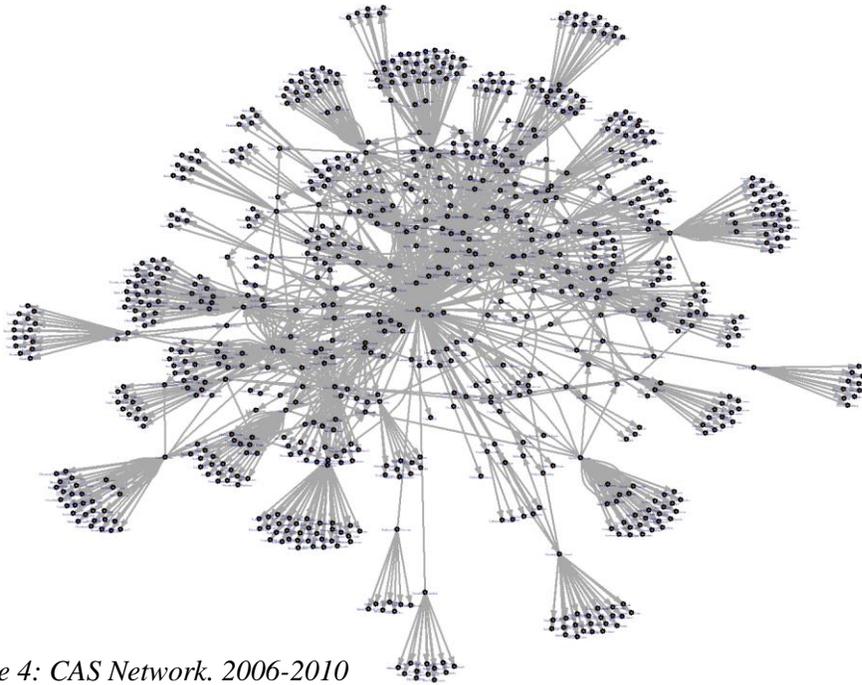


Figure 4: CAS Network, 2006-2010

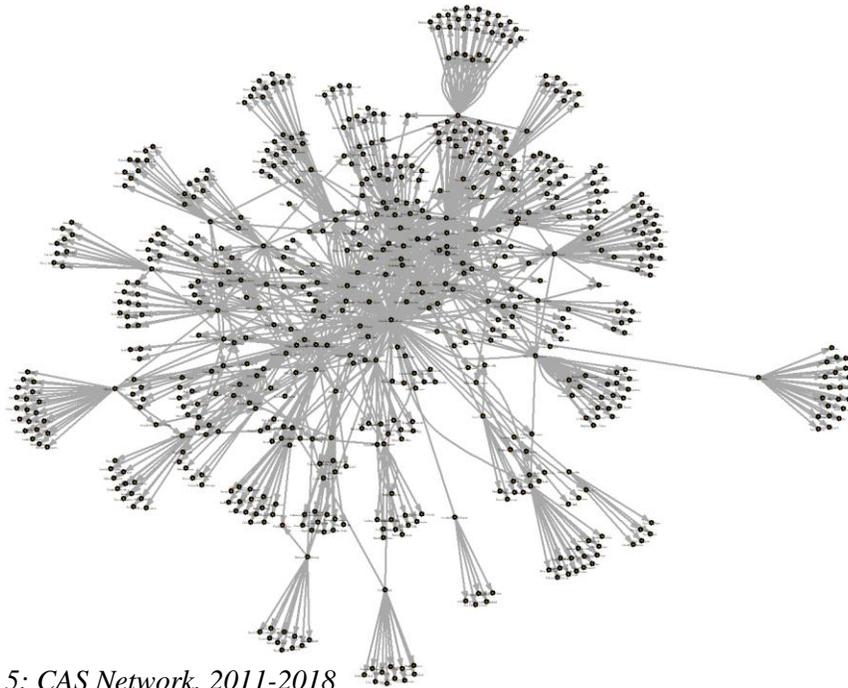


Figure 5: CAS Network, 2011-2018

Tables 3, 4, and 5 report the top six nodes based on degree-centrality (incoming thanks only), closeness-centrality, and betweenness centrality, respectively, for this network.

*Table 3: Degree-Centrality, 1975-2018*

Degree	Node Label
157	Charles A. Sullivan
43	Michael J. Zimmer
40	Tristin K. Green
36	Timothy P. Glynn
25	Kathleen M. Boozang
24	D. Michael Risinger

*Table 4: Closeness-Centrality, 1975-2018*

Closeness	Node Label
0.0002435	Charles A. Sullivan
0.0002179	Timothy P. Glynn
0.0002157	Michael J. Zimmer
0.0002153	Tristin K. Green
0.0002067	Marc R. Poirier
0.0002061	Frank Pasquale

*Table 5: Betweenness-Centrality, 1975-2018*

Between	Node Label
91140.82	Charles A. Sullivan
30208.40	Michael J. Zimmer
25567.50	Timothy P. Glynn
22892.29	Marc R. Poirier
22176.43	Tristin K. Green
15123.52	Brian Sheppard

Again, these centrality scores capture one's network position among the subset of law professors in the U.S. connected to Charlie, not among all law professors in the U.S. (our initial population of interest). This is why Charlie himself ranks highest on all three centrality measures. All the other top-ranked as central in this network do or did work at Seton Hall's law school at the same time as Charlie.<sup>41</sup>

Additional complications arise if we care about the *strength* of ties between nodes in a network. When we say Charlie is "connected" to another law professor, what do we mean? For some network connections, we can operationalize whether the connection exists or not by searching for relationships between the nodes of interest that are determined or recognized by law (e.g., spouse, family member, co-employee), social norms (e.g., a "romantic" partner), or mutual agreement (e.g., a LinkedIn connection, Facebook "friend," emergency contact). Then, we might rank order individuals by type of connection, and within connection type, to capture how "strong" those connections are, either by resorting to someone's perceptions of that strength (as reported in a survey) or a prior theory about the dimensions of tie strength that can be directly measured.

Here, for example, though we did not account for it above, we might believe that, in a network of law professors, co-authors are more strongly connected to each other than those who write separately but exchange thanks. Another example: because of concern about tie strength, we may assign less weight to older thanks in estimating Charlie's current network position. If one is more likely to transmit information to another based on how strong their professional connection is,<sup>42</sup> then that probability may decrease as that connection weakens over time. Accordingly, if we care about Charlie's network position because we want to know how likely he is to learn of industry innovations, we may prefer to estimate Charlie's network position based on his professional connections over a more recent period (say, the past decade) than over his entire career.

Measuring tie strength is further complicated here because thanking norms likely vary by author and over time.<sup>43</sup> For example, as the average

---

<sup>41</sup> See AM. ASS'N LAW SCHS., DIRECTORY OF LAW TEACHERS (2020), <https://dlt.aals.org/> (Boozang: 1989–present; Barnes: 2001–present; Glynn: 1999–present; Green: 2000–2003; Sheppard: 2010–present; Sullivan: 1978–present); AM. ASS'N LAW SCHS., 2011–2012 DIRECTORY OF LAW TEACHERS 1104 (2012) (Poirier: 1991–present); *id.* at 1079 (Pasquale: "since 2004"); *id.* at 1438 (Zimmer: 1978–2008). Poirier and Zimmer died in 2015. See AM. ASS'N LAW SCHS., 2016 PROCEEDINGS 39, 42 (2016).

<sup>42</sup> Minjae Kim & Roberto M. Fernandez, *Strength Matters: Tie Strength as a Causal Driver of Networks' Information Benefits*, 65 SOC. SCI. RES. 268 (2017).

<sup>43</sup> Tietz & Price, *supra* note 40, at 29–30 (finding, in convenience sample of law faculty, "wide variability in standards" for including peers in acknowledgements footnote).

number of thanks per paper increases with the ease and frequency of electronic communication between scholars, authors may tend to thank not only more people, but also more high-status people, regardless of how much or little they and the author interacted with them, so as to signal that author's own status or legitimacy as a bona fide member of a scholarly community.<sup>44</sup> If so, we should expect higher-status scholars to receive more thanks on average, and accordingly expect some measurement error when estimating network position.

Finally, even if we converged on theory or collected better data, this research strategy is extremely limited in capturing worker network position in most fields, because most professions do not regularly assign credit or authorship for work contributions in ways that are publicly available. Rarely does a building display the names of all the engineers, architects, construction workers, and others who made it, let alone how they interacted with each other. The same holds for all the people in the supply chain for a mass-produced product offered for sale. Moreover, most people within a profession or occupation do not have social media presences on platforms that record the kind of information we might use to estimate a worker's professional network position. Accordingly, we are today far from the counterfactual world where most employers can see workers' network positions. As a result, we can do more now to prepare employment law for when that world comes.

#### IV. CONCLUSION

This paper argued that, in the counterfactual world of widely available evidence of workers' network position, we should expect little change in the soundness of Charlie's past legal arguments about hiring algorithms and the partial enforcement of invalid contracts. Then, the paper used Charlie's corpus of scholarship to illustrate some current challenges to using evidence of workers' social network position in employment law.

#### V. COLOPHON

All figures and tables (except table 2) were created using R 3.6.3 and igraph 1.2.5.<sup>45</sup>

---

<sup>44</sup> Sullivan, *supra* note 39, at 1101–08; Tietz & Price, *supra* note 40, at 32–33.

<sup>45</sup> R CORE TEAM, R: A LANGUAGE AND ENVIRONMENT FOR STATISTICAL COMPUTING (2020), <https://www.R-project.org/>; Gábor Csárdi & Tamás Nepusz, *The igraph Software Package for Complex Network Research*, INTERJOURNAL 1695 (2006), <http://igraph.org>.