Behaviour in Networks of Collaborators: Theory and Evidence from the English Judiciary*

Jordi Blanes i Vidal[†] Clare Leaver[‡]

December 2006

Abstract

This paper uses data on judicial citations to explore whether the diffusion and/or application of knowledge within an organisation is affected by worker connectivity. Developing a simple model of discretionary citations, we distinguish between two hypotheses: *knowledge diffusion* whereby connected judges are more likely to be aware of each others' cases than unconnected judges, and *socialisation* whereby judges are more likely to be positively disposed to judges to whom they are more connected. Our empirical strategy exploits three important institutional features: (a) the random allocation of judges to case committees in the English Court of Appeal, (b) the existence of both positive and neutral citations and (c) the fact that connections occur over time. We are able to reject the knowledge diffusion hypothesis in its simplest form. We are unable to reject the socialisation hypothesis, and find strong evidence to support it. The paper concludes with a discussion of implications for other knowledge-based organisations.

Keywords: Networks, Public Sector Organizations, Judicial Citations.
JEL Classification: H1, K4, Z13.

---

*We thank Iwan Barankay, Martin Browning, Marcel Fafchamps, Rafael Hortala Vallve, Marco van der Leij, Margaret Meyer, Victoria Prowse, Jeremy Tobacman and seminar audiences at City University, Oxford, UAB and Warwick for their advice and helpful comments prior to this draft. We are also grateful to the many Oxford MPhil students who helped us collect the data and to the Oxford Economics Department for funding. All errors are our own.

†Nuffield College, Oxford. Email: jordi.blanes-i-vidal@nuffield.ox.ac.uk.

‡University of Oxford, Nuffield College, CEPR and CMPO. Email: clare.leaver@ economics.ox.ac.uk

1 Introduction

A common question in the burgeoning empirical literature on social networks is whether behaviour (e.g. technology adoption, innovation, job search) is shaped by the distance between certain nodes. Notwithstanding the number of papers tackling the issue, progress is generally recognised to have been slow due to problems of identification and interpretation.

To see the subtleties involved, consider the "network of friends" illustrated in Figure 1, in which agents A and C are directly linked to B and indirectly to each other, and D is not connected to any other agent.[1] Assume that C makes a different decision to D. Can we attribute this difference to the fact that C is closer to A? Obviously, the answer is no unless it is safe to assume that C and D are identical in every relevant respect other than their position in the network. While some progress has recently been made, through the use of instrumental variables (Conley and Udry [5], Munshi [18], Bandiera et al. [2]) or field experiments (Duflo and Saez [7]), identification remains a challenging issue.

Even if causality can be established, however, a second issue will require attention: what is the mechanism underlying this differential behaviour? Some papers (Conley and Udry [5], Singh [19], Fafchamps et al [8]) have focused on the potential for networks to facilitate information transmission. Others (Bertrand et al [3], Munshi [18], Bandiera et al [2]) have emphasised the possibility of social pressure and influence. In Figure 1, C's differential behaviour could be caused by A's knowledge flowing through B towards him.[2] Alternatively, it could be caused by C wishing to adopt actions similar to those with whom he is linked. In many contexts (see Duflo and Saez [7] for a discussion) it is difficult to disentangle these two effects.

This paper proposes an empirical strategy that tackles the problems of both identification and interpretation in a particular class of network, which we term "networks of collaborators" (hereafter NOCs). In NOCs, nodes are teams, while links signify that two teams share a common member. Figure 2 provides an illustration. A, B, C and D now refer to teams, each comprised of three individuals. A and B are linked by the fact that they share individual 2, while B and C are linked via individual 4.

There are two important characteristics which, while by no means exclusive to NOCs, are typically associated with them. The first feature is that links are often "directed". Such one-way links occur when teams are formed in sequence, so that an individual cannot be part of two teams contemporaneously. In Figure 2, team A might form (and dissolve) in period t, whereas team B and C might occur in periods t+1 and t+2 respectively. The second feature is that team behaviour often imposes externalities on other nodes in the network. For instance, in the course of fulfilling its function, a team may criticise or even revise another team's decision.

[1] Another way of expressing this is to say that A and C are separated by one unit of network distance from B and by two units from each other, whereas D is separated by infinite units of distance from any other agent.

[2] This would require one conversation between A and B and another between B and C, which is the rationale for saying that A and C are separated by two units of network distance.

The empirical strategy that we propose in this paper draws on these features of NOCs. By exploiting the dynamics of team formation, and hence the existence of directed links, we separate circumstances where knowledge held by one node could diffuse through the network, from circumstances where it could not. Similarly, by exploiting the qualitative nature of any externalities, we separate circumstance where behaviour could be driven by socialising forces (the imposition of positive externalities) from those where this is unlikely (negative externalities).

An explicit analysis of NOCs has largely been overlooked in the economics literature. This is not due to their lack of pervasiveness or importance. NOCs can be found in science (engineers collaborating on inventions), industry (companies forming joint ventures, executives sitting on the boards of companies), government (politicians sitting on legislative committees), the law (judges hearing cases in panels), academia (coauthors writing papers) and indeed in many other occupations which require the pooling of human capital.

While all of these networks are important, we restrict our attention to a particular NOC: the network of judicial collaborators hearing cases in the English Court of Appeal. This network displays both features of NOCs discussed above. Teams (three-judge panels) are formed in sequence as cases arrive with Listing Officers. Actions (rulings dismissing or allowing appeals, citations praising or criticising previous cases) clearly impose positive and negative externalities on other nodes in the network. These institutions, together with the fact that panels are formed via a random process, imply that the judiciary is an ideal environment in which to explore behaviour in networks.[3]

Of course, the merits of studying an institution as fundamental as the judiciary extend beyond its potential as a natural experiment. The proper functioning of the judiciary has been associated with economic growth (Glaser et al [10]) but, to date, there is little micro-level evidence showing what makes a judiciary operate efficiently. A growing body of work is using citations to measure aspects of the legal process, from judicial influence (Landes [16]) to the identification of legal revolutions (Cross et al [6]). In fact, Choi and Gulati [4] have even suggested that citation counts be used as a metric to select candidates for the American Supreme Court. Aside from developing a new empirical strategy, a second contribution of this paper is to highlight that judicial citations can be subject to network effects and hence unsuitable for use as performance measures.

Taking the unit of analysis to be a *dyad* (a pair of cases i and j, where j is the earlier "citable" case), we pose two specific research questions. First, does the distance between i and j in the NOC affect i's decision to cite positively (praising j), neutrally (merely mentioning j), negatively (criticising j) or not at all? Second, to the extent that there is a causal relation between distance and the outcome of this multinomial choice problem, can the mechanism be

[3]The random allocation of judges to cases has been widely used in the legal literature to explore, e.g., the effect of judicial background on case outcomes (Ashenfelter et al. [1], Sunstein et al [20]) and voting behaviour in panels (Fischmann [13]). To our knowledge, this paper is the first to use random panel formation to identify the effect of social connections.

established? Specifically, is the mechanism knowledge diffusion whereby panel members pass on their knowledge of prior cases in the course of collaboration, socialisation whereby judges become positively disposed to each other (or each other's cases) in the course of collaboration, both or "none of the above"?

We start our analysis by developing a theoretical model of discretionary citations.[4] Central to this model is the notion that, while citing a case entails a certain effort, it also allows the panel to signal that its decision is likely to be correct, thus reducing the likelihood of an appeal. We show that, in equilibrium, panels with high confidence in their ruling cite positively, while those with lower confidence cite neutrally or not at all. Negative citations are avoided because they invite appeals *and* impose negative externalities on others in the network.[5] Key comparative statics results are that an increase in knowledge (the probability that i is aware of j) results in substitution towards neutral citations, while an increase in socialisation (i's social preference for j) prompts substitution away from neutral citations. The reason behind the rise is obvious, the fall perhaps less so. The intuition is that, with negative citations ruled out, neutral citations can convey a less than ambivalent message, essentially by "damning with faint praise".

We set out the testable implications of this theory by defining four "treatment" groups. Appealing to the logic of quasi-experimental methods, we consider a dyad to have received a particular treatment in case i and j are indirectly linked via at least one connecting case (B in Figure 2) in a particular, narrow window of time. Institutions (and a direct test) suggest that panel formation is most plausibly random at the start of the legal year in what is known as the Michaelmas Term. Consequently, we look for the presence of connecting cases in the first complete Michaelmas Term immediately: before the end of case j; after the end of case j; and after the end of case i.[6]

Our first treatment group, *Distance 2/Before j*, consists of dyads with connecting cases exclusively in the first Michaelmas Term before case j. The second, *Distance 2/After j*, is identical to the first, except that connecting cases occur exclusively in the first Michaelmas Term after case j. The third and fourth groups lack connecting cases in both the first Michaelmas Term before and after j, and can therefore be described as being at a distance greater than two at the date of case i. The third group, *Distance >2/After i*, consists of dyads with connecting cases in the first Michaelmas Term after i, while the fourth group, *Distance >2/None*, consists of dyads without any connecting cases in any of the three terms.

[4] A panel makes a *discretionary* citation when it refers to a prior case even when not bound to do by precedent (e.g. the facts of the two cases are materially different). For a theoretical model of model of citation behaviour when the principle of precedent applies see Gennaioli and Shleifer [9].

[5] Negative discretionary citations are also absent in the data as we document in Table 3.

[6] Providing we focus on dyads that *could* have received a random link in all three terms, under random panel formtion there should not be any systematic differences between dyads that, say, received a link just before case j and those that received a link just after j.

The essence of our empirical strategy is to compare selection (citation vs. no citation) and outcome (positive vs. neutral) responses across treatment groups. Using this joint comparison we can infer substitution patterns and then, in light of the theory, reject possible mechanisms for any causal effects. In practical terms, this amounts to conducting three tests, each with *Distance >2/None* as the base category. The first is a *Placebo Test* that compares selection and outcome responses across the *Distance >2/After i* group and the base category. Notice that, since an After i link cannot affect citation behaviour itself, a difference could only arise if links are correlated with some other citation relevant characteristic. Consequently, a rejection of the null for either response is a rejection of our key identifying assumption.

The second test is a *Socialisation Treatment Test* that compares responses across the *Distance 2/Before j* group and the base category. Here the important observation is that knowledge could not diffuse from j to i in either group but any socialising forces present in the *Distance 2/Before j* group could produce a substitution *away* from neutral citations. This insight aids interpretation. For instance, a positive difference rejecting the null in the outcome response could not have been driven (entirely or in part) by knowledge diffusion but, in conjunction with a weakly negative difference in the selection response, is consistent with socialisation.

Finally, the third test is a *Combined Treatment Test* that compares responses across the *Distance 2/After j* group and the base category. Now knowledge diffusion from j to i in the *Distance 2/After j* group could produce a substitution *towards* neutral citations. Since socialising forces produce the reverse, there is again the possibility of interpretation. For instance, a negative difference rejecting the null in the outcome response could not have been driven (entirely) by socialisation, but in conjunction with a positive difference in the selection response, is consistent with knowledge diffusion.

To conduct these tests we obtain a sample of potential and actual citations. We achieve this by using the 8,923 Court of Appeal cases reported by Westlaw (an internet based case law provider) between 1980-2005 to construct a population of 23,517,932 dyads. Since the unconditional likelihood of a citation between any two given cases is extremely low, we select dyads with a citation and then draw a random sample of the dyads with no citation. Given random panel formation, the methods used to analyse this data can be simple, namely comparison of means (corrected for choice-based sampling) with robustness checks via independent Probits and a jointly estimated binary sample selection model with controls.

By way of preliminary analysis, we first explore whether treatment status can indeed be taken to be randomly assigned. We offer three pieces of evidence supporting this key identifying assumption. First, the results of a direct test for random panel formation (based on the test for random case allocation in Ashenfelter et al [1]) fail to reject the null of randomness. Second, we find few significant differences in observable case-level characteristics (including time difference and similarity of subject) or judge-level social network characteristics (including attending the same school or university and being members of the same chambers or Gentlemen's Club) across groups. In fact, the only characteristic where there is a systematic difference across groups is in the size of the panels in the two cases. Third, we fail to reject the null in the *Placebo Test* in every specification.

Turning to our main results, in the *Socialisation Treatment Test* we find a positive difference in the outcome response and a small, statistically insignificant negative difference in the selection response. Taken together, these results imply a substitution *away* from neutral citations towards positive citations. The point estimate of marginal effect in the outcome response is stable across specifications to within .01 of .18. Given such stability, the simple proportions give a reliable indicator of the economic magnitude of the effect. The proportion of citing dyads that are positive in the base group is a little over .40; while the proportion of citing dyads that are positive in Before j connected treatment group is just under 50% higher at .58. To the extent that one reads this as an "effect" (given a $p-value$ of .09), it cannot be attributed to (entirely or in part) knowledge diffusing in the process of collaboration. Instead, while all bar one set of results supporting a causal effect would have rejected socialisation as the mechanism, the results we obtain are consistent with the existence of such a force.

The results of the *Combined Treatment Test* tell a very similar story. Here we find a positive difference in the outcome response and no difference in the selection response. The point estimate for the outcome response is stable at .19. There is now less doubt that this can genuinely be read as an effect ($p-value$ of .04) and, again, it cannot be attributed solely to knowledge diffusion. These results are, however, consistent with socialisation alone or the presence of both forces.

In conclusion, we find that socialisation appears to be a strong economic force influencing citations among the English judiciary. This finding suggests that extreme caution should be used when interpreting citations as a measure of judicial ability. As we discuss in the Conclusion it also implies that a different allocation of judges to panel cases cases $-$ a different network formation process $-$ woul be able to increase the knowledge diffusion among judges while reducing the socialising forces presently at work.

2 The English Judiciary

Our study focuses on cases in the Civil Division of the Court of Appeal 1980 $-$ 2005. This Section briefly summarises the structure of the English court system, including a description of the tasks that are performed by judges and the mechanisms by which they are allocated to panels dealing with particular cases.

2.1 Organisational Structure

The Court of Appeal is the second most senior court in England (and Wales), immediately above the High Court and below the House of Lords. It consists of two divisions: Criminal and Civil. The Criminal Division hears trials from indictment in the Crown Court, appeals from Courts Martial and references from the Attorney General. The Civil Division hears appeals from the High Court, appeals from the county courts in civil matters and appeals from various appeal tribunals (e.g. the Immigration Appeals Tribunal).

The House of Lords has 12 judges (Law Lords) and cases are usually heard by at least five judges which together ensure that there is little variation in on-the-job interaction between Law Lords. The High Court has 107 judges (Justices) and cases are usually heard by a single judge, largely precluding on-the-job interaction between Justices. In contrast, the Court of Appeal has 37 judges (Lord Justices) and the overwhelming majority of cases are heard by a panel of three judges. There is therefore on-the-job interaction but, equally, two individual Lord Justices can spend years without coming into direct contact with each other.

We restrict our analysis to the Civil Division of the Court of Appeal. Panels in the Criminal Division are usually staffed by a Lord Justice and two Justices. While we have some understanding of how judges are allocated to panels *within* courts, it is less clear what determines the allocation of justices between the High Court and the Criminal Division. In contrast, the Civil Division is overwhelmingly staffed by Lord Justices and most Lord Justices spend most of their time sitting in the Civil Division. The self-contained nature of the Civil Division makes it ideal for our analysis, as the movement of judges across courts is unlikely to introduce non-randomness into the formation of panels.

2.2 Judicial Tasks and Citations

The main task facing a panel sitting in the Court of Appeal is, of course, to make a ruling allowing or dismissing the appeal. When dealing with non-trivial cases, the panel is also expected to publish its "opinion" giving reasons for its decision. Published opinions usually refer to the facts of the case, the different laws being applied, the legal issues being considered and resolved and, crucially for our purposes, citations to prior cases.

In our empirical analysis we exploit the existence of different types of citations. An important initial distinction is between discretionary and non-discretionary citations. When the facts of the case being considered are identical to those of a previous case, the rule of precedent applies and the panel has no alternative but to "follow" the previous decision.[7] If the facts are essentially the same, except for some minor but important distinction motivating a different decision, the panel can choose to "distinguish" the case from the prior case, but must

[7]For the rule of precedent to apply it is also necessary that the previous court is not subordinate to the current court.

still refer to it. Lastly, when the facts are identical, but the panel regards the previous decision as unsound (and is sitting in a higher court), it can choose to "overrule" the previous decision, but again must still mention it. We choose not to focus on these non-discretionary citations for two reasons. First, the existence of a case with *identical* facts is unlikely to be missed by the panel or the party whom following the previous decision might benefit. This implies that there is little scope for network distance to have a knowledge diffusion effect. Second, the all but compulsory nature of the citation reduces the possibility socialising forces affecting behaviour.

Instead, our focus in on discretionary citations. When the facts of the case do not have a pre-existing equivalent, there may still be a prior case dealing with a similar legal issue. Citing this case is, however, not compulsory. We distinguish between positive, neutral and negative discretionary citations. If the previous case has dealt with the shared legal issue in a manner consistent with that of the current case, the citing panel can highlight this fact in its published decision (the phrase case x "applied" appears in the opinion). We treat this explicit and public statement of agreement with a previous decision as a "positive citation". A citation is "neutral" if the previous case is mentioned as dealing with a similar legal issue, but the panel does not express agreement or disagreement with it (case x "considered"). Lastly, a citation is "negative" if the panel highlight that a previous case has reached a conclusion that is contradicts its view on the common legal issue (case x "not applied").

2.3 Panel Formation

To conclude this Section, we outline the procedures governing the allocation of judges to panels. The first stage is for the case to be allocated to a particular court according to the guidelines in the relevant statute. When the court to which a case is allocated can have panels of different sizes it is also statute that determines panel size. Once the court is determined, the task of forming the panel falls to the court's Listing Officer.[8]

The Civil Division Listing Officer is an employee of the court - a bureaucrat rather than a judge. Our conversations with the Listing Officer in the Civil Division suggest that the principles underlying panel formation run as follows. First, only "ticketed" judges (i.e. only serving, or occasionally retired / promoted Lord Justices) can be chosen. Second, whenever possible, allocation follows the "cab-rank principle". That is, as judges dispose of their cases, they join the back of the rank and wait to receive a new case; as a case requiring a panel of size n arrives, the Listing Officer matches it to the first n judges in the rank. Third, in the event of a tie (more than n judges joined the rank at the same time), the panel is formed at random.

Strict application of the cab-rank principle suggests that it is the duration of the previous case that determines whether or not a judge will be allocated to a particular case. Since we

[8]In principle the judge in charge of the division can exercise a supervisory role. Our conversations with one of the judges with experience of running a division suggest that the prerogative to determine the allocation of committee formation is not exercised in more than 2-3% of cases.

are interested in the connections between judges rather than in the matching of judges to cases per se, an important caveat applies. Suppose that three judges are already working on a case together. As they dispose of that case, the three of them will join the cab-rank at the same time. As a result, they are more likely to be matched with each other in the immediate future than to be matched with other judges dealing with other cases. In short, in steady state there will be serial correlation in on-the-job interaction *even if the initial / tie breaking process is truly random.*

For our purposes, serial correlation is not a problem providing the initial / tie breaking process *is* truly random. Ascertaining whether this is true is not as hard as it sounds. Figure 1 shows the number of cases brought to a close on each day of the judicial year between October 1, 2001- October 1, 2005. The legal year, much like the academic one, is structured in to terms and vacations. In each plot the $y-axis$ represents the start of the first legal term (Michaelmas), the second line the end of Michaelmas term, the third line the start of the second legal term (Hilary) and so on. Two regularities are immediate from Figure 1. First, cases are almost exclusively decided during term-time, with, on average, only 15 cases decided during the long summer vacation. Second, there is always a large spike on the final day of the judicial year. For instance, six percent of the 2003-2004 decisions were issued on this day (July 31) alone. Taken together, these pieces of evidence suggest that the panel formation process starts afresh each legal year. In Section 4, we treat the start of the legal year as an initial condition and test for randomness of panel formation in Michaelmas Term.

3 Theory: A Model of Discretionary Citations

This Section develops a simple model of discretionary citations. The purpose of this model is two-fold. First, it enables us to state precisely how the knowledge diffusion and socialisation mechanisms might apply in the context of a network of judicial collaborators. Second, by pinning down a rationale for discretionary citations we can give equilibrium meaning to the citation terms discussed in Section 2 and, in turn, derive comparative statics results (driven by these mechanisms) for use in our empirical analysis.

3.1 Model Description

There are two strategic players: a panel (it) hearing a case i in the Court of Appeal and a losing party (she, the appellant if the appeal is dismissed or the respondent if the appeal is allowed). Motivated by the ensuing empirical analysis, our primary interest is in the forces determining whether/how the panel in case i cites some previous case j (that is, on the *dyad* between case i and j).

Our starting assumption is that there is a correct ruling, a "state of the world" $x \in \{0, 1\}$. For concreteness, we let $x = 0$ denote the state where the appeal should be dismissed (the first instance ruling was right) and $x = 1$ the state where the appeal should be allowed (the first instance ruling was wrong). The panel in case i cannot observe x but can combine its own legal knowledge with the facts of the case to deduce what this state is likely to be. We equate this process of deduction with the generation of a private non-verifiable signal on x, $s \in \{0, 1\}$, with precision $p \in (1/2, 1]$. Having generated s, the panel (i) makes a ruling $r \in \{0, 1\}$ dismissing or allowing the appeal and (ii) sends a message M via its "published opinion".

After the panel has made its ruling and published its opinion, the losing party decides whether to lodge an appeal to the House of Lords. We normalise the losing party's net benefit from a successful appeal to 1 and assume that her cost of appealing, a, is private information but is commonly known to be distributed $U[0, 1]$. If the losing party does not appeal, the panel's ruling is implemented immediately. Otherwise, case i moves to the House of Lords where, after some delay, Law Lords implement the correct ruling $r = x$.[9]

Private Preferences (Confidence in i's Ruling) The panel in case i wants to see "justice done", receiving a private payoff of 1 when the correct ruling is implemented (either by itself or by Law Lords on appeal) and 0 otherwise. However, it is also keen to avoid unnecessary delays in the legal process and will therefore strive to avoid an appeal when it is sufficiently confident that its ruling is correct. Formally, we write the panel's expected private utility from taking ruling r and sending message M as:

$$E[U(r, M) \mid I] = \Pr(\textit{not appealed}|r, M, I) \Pr(r = x|r) + \delta \Pr(\textit{appealed}|r, M, I), \qquad (1)$$

where $\delta < 1/2$ is a discount factor applied to future payoffs and I is the panel's information set defined immediately below.

Information and the Published Opinion Aside from generating s, the panel can potentially improve its estimate of x by consulting prior cases. By assumption, the panel can only consult cases in its *knowledge set* (the subset of the population of prior cases that it is aware of). Since our objective is to derive predictions at dyad-level, we focus our attention on the panel's behaviour towards a single prior case, case j, that may or may not be in case i's knowledge set.[10]

We denote the probability that case j is in case i's knowledge set by t. It is important to stress that t is an exogenous parameter rather than a choice variable. As we discuss in detail in Section 3.3, under the "knowledge diffusion" mechanism, t is decreasing in the distance between case i and j in the network of judicial collaborators. With probability $1 - t$, the panel

[9] The important (and one would hope accurate) assumption is that, given their superior legal knowledge, Law Lords are better able to deduce x than the panel in the Court of Appeal.

[10] The panel may consult multiple cases in this set but, assuming behaviour is independent across prior cases, we leave this unmodelled. The fact that panels in the Court of Appeal often make multiple citations (even within categories) suggests that an assumption of independence is not unduly restrictive.

is unaware of case j and necessarily makes no citation. Otherwise with probability t, the panel consults case j and, as a first step, learns whether it tackles the same legal issue (an event that occurs with probability β) or is unrelated (an event that occurs with probability $1-\beta$). If the two cases do tackle the same legal issue the panel receives a signal $s_l = \{same\}$ that provides hard evidence that case j is informative about x.

The sense in which case j might be informative about x is obvious: observing that a panel in a previous case *tackling the same legal issue* took a particular ruling provides evidence that the same ruling should be applied in the current case. To capture this idea, we assume that, conditional on case j tackling the same legal issue, the panel in case i receives a second signal on x, $s_r \in \{0,1\}$. (The subscript r is used to signify "ruling" as distinct from "legal issue".) If the panel in case j allowed its appeal, then $s_r = 1$. Conversely, if the panel in case j dismissed its appeal, then $s_r = 0$. For simplicity, we assume that the precision of s_r (as privately perceived by the panel in case i) is $q \in [1/2, p)$ since this ensures that this second signal is informative but not pivotal.[11]

To summarise, for the purposes of exploring dyad-level citation behaviour, we can think of the panel in case i as being one of three different "types":

$$type = \begin{cases} 1 \text{ if } I = \{same, r \neq s_r\} \\ 2 \text{ if } I = \{\varnothing\} \\ 3 \text{ if } I = \{same, r = s_r\} \end{cases} . \tag{2}$$

Notice that the panel's confidence that its ruling is correct is increasing in type. $Type-1$ has low confidence, having received a contradictory signal, $type-2$ has intermediate confidence having received no new information and $type-3$ has high confidence, having received a confirmatory signal.

It follows from (2) that, by disclosing (or indeed hiding) the existence of the previous case j in its published opinion, the panel may be able to influence the likelihood of an appeal. We assume, however, that this signalling comes at a cost. First, there is a cost c for simply making a reference to case j (i.e. revealing that $s_l = \{same\}$ which can be equated with the citation term "considered"). This cost reflects the effort necessary to enter the citation in the correct place in the opinion, cite the original source accurately and clarify why the legal issues are the same. Second, there is an extra cost κ for making a reference to the ruling in case j (i.e. revealing that $r = s_r$ which can be equated with the citation term "applied" and or $r \neq s_r$ which can be equated with the citation terms "not applied/disapproved"). This cost reflects the effort required to justify in writing why the previous case reaches a conclusion which lends credibility to the panel's decision ($r = s_q$) or why it is contradictory with it ($r \neq s_r$). We assume that these costs are the panel's private information but are commonly known to be *iid* $U[0,1]$.

[11]Assuming that $q < p$ enables us to focus on the signalling aspect of judicial citations, abstracting from any issues arising from potential distortions in judicial rulings (for a discussion of this, see Levy [17]).

To recap, the information that could be disclosed by the panel in its published opinion is $M \in \{\oslash, same, r = s_r, r \neq s_r\}$. In what follows we will say that the panel makes *no citation* if $M \in \{\oslash\}$, a *neutral citation* if $M \in \{same\}$, a *positive citation* if $M \in \{r = s_r\}$ and a *negative citation* if $M \in \{r \neq s_r\}$.

Social Preferences (Confidence in j's Ruling) As alluded to above, making a positive citation increases the common posterior that case i is correct and therefore discourages an appeal by the losing party. Likewise making a negative citation reduces the posterior that case i is correct and therefore encourages an appeal by the losing party. However, the same logic ensures that a positive (respectively negative) citation also increases (reduces) the posterior that case j is correct. In short, making a positive citation casts *both* cases in a good light, while making a negative citation casts *both* cases in a bad light.

We complete the model with the final assumption that, as well as wanting to see justice done (quickly), the panel in case i potentially cares about the public perception of case j. Specifically, we assume that the panel in case i receives utility $\alpha \in [0, 1)$ from casting the panel in case j in a good light via a positive citation.[12] As we discuss in detail in Section 3.3, under the "socialisation" mechanism, α is decreasing in the distance between case i and j in the network of collaborators. The panel's total (private plus social) expected utility from taking ruling r and sending message M given information I is therefore given by:

$$E[U(r, M) \mid I] + \alpha K(M), \tag{3}$$

where $K(M)$ is an indicator variable that takes the value 1 if $M \in \{r = s_q\}$ and 0 otherwise.

Timing The timing of the game is as follows: (1.a) The panel in case i in the Court of Appeal observes s and (with probability $t\beta$) s_l, s_r, c and κ. It then takes a ruling r and sends message M in its published opinion. (1.b) The losing party observes r, M and a and decides whether to appeal the case. If the case is not appealed, the panel's ruling is implemented. If the case is appealed, no ruling is implemented and the case passes to the House of Lords. (2) If the case reaches the House of Lords, Law Lords implement the correct ruling $r = x$.

3.2 Analysis

Our solution concept is Perfect Bayesian Equilibrium. This requires the losing party and each type of panel $(1, 2$ or $3)$ to maximise expected utility given the other players' strategies, and the losing party to adjust her beliefs using Bayes' rule whenever possible. In solving the model, we make the simplifying assumption that the *type* -3 panel always makes a citation. Although

[12]Obviously, one might also assume that the panel looses some utility from casting panel j in a bad light. Since negative citations are never sent in equilibrium (or indeed in practice as we show in Section 4) the restriction to utility gains is without loss of generality.

this is a strong assumption, it has the advantage of simplifying the comparative statics that are a key objective of the modelling exercise.[13]

Before solving the model, it will be helpful to derive some key probabilities that underpin the analysis. First, the common prior beliefs over type are

$$\Pr[type = 1] = t\beta\left(p(1-q) + (1-p)q\right) \tag{4}$$

$$\Pr[type = 2] = 1 - t\beta \tag{5}$$

and

$$\Pr[type = 3] = t\beta\left(pq + (1-p)(1-q)\right). \tag{6}$$

Second, the probabilities of each type's ruling being correct are

$$\Pr[r = x|type = 1] = \tfrac{p(1-q)}{p(1-q)+(1-p)q} < \Pr[r = x|type = 2] = p$$
$$< \Pr[r = x|type = 3] = \tfrac{pq}{pq+(1-p)(1-q)}. \tag{7}$$

Losing Party Behaviour Naturally, when deciding whether to appeal to the House of Lords, the losing party weighs her belief that the ruling was incorrect against the appeal cost a. Letting σ denote the panel's strategy, the losing party appeals if[14]

$$\Pr[r \neq x|M, \sigma] > a \tag{8}$$

where

$$\Pr[r \neq x|M, \sigma] = \sum_{type \in \{1,2,3\}} \Pr[r \neq x|type] \Pr[type|M, \sigma]. \tag{9}$$

As the final term on the RHS in (9) indicates, to apply this decision rule the losing party must update her beliefs over type given M and σ. This updating is much simpler than it sounds. $Type - 1$ will either make no citation or a neutral citation (a negative citation is possible but, from the LHS of (7), is strictly dominated as it invites an appeal); $type - 2$ can only make no citation; while $type - 3$ will, by assumption, either make a neutral citation or a positive citation. As a result, the panel's strategy space collapses to the unit square.

We write a strategy for the panel as $\sigma = (b_1, b_3)$, where b_1 denotes the probability $type - 1$ makes a neutral (rather than no) citation and b_3 the probability $type - 3$ makes a positive citation (rather than a neutral) citation. Applying Bayes' rule to obtain $\Pr[type|M, \sigma]$ and using $\Pr[r = x|type]$ given in (7), the losing party's assessment that the ruling is correct given M and beliefs over σ can then be written as

$$\Pr[r = x|M \in \{r = s_r\}, \sigma] = \tfrac{pq}{pq+(1-p)(1-q)} \tag{10}$$

[13] To be relaxed in the next draft.

[14] Note that the model is symmetric in relation to the state of the world and signal structures. As a result r is uninformative in itself about the likelihood that the appeal will be successful.

$$\Pr(r = x|M \in \{same\}, \sigma) = \frac{b_1 p(1-q)+(1-b_3)pq}{(1-b_3)(pq+(1-p)(1-q))+b_1(p(1-q)+(1-p)q)} \tag{11}$$

and

$$\Pr(r = x|M \in \{\varnothing\}, \sigma) = \frac{t\beta(1-b_1)p(1-q)+(1-t\beta)p}{t\beta(1-b_1)(p(1-q)+(1-p)q)+(1-t\beta)}. \tag{12}$$

Obviously, the losing party's assessment that the ruling is incorrect (i.e. what is needed to apply the decision rule in 8) is just the complement of these probabilities. Note that, since $a \sim U[0,1]$, (10)-(12) also give us the panel's belief that the case will not be appealed (i.e. $\Pr[a > 1 - \Pr[r = x|M]] = \Pr[r = x|M]$).

Citation Behaviour As noted above, a $type-2$ panel cannot make a citation because it has no information. Instead, consider $type-1$ behaviour. From (1), the net benefit to a $type-1$ panel of a making a neutral citation rather than no citation (via the reduction in the probability of an appeal) is

$$(\underbrace{\Pr[r = x|M \in \{same\}, \sigma]}_{\text{depends on } p,\ q,\ b_1 \text{ and } b_3} - \underbrace{\Pr[r = x|M \in \{\varnothing\}, \sigma]}_{\text{depends on } t,\ \beta,\ p,\ q \text{ and } b_1}) \cdot (\underbrace{\Pr[r = x|type = 1]}_{\text{depends on } p \text{ and } q} - \delta). \tag{13}$$

For any level of this net benefit, we can define a threshold level of the neutral citation cost, $\bar{c}$, equal to (13) such that all $type-1$ panels with $c > \bar{c}$ make no citation and all $type-1$ panels with $c < \bar{c}$ make a neutral citation. The probability $type-1$ makes a neutral citation, b_1, is then obtained by evaluating the uniform CDF at this threshold $\bar{c}$. Notice that the behaviour of $type-1$ depends on the behaviour of $type-3$ via the losing party's inference following a neutral citation (the first term in (13) is a function of b_3).

Now consider $type-3$ behaviour. From (3), the net benefit to a $type-3$ panel of making a positive rather than neutral citation (via the reduction in the probability of an appeal *and* the opportunity to cast case j in a good light) is

$$(\underbrace{\Pr[r = x|M \in \{r = s_r\}, \sigma]}_{\text{depends on } p \text{ and } q} - \underbrace{\Pr[r = x|M \in \{same\}, \sigma]}_{\text{depends on } p,\ q,\ b_1 \text{ and } b_3}) \cdot (\underbrace{\Pr[r = x|type = 3]}_{\text{depends on } p \text{ and } q} - \delta) + \alpha. \tag{14}$$

Again, for any level of this net benefit, we can define a threshold level of the positive citation cost, $\bar{\kappa}$, equal to (14) such that all $type-3$ panels with $\kappa > \bar{\kappa}$ make a neutral citation and all $type-3$ panels with $\kappa < \bar{\kappa}$ make a positive citation. The probability $type-3$ makes a positive citation, b_3, is then obtained by evaluating the uniform CDF at this threshold $\bar{\kappa}$. Again, notice that the behaviour of $type-3$ depends on the behaviour of $type-1$ via the losing party's inference following a neutral citation (the second term in (14) is a function of b_1).

Equilibrium and Comparative Statics An equilibrium of this game between the panel and losing party is a pair of thresholds $\bar{c}$ and $\bar{\kappa}$ that imply losing party beliefs over σ that are correct.

Proposition 1. *A PBE exists and is unique. In equilibrium, an uninformed panel ($type-2$) makes no citation, an informed panel with low confidence ($type-1$) makes no citation if $c > \bar{c}^*$ and a neutral citation if $c < \bar{c}^*$, and an informed panel with high confidence ($type-3$) makes a neutral citation if $\kappa > \bar{\kappa}^*$ and a positive citation if $\kappa < \bar{\kappa}^*$.*

A proof of this result is provided in the Appendix. To clarify the intuition, we start by considering how the $type-1$ threshold $\bar{c}$ varies with the $type-3$ threshold $\bar{\kappa}$. When $\bar{\kappa} = 1$, no $type-3$ panel is issuing a neutral citation. From a $type-1$ panel's perspective, making a neutral citation therefore entails paying a cost to increase the likelihood of an appeal. This is because a neutral citation will be read as cast iron evidence that the panel in case j tackled the same legal issue but made a different ruling to the panel in case i. Consequently, no $type-1$ panel finds it worth while to make a neutral citation ($\bar{c} = 0$). In contrast, when enough $type-3$ panels issue neutral citations (requiring lower $\bar{\kappa}$), a neutral citation is better news and at least some $type-1$ panels are prepared to pay a cost to mimic $type-3$ (implying an increase in $\bar{c}$). In sum, $\bar{c}$ is decreasing in $\bar{\kappa}$.

Now consider how the $type-3$ threshold $\bar{\kappa}$ varies with the $type-1$ threshold $\bar{c}$. When $\bar{c} = 0$, no $type-1$ panel is issuing a neutral citation. From a $type-3$ panel's perspective, there is no private preference gain from issuing a positive citation. This is because a neutral citation will now be read as cast iron evidence that the panel in case j tackled the same legal issue and made the same ruling as the panel in case i. Instead, the only benefit from issuing a positive citation stems from the social preference term α. Consequently, only those $type-3$ panel's with a cost less than α make a citation ($\bar{\kappa} = \alpha$). In contrast, when at least some $type-1$ panels issue neutral citations (requiring higher $\bar{c}$), making a neutral citation is worse news and at least some $type-3$ panels are prepared to pay a cost to signal their high confidence. In sum, $\bar{\kappa}$ is increasing in $\bar{c}$.

It is easy to show that the (reaction) functions $\bar{c}$ and $\bar{\kappa}$ cross once, giving us a unique PBE. Having characterised this equilibrium, we can proceed to the main objective of the exercise: comparative statics with respect to "knowledge" t and "social preference" α.

Proposition 2. *The citation threshold $\bar{c}^*$ is increasing in "knowledge" t and decreasing in the "social preference" α. The positive citation threshold $\bar{\kappa}^*$ is increasing in both t and α. Furthermore, $\bar{c}^* + (1 - \bar{\kappa}^*)$ is increasing in t.*

The intuition behind Proposition 2 is straightforward. Recall that the parameter t reflects the probability that case j is in panel i's knowledge set. As t increases, making no citation becomes more indicative of being a $type-1$ panel and less indicative of being a $type-2$ panel.

This increases the likelihood of appeal under no citation, thus strengthening the incentive for a $type-1$ panel to issue a neutral citation. This in turn increases the likelihood of an appeal under a neutral citation and so a $type-3$ panel's benefit from making a positive citation rises. In the new equilibrium both $\bar{c}^*$ and $\bar{\kappa}^*$ have increased. We will use the observation that $\bar{c}^* + (1-\bar{\kappa}^*)$ is increasing in t since it implies that the probability (from the econometrician's perspective) of case i issuing a neutral citation is strictly increasing in t. The reason stems from the fact that the $type-3$ reaction function has a slope lower than one. Consequently, the tendency for a $type-1$ panel to substitute towards a neutral citation dominates the tendency for a $type-3$ panel to substitute away from a neutral citation.

As α increases, making a positive citation becomes more attractive for a $type-3$ panel. As a result the $\bar{\kappa}$ reaction function shifts upwards, implying that a smaller proportion of $type-2$ panels issue neutral citations. This, in turn, increases the probability of an appeal under a neutral citation, thus reducing the incentives for $type-1$ panels to issue them. In the new equilibrium $\bar{\kappa}^*$ has increased and $\bar{c}^*$ has decreased.

Observable Citation Behaviour For the purposes of the empirical analysis, we now draw out two predictions of the theory that are couched in terms of observable citation behaviour. To do so, we first use Proposition 1 to derive the probability that the econometrician (unaware of panel type) will observe each of the four possible citations. Since negative citations never occur in equilibrium, the three relevant probabilities for no neutral and positive citation:

$$\Pr[M \in \{\varnothing\}] = \Pr[type = 1] \cdot (1 - \Pr[c < \bar{c}^*]) + \Pr[type = 2] \tag{15}$$

$$\Pr[M \in \{same\}] = \Pr[type = 1] \cdot \Pr[c < \bar{c}^*] + \Pr[type = 3] \cdot (1 - \Pr[\kappa < \bar{\kappa}^*]) \tag{16}$$

and

$$\Pr[M \in \{r = s_r\}] = \Pr[type = 3] \cdot \Pr[\kappa < \bar{\kappa}^*]). \tag{17}$$

Using (4)-(6) and Proposition 2 to establish how (15) - (17) depend on t and α, we immediately obtain the following two predictions:

Prediction 1. *An increase in "knowledge" t reduces the probability of no citation and increases the probability of neutral and positive citation from case i to j.*

Prediction 2. *An increase in "social preference" α increases the probability of no citation and positive citation and decreases the probability of neutral citation from case i to j.*

Notice that t and α have contrasting predictions relating to the probability of no citation and neutral citation. Since we will make use of this difference in the empirical analysis, it is worth emphasising the intuition. The reason why t decreases the probability of no citation should be obvious: $type-2$ (the type that is not aware of case j) is less likely. Similarly, the probability of neutral citation increases because $type-1$ (the type inclined to cite neutrally) is more likely. True, $type-3$ is also more likely, and further, is more likely to issue positive citations. But, for the reason stated in Proposition 2 ($\bar{c}^{*}+(1-\bar{\kappa}^{*})$ is increasing in t), the first effect dominates.

The reason why α increases the probability of no citation and decreases the probability of neutral citations is more subtle. The intuition is that, with negative citations ruled out, neutral citations can convey a less than ambivalent message, essentially by "damning with faint praise". As α increases, the panel in case i gives more thought to the externality this imposes on case j. Consequently, there is a tendency for a $type-1$ panel to substitute towards no citation and a $type-3$ panel to substitute towards positive citation.

3.3 The Impact of Network Distance

In Section 3.1 we asserted that both t and α are decreasing in network distance. While nothing in the theoretical model relied on it, the precise nature of this dependence forms an important part of our empirical strategy. In this final sub-Section before the empirics we give precise definitions of the knowledge diffusion and socialisation mechanisms which, when combined with Predictions 1 and 2, give us a set of empirical tests to apply to the data.

The Knowledge Diffusion Mechanism Recall that in the network of judicial collaborators a node is a case, a link between two nodes is a judge who sits on the panel in both cases and (given the dauntingly large population of previous cases) knowledge is awareness of the existence of other nodes. The simplest way for knowledge to diffuse through such a network is therefore for a judge to pass on information gained in the course of collaboration in one node (the case exists, the cases discussed by other panel members exist) to another node, again in the course of collaboration (writing the published opinion).[15]

Prompted by this observation, we specify the knowledge diffusion mechanism as follows. Panel members can learn of a case through two channels: experience (sitting on the panel in the case) and/or discussion (being informed of the case by a fellow panel member in the process of writing a published opinion). While the experience channel is uncontroversial, the discussion channel requires some explanation. Our specific assumption is that in the process of collaboration, each panel member informs his current colleagues of his past cases and the past cases of his past colleagues but not his past colleagues' past colleagues' past cases.

Two points are worth noting here. The first is that knowledge decays. Returning to Figure 2, in period $t+2$ Judge 4 tells Judges 6 and 7 of the existence of his past case B and his past colleague Judge 2's past case A. He would not, however, pass on information about the existence of, say, case Z heard by Judges 1, 11 and 12 in period $t-1$ since this is his past colleague Judge 2's past colleague Judge 1's past case. Consequently, cases at a distance

[15]Note that judges have an incentive to share information give the common preference to avoid appeals.

greater than 2 are unconnected. The second point is that timing is crucial. Case A must occur *before* case C for a citation of A by C to be considered. But, equally for knowledge diffusion to take place, case B must occur *after* case A. If case B occurred before case A, then Judge 2 cannot pass on knowledge of A to Judge 4 (since A cannot possibly be in 2's knowledge set) and so, by implication, Judge 4 cannot pass on knowledge of case A to case C.

We are now in a position to state precisely how, under the knowledge diffusion mechanism, t depends on distance at the date of case i. Case j is in case i's knowledge set if the two cases are directly linked with j occurring before i (Distance 1) or indirectly linked via at least one case occurring after j but before i (what we will term Distance 2/After j). In contrast, case j is not (or in practice is less likely to be) in case i's knowledge set if the two cases are indirectly linked via a case occurring before j (Distance 2/Before j), indirectly linked via a case occurring after i (Distance >2/After i) or never linked via a third case (Distance 2/None). Since t is the probability that any given prior case is in a panel's knowledge set, we immediately obtain:

Definition 1 (Knowledge Diffusion). *Dyads at Distance 1 and Distance 2/After j have a higher t than dyads at Distance 2/Before j, Distance >2/After i or Distance >2/None.*

Socialisation Mechanism We now turn to the dependence of the "social preference" parameter α on distance. The idea here is that judges become positively disposed to each other (or each other's rulings) through personal contact. Our specific assumption is that the personal contact implicit in collaboration imbues each panel member with a social preference α for his current colleagues (or their rulings) but not his current colleague's past colleagues.

There are again two points worth noting. First, in common with knowledge diffusion socialisation is also subject to decay.[16] The panel in case C has a social preference for case A because Judge 4 previously collaborated with Judge 2 (in case B). However, this panel will not have a social preference for case Z mentioned above since no member of the panel in case C has ever collaborated with a member of the panel in case Z. True, Judge 2 previously collaborated with Judge 1 (in case A) but, given decay, this social preference is not transmitted to Judge 4. Second, in contrast to knowledge diffusion, timing is less crucial. While case A must obviously still occur before case C to allow for a citation, case B could now have taken place *before* A. Since no knowledge is being transferred, all that matters is whether Judges 2 and 4 have ever collaborated before case C. Using these observations, we immediately obtain:

Definition 2 (Socialisation). *Dyads at Distance 1, Distance 2/After j and Distance 2/Before j have a higher α than dyads at Distance >2/After i or Distance >2/None.*

[16]It worth noting the difference between network decay and time decay. Network decay implies that α is lower at Distance >2 than Distance 2. In contrast, time decay implies that α is lower at *Distance 2/Before j* than *Distance 2/After j*. Since it is easy to argue that old "friendships" matter more than new "friendships", but also vice versa, we view the issue of time decay/growth as an empirical question best left to the data.

Empirical Tests We are now in a position to state the tests that we will use to answer the research questions posed in the Introduction. Namely, the network distance between i and j affect i's decision to cite j positively, neutrally, negatively, or not at all? And, to the extent that a causal effect exists, is the mechanism knowledge diffusion, socialisation, both or neither?

The essence of this exercise is to compare selection (citation vs. no citation) and outcome (positive vs. neutral) responses across the different groups of dyads noted in Definitions 1 and 2. Using this joint comparison we can infer substitution patterns and then, in light of Predictions 1 and 2, reject possible mechanisms for any causal effects. In practical terms, this amounts to conducting three tests − summarised in Table 2 − each with *Distance >2/None* as the base category.

The first of these tests, the *Placebo Test*, that compares selection and outcome responses across the *Distance >2/After i* group and the base category. Notice that, since an after i link cannot affect citation behaviour itself, a difference could only arise if links are correlated with some other citation relevant characteristic. Consequently, a rejection of the null for either response (rows 2-9 in Table 2) is a rejection of our key identifying assumption.

The second test, the *Socialisation Treatment Test*, compares responses across the *Distance 2/Before j* group and the base category. Here the important observation is that knowledge could not diffuse from j to i in either group but any socialising forces present in the *Distance 2/Before j* group could produce a substitution *away* from neutral citations. This insight aids interpretation. A positive or negative difference in either response (rows 2-9) could not have be driven entirely, or even in part, by knowledge diffusion.[17] However, a positive difference in the outcome response combined with a weakly negative difference in the selection response (rows 3 or 6) is consistent with socialisation.

The final test, the *Combined Treatment Test*, compares responses across the *Distance 2/After j* group and the base category. Now knowledge diffusion from j to i in the *Distance 2/After j* group could produce a substitution *towards* neutral citations. Since socialising forces produce the reverse, there is again the possibility of interpretation. For instance, a negative difference rejecting the null in the outcome response (rows 2, 5 and 8) could not have been driven entirely by socialisation, but in conjunction with a positive difference in the selection response (row 8), is consistent with knowledge diffusion.

4 Empirical Strategy

Table 2 describes the empirical tests suggested by the theoretical model and our discussion of the economic forces at work in each of the dyads groups. In this Section we describe how we conduct these tests using data on judicial citations and panel membership within the Civil

[17] The entries under the Socialisation Test in rows 2-9 signify that it would be incorrect to interpret the causal effect as knowledge diffusion (KD) or some combination of knowledge diffusion and socialisation (Both).

Division of the English Court of Appeal. This exercise is complicated by two issues. First, while it is straightforward to identify dyads where a citation occurred (*uncensored dyads*), it is less obvious how to identify *censored* dyads where a citation could have, but did not, occur. We discuss the methodology used to draw a choice based sample of censored dyads in Section 4.1. Second, identification clearly requires our measure of network distance to be orthogonal to other forces determining citation behaviour. We discuss the quasi-experimental method we use to approximate random assignment in Section 4.2. Finally, the point estimate and variance corrections needed to account for our choice based sampling design are briefly reviewed in Section 4.3.

4.1 A Choice Based Sample of Dyads

We construct our data set in two stages: first identifying a population of dyads and then drawing a choice based sample. To construct the population, we start with the 8,923 cases heard in the Civil Division of the English Court of Appeal between 1 January 1980 and 31 December 2005 that are available on Westlaw UK. Without loss of generality, we think of these 8,923 cases as putative citing cases and index them (in descending chronological order) by $i = 1, ..., 8,923$. Taking case 1, we then return to the remaining 8,922 cases and find every judgement given within the preceding 5 years. Matching each of these 'case j's' with case 1 gives us the population of *Within 5 year Court of Appeal (Civil Division) Westlaw Reported* case $1-$ case j dyads. Repeating this exercise for $i = 2, ..., 8923$ cases we obtain a population of $23,551,316$ *Within 5 year Court of Appeal (Civil Division) Westlaw Reported* case $i-$ case j dyads (hereafter simply dyads).

It is helpful to divide this large population into 'year of citing case' strata. Following standard notation, we let $h = 1, ..., L$ denote the 'year of citing case' where, in the full sample, $L = 26$. Further, we let $n = 1, ..., U_h$ index the uncensored dyads within strata h and $n = U_h + 1, .., M_h$ the censored dyads within strata h. Here M_h is number of dyads (observations) within strata h, with U_h the number of uncensored dyads and $C_h = M_h - U_h$ the number of censored dyads. The size of the population, M, can therefore be written as,

$$M = \sum_{h=1}^{L} M_h = \sum_{h=1}^{L} U_h + C_h = 23,551,316.$$

For each strata, we first select all U_h uncensored dyads and then draw (with replacement) a sample of censored dyads of size $c_h = U_h \cdot 10$. The size of our sample, m, is therefore

$$m = \sum_{h=1}^{L} m_h = U_h + c_h = 1,412 + 14,120 = 15,532.$$

Under this sample design, an uncensored dyad n in strata h has a sampling weight $w_{hn} = U_h/U_h = 1$, while a censored dyad n in strata h has a sampling weight $w_{hn} = C_h/c_h = (M_h - U_h)/U_h \cdot 10$. Since Westlaw appeared to change its own sampling methodology in 1993 (see Table A1), rather than use all 26 strata we focus only on 13 strata between 1993 and 2005. This gives us a marginally smaller population size of 23,517,932 and, as Table 3 shows, a smaller sample size of 15,445 observations, 14,050 of which are uncensored. We further exclude two groups of dyads. The first group includes the 279 dyads which display a non-discretionary citation. As we argued above, this type of citations are unlikely to be choice variables of the case panel and need to be excluded from our empirical analysis. The second group refers to dyads in which case i occurs after October 2004. Since we lack information on the first Michaelmas Term following i these dyads cannot be allocated to the *Distance>2/After i* group. As a result the orthogonality of network distance to other case-level characteristics, which we need for our empirical strategy is violated. To ensure that our findings are not contaminated by omitted variable bias, we drop these 2,144 observations. We therefore have 13,032 observations.

4.2 Random Assignment

Our empirical strategy relies strongly on our ability to obtain an exogenous measure of network distance. To have any hope of establishing causality, it is essential that distance be uncorrelated with (unobservable as well as observable) judge and case-level characteristics likely to influence citation behaviour.

To understand why correlation between our measure of distance and a judge-level characteristic would bias our results, suppose that in half of our dyads a judge sitting on the committee in case i is an old university friend of a judge sitting in the committee in case j but that the judges in the other half of our dyads have no such social connections. If 'old friends' are able to engineer meetings on panels, then the socially connected dyads are likely to be over-represented in group of dyads at Distance ≤ 2, while the socially unconnected dyads are likely to be over-represented in group of dyads at Distance >2. Since 'old friends' are likely to gain pleasure from casting each in a positive light, the socially connected dyads are also likely to engage in systematically different citation behaviour to the socially unconnected dyads. This omitted judge-level characteristic would bias our tests, leading us to wrongly conclude that network distance has a causal effect on citation behaviour.

The reason why case-level characteristics might matter is more mechanical but equally important. Suppose that in half of our dyads case i took place 4 years after case j, while in the other half of our dyads case i took place just 1 year after case j. Clearly, the dyads with the longer time difference are more likely to belong to the Distance 2 After j group while the dyads with the shorter time difference are more likely to be at *Distance 2/Before j*. The reason is that, for the 4 years time difference dyads, there is plenty of time for the connecting case to have taken place after case j but before case i. For the 1 year time difference dyads there is little time for this to occur. Since recent cases are more likely to be relevant and hence cited, our tests would now be biased due to the existence of an omitted case-level characteristic.

We address the issue of possible correlation between distance and judge-level characteristics in the following way. First, in Section 4.2.1 we test for random panel formation in initial judge matches at the start of the legal year. The results of this test substantiate our hypothesis that, at the start of the legal year, panels within the Civil Division of the Court of Appeal *are indeed* formed via random draws from the pool of ticketed judges. As we discuss in Section **??**, data limitations prevent us from testing the notion that the cab-rank principle is strictly applied during the remainder of the legal year. Since our hypothesis that initial matches are random can be tested and fails to be rejected by the data, we use these initial (Michaelmas Term) matches to construct our measures of distance.

To avoid mechanical correlation with case characteristics (principally time difference), we use these Michaelmas Term matches only at pre-specified points in time. As we discuss in Section 4.2.2, this quasi-experimental approach – that is in the spirt of regression discontinuity design – approximates random assignment. This is confirmed, in Section **??**, by our finding that observable judge-level and case-level characteristics are largely uncorrelated with our measures of network distance. However, to be absolutely certain that omitted variables are not biasing our results we conduct both simple comparison of means tests and regression tests. In our regression tests we find that our point estimates are unchanged by the introduction of judge and case-level controls. As Bertrand et al [3] argue, this represents further evidence that our results are unlikely to be caused by the existence of omitted variable bias.

4.2.1 Testing for Random Panel Formation

In Section 2 we discussed how Court of Appeal Listing Officers form committees at the start of the legal year. We now test whether these initial matches (not dictated by the cab-rank principle) can realistically be viewed as random draws.

Between January 1 1980 and December 31 2005, 123 individuals held the title Lord Justice of Appeal. We have sufficient case data to undertake a test of random committee formation for 91 of these 123 appeal judges. Our procedure echoes Ashenfelter et al's [1] test of random case assignment. Specifically, we compare the proportion of *initial cases* matched with a given judge across *contemporaries*. Before illustrating the steps involved, it will be helpful to define precisely what we mean by initial cases and contemporaries.

Since our objective is to test for random initial conditions, the ideal strategy would be to restrict attention to committees formed on or shortly after October 1. Unfortunately, since we observe judgement but not commencement dates, this is not possible. Instead, we deem a case to be an *initial case* if its judgement date falls between October 1 and December 21 (i.e. in the first term of the legal year).[18] In view of the end of term effects apparent in Figure 1, the committees in these cases will plausibly be filled early in Michaelmas Term before the

[18]If October 1 falls on a weekend, the term begins on the following Monday; if December 21 falls on a weekend, the term ends on the proceeding Friday.

cab-rank principle (and the concomitant serial correlation) is at work.

Turning to the notion of contemporaries, it clear that a match can only occur if the judges concerned have overlapping periods of service. We account for this potential source of non-randomness in two stages. First, we identify the set of appeal judges whose periods of service overlap with a given judge x. Then, given our focus on Michaelmas Term, we exclude judges who were either appointed before judge x's first Michaelmas Term or after his/her last Michaelmas Term. We define the remaining individuals as judge x's *contemporaries*; the idea being that these are judges who could have been matched with judge x on an initial case.

Of course, adopting the above definition of an initial case, raises the issue of whether to use all of a judge's cases in a Michaelmas Term when calculating his/her proportion matched with judge x. To mitigate the possibility of within-term serial correlation, we use a binary rather than continuous match variable. Specifically, we construct a dummy variable $matched_{nt}$ that takes a value of 1 if contemporary n is matched with judge x in at least one case in Michaelmas Term t and 0 if contemporary n is not matched with, but still serves alongside, judge x in the same Michaelmas Term t.

With these definitions in place, our test of random committee formation proceeds as follows. For each judge $x = 1, ..., 91$ we:

1. Identify his/her set of contemporaries (and denote its cardinality by N_x).

2. For each contemporary $n = 1, ...N_x$, construct the dummy variables $matched_{nt}$ and *contemporary*$_n$ (an individual fixed effect).

3. Estimate the following equation,

$$matched_{nt} = \alpha + \beta_1 contemporary_1 + ... + \beta_{N_x - 1} contemporary_{N_x - 1} + \varepsilon_{nt},$$

   and obtain the $p - value$ from an $F - test$ for the joint significance of the $N_x - 1$ included contemporary dummies.

If our null hypothesis of random committee formation is correct, the $p - values$ obtained from these 91 regressions, should be distributed $U[0, 1]$. That is, the fraction of the regressions with $p - values$ of .05 or less should be .05, the fraction of the regressions with $p - values$ of .10 or less should be .10 and so on.

One final outstanding issue is time allocation across divisions. For instance, a criminal specialist who spends most of his/her time in the Criminal Division of the Court of Appeal is unlikely to be matched with a civil specialist who spends most of his/her time in the Civil Division of the Court of Appeal (particularly relative to another civil specialist). To explore this possibility, Figure 4 reports our results disaggregated by time spent in the two respective divisions.

Our findings are as follows. For the 27 judges with 100% of their cases in the Civil Division of the Court of Appeal cases, there is strong evidence of random committee formation. The fraction of regressions with $p-values$ of .05 or less is .07 (2 judges). Moreover, as Figure 4 illustrates, the entire distribution closely approximates a uniform CDF. The approximation is even closer if we further exclude the Heads of Division from our analysis. The evidence is more mixed for the judges who also hear cases in the Criminal Division of the Court of Appeal, due to the non-randomness introduced by movements across courts. For the 58 judges with 50-99% of their cases in the Civil Division of the Court of Appeal cases, the fraction of regressions with $p-values$ of .05 or less is .15 (9 judges). The overall distribution is further from the 45 degree line. Finally, for the remaining 6 judges with 10- 50% of their cases in the Civil Division of the Court of Appeal cases, the fraction of regressions with $p-values$ of .05 or less is .33 (2 judges). On the basis of these results, it seems fair to conclude that initial matches *within our court of interest*, namely the *Civil Division* of the Court of Appeal, can realistically be viewed as random draws.

4.2.2 Assigning Dyads to Network Distance Groups

The above test substantiates our presumption that, at the start of the legal year, committees within the Civil Division of Court of Appeal are formed at random from the pool of ticketed judges. Ideally, we would also test whether the cab-rank principle was strictly applied during the remainder of the legal year. Clearly, if this second claim was true, every case committee throughout the year would be formed at random and we could safely use all committee connections to assign dyads to Network Distance Groups. Unfortunately we have judgement but not commencement dates. As a result such a test is not possible. For this reason, our measure of network distance uses only initial matches.

To construct this measure, we assign each observation (dyad) to a different Group, depending on the network distance between the two cases in the dyad. We can observe the four groups in which we base our empirical tests (our 'treatment groups') in Figure 5. In order to qualify to be assigned to one of our four treatment groups, a dyad must meet two conditions. First, there can be no common judges. In other words, i and j cannot be directly connected, and the network distance between the two cases must be at least 2. Second, there must be a least a pair of $i-j$ judges that are *contemporaries* (as defined in the previous Section) at some point.

Consider a dyad meeting both conditions, with a case panel i comprising of judges $i1$, $i2$ and $i3$ and a case panel j comprising of judge $j1$, $j2$ and $j3$. These two cases are indirectly connected when there is (at least) one other case including at least one judge from panel i and at least one judge from panel j. In Figure 5 we have displayed an example of such connecting case, with a panel comprising of judges $i1$, $j1$ and a third judge, $h3$. Our assignment of a dyad to the four treatment groups depends both on the existence and on the timing of such connecting case. We assign dyad ij to our first treatment group, *Distance 2/Before* j, if i and j are connected *exclusively* by other cases which take place during the first Michaelmas Term before case j. We assign dyad ij to the *Distance 2/After* j group if there are cases connecting i and j and they occur exclusively in the first Michaelmas Term after case j.

In our third and fourth treatment groups there are no cases connecting i and j which occur either in the first Michaelmas Term before or in the first Michaelmas Term after j. As a result the network distance between i and j is greater than two at the date of case i. We assign dyad ij to the *Distance >2/After* i group if i and j are connected by cases which occur in the first Michaelmas Term after i. Lastly, we assign dyad ij to the *Distance >2/None* group if i and j are not connected by any case which takes place in any of the three bordering Michaelmas Terms (immediately before j, immediately after j and immediately after i) described above.

There are of course other potential network distance groups other than the four 'treatment groups' in which we base our tests. Table 3 splits the database into several other groups, including the Distance 1 group, the group in which judges in case i and judges in case j are never each other contemporaries, and different combinations of Distance 2 dyads.

Why do we use only small temporal windows (i.e. one Michaelmas Term) around i and j when allocating dyads into network distance groups? At the beginning of this Section we argued that a mechanical correlation could easily arise between case-level characteristics (in particular time difference between i and j) and our measure of connectivity. By restricting our attention to connecting cases which occur shortly before or after the dyad cases, we reduce the disparities in the time differences between i and j across groups and therefore the importance of this mechanical correlation. This strategy is similar to the regression discontinuity technique. To see this, consider two dyads, ij_1 and ij_2. Assume that ij_1 belongs to the *Distance2/Before* j group, while ij_2 belongs to the *Distance2/After* j group[19]. In the first dyad a connection happened shortly before ij, whereas in the second dyad it happened shortly after ij. This fact, together with the conclusion in Section 4.2.1 that the process generating connections can be regarded as random, implies that dyads ij_1 and ij_2 are statistically identical in terms of observables (including time difference) and, we should hope, unobservables.

4.3 Estimation

In Section 3 we concluded that the *Placebo Test*, *Socialisation Treatment Test* and *Combined Treatment Test* allow us to identify the knowledge diffusion and socialisation effects of connectivity. These three tests require comparing the proportion of dyads with a citation and the proportion of cited dyads with a positive cite across the four treatment groups described in 4.2.2. In this Section we outline the estimation procedures which we follow to perform these

[19]Our tests do not include the direct comparison of *Distance 2/Before* j and *Distance2/After* j. However, this technique also limits any difference between these groups and the *Distance>2/None,* to which they are compared.

comparisons. First, we discuss the advantages and drawbacks of performing a simple comparison of means, together with the explicit formulas of the estimators used in this comparison. Second, we discuss the logic of complementing these findings with probit and sample selection models.

4.3.1 Comparisons of Means

The simplest estimator of the difference in the proportions of dyads with a citation across two groups is the difference in their sample proportions. A test based on the sample difference is valid as long as we are certain that the location of a dyad into the alternative network distance groups is uncorrelated with any other variable determining citation behaviour. Apart from its simplicity, a comparison of means (or proportions) has the advantage of not imposing any particular functional form in the estimation of the relation between network distance and citation behaviour.

We calculate the difference in the sample proportions as follows. Let $X_{hn} \in \{0,1\}$ denote a binary citation outcome variable for dyad n in strata h, and x_{hn} the corresponding variable for *sampled* dyads n in strata h. Moreover, let S denote the set of dyads that belong to the subpopulation of interest, where the variable I^S_{hn} indicates whether a dyad belongs to this subpopulation. Recalling that the population size is M, the population proportion for this binary citation outcome variable is

$$P^S = \frac{X^S}{M^S} = \frac{\sum_{h=1}^{L}\sum_{n=1}^{M_h} I^S_{hn} X_{hn}}{\sum_{h=1}^{L} I^S_{hn} M_h}.$$

Using the sample weights, w_{hn}, defined in Section 4.1, the estimator for the population proportion, $\widehat{P}$, can be written as[20]

$$\widehat{P}^S = \frac{\widehat{X}^S}{\widehat{M}^S} = \frac{\sum_{h=1}^{L}\sum_{n=1}^{m_h} I^S_{hn} w_{hn} x_{hn}}{\sum_{h=1}^{L}\sum_{n=1}^{m_h} I^S_{hn} w_{hn}}.$$

4.3.2 Binary Response Models

We complement our comparisons of means with the estimation of binary response nonlinear models. In each of these models we include a set of dummies which account for our different treatment groups. For instance, we include a dummy $D_{ij}^{Distance2/Before\ j}$ which takes value 1 if the dyad ij belongs to the *Distance2/Before j* group and value 0 if it does not.

[20] The variance of this estimator, $V(\widehat{P})$, can be approximated as follows. First define the score variable by $z_{hn}(\widehat{P}^S) = \frac{\partial \widehat{P}^S}{\partial w_{hn}} = I^S_{hn}\frac{x_{hn}-\widehat{P}^S}{\widehat{M}^S} = I^S_{hn}\frac{\widehat{M}^S x_{hn}-\widehat{X}^S}{\left(\widehat{M}^S\right)^2}$ and the weighted total of the score variable by $\widehat{Z} = \sum_{h=1}^{L}\sum_{n=1}^{m_h} I^S_{hn} w_{hn} z_{hn}$. Then obtain the variance $V(\widehat{Z})$ as $V(\widehat{Z}) = \sum_{h=1}^{L}\frac{m_h}{m_h-1}\sum_{n=1}^{m_h}\left(I^S_{hn} w_{hn} z_{hn} - \overline{z}_h\right)^2$, where $\overline{z}_h = 1/m_h\left(\sum_{n=1}^{m_h} I^S_{hn} z_{hn}\right)$. This variance estimator is an approximation to $V(\widehat{P})$ and is algebraically equivalent to the variance estimator derived from directly applying the delta method.

The excluded category in each of the regressions is the *Distance>2/None* group. Comparing groups *Distance>2/Before j* and *Distance>2/None* (the *Socialisation Treatment Test*) is therefore equivalent to examining the coefficient of $D_{ij}^{Distance2/Before\ j}$. Similarly, the coefficient of $D_{ij}^{Distance2/After\ j}$ represents the *Combined Treatment Test*, while the coefficient of $D_{ij}^{Distance>2/After\ i}$ captures the *Placebo Test*.

Our objective is to compare the proportion of dyads with a citation and the proportion of cited dyads with a positive cite. We estimate two sets of binary models. First we run two independent Probits for each of these two variables. We then run a jointly estimated binary sample selection model. The advantage of this second model is that it allows for cross-correlation in the error terms of the two binary models, thus estimating jointly the decisions of whether to cite positively, neutrally or not at all. The formulas for the estimators based on these models are standard (see, for example, Greene [12] Ch. 21) and we do not reproduce them here. We are, however, careful to adjust the estimators of the coefficients and standard errors to account for our choice-based stratified sampling strategy, using the sampling weights, w_{hn}, defined in Section 4.1.

The purpose of including these two binary response models is to provide a robustness test to the comparison of means estimators. Unfortunately, the likelihood function implied by our theoretical model is complex and does not correspond with standard non-linear models for binary response. Notwithstanding this, standard nonlinear models do have the advantage that they provide a simple way to include a set of case and judge level controls and to examine whether our estimates for the dummy variables of interest vary with the introduction of these controls. A finding that these estimates remain unchanged provides additional evidence that the relation between network distance and citation behaviour is unlikely to be contaminated by the presence of omitted variables.

5 Results

In Table 3 we report the number of observations in each of the network distance group cells. Note that this table reports sample rather than population sizes. While the number of population dyads is 23,517,932 our "treatment" sample only contains 13,032 dyads. Out of these, 3,249 dyads belong to one of the four "treatment" groups which we use in our empirical tests. The number of dyads displaying a positive citation is 423, while 471 dyads display a neutral citation. It is comforting to observe that the number of negative citations is extremely low (just 5 citations). This is consistent with our theoretical model in which negative citations are both costly *and* increase the likelihood of appeal and are thus avoided. In the remainder of this Section we discuss our empirical findings. We first show that our network distance groups are indistinguishable in terms of a number of judge and case observable characteristics. We then present the results of our main empirical tests.

5.1 Case and Judge Characteristics by Network Distance Group

Table 4 displays the average number of cases connecting i and j in a dyad by network distance group. In the average dyad of the *Distance 2/Before j* group there are 4.03 cases which connect i and j in the Michaelmas Term before j. By definition, there are 0 connecting cases in the Michaelmas Term after j and also 0 connecting cases in the Michaelmas Term after i. A similar discussion can be made of the average dyads of the *Distance 2/After j* group and the *Distance>2/After i* group. By definition, there are no connecting cases in the dyads of the *Distance>2/None* group. Interestingly, note that the number of *total* connecting cases is remarkably similar across the first three treatment groups: 4.03, 3.85 and 3.89. This finding is confirmed by Figure 6, which plots the connecting case distribution for each of these three groups. With the exception of the *Distance 2/Before j* at very low numbers of connecting cases, the distributions for the three groups are virtually identical.

In Table 5 we compare a number of case characteristics across different network distance groups. Our first set of variables are the time difference between i and j, whether i and j share the same legal subject, whether j has appeared in the Times Law Reports and the number of times that j has been mentioned in Legal Journal Articles. We find that none of the first three treatment groups are statistically different in terms of their mean from the *Distance>2/None.* Figures 7 and 8 confirm that a similar conclusion can be drawn in terms of their overall distributions. In Figure 7 we plot the distribution of the time difference between case i and case j in numbers of years. With the partial exception of the *Distance>2/After i*, we find that the percentages of dyads at each interval of the time difference distribution are effectively the same. In Figure 8 we plot the distribution of the number of Legal Journal Articles mentioning case j, and reach the same conclusion.

The last two columns of Table 5 display the panel sizes of the case j and case i. In terms of these variables we find meaningful differences across network distance groups. The panel sizes of the i and j cases in the *Distance>2/None* group are a bit smaller than those of the other groups and the differences are statistically significant. While this is potentially troubling, note there is a chance that this may be a purely spurious finding: the higher the number of comparisons that we perform, the more likely it is that one of these comparisons will appear as statistically significant. More reassuring is the fact that when we include panel sizes as controls, the coefficients of our regressions below are virtually unchanged.

Lastly, we compare several judges characteristics across treatment groups in Table 6. Our variables are the following: whether at least one judge pairs attended the same secondary school at the same time, whether at least one judge pairs attended the same university (or college in case of Oxford and Cambridge) at the same time, whether at least one judge pairs practised as barristers from the same chambers prior to becoming judges and whether at least one judge pairs are members of the same (Gentleman's or sports) club. We fail to reject the hypothesis that the first three treatment groups have the same mean as the *Distance>2/None* group in all but one occasion (the *Distance>2/After i* groups seems to have a disproportionately large number of dyads in which at least one judge pairs practised as barristers from the same chambers). In sum, it seems reasonable to conclude that our treatment groups are statistically equal in terms of (almost) every observable variable.

5.2 Citation Outcomes by Network Distance Group

In Table 7 we report the proportion of dyads with a citation across treatment groups. The proportions are lowest for the *Distance 2* groups and largest for the *Distance>2* groups. However, the comparison of means and the probit regression with dummies outlined in Section 4.3 and displayed in Column 1 of Table 7 are unable to reject the hypothesis that these proportions are identical across the four groups. In Column 2 we add the case-level controls discussed in the preceding Section, all of which appear strongly significant at explaining the likelihood of a citation. In Column 3 we add the judge social network controls also discussed above. The coefficients of the *Distance 2/Before j*, *Distance 2/After j* group and *Distance>2/After i* dummies remain insignificant at conventional levels.

Table 8 displays the proportion of cited dyads with a positive cite by treatment groups. The proportions are highest for the *Distance 2* groups and lowest for the *Distance>2* groups. A comparison of means and a probit regression without controls in Column 4 are unable to reject the hypothesis that this proportion is identical across the *Distance>2/After i* and *Distance>2/None* groups. The difference between *Distance 2/Before j* and *Distance>2/None* is of 18 percentage points, and statistically significant at the 10% level. The difference between *Distance 2/After j* and *Distance>2/None* also economically large (19 percentage points) and significant at the 5% level. These findings are robust to the introduction of case-level controls (in Column 5) and of judge social network controls (in Column 6). A comparison of Columns 4-6 reveals that the coefficients for the *Distance 2/Before j* and *Distance 2/After j* group dummies are virtually the same in the three regressions. This finding strengthens our belief that our empirical findings are not due to the presence of omitted variable biases.

We proceed now to present the results of estimating a binary sample selection model. Remember that in the selection equation of this model the endogenous variable is whether the dyad displays a citation. The outcome equation captures, when a citation occurs, whether the citation is positive or neutral. Table 9 displays the coefficients and standard errors arising from estimating an outcome equation jointly with a selection equation, as part of a binary sample selection model[21]. The most important finding in Table 9 is that the coefficients of the outcome equation in the two-stage model are very similar to those in Table 8, where whether a citation is positive (conditional on a citation occurring) is estimated as an independent probit model. The coefficients for the *Distance 2/Before j* and *Distance 2/After j* group dummies are

[21]The selection equation of this model is equivalent to a standard probit estimation and therefore be found in Table 7.

statistically significant at the 10% and 5% respectively. Similarly to Table 8 the introduction of case-level and judge social network controls does not vary the value of the coefficients or their statistical significance.

5.2.1 Identification and interpretation of our results

How can we interpret our results? Note first that the *Placebo Test* fails to reject the hypothesis that *Distance>2/After i* and *Distance>2/None* are identical in terms of their citation behaviour and thus our main identifying assumption of no omitted variable bias.

The interpretation of our findings must be done by combining the coefficients from the selection equation and the outcome equation. Consider first the Socialisation Test, which arises from comparing the *Distance 2/Before j* and *Distance>2/None* groups. Our results can be located in row 3 of Table 2. The fact that the overall number of citations remains unchanged while the ratio of positive versus neutral citations increases implies *a fortiori* that the number of positive citations increases while the number of neutral citations decreases. As Table 2 shows, this finding rejects the notion that knowledge diffusion is the only economic effect of network connectivity, while being consistent with the hypothesis that socialisation is a force at work.

Consider now the Combined Test, which arises from comparing the *Distance 2/After j* and *Distance>2/None* groups. Again, our results are located in row 3 of Table 2, which implies that neutral citations decrease and positive citations increase as network distance decreases. This finding is again consistent with the presence of socialisation, while strongly rejecting the notion that the only effect of stronger connectivity across case panels is to facilitate the diffusion of knowledge between them.

6 Concluding Remarks

In this paper we have attempted to understand how distance affects team decisions in a network of collaborators (NOCs). In particular our focus has been on the citation decisions among judges in the Civil Division of the English Court of Appeal. To identify and interpret adequately the effect of network distance, we have exploited two characteristics typical of NOCs − the existence of directed links and the qualitative nature of decision externalities − and one specific to our institutional context − the fact that links among team of judges are generated randomly.

We have found that network distance does indeed determine citation decisions among English senior judges. Our emphasis has been on understanding the economic mechanism behind this effect, in particular the relative importance of knowledge diffusion versus socialisation. We have found no evidence suggesting that knowledge diffusion is a force at work and strong evidence contradicting the notion that it is the only force through which network distance affects citation behaviour. Our findings are instead consistent with a socialising process occurring among judges who work together in case committees.

While our paper is the first to discuss how features typical of NOCs can help establish causality and interpret the effect of network distance, we are not the first to examine these networks. To the best of our knowledge previous work (see, for instance, Fafchamps et al [8] for academic networks, Singh [19] for patent co-inventors and Gomes-Caseres et al [11] for joint ventures) has exclusively emphasised the knowledge diffusion effects of connectivity in NOCs. The reasons why socialising forces are prevalent among senior judges probably do not include their omniscience. Remember that a judge in our data can expect to sit in cases of any kind of (civil) law. The issue of imperfect knowledge is probably at least as prevalent among senior judges as in other contexts. In other ways, however, the judiciary might be different from other institutions. In particular, the formal knowledge management systems – including internet case law providers like Westlaw – now available to senior judges might have decreased the costs of finding past cases on related legal issues, thus limiting the role of personal connections.

In a more subtle way, the random nature of the link formation process in the Court of Appeal, which we use to establish causality, might itself be responsible for the socialising effect of connectivity. A recent paper, (Jackson and Rogers [14]) suggests that random link formation dominates friendship networks while a more systematic link formation process is prevalent in other types of networks. While we find that socialising forces are the effect of connectivity in the current (random) network, we cannot discard the possibility that different economic forces might dominate under other case assignment procedures.

Our findings have implications for the study and design of judicial institutions. First, they indicate that citation measures are indeed contaminated by the network of social relations. Extreme caution is therefore necessary when using citation counts as measures of judicial influence or benchmarks to evaluate the ability of different judges. Second, our findings suggest that little knowledge is being transmitted among randomly matched judges. Whether this would be different under other case allocation processes – such as matching judges with different expertise – is of course not assured. As noted above, Jackson and Rogers' work, together with the body of evidence emphasising the potential role of NOCs for social learning, suggests that this might be the case.

References

[1] Ashenfelter, Orley, Theodore Eisenberg and Stewart J. Schwab "Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes", *Journal of Legal Studies,* 1995, 24:2 257-281.

[2] Bandiera, Oriana, Iwan Barankay and Imran Rasul "Social Connections and Incentives in the Workplace: Evidence from Personnel Data", 2006 mimeograph.

[3] Bertrand, Marianne, Erzo Luttmer and Sendhil Mullainathan "Network Effects and Welfare Cultures", *Quarterly Journal of Economics,* 2000, 115:3, 1019-1055.

[4] Choi, Stephen and Mitu Gulati "A Tournament of Judges?", *California Law Review,* 2004, 92, 299-325.

[5] Conley, Tim and Chris Udry "Learning about a new technology: Pineapple in Ghana", 2005 mimeograph.

[6] Cross, Frank B., Thomas A. Smith and Antonio Tomrachio "The Reagan Revolution in the Network of Law", 2006 mimeograph.

[7] Duflo, Esther and Emmanuel Saez "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment", *Quarterly Journal of Economics,* 118(3), 2003, pp.815-842.

[8] Fafchamps, Marcel, Marco van der Leij and Sanjeev Goyal "Matching and Network Effects", mimeograph 2006.

[9] Gennaioli, Nicola and Andrei Shleifer "The Evolution of Common Law", mimeograph 2006.

[10] Glaeser, Edward, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer "Do Institutions Cause Growth?" *Journal of Economic Growth,* 2004, 9:3, 271-303.

[11] Gomes-Caseres, Benjamin, John Hagedoorn and Adam Jaffe "Do Alliances Promote Knowledge Flows?", *Journal of Financial Economics*, 2006, 80:5, 1-33

[12] Greene, William *Econometric Analysis*, 2000, Prentice Hall

[13] Fischmann Joshua "Decision Making Under a Norm of Consensus: An Analysis of Three-Judge Panels", 2006, mimeograph.

[14] Jackson, Matthew and Brian Rogers "Meeting Strangers and Friends of Friends: How Random are Social Networks?", *American Economics Review*, forthcoming

[15] Kramarz, Francis and David Thesmar "Beyond Independence: Social Networks in the Boardroom", 2004 mimeograph.

[16] Landes, William M., Lawrence Lessig and Michael Solimine, "Judicial Influence: A Citation Analysis of Federal Courts of Appeal Judges", *Journal of Legal Studies,* 1998, 27:2, 333-372.

[17] Levy, Gilat "Careerist Judges and the Appeals Process", *Rand Journal of Economics*, 36:2, 2005, 275-297.

[18] Munshi, Kaivan "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market", *Quarterly Journal of Economics,* 2003, 118:2, 549-598.

[19] Singh, Jasjit "Collaborative Networks as Determinants of Knowledge Diffusion Patterns", *Management Science,* 51(5), 2005, 756-770.

[20] Sunstein, Cass, David Schkade, Lisa Ellman and Andres Sawicki, "Are Judges Political? An Empirical Investigation of the Federal Judiciary", Brookings Institution Press, 2006.

Appendix

Proof of Proposition 1. First, note that (13) and (14) are decreasing in c and κ respectively. This implies that if cost c' is equal to (13), then every $c < c'$ strictly prefers to issue a citation than no citation at all, and similarly for a cost κ' equal to (14). This proves the second part of the proposition.

Now consider the existence of equilibrium. To save notation, let $p_{\oslash} = \Pr(d = x|citation_{ij} = 0)$, $p_n = \Pr(d = x|neutral_{ij} = 1)$, $p_p = \Pr(d = x|\,\text{positive}_{ij} = 1)$, $\delta_1 = [\Pr(d = x|type_{ij} = 1) - \delta]$ and $\delta_2 = [\Pr(d = x|type_{ij} = 2) - \delta]$. Using the fact that the costs follow a uniform distribution, we have that $b_1 = \Pr(c < c') = c'$ and $b_2 = \Pr(\kappa < \kappa') = \kappa'$, which we can substitute in (11) and (12). An equilibrium is then characterised by the following two conditions:

$$[p_n(c', \kappa') - p_{\oslash}(c')]\,\delta_1 - c' \;=\; 0 \tag{18}$$

$$[p_p - p_n(c'', \kappa'')]\,\delta_2 + \alpha - \kappa'' \;=\; 0 \tag{19}$$

An equilibrium exists if $c^* = c' = c''$ and $c^* = \kappa' = \kappa''$.

Note that $\frac{\partial[p_n(c',\kappa') - p_{\oslash}(c')]}{\partial c'}\delta_1 - 1 < 0$, $\frac{\partial p_n(c',\kappa')}{\partial \kappa'}\delta_1 < 0$ and $-\frac{\partial p_n(c'',\kappa'')}{\partial c''}\delta_2 > 0$. It can also be shown (after an algebraic inferno) that $-\frac{\partial p_n(c'',\kappa'')}{\partial \kappa''}\delta_2 - 1 < 0$.

Differentiating (18) we have then that $\frac{d\kappa'}{dc'} = -\frac{\frac{\partial[p_n(c',\kappa') - p_{\oslash}(c')]}{\partial c'}\delta_1 - 1}{\frac{\partial p_n(c',\kappa')}{\partial \kappa'}\delta_1} < 0$. Differentiating (19) we have that $\frac{d\kappa''}{dc''} = \frac{-\frac{\partial p_n(c'',\kappa'')}{\partial c''}\delta_2}{\frac{\partial p_n(c'',\kappa'')}{\partial \kappa''}\delta_2 + 1} > 0$. Their monotonicity implies that (11) and (10) intersect at most once. As a result an equilibrium, if it exists, must be unique. Do (11) and (10) intersect at all? Note that $\kappa''(c'' = 0) = \alpha < 1 = \kappa'(c' = 0)$. Lastly, note that $c'(\kappa' = 0) < 1$, which derives from the fact that $[p_n(c', \kappa') - p_{\oslash}(c')] < 1$ and $\delta_1 < 1$. This implies that c and κ do intersect in the $[0, 1] \times [0, 1]$ space. ■

Proof of Proposition 2. Consider the equilibrium given by the conditions:

$$[p_n(c^*, \kappa^*) - p_{\oslash}(c^*)]\,\delta_1 - c^* \;=\; 0 \tag{20}$$

$$[p_p - p_n(c^*, \kappa^*)]\,\delta_2 + \alpha - \kappa^* \;=\; 0 \tag{21}$$

Differentiating with respect to α, c^* and κ^* and using the inequalities derived in the previous subsection we have that

$$\frac{dc^*}{d\alpha} = -\left[-\frac{\partial p_n(c^*, \kappa^*)}{\partial c^*}\delta_2 + \left(\frac{\partial\,[p_n(c^*, \kappa^*) - p_{\oslash}(c^*)]}{\partial c^*}\delta_1 - 1\right)\frac{1 + \frac{\partial p_n(c^*,\kappa^*)}{\partial \kappa^*}\delta_2}{\frac{\partial p_n(c^*,\kappa^*)}{\partial \kappa^*}\delta_1}\right]^{-1} < 0$$

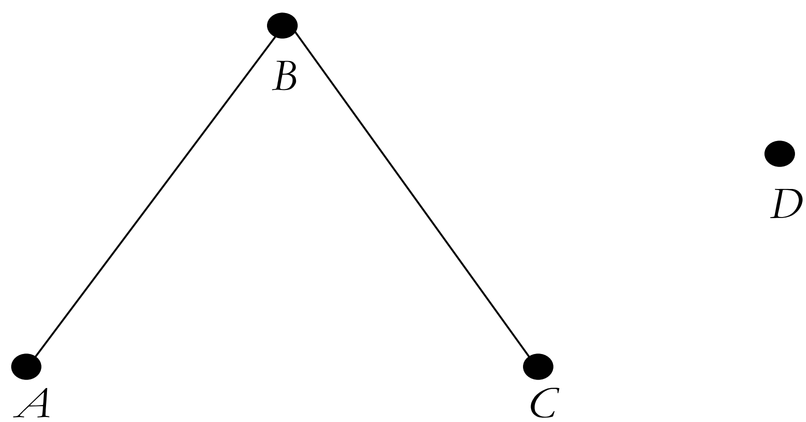
$$\frac{d\kappa^*}{d\alpha} = \left[1 + \frac{\partial p_n(c^*, \kappa^*)}{\partial \kappa^*}\delta_2 + \frac{\partial p_n(c^*, \kappa^*)}{\partial \kappa^*}\delta_1 \frac{-\frac{\partial p_n(c^*,\kappa^*)}{\partial c^*}\delta_2}{\frac{\partial[p_n(c^*,\kappa^*) - p_{\oslash}(c^*)]}{\partial c^*}\delta_1 - 1}\right]^{-1} > 0$$

Similarly, differentiating with respect to t, c^* and κ^* we have that

$$\begin{aligned}
\frac{dc^*}{dt} &= \frac{\partial\left[p_{\oslash}(c^*)\right]}{\partial t}\left[\frac{\partial\left[p_n(c^*,\kappa^*)-p_{\oslash}(c^*)\right]}{\partial c^*}\delta_1 - 1 + \frac{\partial p_n(c^*,\kappa^*)}{\partial c^*}\delta_2\frac{\frac{\partial p_n(c^*,\kappa^*)}{\partial \kappa^*}\delta_1}{1+\frac{\partial p_n(c^*,\kappa^*)}{\partial \kappa^*}\delta_2}\right]^{-1} > 0 \\
\frac{d\kappa^*}{dt} &= \frac{\partial\left[p_{\oslash}(c^*)\right]}{\partial t}\left[\frac{\partial p_n(c^*,\kappa^*)}{\partial \kappa^*}\delta_1 + \left[1+\frac{\partial p_n(c^*,\kappa^*)}{\partial \kappa^*}\delta_2\right]\frac{\frac{\partial\left[p_n(c^*,\kappa^*)-p_{\oslash}(c^*)\right]}{\partial c^*}\delta_1 - 1}{-\frac{\partial p_n(c^*,\kappa^*)}{\partial c^*}\delta_2}\right]^{-1} > 0
\end{aligned}$$

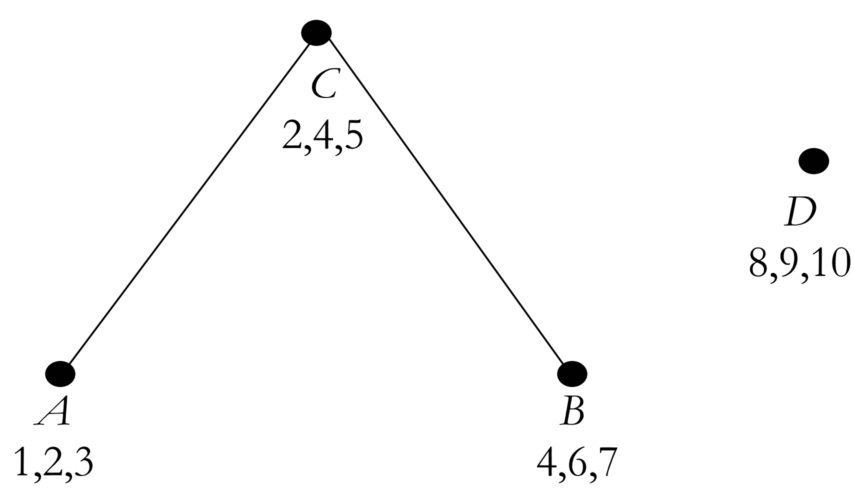
By checking the signs of the partial derivatives, the sign of the equilibrium total derivatives follows immediately. ■

Figure 1. A Simple "Network of Friends"



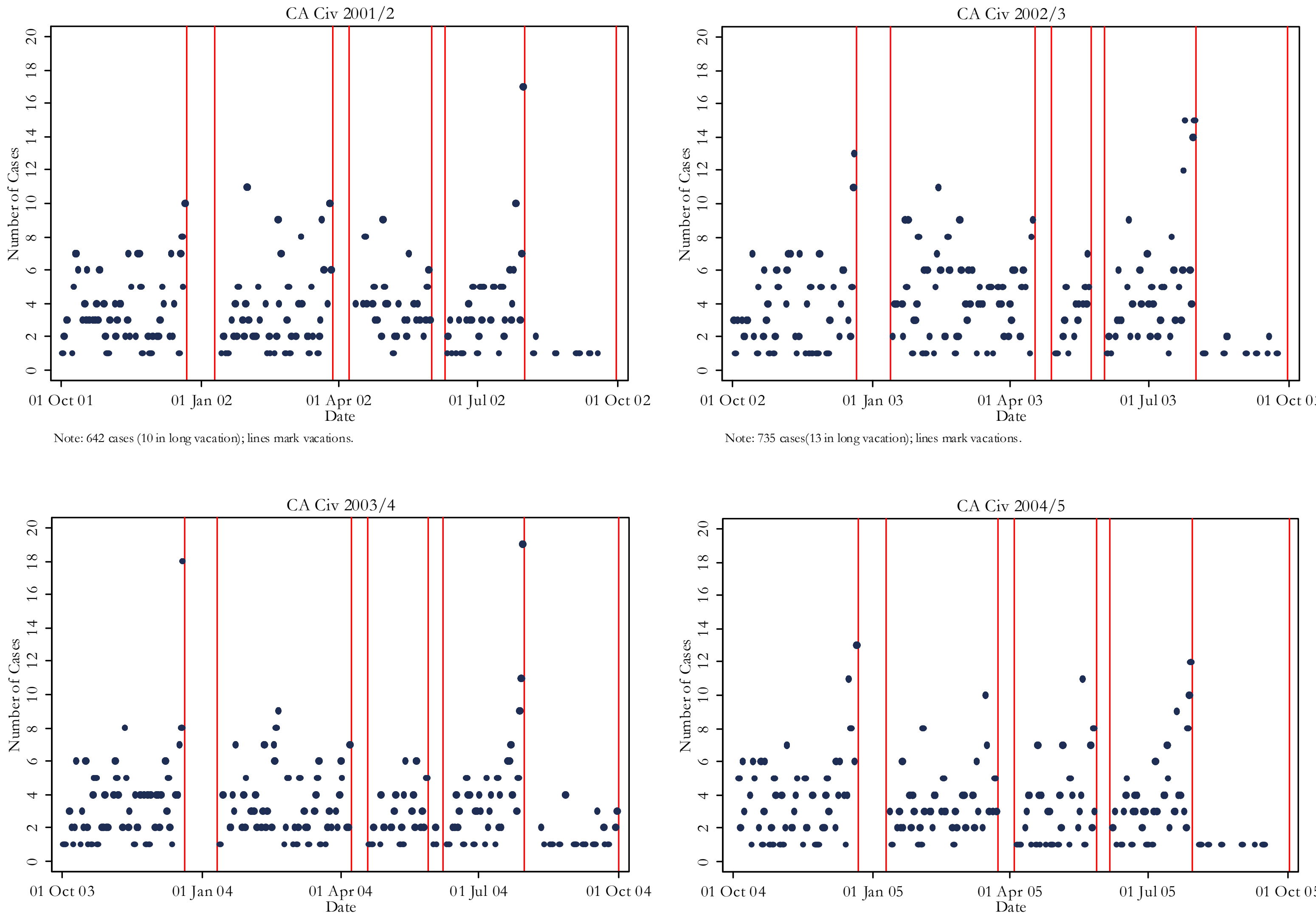
Individual *C* is distance 1 from *B* and distance 2 from *A*
Individual *D* is distance ∞ from *A* (and *B* and *C*)

Figure 2. A Simple "Network of Collaborators"



Team *C* is distance 1 from *B* and distance 2 from *A*
Team *D* is distance ∞ from *A* (and *B* and *C*)

Figure 3. Number of Cases Decided on Each Day of the Judicial Year (2001-2005)



Note: 642 cases (10 in long vacation); lines mark vacations.

Note: 735 cases(13 in long vacation); lines mark vacations.

Note: 635 cases(28 in long vacation); lines mark vacations.

Note: 593 cases (8 in long vacation); lines mark vacations.

Figure 4. Results of Random Panel Formation Tests

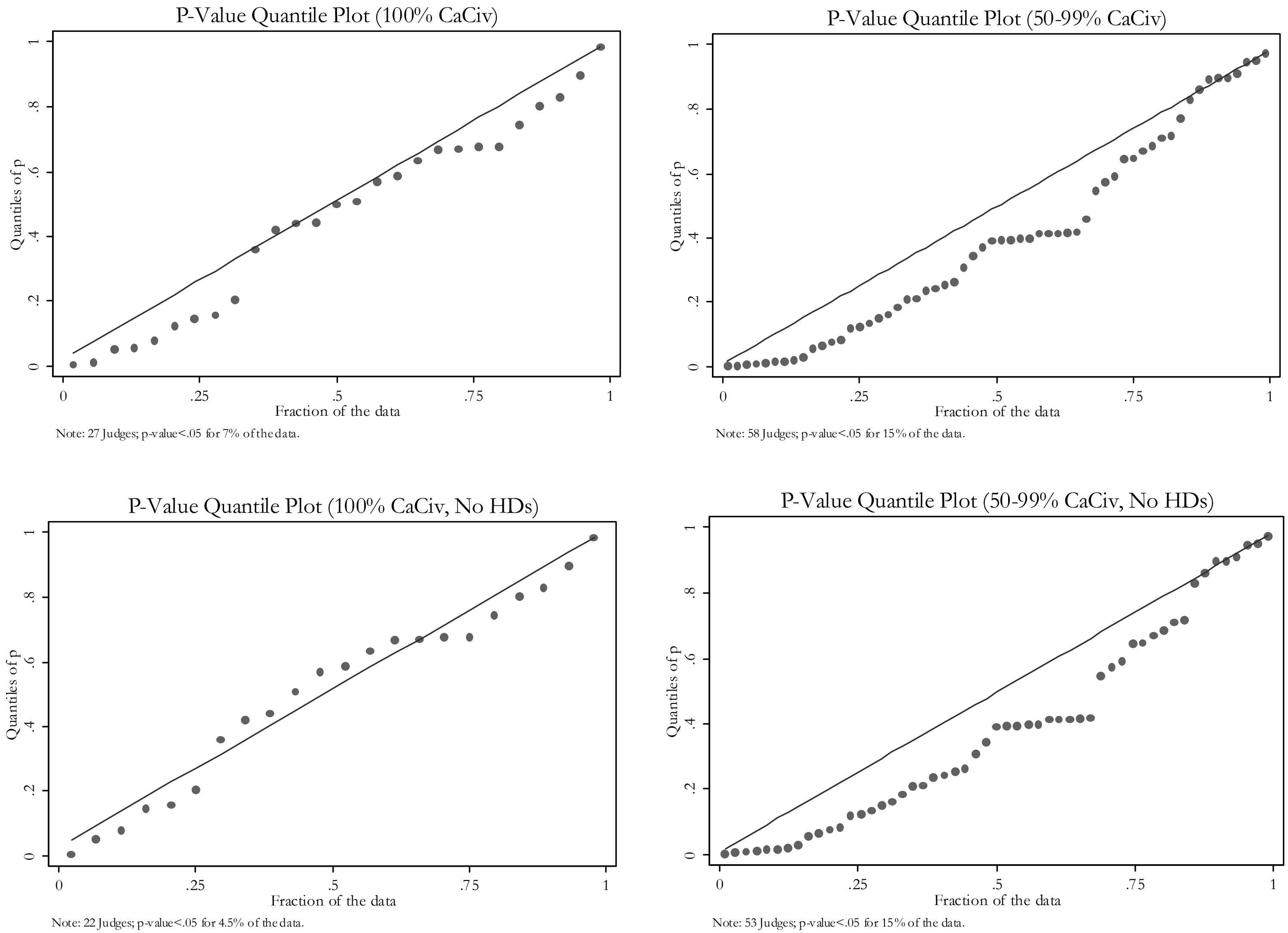


Figure 5. Assignment of Dyads to Network Distance Groups

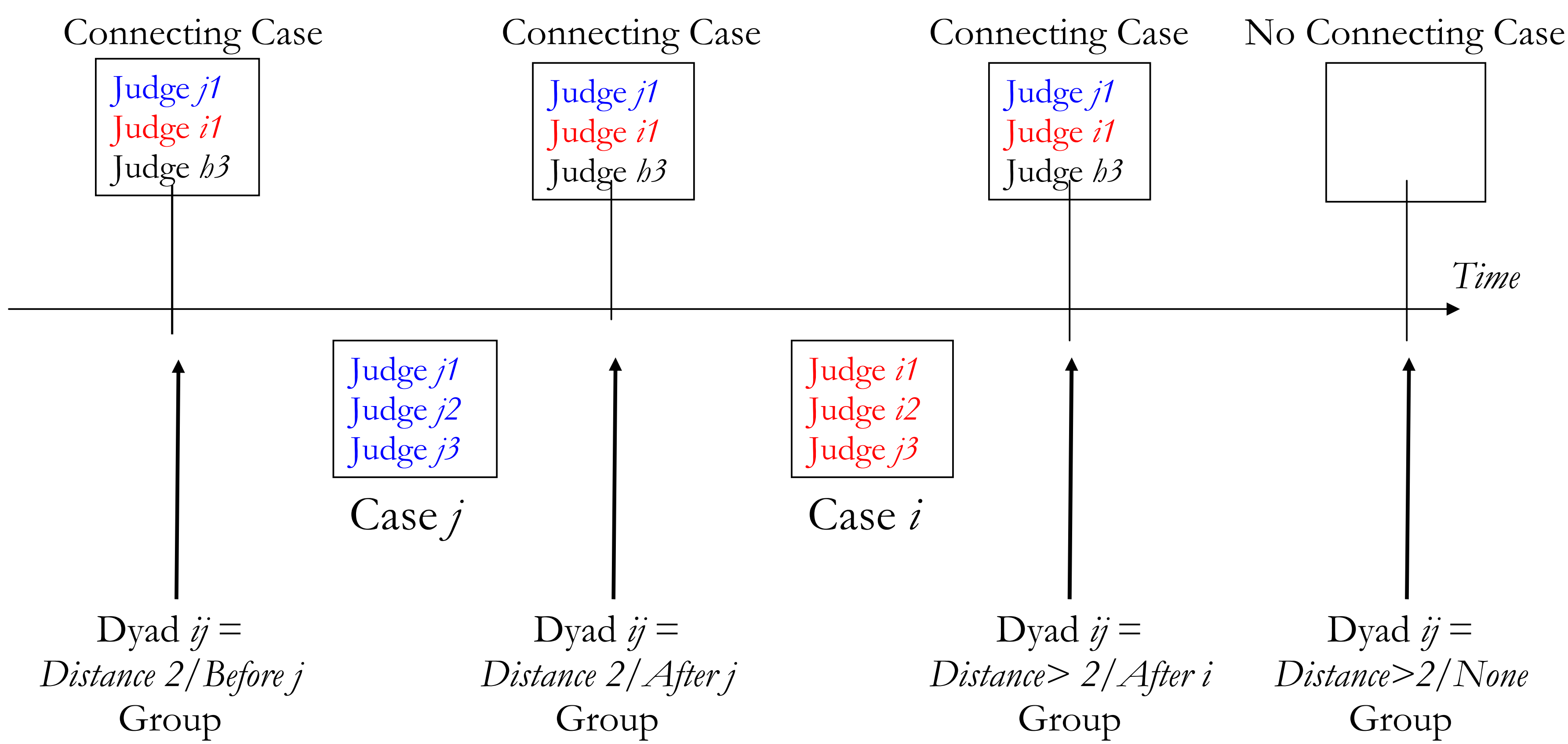
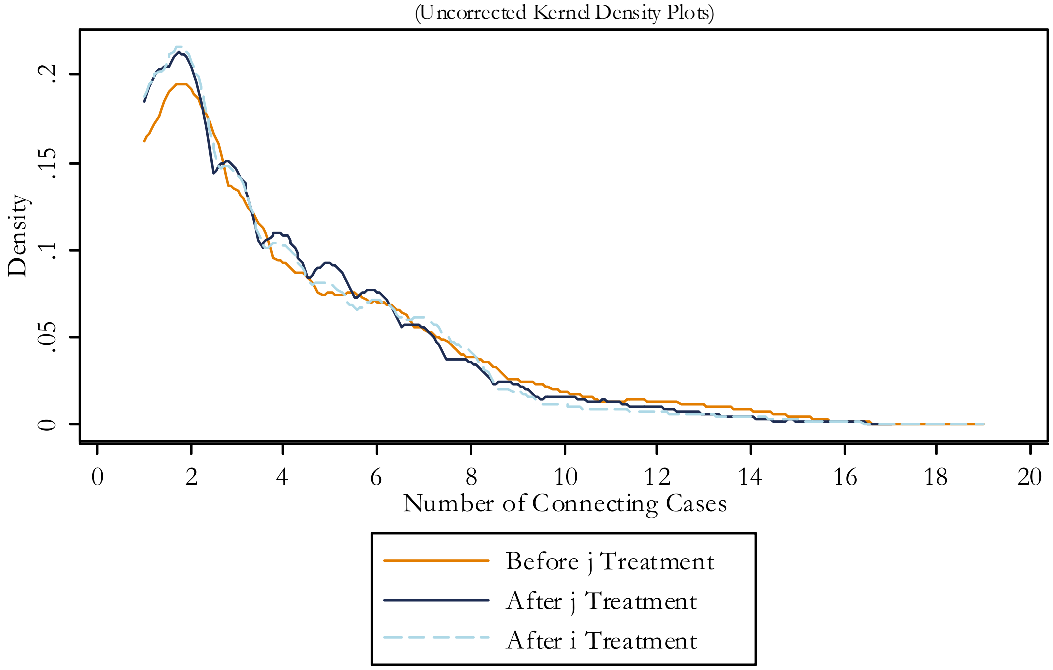


Figure 6. Connecting Case Distribution by Treatment

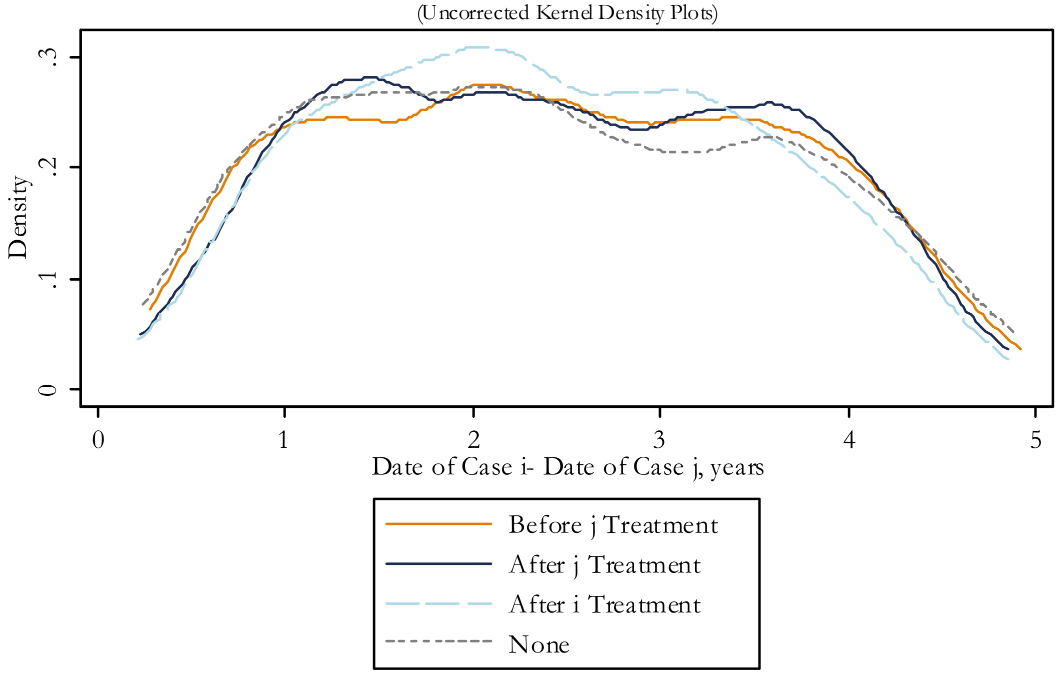
(Uncorrected Kernel Density Plots)



Note: 2425 dyads.

Figure 7. Time Difference Distribution by Treatment

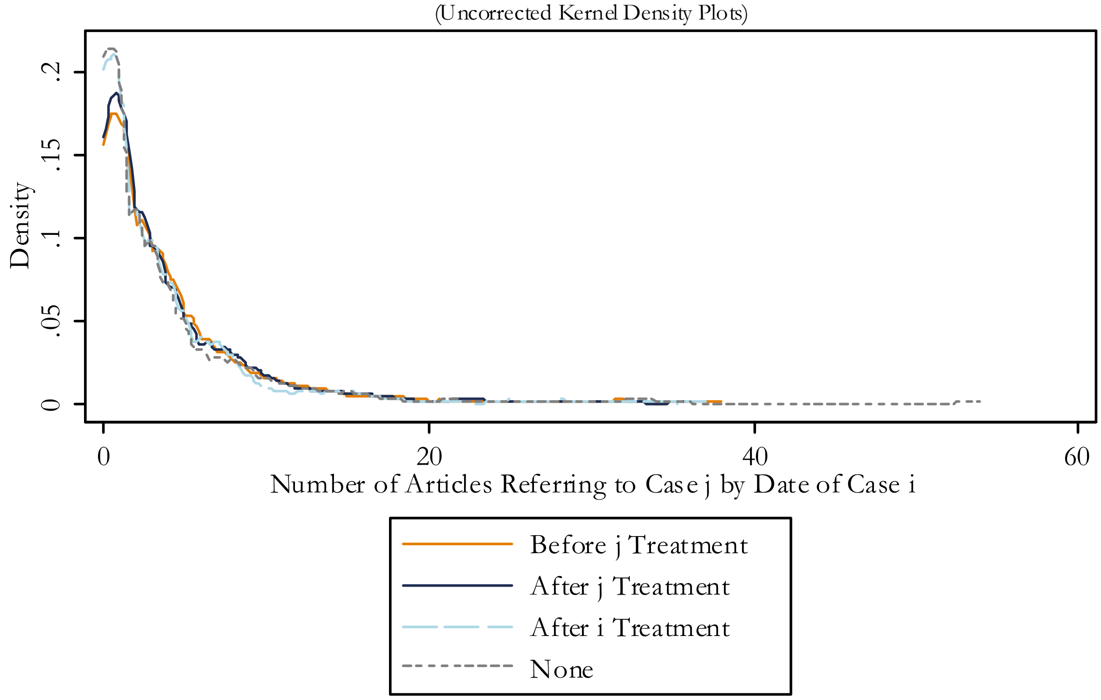
(Uncorrected Kernel Density Plots)



Note: 3249 dyads.

Figure 8. Legal Journal Articles Distribution by Treatment

(Uncorrected Kernel Density Plots)



Note: 3249 dyads.

Table 1. Summary of Comparative Statics Results

Effect on:	Parameter Change: **Knowledge of *j*** Increase in t (probability i aware of j)	Parameter Change: **Socialisation towards *j*** Increase in α (i's social preference for j)
Unobservable Types		
Ex ante probability of *type-1*	Negative	Zero
Ex ante probability of *type-2*	Positive	Zero
Ex ante probability of *type-3*	Positive	Zero
Probability of *type-1* given citation	Positive	Negative
Probability of *type-3* given citation	Negative	Positive
Unobservable Thresholds		
type-1 citation threshold: $\bar{c}^*$	Positive	Negative
type-3 positive citation threshold: $\bar{\kappa}^*$	Positive	Positive
Observable Citation Choices		
Probability of no citation: $M \in \{\varnothing\}$	Negative	Positive
Probability of neutral citation: $M \in \{same\}$	Positive	Negative
Probability of positive citation: $M \in \{r = s_r\}$	Positive	Positive
Probability of negative citation: $M \in \{r \neq s_r\}$	Zero	Zero

Table 2. Summary of Empirical Tests

	Difference in Response Across Treatments: Selection Response	Difference in Response Across Treatments: Outcome Response	Implied Substitution Patterns [1]	**Socialisation Test** [2] Dist 2/Before *j minus* Dist2/None	**Combined Test** Dist 2/After *j minus* Dist2/None	**Placebo Test** Dist 2/After *i minus* Dist2/None
				Rejects	*Rejects*	*Rejects*
1	=0	=0	No Change			
2	=0	<0	N ↑ P ↓	KD, SOC, Both	KD, SOC, Both	Findings 2 – 9 all reject identifying assumption of no omitted variables
3	=0	>0	N ↓ P ↑	KD, Both	KD	
4	<0	=0	None ↑ N ↓ P ↓	KD, SOC, Both	KD, SOC, Both	
5	<0	<0	None ↑ N ↓ P ↓ None ↑ N ↑ P ↓	KD, SOC, Both	KD, SOC, Both	
6	<0	>0	None ↑ N ↓ P ↑ None ↑ N ↓ P ↓	KD, Both	KD	
7	>0	=0	None ↓ N ↑ P ↑	KD, SOC, Both	SOC	
8	>0	<0	None ↓ N ↑ P ↑ None ↓ N ↑ P ↓	KD, SOC, Both	SOC	
9	>0	>0	None ↓ N ↑ P ↑ None ↓ N ↓ P ↑	KD, SOC, Both	SOC	

Notes:

1. None denotes no citation, N denotes neutral citation and P denotes positive citation.
2. "KD" refers to the assertion that there is a causal effect and the only mechanism at work is knowledge diffusion as stated in Definition 1.
   "SOC" is as above but the claim is that the only mechanism at work is socialisation as stated in Definition 2.
   "Both" is as above but the claim is that the only mechanism at work is some combination of knowledge diffusion and socialisation.

Table 3. Sample Size by Treatment

		Number of Dyads	Number of Discretionary Citations				No of Non-Discretionary Citations		
			Neutral	Negative	Positive	All	Negative	Positive	All
Treatment									
Distance 1	Same Judge	1,609	99	2	98	199			
Distance 2	Before *j*	721	15	0	21	36			
	After *j*	975	21	0	31	52			
Distance > 2	After *i*	729	21	0	22	43			
	None	824	31	0	21	52			
Other Distance 2	Before *j* and After *j*	1,524	40	0	53	93			
	After *j* and After *i*	1,152	35	2	34	67			
	Before *j* and After *i*	779	21	1	22	42			
	All	2,175	87	0	78	165			
No Contemporaries		2,544	101	0	43	165			
	Total Treatment	**13,032**	**471**	**5**	**423**	**899**			
Excluded									
	No After *i* MT	2,144	90	0	87	177	23	27	50
	After *i* MT but Discretionary Cite	279	0	0	0	0	111	168	279
	Total Excluded	2,423	195	1	136	177	134	195	329
	Total Treatment and Excluded	**15,455**	**561**	**5**	**510**	**1,076**	**134**	**195**	**329**

Notes: The table reports sample rather than population sizes; i.e. there is no correction for choice based sampling.

Table 4. Number of Connecting Cases in Before *j*, After *j* and After *i* Terms, by Treatment

Treatment		Response: # Cases in Before *j* MT Mean	Response: # Cases in After *j* MT Mean	Response: # Cases in After *i* MT Mean
Distance 1	Same Judge	15.57	17.09	13.33
Distance 2	Before *j*	4.03	0	0
	After *j*	0	3.85	0
Distance > 2	After *i*	0	0	3.89
	None	0	0	0
Other Distance 2	Before *j* and After *j*	4.39	4.58	0
	After *j* and After *i*	0	4.34	3.88
	Before *j* and After *i*	4.49	0	3.86
	All	4.97	5.17	4.28
No Contemporaries		.68	1.32	.94
	Subpopulation	13,032	13,032	13,032
	Number of PSUs	13,311	13,311	13,311
	Number of Strata	13	13	13

Notes: Corrected means are corrected for choice based sampling using STATA 9's suite of survey commands. See also Figure 5.

Table 5. Case Characteristics, by Treatment

Treatment		Response:[1] Time Diff Mean		p-value	Response:[2] 1[Same Subject] %		p-value	Response:[3] 1[Times Law Report] %		p-value	Response:[4] Legal Journal Articles Mean		p-value	Response:[5] Case *j* Panel Size Mean		p-value	Response:[6] Case *i* Panel Size Mean		p-value
Dist 1	Same Judge	2.22	***	.00	6.11		.56	51.61		.56	3.73	**	.05	2.77	***	.00	2.83	***	.00
Dist 2	Before *j*	2.40		.80	6.34		.75	49.58		.81	3.47		.41	2.71	***	.00	2.66	***	.00
	After *j*	2.50		.17	6.33		.73	51.69		.56	3.23		.98	2.62	***	.00	2.68	***	.00
Dist > 2	After *i*	2.44		.66	6.18		.66	49.31		.78	3.08		.58	2.52	***	.01	2.71	***	.00
	None	**2.42**		***Excluded***	**6.80**		***Excluded***	**50.24**		***Excluded***	**3.24**		***Excluded***	**2.43**		***Excluded***	**2.56**		***Excluded***
Other Dist 2	Before *j* & After *j*	2.28	**	.02	5.90		.45	54.48	*	.07	3.71	**	.05	2.78	***	.00	2.78	***	.00
	After *j* &After *i*	2.25	***	.00	4.20	**	.03	49.98		.92	3.49		.34	2.69	***	.00	2.80	***	.00
	Before *j* & After *i*	2.19	***	.00	5.23		.22	48.62		.55	3.41		.52	2.67	***	.00	2.78	***	.00
	All	2.07	***	.00	4.12	***	.01	53.66		.13	3.77	**	.02	2.82	***	.00	2.85	***	.00
No Contemporaries		2.91	***	.00	7.34		.63	55.21	***	.02	3.42		.43	2.08	***	.00	2.48	***	.00
	Subpopulation	13,032			13,032			13,032			13,032			13,032			13,032		
	No of PSUs	15,455			15,455			15,455			15,455			15,455			15,455		
	No of Strata	13			13			13			13			13			13		

Notes: All means and proportions are within treatment and are corrected for choice based sampling using STATA 9's suite of survey commands.
Point estimates for proportions multiplied by 100 to give percentages.
P-values (and ***, **, * denoting significance at 1%, 5% and 10% levels) are from an unconditional test for equality of means/proportions with Distance 2/None. Given the absence of conditioning variables, svy: probit and svy: prop naturally give identical results.

1. Time Difference is calculated as the number of days (divided by 365.25) between the dates of case *i* and case *j*. See also Figure 6.
2. Same Subject is equal to 1 if case *i* and case *j* share the same subject based on Westlaw's 100 part classification of legal subjects.
3. Times Law Report is equal to 1 if case *j* (cited) is reported in newspaper The Times Law Report in the week preceding the case.
4. Legal Journal Articles is calculated as the number of articles in legal journals that have referred to case *j* by the date of case *i*. See also Figure 7.
5. Case *j* Panel Size is calculated as the number of judges (1, 2 or 3) hearing case *j* (cited).
6. Case *i* Panel Size is calculated as the number of judges (1, 2 or 3) hearing case *i* (citing).

Table 6. Judge Social Network Characteristics, by Treatment

Treatment		Response:[1] 1[At School Together] Mean	*p*-value	Response:[2] 1[At Univ Together Mean	*p*-value	Response:[3] 1[Same Chambers] Mean	*p*-value	Response:[4] 1[Same Club] Mean	*p*-value
Dist 1	Same Judge	3.13	.20	11.40 **	.05	8.98	.29	13.56	.67
Dist 2	Before *j*	3.52	.49	11.15	.12	7.94	.80	14.38	.44
	After *j*	3.49	.42	10.04	.33	9.51	.18	13.58	.69
Dist > 2	After *i*	3.13	.26	9.31	.64	11.11 ***	.03	15.88	.13
	None	**4.26**	***Excluded***	**8.55**	***Excluded***	**7.57**	***Excluded***	**12.87**	***Excluded***
Other Dist 2	Before *j* & After *j*	4.59	.76	14.24 **	.00	7.72	.90	15.80	.08
	After *j* &After *i*	4.40	.90	13.81 ***	.00	7.81	.85	17.69 ***	.01
	Before *j* & After *i*	3.65	.56	14.04 ***	.00	8.57	.49	13.12	.89
	All	5.59	.20	17.54 ***	.00	7.57	.99	17.92 ***	.00
No Contemporaries		3.07	.15	5.56 ***	.01	7.72	.90	11.93	.52
	Subpopulation	13,032		13,032		13,032		13,032	
	No of PSUs	15,455		15,455		15,455		15,455	
	No of Strata	13		13		13		13	

Notes: All proportions are within treatment and are corrected for choice based sampling using STATA 9's suite of survey commands.
Point estimates for proportions multiplied by 100 to give percentages.
P-values (and ***, **, * denoting significance at 1%, 5% and 10% levels) are from an unconditional test for equality of means/proportions with Distance 2/None.

1. At School Together is equal to 1 if at least one judge pairs (one from case *i* and one from case *j*) attended the same secondary school at the same time (measured as same date of birth plus or minus five years).
2. As 1 but where both attended the same university (or college if Oxford or Cambridge) at the same time (measured as same date of birth plus or minus three years).
3. As 1 but where both practised as a barrister from the same chambers prior to becoming a judge.
4. As 1 where both are members of the same (Gentleman's or sports) club.

Table 7. Proportion of Dyads with a Citation, by Treatment

		Comparison of Proportions (1)					Plus Case-Level Controls (2)				Plus Judge Social Network Ctrls (3)			
		Proportion	Probit Coef	Probit SE	*p*-value	MEffect	Probit Coef	Probit SE	*p*-value	MEffect	Probit Coef	Probit SE	*p*-value	MEffect
Treatment														
Dist 1	Same Judge	.0084	.1710	.0395***	.00	**.0042**	.1291	.0581**	.03	**.0010**	.1327	.0586**	.02	**.0010**
Dist 2	Before *j*	.0030	-.0764	.0532	.15		-.0878	.0748	.24		-.0786	.0747	.29	
	After *j*	.0034	-.0502	.0485	.30		-.0661	.0705	.35		-.0683	.0716	.34	
Dist > 2	After *i*	.0037	-.0284	.0511	.58		-.0004	0691	.99		.0051	.0692	.94	
	None	.0042	*Excluded Category*				*Excluded Category*				*Excluded Category*			
Oth Dist 2	Before *j* & After *j*	.0039	-.0139	.0431	.75		-.0535	.0611	.38		-.0503	.0618	.42	
	After *j* &After *i*	.0039	-.0181	.0454	.69		-.0080	.0653	.90		-.0033	.0660	.96	
	Before *j* & After *i*	.0035	-.0434	.0507	.39		-.0666	.0800	.41		-.0712	.0821	.38	
	All	.0047	.0272	.0398	.49		.0291	.0578	.62		.0409	.0583	.48	
No Contemporaries		.0037	-.0317	.0402	.43		-.0004	.0567	.99		-.0060	.0572	.92	
Control Variables														
	Time Difference, yrs						-.0406	.0108***	.00	-.0002	-.0402	.0108***	.00	-.0002
	1[Same Subject]						.9265	.0272***	.00	**.0442**	.9263	.0274***	.00	**.0438**
	1[Times Law Report]						.2046	.0276***	.00	.0013	.2036	.0276***	.00	.0013
	Legal Journal Articles						.0212	.0018***	.00	.0001	.0213	.0018***	.00	.0001
	Cited Panel Size						.0709	.0214***	.00	.0004	.0687	.0217***	.00	.0004
	Citing Panel Size						.0491	.0235**	.04	.0003	.0483	.0239**	.04	.0003
	Social Network Vars?	No					No				Yes			
	F-Statistic (*p*-value)	9.72 (.00)					82.51 (.00)				64.28 (.00)			
	Subpopulation	13,032					13,032				13,032			
	No of PSUs, Strata	15,455, 13					15,455, 13				15,455, 13			

Notes: All columns report Probit coefficient and (linearized robust) SEs corrected for choice based sampling using STATA 9's suite of survey commands. In column (1), proportions are obtained via STATA's proportion estimator for survey data (svy: prop) and multiplied by 100.

Table 8. Proportion of Cited Dyads with a Positive Cite, by Treatment (Without Selection Model)

		Comparison of Proportions (4)					Plus Case-Level Controls (5)				Plus Judge Social Network Ctrls (6)			
		Proportion	Probit Coef	Probit SE	*p*-value	MEffect	Probit Coef	Probit SE	*p*-value	MEffect	Probit Coef	Probit SE	*p*-value	MEffect
Treatment														
Dist 1	Same Judge	49.25	.2245	.1969	.25		.1975	.2013	.33		.1984	.2015	.33	
Dist 2	Before *j*	58.33	.4538	.2742 *	.10	**17.95**	.4415	.2731 *	.10	**17.39**	.4401	.2741	.11	**17.35**
	After *j*	59.62	.4868	.2485 **	.05	**19.24**	.4924	.2502 **	.05	**19.32**	.4949	.2510 **	.05	**19.43**
Dist > 2	After *i*	51.16	.2726	.2597	.29		.3014	.2627	.25		.3115	.2638	.24	
	None	40.38	*Excluded Category*				*Excluded Category*				*Excluded Category*			
Oth Dist 2	Before *j* & After *j*	56.99	.4195	.2190 *	.06	**16.56**	.3826	.2223 *	.09	**15.15**	.3926	.2234 *	.08	**15.54**
	After *j* &After *i*	47.89	.1904	.2303	.40		.1875	.2319	.42		.2020	.2347	.39	
	Before *j* & After *i*	50.00	.2434	.2580	.35		.2165	.2597	.41		.2216	.2604	.40	
	All	47.27	.1750	.2010	.38		.1403	.2060	.52		.1406	.2062	.50	
No Contemporaries		29.86	-.2850	.2072	.17		-.1367	.2140	.21		-.1272	.2153	.56	
Control Variables														
	Time Difference, yrs						.0146	.0397	.71		.0151	.0397	.70	
	1[Same Subject]						-.0883	.0913	.33		-.0882	.0916	.34	
	1[Times Law Report]						.0744	.1054	.48		.0668	.1062	.53	
	Legal Journal Articles						.0071	.0056	.21		.0073	.0056	.20	
	Cited Panel Size						.2252	.0807 ***	.00	**8.93**	.2295	.0823 ***	.00	**9.10**
	Citing Panel Size						-.0103	.0850	.90		-.0055	.0853	.95	
	Social Network Vars?	No					No				Yes			
	F-Statistic (*p*-value)	3.07 (.00)					2.75 (.00)				2.21 (.00)			
	Subpopulation	899					899				899			
	No of PSUs, Strata	15,455, 13					15,455, 13				15,455, 13			

Notes: All columns report Probit coefficient and (linearized robust) SEs corrected for choice based sampling using STATA 9's suite of survey commands. In column (4), proportions are obtained via STATA's proportion estimator for survey data (svy: prop) and multiplied by 100

Table 9. Proportion of Cited Dyads with a Positive Cite, by Treatment (<u>With</u> Selection Model)

		No Controls in Outcome Equation (7)				Plus Case Controls in Outcome Eqn (8)			Plus Social Network Controls in Outcome Eqn(9)		
		Probit Coef	HProbit Coef	HProbit SE	*p*-value	HProbit Coef	HProbit SE	*p*-value	Probit Coef	Probit SE	*p*-value
Treatment											
Dist 1	Same Judge	.2245	.2287	.1990	.25	.1330	.1945	.49	.1337	.1948	.49
Dist 2	Before *j*	.4538	.4521	.2745 *	.09	.4366	.2611 *	.09	.4333	.2620 *	.10
	After *j*	.4868	.4860	.2485 **	.05	.4762	.2379 **	.04	.4810	.2386 **	.04
Dist > 2	After *i*	.2726	.2742	.2600	.29	.2716	.2502	.28	.2805	.2515	.26
	None		*Excluded Category*			*Excluded Category*			*Excluded Category*		
Oth Dist 2	Before *j* & After *j*	.4195	.4196	.2190 *	.06	.3647	.2120 *	.09	.3741	.2129 *	.08
	After *j* &After *i*	.1904	.1918	.2305	.40	.1661	.2204	.45	.1795	.2232	.42
	Before *j* & After *i*	.2434	.2442	.2580	.35	.2106	.2470	.39	.2172	.2478	.38
	All	.1750	.1779	.2021	.38	.1051	.1953	.59	.1023	.1958	.60
No Contemporaries		-.2850	-.2840	.2073	.17	-.1235	.2035	.54	-.1118	.2047	.58
Control Variables											
	Time Difference, yrs					Instrument in Selection Equation					
	1[Same Subject]					-.3638	.1723 **	.04	-.3630	.1731 **	.04
	1[Times Law Report]					Instrument in Selection Equation					
	Legal Journal Articles					Instrument in Selection Equation					
	Cited Panel Size					.2007	.0829 ***	.01	.2058	.0840 ***	.01
	Citing Panel Size					-.0179	.0815	.83	-.0126	.0819	.88
	Social Network Vars?		No			No			Yes		
	Athrho (*p*-value)		.0975 (.88)			.2061 (.12)			.2071 (.12)		
	F-Statistic (*p*-value)		3.06 (.00)			4.01 (.00)			3.05 (.00)		
	Subpopulation		13,032			13,032			13,032		
	No of PSUs, Strata		15,455, 13			15,455, 13			15,455, 13		

Notes: All columns report coefficients from a binary sample selection model (Heck Probit) and (linearized robust) SEs corrected for choice based sampling using STATA 9's suite of survey commands. Column (7) reports estimates from Table 8 Column (4) for comparison purposes.

Appendix Table A1. Population and Sample Sizes, by Strata (Case *i* year)

Out of Sample					**In Sample**				
Strata	Westlaw Sample	Dyad Population Size (M_h)	Uncensored Dyads (U_h)	Censored Dyad Sample Size	Strata	Westlaw Sample	Dyad Population Size (M_h)	Uncensored Dyads (U_h)	Censored Dyad Sample Size ($m_h - U_h$)
1980	28	375	0		1993	355	111,404	17	170
1981	29	1,209	0		1994	427	295,749	37	370
1982	20	1,328	0		1995	633	764,641	73	730
1983	33	3,047	2		1996	709	1,313,687	101	1010
1984	36	4,585	1		1997	710	1,758,086	97	970
1985	22	2,826	0		1998	678	1,908,762	96	960
1986	22	2,672	2		1999	801	2,505,987	112	1,120
1987	21	2,582	2		2000	958	3,231,716	126	1,260
1988	14	1,504	0		2001	891	3,198,076	114	1,140
1989	12	1,014	0		2002	621	2,257,840	138	1,380
1990	14	1,056	0		2003	745	2,712,193	219	2,190
1991	26	1,908	0		2004	628	2,214,609	168	1,680
1992	86	9,287	0		2005	404	1,246,173	107	1,070
Total	363	33,384	7	0	**Total**	**8,560**	**23,517,932**	**1,405**	**14,050**

Notes: Westlaw Sample refers to the number of Court of Appeal (Civil Division) cases in the given year available on Westlaw UK. Dyad population size refers to the number of *Within 5 year Court of Appeal (Civil Division)* dyads.