Advance Access publication May 9, 2012

Political Analysis (2012) 20:316–328
doi:10.1093/pan/mps016

Estimating the Causal Effects of Social Interaction with Endogenous Networks

Jon C. Rogowski
Department of Political Science, University of Chicago, 5828 S. University Ave., Pick Hall, 4th Floor, Chicago, IL 60637
e-mail: jrogowski@uchicago.edu (corresponding author)

Betsy Sinclair
Department of Political Science, University of Chicago, 5828 S. University Ave., Pick Hall 415, Chicago, IL 60637
e-mail: betsy@uchicago.edu

Edited by Neal Beck

Identifying causal effects attributable to network membership is a key challenge in empirical studies of social networks. In this article, we examine the consequences of endogeneity for inferences about the effects of networks on network members' behavior. Using the House office lottery (in which newly elected members select their office spaces in a randomly chosen order) as an instrumental variable to estimate the causal impact of legislative networks on roll call behavior and cosponsorship decisions in the 105th–112th Houses, we find no evidence that office proximity affects patterns of legislative behavior. These results contrast with decades of congressional scholarship and recent empirical studies. Our analysis demonstrates the importance of accounting for selection processes and omitted variables in estimating the causal impact of networks.

1 Introduction

Scholarly interest in social networks has increased dramatically in recent years. A particularly prominent line of inquiry centers on the capacity of networks to transmit information and influence, often shaping some behavioral outcome of interest among a target population. For instance, legislative scholars have invoked networks to explain how influence is transmitted between legislators. According to David Truman (1956, 1024), such networks are often "created by the proximity of offices or residences in Washington." Young's (1966) classic qualitative study of the nascent congressional years shows that living in congressional boardinghouses helped legislators form networks with other geographically dispersed members, and recent work by Masket (2008) shows that legislators in the California Assembly voted together more frequently when they shared desks on the chamber floor.

Identification of causal effects that result from network processes, however, is often difficult. The endogeneity that often accompanies network membership makes it difficult to separate the effects of the network from the ways in which network membership was determined. This problem is especially rampant in studies of political attitudes and behavior, in which researchers often have few alternatives but to use observational data to examine how one's network peers affect her behavior. Without information about how individuals select into such networks, however, it is nearly impossible to make causal inferences about network effects (see Manski 1993).

Authors' note: We thank Sara Brady, James Fowler, Keith Poole, Keith Krehbiel, and Jonathan Woon for providing data used in this project. We thank Kevin Collins, David Darfomal, James Fowler, Rob Franzese, Donald Green, Greg Huber, Greg Koger, Seth Masket, Eric Oliver, Jennifer Victor, Nick Weller, and participants of the Fourth Annual Political Networks Conference and the Harris School Political Economy Lunch for lively discussion and helpful comments. The editors and two anonymous reviewers provided especially conscientious feedback. We are most indebted to John J. Pitney, who used his own social networks to help us locate the House lottery data. This research protocol (H10171) was approved by the University of Chicago IRB. Supplementary materials for this article are available on the *Political Analysis* Web site.

© The Author 2012. Published by Oxford University Press on behalf of the Society for Political Methodology.
All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

In this article, we address this problem in the context of legislative behavior. We use the House office lottery—in which newly elected members select their offices in a randomly assigned order—and instrumental variables to examine the impact of physical proximity on roll call behavior and cosponsorship activity in the 105th–112th Congresses. In contrast with existing scholarship, we find no evidence that office proximity affects legislative behavior. Moreover, the inferences that this approach enables us to draw differ substantially from the inferences we would make if we were to use alternative research designs that attempt to control for homophily by including additional covariates or use a natural experiment to identify the effects by examining those legislators who switched offices between Congresses. Our analysis demonstrates the importance of accounting for selection processes and omitted variables in estimating the causal impact of social relationships.

The article proceeds as follows. We first review existing evidence about network effects and enumerate the particular identification concerns that are often found in this literature. We then turn to our study of legislative behavior in the U.S. House of Representatives, first describing existing arguments about the importance of physical proximity for influencing legislative behavior and then introducing the House office lottery as a means of accounting for how legislators select into their office networks. After presenting our main findings, we further explore them in the context of local average treatment effects and randomization inference. Finally, we demonstrate how conventional research designs that attempt to account for the selection of legislators into office networks via alternative research designs without true randomization generate a different set of inferences.

2 Social Networks and Causality

In studying the ways that networks affect political outcomes of interest, the presence of homophily—the propensity of individuals to select social network members with shared characteristics (McPherson, Smith-Lovin, and Cook 2001)—often makes it difficult to establish causal evidence for social network effects. The principal concern is that individuals choose their networks on the basis of their similarity with other network members. When evaluating attitudinal or behavioral outcomes, then, it is difficult to evaluate whether an individual has been influenced by the network to which they belong or whether the characteristics that led them to select their network also led them to exhibit the outcome of interest.

Legislative scholars have long argued that legislators often look to each other for cues about how to vote. Kingdon's (1989, 22) landmark study of Congress concludes that ''fellow congressmen appear to be the most important influence on voting decisions,'' possibly because, as Matthews and Stimson (1975, 25) argue, ''[t]here are too many decisions to be made across too wide a span of subjects; the issues involved are too complex for quick decision, and there is too little time for anything else.'' Moreover, physical proximity plays a key role in generating legislative networks that then facilitate the formation of cue-taking relationships (Caldeira and Patterson 1987).

The possibility of these kinds of network effects in legislatures has both empirical and theoretical implications. Because legislators are unlikely to select their cue sources at random, the error terms are likely to be correlated across legislators in any regression that treats legislators as units, and legislative scholars should explicitly account for this interdependence using (for instance) spatial lag terms (Beck, Gleiditsch, and Beardsley 2006; Franzese and Hays 2007). Moreover, as earlier scholars have argued, evidence for the importance of these kinds of relationships between legislators should revise our *theories* of legislative behavior. Thus, identifying these effects has significant theoretical and empirical implications for legislative scholarship.

Empirical studies of legislative networks conclude that congressional networks that are based upon physical proximity exert observable influences on legislator behavior (Young 1966; Fowler 2005, 2006; Masket 2008). For instance, in the study most like ours, Masket (2008) investigates patterns of legislative voting behavior in the California Assembly and finds that pairs of legislators that are assigned by the Speaker to share desks in the Assembly chamber exhibit greater similarity in voting behavior. As a practical matter, though, it is nearly impossible to causally estimate the effects of legislators' social networks in observational analyses due to homophily, common external stimuli, and two-way causation (Shalizi and Thomas 2011). However, due to the research designs used in these studies, it is difficult to evaluate whether these networks exhibited a *causal* effect on member behavior because of the possibility, for instance, that members select into cosponsorship networks on the basis of shared policy concerns or because of the

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

California Speaker's strategic interest in pairing legislators. The most promising identification strategy to estimate these effects is to randomly assign the social network (Sacerdote 2001; Shalizi and Thomas 2011), yet legislators' social networks are seldom randomly assigned and are difficult to manipulate.

The House office lottery (which we explain in greater detail in the next section) characterizes the process by which legislators have the opportunity to form congressional networks on the basis of office proximity. By accounting for the formation of these networks, then, we can clearly identify their causal impact. This approach provides two further advantages. First, in contrast to other work that infers networks on the basis of common behavioral patterns (i.e., Fowler 2005, 2006), our data clearly reveal the structure of House office networks. Second, our empirical approach tests a specific theoretical account of how networks affect behavior, and the results that emerge enable us to make clear inferences in reference to this account.

3 Instrumental Variables, the House Office Lottery, and Causal Inference

We use the House office lottery and instrumental variables to estimate the causal effect of proximity—a proxy for legislative social networks—on legislative behavior. Soon after each congressional election, newly elected members of the House gather in Washington, DC, for a week of new-member orientation. Near the end of the week, members choose their offices for the coming term via the House office lottery. Importantly, the order in which members may choose their offices is randomly assigned.

New members select from the offices that are vacated by members who are not returning due to retirement or defeat or who are upgrading to more desirable locations. There are three House office buildings: Cannon, Longworth, and Rayburn. Freshmen almost never get to choose from offices located in Rayburn, which has the best amenities and largest offices and is located closest to the Capitol. Cannon is the oldest House office building and is located farthest from the Capitol but is most accessible to the stop for Washington's subway system.[1] Longworth is located between Cannon and Rayburn and has offices on eight floors. The basement of Longworth houses a large cafeteria, and the main House mailroom and banking offices are also located in Longworth.

House offices vary substantially in desirability both between and within floors. Lower floors are generally preferred to higher floors because they are more accessible to visitors, require members and staff to navigate fewer stairs, are closer to the underground trams that go to the Capitol, and are more proximate to the amenities (such as cafeterias, post offices, banking facilities, and gyms) located in the basement floors. Square footage is an important consideration, but members dread the "split suite," in which part of the office is accessible only by exiting through the main door and reentering through another. Some offices have nicer views than others or are located closer to the District of Columbia Metro, which permits easier access for staff and constituents. Members also tend to prefer offices that are located closer to the elevators and committee rooms but farther from the restrooms.

The lottery is overseen by the Superintendent of House Office Buildings and is performed in two steps. First, members randomly draw lottery numbers based upon the alphabetical order of their last names. The numbers they select determine the order in which they choose their office locations. Members are then provided maps of the House office building complex and given several hours to consult with staff and determine which offices they would like to choose. Later that day, members select offices in the order determined by the lottery.

Our measure of proximity indicates whether or not two first-term members of Congress have offices in similar locations. We characterize this in three different ways. First, we indicate whether or not their offices are located in the same building. Second, we indicate whether they have offices located on the

[1]Particularly relevant for this study, the top (fifth) floor of Cannon, which was added some five years after the building was initially constructed, is often labeled the "freshman dorm" because of its cramped offices and low ceilings, and because many of the Cannon elevators only go as high as the fourth floor.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

Table 1 New members and legislation in the 105th–112th Houses

Congress	*No. of legislators*	*No. of roll calls*	*No. of cosponsorships*
105th	64	1166	2695
106th	39	1209	3000
107th	40	890	2954
108th	50	1218	3189
109th	37	1210	2943
110th	51	1865	4503
111th	52	1647	—
112th	85	412	—

Note. Number of cosponsored bills indicates the number of bills that were cosponsored by at least one first-term member of Congress.

same floor of the same building. And third, we use a more fine-grained measure that indicates whether or not two members have offices located on the same wing of the same floor (and in the same building).[2]

The House office lottery functions as a valid instrument in this analysis for three key reasons. First, lottery numbers are wholly exogenous to shared preferences between legislators. Thus, using the lottery number as an instrument is akin to random assignment in experimental settings.[3] Second, satisfying the exclusion restriction, lottery numbers are correlated with legislative behavior only through two members' spatial proximity. And third, newly elected members of Congress appear to evaluate potential House offices using similar criteria. Journalistic accounts describe members' interest in choosing offices that are proximate to entryways, elevators, restrooms, and the gym; have a view of the Mall or the Capitol dome; or are large enough for sleeping.[4] Thus, this instrumental variables approach enables us to estimate the effect of congressional office networks on legislative behavior while accounting for the process by which members selected their networks.

4 Data and Methods

We examine the effect of office networks using two different forms of legislative behavior in the 105th–112th Houses. Our unit of analysis is the legislator pair, and our key dependent variables are legislator agreement scores, which are generated using roll call and cosponsorship data.[5] These scores express the number of instances in which legislators i and j voted the same way, $\forall\ i \neq j$.[6] This generates an adjacency matrix, A, in which each entry in the matrix describes the number of times i and j voted similarly. We normalize these entries by the number of votes cast such that each entry a_{ij} represents the percentage of the time that legislators i and j exhibited identical behavior. Higher values of these entries indicate high degrees of similarity in legislative behavior. We generate agreement scores for each possible pair of House members for all legislators who served full terms.

In each Congress, this procedure generates a vector of agreement scores of length $\binom{N}{2}$, $\forall$ legislators $i \neq j$, where N is the number of first-term members of Congress. Table 1 shows the number of newly elected members of Congress, roll call votes, and cosponsorship opportunities in each of the Congresses used in this analysis. Pooling together the 105th–112th Congresses results in 11,599 pairs of legislators.

[2]Both Cannon and Longworth are roughly rectangular in shape, and by "wing," we simply mean one of the four edges. For instance, in the 112th House, Representatives Cory Gardner (213), Jeff Landry (206), and Scott Tipton (218) all have offices on the second floor of Cannon. Gardner's and Tipton's offices are located on the southeast edge of Cannon, whereas Landry's office is located along the southwest edge. Thus, because Gardner and Tipton are more likely to use the same elevator and to pass each other in the hall more frequently, the indicator for "same wing" may more closely capture the extent to which pairs of legislators interact based upon office proximity.

[3]By using this instrument, then, our study resembles other research that has randomly assigned participants to roommates and examined the consequences of their interaction (e.g., Sacerdote 2001).

[4]See, for instance, Durbin (2010) and Phillips and Yadron (2010).

[5]Cosponsorship data are available through the 110th Congress, and we include only those bills sponsored or cosponsored by at least one first-term legislator.

[6]For cosponsorship, we calculate agreement scores based upon the number of times in which i and j were sponsors or cosponsors on the same item. We do not distinguish sponsors and cosponsors.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

Table 2 Variable summaries

Variable name	*Min*	*Max*	*Mean*	*SD*
Roll call agreement rate (percentage points)	11.65	96.72	70.89	21.48
Cosponsorship agreement rate (percentage points)	0.04	10.44	1.91	1.17
Difference in presidential vote share (percentage points)	0	62.70	12.53	10.63
Same party	0	1	0.61	0.49
Same state	0	1	0.03	0.17
Shared building	0	1	0.50	0.50
Shared floor	0	1	0.09	0.29
Shared wing	0	1	0.02	0.15
Lottery distance (untransformed)	1	84	21.10	16.19

Our key independent variable is whether the pair of legislators have offices located in similar locations. We characterize this using the three sets of indicators that correspond to whether the pair has offices located in the same building, on the same floor of the same building, and in the same wing of the same building. To instrument for this, we use the absolute value of the difference in members' lottery numbers. For ease of interpretation, we take the inverse of the absolute difference in lottery numbers, so that larger values correspond to pairs that drew similar lottery numbers.

We include several exogenous regressors to control for other factors that influence the degree to which two members of Congress exhibit similar patterns of behavior. We account for the degree of preference similarity between two members using district presidential vote share in the most recent presidential election and include the absolute value of the difference between the Democratic presidential vote share in member *i*'s district from the presidential vote share in member *j*'s district in the most recent election. Thus, larger values indicate that the members represent increasingly dissimilar districts, which will be reflected in their voting patterns. We also include indicators for whether the pair of legislators are members of the same party or the same state, and we expect both of these coefficients to be positive. We also include Congress fixed effects to account for baseline differences in agreement rates between Congresses. Table 2 presents the summary statistics for all the independent and dependent variables described here. Our instrument is the fitted value of office proximity, based on a regression of proximity on the difference in lottery positions and the exogenous variables.[7] Finally, because a legislator *i*'s relationship with legislator *j* may also be informative about *i*'s relationship with legislator *k*, within each Congress, we cluster SEs by legislator *i* to account for unmodeled correlations between a legislator *i* and all other legislators $j \neq i$.[8]

5 Causal Estimates of the Effect of Office Proximity

Table 3 presents the estimates from the first-stage regressions. We are primarily interested in examining the strength of the relationship between lottery distance and proximity. The coefficients for lottery distance are positive and statistically significant in each of the six regressions, indicating that lottery distance and office proximity are indeed related. More crucially, we are interested in the *F* statistic that tests the null hypothesis that the coefficient for lottery distance is equal to zero. For all six regressions, *F* is substantially larger than 10, which allays concerns about weak instruments (Stock and Watson 2007). Though weak instruments can increase bias in two-stage least squares (2SLS) estimates (Staiger and Stock 1997), the bias falls as sample size increases. With our very large sample size and single endogenous regressor, 2SLS provides consistent estimates of the causal effect of office proximity, and these estimates contain significantly less bias compared to ordinary least squares (Murray 2006, 123–24).

[7] All models are estimated using the ivregress command in *Stata 12*.

[8] We also estimated these equations with legislator-specific random effects. The coefficients from these models are virtually identical to those reported here, and the substantive conclusions do not change.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

Table 3 First-stage results of instrumental variables estimation

Independent variables	*Roll calls*			*Cosponsorship*		
	Building	*Floor*	*Wing*	*Building*	*Floor*	*Wing*
Lottery distance	0.13 (0.02)	0.17 (0.02)	0.06 (0.01)	0.15 (0.03)	0.19 (0.02)	0.08 (0.02)
Ideological similarity	−0.04 (0.05)	−0.03 (0.03)	−0.00 (0.01)	−0.08 (0.06)	−0.04 (0.04)	−0.02 (0.02)
Same party	0.00 (0.01)	−0.00 (0.01)	−0.01(0.00)	0.01 (0.01)	−0.00 (0.01)	0.00 (0.00)
Same state	−0.05 (0.01)	−0.00 (0.01)	0.00 (0.00)	−0.06 (0.03)	0.01 (0.02)	−0.01 (0.01)
(Constant)	0.54 (0.01)	0.09 (0.01)	0.02 (0.00)	0.54 (0.02)	0.09 (0.01)	0.02 (0.00)
N	11,599	11,599	11,599	6703	6703	6703
MSE	0.50	0.29	0.15	0.50	0.30	0.15
F (lottery distance $\neq 0$)	20.15	59.98	23.61	16.98	42.83	23.78
Congress fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Clusters	403	403	403	269	269	269

Note. Entries are first-stage estimates and clustered SEs from instrumental variables estimation. The dependent variable is whether a pair of legislators has offices in similar locations. MSE, mean squared error.

5.1 *Office Proximity and Roll Call Voting Behavior*

We now present our estimates for the effect of office proximity on roll call behavior using each of the three indicators for proximity (building, floor, and wing). The results are shown in Table 4. Though the coefficient estimates for the indicators for proximity are all positive, none of these estimates approaches standard levels of statistical significance. The SEs are larger than the estimates for each of the three indicators. Thus, we cannot reject the null hypothesis that office proximity has no effect on roll call voting behavior.

The other coefficients, however, are quite sensible. Legislators agree with other legislators at higher rates when they represent more similar districts. Members also agree more frequently with their copartisan colleagues. The estimates for same state are all positive, suggesting that members from the same state exhibit more similar voting records, but none of these estimates is statistically significant.

But legislators vote on a wide range of issues in each Congress, and informal social ties—such as those afforded through office proximity—may be more likely to affect behavior on some votes than on others.

Table 4 Office proximity and roll call voting behavior

Independent variables	*(1)*	*(2)*	*(3)*
Same building	0.04 (0.07)		
Same floor		0.03 (0.05)	
Same wing			0.08 (0.14)
Ideological similarity	−0.94 (0.13)	−0.94 (0.13)	−0.94 (0.13)
Same party	0.62 (0.02)	0.62 (0.02)	0.62 (0.02)
Same state	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
(Constant)	−0.69 (0.04)	−0.67 (0.02)	−0.67 (0.02)
N	11,599	11,599	11,599
MSE	0.20	0.20	0.20
Congress fixed effects	Yes	Yes	Yes
Clusters	403	403	403

Note. Entries are second-stage coefficients and clustered SEs from instrumental variables estimation. The dependent variable is the logged roll call agreement rate. MSE, mean squared error.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

Table 5 Office proximity and cosponsorship

Independent variables	*(1)*	*(2)*	*(3)*
Same building	0.14 (0.19)		
Same floor		0.11 (0.15)	
Same wing			0.24 (0.34)
Ideological similarity	−0.61 (0.12)	−0.62 (0.12)	−0.62 (0.12)
Same party	0.65 (0.03)	0.65 (0.03)	0.65 (0.03)
Same state	0.35 (0.04)	0.34 (0.03)	0.35 (0.03)
(Constant)	−4.58 (0.12)	−4.51 (0.05)	−4.51 (0.05)
N	6703	6703	6703
MSE	0.47	0.47	0.47
Congress fixed effects	Yes	Yes	Yes
Clusters	269	269	269

Note. Entries are second-stage coefficients and clustered SEs from instrumental variables estimation. The dependent variable is the logged cosponsorship agreement rate. MSE, mean squared error.

We investigate the possibility that members are more likely to rely upon cue sources when information is scarce, as Matthews and Stimson (1975) argue, by examining roll call voting behavior on the first 150 roll call votes cast in a new Congress. The first 150 votes correspond roughly to the first 2 months of a new Congress, during which time new members are assembling and learning how to manage an office staff, responding to constituents, adjusting to a new city, and getting acquainted with the complex world of policymaking. If legislators are ever in need of guidance from their colleagues, this initial period would seem to be a good place for informal ties to influence voting behavior. To address these possibilities, we replicated the analyses shown in Table 4 but investigated the first 150 votes of each of the Congresses we analyze. We again find no evidence for the effect of office proximity on roll call voting behavior. The estimates for office proximity are all positive, but none of them approaches conventional levels of statistical significance. The estimates and SEs for the first 150 votes are nearly identical to those from Table 4 shown above.[9]

To better distinguish whether bill importance or salience affects legislators' use of social ties for voting decisions, we used a data set assembled by Keith Krehbiel and Jonathan Woon (2005) to examine roll call voting behavior in the 105th–107th Congresses. We used two criteria to distinguish high- and low-salience bills, across which the effects of office proximity may vary. First, we identified whether the bills received coverage in either *The New York Times* or the *Wall Street Journal*. Second, we distinguished bills on which the president took a publicly stated position from those in which he did not. Legislation that meets either or both of these criteria are the kinds of votes for which party leaders, interest groups, and constituents are most likely to attempt to persuade members of Congress to vote in a particular way. If the effect of office proximity varies by the importance or salience of the vote, we expect that the coefficient estimates for office proximity should be smaller in magnitude for the bills that did not receive media coverage or on which the president did not take a public position than the estimates for the bills that did receive media coverage or on which the president did take a position, respectively. However, we again find no causal evidence for the impact of proximity on roll call voting behavior using either criteria. Office proximity does not exert a greater influence on roll call behavior depending upon whether a given roll call receives media coverage or on which the president has taken a public position.

5.2 *Office Proximity and Cosponsorship Activity*

We find similar results when examining cosponsorship behavior. The results of our analysis for mutual cosponsorship in the 105th–110th Houses are shown in Table 5 above. The estimates for same building, same floor, and same wing are all positive, but none of them are statistically significant. Even though

[9]These and other supplementary results can be accessed through the Supplementary Appendix.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

cosponsorship is a much less visible and costly form of legislative behavior, we find no evidence that physical proximity affects systematic patterns of cosponsorship activity.

The estimates for the other covariates, however, are again quite sensible. Across the models for both roll call and cosponsorship, legislators exhibit more dissimilar patterns of behavior as the ideological differences of their constituencies increase, and copartisans behave more similarly than legislators from across the partisan aisle. The main difference between roll call voting behavior and cosponsorship, however, is that legislators from the same state are much more likely to cosponsor the same pieces of legislation but are no more likely to vote the same way on roll call items.

In sum, the results we have presented in this section provide no evidence that legislators' office networks affect roll call behavior or cosponsorship activity. Though physical proximity is commonly believed to facilitate cue taking and give rise to the development of relationships that would generate shared patterns of legislative behavior, we find no persuasive causal evidence in support of these claims. Though legislators may indeed influence their colleagues' behavior, this does not appear to be caused by physical proximity.

6 Extensions

In this section, we explore two extensions to the analyses presented above. First, we reestimate the models shown above to account for the possibility of local average treatment effects or the effects of office proximity among legislators plausibly thought to be most affected by the lottery assignment. Second, given the consistently positive coefficient estimates we showed in the tables above, we check for the presence of insufficient statistical power using randomization inference.

6.1 *Estimates of LATE*

Local average treatment effects (LATE) are often estimated via instrumental variables in the context of randomized experiments. In these instances, Angrist, Imbens, and Rubin (1996) define LATE as the effect of the treatment (X^{T}) on an outcome (y) for those who are induced by higher values of the instrument (Z) to have higher values of X^{T}. Thus, treatment assignment is used as an instrument for whether an individual actually received the treatment, which provides an estimate of the effect of the treatment among compliers.

Because our instrument takes on a range of values, we do not characterize compliance in a dichotomous fashion. In our case, estimating LATE focuses our attention on the subset of individuals for whom small values of the instrument induced them to select more proximate offices. All legislators are essentially assigned to some varying intensity of treatment, and similarly, we have varying intensities of compliance. A simple example illustrates this clearly. Suppose there are 80 new members of Congress, from which we select two pairs of legislators. Both pairs' lottery numbers are separated by five numbers. One pair received lottery numbers 3 and 8, and the other pair received lottery numbers 73 and 78. In general, members should be highly likely to choose offices in similar locations when their lottery assignments are separated by only five places. However, we expect the pair of legislators with the lower numbers to be more likely to choose offices in similar locations because nearly all the office locations that were available before any offices were chosen remain available. However, whether the pair of legislators with high lottery numbers also chooses proximate offices is largely a function of what the legislators before them have chosen. The legislators with low lottery numbers do not face this constraint to the same degree, and in the language of LATE, we term such pairs ''compliers.''

We reestimate our models for shared rates of agreement in roll call voting behavior and mutual patterns of cosponsorship but limit the sample to those legislators who might plausibly be considered compliers. In particular, we investigate the effects of shared office proximity among legislators who drew lottery numbers that were in the lowest 20%, 30%, 40%, and 50% of their entering cohort. For instance, 85 new members were elected to the 112th Congress, and thus, we limit our analysis to those members with lottery numbers less than or equal to 17, 26, 34, and 43.

We again find no evidence that shared office location has a meaningful impact on shared patterns of behavior. None of the 24 coefficients from these models is statistically significant. Thus, examining our results in the context of LATE also fails to provide convincing evidence that office proximity has a meaningful effect on legislative behavior, even among the subset of legislators for whom these effects are likely to be strongest.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

Table 6 Randomization inference results

Office proximity measure	*Source*	*Coefficient estimate*	*Exact p-value*
Same building	Table 4	0.04	0.462
Same floor	Table 4	0.03	0.472
Same wing	Table 4	0.08	0.425
Same building	Table 5	0.14	0.374
Same floor	Table 5	0.11	0.399
Same wing	Table 5	0.24	0.301

Note. The column of coefficient estimates contains estimates for the effects of office proximity on legislative behavior presented in previous tables and estimated via instrumental variables. The last column of entries shows the exact *p*-values for the coefficients as estimated from the distribution of coefficients obtained via instrumental variables from 100,000 randomizations of the lottery instrument.

6.2 *Statistical Power*

The results reported above present limited evidence to support the claim that office proximity affects legislator behavior. However, note that most of the coefficient estimates for office proximity shown in these tables are positive. If we were able to increase our sample size, and thus our statistical power, perhaps the SEs would shrink, and these coefficients could be judged to be statistically significant against a null hypothesis of no effect.

We evaluate the magnitude of our results with respect to the permutation distribution of the randomly assigned instrument so as to speak to these issues of statistical power. In particular, we are interested in comparing our estimated coefficients to a distribution of coefficients that would result if the null were true (Fisher 1935; Hansen and Bowers 2009). This procedure is commonly done in studies using randomized experiments, in which researchers compare the estimated effect sizes against what would result if subjects were rerandomized to treatment and control conditions (Hansen and Bowers 2009). Let Ω be the set of all possible arrangements of the lottery numbers, and let z be a draw from Ω. Taking many draws of z, each time substituting the simulated values of z for the lottery numbers and then reestimating the 2SLS models described above, produces a distribution of estimated coefficients when the instrument is completely uninformative. This distribution is determined solely by the randomization of the lottery number assignment. We can then compare our coefficients from the above table to these simulated distributions to determine where in the distribution they fall. If they are sufficiently infrequent with respect to the distribution of coefficients under an uninformative instrument, then we can conclude that our test simply lacks sufficient statistical power to allow us to reject the null. However, if their occurrence is not sufficiently rare, then we conclude that we can reject the alternative hypothesis that office proximity exerts a causal effect on legislative behavior.

Suppose, for instance, that we observe a distribution of coefficients such that 20% of the possible coefficients are larger than the observed value, 0.14 (from the first column of Table 5). We would conclude that the *p*-value associated with this coefficient is .20, which quantifies how surprising it is that we observe a coefficient of at least 0.14 under the null hypothesis. Using the models we have estimated throughout this article, we rerandomized the lottery numbers within each Congress 100,000 times, each time using the randomized instruments to estimate the effect of office location on shared behavior and saving each of the coefficient estimates for office location. The distribution of coefficients is approximately normal with very long tails, with mean -0.03, median -0.00003, and SD 42.14.

Table 6 below compares the coefficients obtained in Tables 4 and 5 to this distribution of coefficients when the lottery instrument is randomly assigned. As the table indicates, none of the recovered coefficients from the earlier models approach standard levels of statistical significance when compared to the distribution of coefficients using a completely uninformative instrument. The smallest *p*-value across these six models is .301. Indeed, even when comparing the largest coefficient obtained across all the models shown in this article -0.78, which resulted from our analysis of LATE discussed above, the corresponding *p*-value is .130. This randomization inference procedure demonstrates that the null findings shown in this article are not an artifact of insufficient statistical power and indicate that we cannot reject the null hypothesis that office proximity has no effect on shared patterns of legislative behavior.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

6.3 *Estimating Network Effects Using Conventional Empirical Strategies*

Throughout this article, we have argued that investigating network effects when the networks are endogenous requires researchers to account for the ways in which individuals have chosen their network. In this section, we examine the results that emerge from more conventional empirical strategies and contrast the inferences that we have made using the above analyses with the inferences that result from these alternative strategies.

First, we examine changes in roll call behavior among members who served in both the 110th and 111th Congresses to see whether changes in roll call behavior are attributable to changes in office locations. This approach allows us to implicitly control for any time-invariant factors that might affect patterns of agreement between legislators. Second, we examine roll call voting behavior as a function of shared office location among all members of the 105th House and estimate a series of cross-sectional regressions. These research designs provide viable means of identifying network effects under certain conditions. For instance, in the office changers analysis, identification is provided if a legislator who changes offices chooses to do so for reasons that are uncorrelated with the legislators that occupy the offices nearby to the one that she is switching into (or out of). In the simple cross-sectional analysis, accounting for all possible factors that could indicate the presence of relationships between legislators could also help provide causal identification.

In both cases, we use standard ordinary least squares (OLS) to estimate the effects of office proximity on shared legislative behavior. We focus on the indicator for whether legislator pairs have offices in the same House office building, which is the coarsest proximity measure used in the above set of analyses and the measure for which it should be the most difficult to recover positive and significant estimates. Across both sets of analyses, we find positive and statistically significant effects. This exercise illustrates the very real possibility of inferring network effects without accounting for the process by which the networks were chosen.

6.4 *Examining Changes in Office Location*

We calculated roll call agreement scores for all 359 legislators who served in both the 110th and 111th Congresses, approximately 30% (109) of whom changed office locations between Congresses. Thus, we model the change in the agreement between the Congresses as a function of changes in office proximity. This is a strong empirical technique because it implicitly controls for all the time-invariant characteristics (such as personal ideology, party, state, and constituency characteristics) that may affect patterns of shared behavior.

The dependent variable is the change in agreement between the 110th and 111th Congresses. Positive values of the dependent variable indicate that legislators agreed at higher rates in the 111th Congress than in the 110th Congress.[10] The main independent variable takes a value of $+1$ if the legislators were not housed in the same building in the 110th Congress but had offices in the same building in the 111th Congress. It has a value of -1 if they were in the same building in the 110th, but were not in the 111th, and has a value of zero if there is no change in their office proximity across the Congresses. The only other covariate we include is the agreement rate in the 110th Congress because legislators that already agree at very high rates may not agree at any higher levels as a function of changes in office proximity due to possible ceiling effects.

Across both specifications, we find positive and statistically significant evidence of an ''office effect.'' The first set of estimates indicates that legislators with offices in the same building vote together approximately a quarter of a percentage point more frequently than legislators whose offices are not located in the same building, whereas the second set of estimates shows that the effect is closer to a percentage point. Both of these coefficients are statistically significant. Were we to rely solely on analyses of this type, we would infer (incorrectly) that office proximity has a causal effect on legislative behavior.[11]

[10] As a second specification, we also take the natural log of the difference in agreement rates and transform them by adding $+0.24$ to each so that all the untransformed agreement rates are greater than zero.

[11] We note, however, that these coefficients do not appear to be statistically distinguishable from the coefficients we estimated in Tables 4 and 5.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

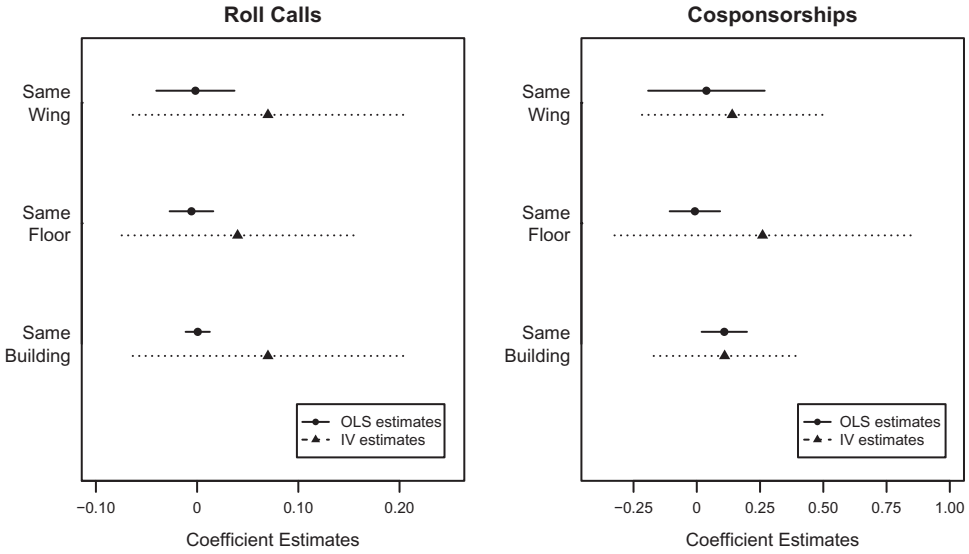


Fig. 1 Comparing ordinary least squares (OLS) and instrumental variables (IV) results. The plotted points indicate the coefficients for the effect of shared office location on roll call and cosponsorship agreement scores, and the horizontal lines are the 95% confidence intervals. The IV estimates are from the coefficients for office proximity reported in Tables 4 and 5. The OLS coefficients are from naive linear regressions run on the same data but without the inclusion of the House lottery instrument.

6.5 *Cross-Sectional Estimates and the 105th Congress*

Rather than focusing on just the newly elected members of Congress, we generate roll call and cosponsorship agreement scores for all possible pairs of legislators (90,525) who served complete terms in the 105th House. We regress the agreement scores on a similar set of covariates used in the previous section and again include indicators for whether a pair of legislators has offices in the same building, whether they represent the same state, and whether they are members of the same party. To also account for any other possible influences that might affect both members' choices of office and their propensity to agree with other legislators, we include a large battery of additional covariates. We include variables that indicate whether a pair entered the House as part of the same cohort, how many years they have served together, and whether they serve on a common committee, are of the same religion, attended the same college, both served in the military, and are of the same gender or racial/ethnic minority group. We also account for common constituency concerns by denoting the differences in the percentage of the members' districts dedicated to farming and military operations.

Here, too, we find statistically significant evidence of an office effect. Members whose offices are located in the same building vote together about 1% more of the time than legislators whose offices are located in different buildings. The effect is even larger for cosponsorship: legislators whose offices are in the same building cosponsor together 3% more frequently than legislators whose offices are less proximate. Should we rely solely on this table of results, we would erroneously conclude that office proximity exerts a causal effect on legislative behavior.

Both these supplementary analyses provide support for the contention that physical proximity plays an important role in affecting patterns of legislative behavior. Because the coefficients for office proximity shown in this section all reach conventional levels of statistical significance, scholars using these conventional—and potentially quite reasonable—empirical strategies would be inclined to reject the null hypothesis that office proximity exerts no effect on shared patterns of voting behavior.

For the most part, however, our substantive conclusions do not depend on whether the effect of office proximity is estimated using instrumental variables (IV) or OLS. Figure 1 above plots the recovered coefficients for the effect of office proximity on roll call and cosponsorship activity. The IV estimates and SEs are those shown in Tables 4 and 5. The OLS estimates result from a simple regression in which we do not

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

use the lottery instrument. The one case in which OLS provides statistically significant evidence of an office effect is in accounting for the effect of shared building on cosponsorship behavior. Generally, though, all the OLS coefficients are closer to zero than the IV estimates.

The limited differences between the IV and OLS estimates bolster our claim that office proximity does not affect shared patterns of legislative behavior. They also stand in stark contrast to the other reasonable empirical approaches that researchers might use to estimate these effects. Identifying these effects is especially challenging when homophily is a potential concern, though it is possible to do so using repeated-measures designs or cross-sectional regressions that include a range of controls. Though the choice of using these approaches or techniques such as instrumental variables may not result in meaningfully different effect estimates, failing to account for the process by which individuals choose their networks may result in Type I errors.

Instead, our approach to estimating network effects when network membership is endogenous highlights the differences in the inferences that one might draw. In particular, if we are concerned that legislators self-select into particular office locations on the basis of characteristics of the other legislators that also have offices in those locations—and, more generally, if we are concerned that network membership is endogenous—then we ought to worry about whether standard empirical strategies provide clear identification of network effects. In the absence of randomized experiments, research designs that carefully account for the ways in which individuals select their memberships are the best way to precisely identify the causal effects of network membership.

7 Discussion and Conclusion

In this article, we have argued that endogenous networks present empirical challenges for identifying causal estimates of network membership. The House office lottery enables us to account for the process by which legislators select into congressional networks that form on the basis of office proximity. Accounting for this selection process allows us to obtain causal estimates of network effects that are not biased due to homophily or endogeneity.

The causal evidence that we have presented shows that the apparent effect of office proximity on legislative behavior is limited. Across a variety of model specifications and subsets of bills, we find no significant evidence that office proximity exerts a causal impact on shared patterns of legislative behavior. Using two common empirical strategies, we demonstrate that standard techniques used in observational network analyses may generate misleading inferences. However, though we do not find persuasive evidence for the effects of legislative office networks on behavior, we do not conclude that legislative behavior is not influenced by a variety of factors—including their peers and cue sources—but simply that causal social structures, including those afforded by office proximity, are not important explanations for legislative behavior.

Network researchers pay a great deal of attention to structure when developing theoretical and empirical explanations for relational ties. As scholarly interest in this research area has increased, social networks have been said to be influential in explaining a range of empirical phenomena. Throughout this article, we have emphasized the empirical challenges that come with assessing these claims. Paying attention not only to the *structure* of networks but also to *how* that structure came to be can help remedy many of the difficulties in providing causal evidence for network effects.

Funding

Dirksen Congressional Center; Social Sciences Division Grant at the University of Chicago.

References

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91:444–55.

Beck, Nathaniel, Kristian Skrede Gleditsch, and Kyle Beardsley. 2006. Space is more than geography: Using spatial econometrics in the study of political economy. *International Studies Quarterly* 50:27–44.

Caldeira, Gregory A., and Samuel C. Patterson. 1987. Political friendship in the legislature. *Journal of Politics* 49:953–75.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press

Durbin, Kathie. 2010. Herrera feels like big winner for D.C. office. *The Columbian,* November 19. http://www.columbian.com/news/2010/nov/19/herrera-winner-office-lottery (accessed February 23, 2011).

Fisher, Ronald A. 1935. *Design of experiments*. Edinburgh, UK: Oliver and Boyd.

Fowler, James H. 2005. Legislative cosponsorship networks in the U.S. House and Senate. *Social Networks* 28:454–65.

———. 2006. Connecting the Congress: A study of cosponsorship networks. *Political Analysis* 14:456–87.

Franzese, Robert J., and Jude C. Hays. 2007. Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data. *Political Analysis* 15:140–64.

Hansen, Ben B., and Jake Bowers. 2009. Attributing effects to a cluster-randomized get-out-the-vote campaign. *Journal of the American Statistical Association* 104:873–85.

Kingdon, John. 1989. *Congressmen's voting decisions*. 3rd ed. Ann Arbor: University of Michigan Press.

Krehbiel, Keith, and Jonathan Woon. 2005. *Selection criteria for roll call votes*. Presented at the annual meeting of the American Political Science Association, Washington, DC.

Manski, Charles F. 1993. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies* 60:531–42.

Masket, Seth. 2008. Where you sit is where you stand: The impact of seating proximity on legislative cue-taking. *Quarterly Journal of Political Science* 3:301–11.

Matthews, Donald R., and James A. Stimson. 1975. *Yeas and nays: Normal decision-making in the U.S. House of Representatives*. New York: Wiley InterScience.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27:415–44.

Murray, Michael P. 2006. Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives* 20:111–32.

Phillips, Michael M., and Danny Yadron. 2010. *Pajama party: New to Congress, many members plan to sleep over. Wall Street Journal,* November 29. http://online.wsj.org/article/SB10001424052748703559504575630661395762460.html (accessed February 23, 2011).

Sacerdote, Bruce. 2001. Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics* 116:681–704.

Shalizi, Cosmo Rohilla, and Andrew C. Thomas. 2011. Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods and Research* 40:211–39.

Staiger, Douglas, and James H. Stock. 1997. Instrumental variables regressions with weak instruments. *Econometrica* 65:557–86.

Stock, James H., and Mark W. Watson. 2007. *Introduction to econometrics*. Boston, MA: Addison-Wesley.

Truman, David B. 1956. The state delegations and the structure of party voting in the United States House of Representatives. *American Political Science Review* 50:1023–45.

Young, James Sterling. 1966. *The Washington community, 1800–1828*. New York: Harcourt Brace Jovanovich.

https://doi.org/10.1093/pan/mps016 Published online by Cambridge University Press