Appendix

A   Additional Analyses and Robustness Checks

A.1   Permutation Tests

We construct 500 datasets of 120 stores each, but within each store of N_j workers, we draw N_j workers *with replacement*, to create a placebo store. To create a within-store network, we first include all the edges (relationships) from the original store where both nodes are in the placebo sample. We then create new edges assuming each worker is connected to any replicates of themselves. We then recalculate all our network statistics in the placebo store, including eigenvector centrality, degree centrality, and the correlation of organizer effort and worker centrality. We then estimate equation 5, including controls for each of these datasets, obtaining 500 placebo β.

$$log(Cards_j) = \beta NDO_j + X_j'\gamma + \epsilon_j \tag{5}$$

Panel A of Figure A2 shows the resulting histogram of the placebos, with a vertical line indicating the empirically-derived coefficient. The observed coefficient for the relationship between network-driven organizing and cards signed is well beyond two standard deviations higher than the mean (or 0).

Panel B of Figure A2 shows the same placebo distribution when the number of organizer notes, rather than the network structure, is randomly reshuffled across workers in a store. We create 500 datasets, holding the network fixed but reshuffling the number of organizer notes across workers within a store. While the true coefficient is substantially greater than the mean, it is less than two standard deviations away, and so is just under the magnitude required for statistically significance at conventional levels.

Panel C of Figure A2 shows a more traditional permutation test result, where the network-driven organizing scores are randomly reshuffled across stores in each of the 500 datasets. In this placebo distribution, the true coefficient is well into the tails of the distribution and looks quite similar to the main OLS specification. Generally, these tests suggest the robustness of the observed relationship between network-driven organizing and card signing.

A.2   Which Measure of Centrality?

In the OLS and instrumental variable regressions, we focused on two measures of centrality, degree centrality and eigenvector centrality. Degree centrality ignores second-order connections (neighbors of neighbors), and only counts centrality as measured by direct links. Alternatively, eigenvector centrality measures centrality globally, assessing connections to other connected individuals.

However, there are numerous other measures of centrality. The Friedkin and Johnsen (1999) model of social learning, where there are "stubborn" agents who update their beliefs taking a weighted average of their neighbors and their own past belief with weight α, gives rise to a centrality measure that is proportional to the well-known Bonacich-Katz centrality measure, parameterized by α.

In Appendix Figure A3 we compare our OLS and instrumental variable estimates across measures of centrality, looking across Bonacich-Katz measures with different parameters. Reassuringly, across the bulk of the parameterizations, the results look numerically quite similar to both eigenvector and degree centrality based estimates. The exceptions come from Bonacich-Katz centralities with $\alpha = 0.8$ and $\alpha = 0.9$, where increasing weight is put on distant ties. The OLS estimates are attenuated towards zero, while the IV estimates become more imprecise, consistent with increased measurement error. While our sparse network data is not ideal for examining details of dynamic percolation of beliefs through networks that would precisely recover α, Appendix Figure A3 suggests that our results are not driven by any idiosyncratic choice of network centrality.

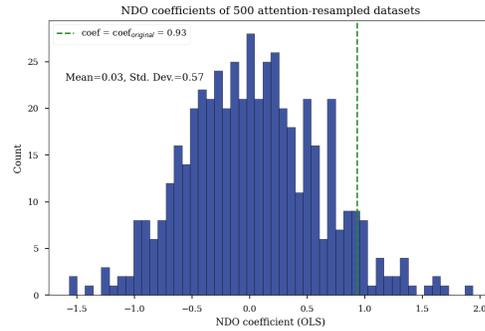Figure A1: Variable Correlation Heatmap
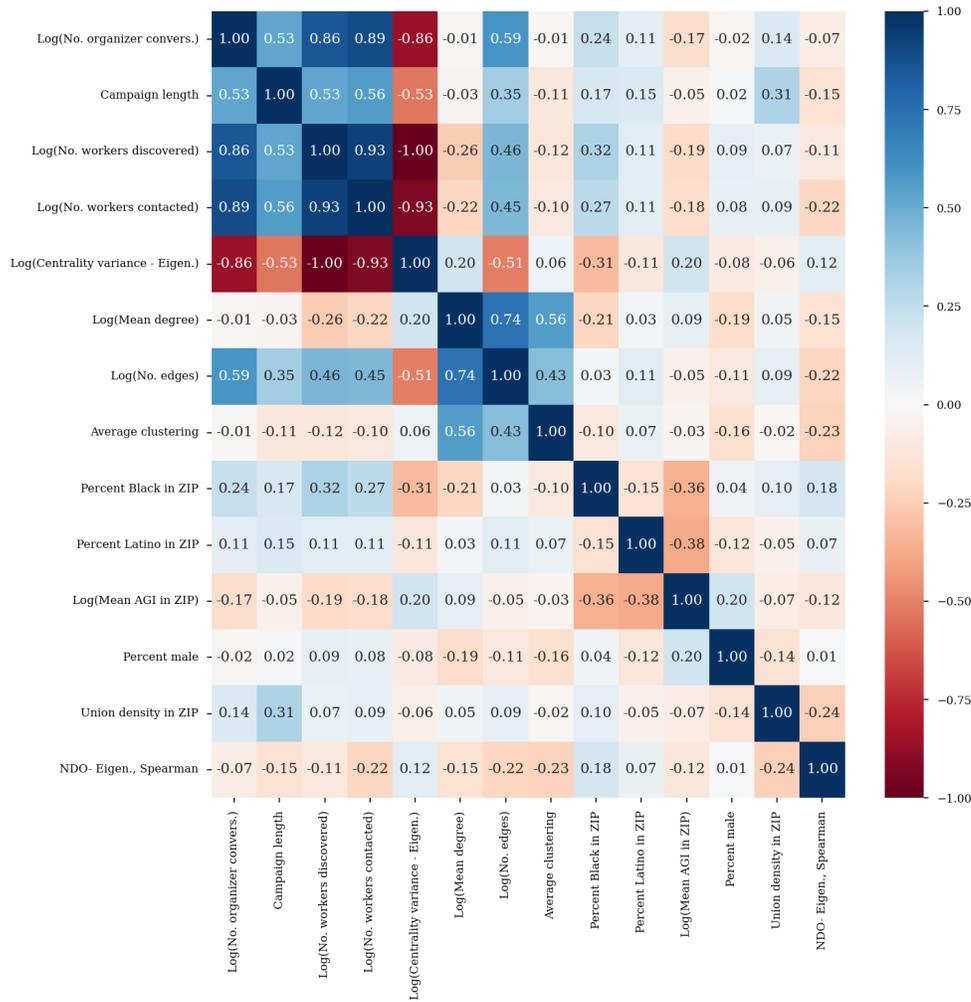
Pearson correlation between pairs of indicator variables used in regression analysis.

Figure A2: Permutation Tests

NDO coefficients of 500 network-resampled datasets

Mean=0.02, Std. Dev.=0.32
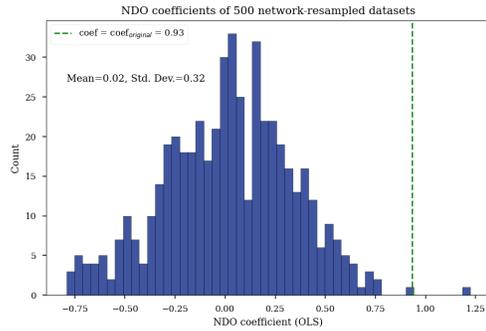
Panel A: Distribution of coefficients in regressions with re-sampled networks.

NDO coefficients of 500 attention-resampled datasets

Mean=0.03, Std. Dev.=0.57

Panel B: Distribution of coefficients in regressions with re-sampled organizer attention.

NDO coefficients of 500 NDO-shuffled datasets

Mean=0.01, Std. Dev.=0.33
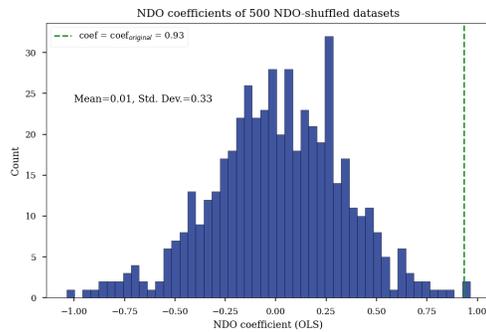
Panel C: Distribution of coefficients in regressions with re-sampled NDO scores.

Distributions of OLS coefficients, regressing logged number of cards signed on network-driven organizing, conditional on all controls in Table 1. Panel A is based on 500 datasets generated by randomly sampling (with replacement) nodes from the original networks to produce networks with the same number of workers and organizer conversations, with random network structures. Panel B is based on 500 datasets generated by holding the network structure constant in each store but randomly permuting the number of organizer conversations across workers. Panel C is based on 500 datasets generates by reshuffling the network-driven organizing score across workers.

Figure A3: Robustness to Different Centrality Measures
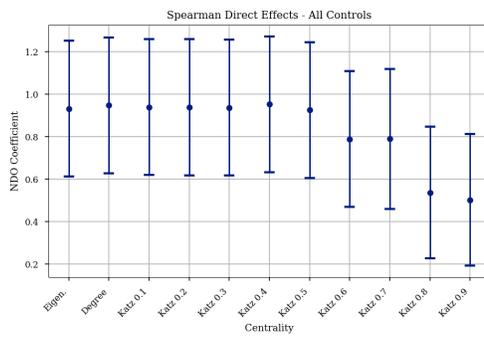
OLS (NDO = rank correlation) with all controls.
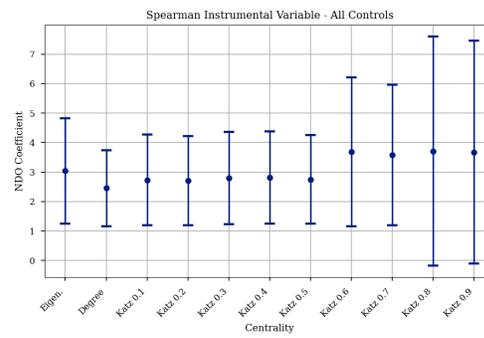
IV (NDO = rank correlation) with all controls.

OLS (NDO = Pearson correlation) with all controls.

IV (NDO = Pearson correlation) with all controls.

Table A1: OLS Results (Pearson correlation)

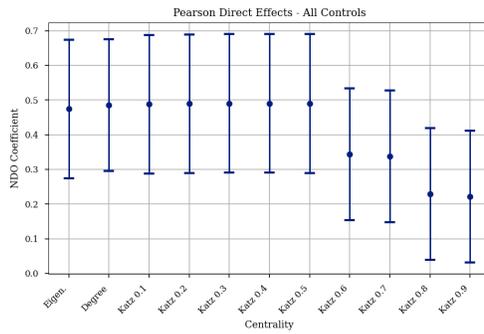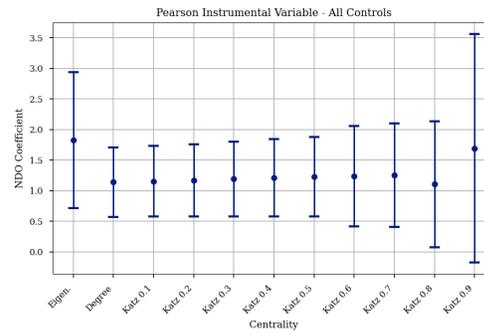|	1	2	3	4	5	6	7	8
NDO	**0.58**	**0.47**	**0.47**	0.38	**0.50**	**0.49**	**0.56**	0.41
	(0.21)	(0.20)	(0.20)	(0.22)	(0.18)	(0.19)	(0.19)	(0.23)
Log(No. organizer convers.)	**0.81**	**0.47**	**0.47**	**0.43**	**0.80**	**0.42**	**0.42**	0.38
	(0.08)	(0.18)	(0.18)	(0.22)	(0.08)	(0.18)	(0.19)	(0.22)
Log(Mean degree)	**0.31**	90.49	90.49	116.73	**0.31**	99.56	64.44	129.60
	(0.05)	(89.65)	(89.65)	(103.95)	(0.05)	(88.49)	(87.63)	(102.99)
LASSO-selected	N	N	Y	N	N	N	Y	N
Centrality metric	Eigen.	Eigen.	Eigen.	Eigen.	Degree	Degree	Degree	Degree
No. Worker Controls	N	Y	Y	Y	N	Y	Y	Y
Campaign Length Controls	N	Y	Y	Y	N	Y	Y	Y
Other Network Statistics	N	Y	Y	Y	N	Y	Y	Y
Demographic-Union Controls	N	N	Y	N	N	N	Y	N
Team Fixed Effects	N	N	N	Y	N	N	N	Y
Adjusted R_{sq}	0.56	0.65	0.65	0.69	0.55	0.65	0.65	0.69
N_{obs}	120	120	120	120	120	120	120	120

Results of regressions showing the effects of network-driven organizing metrics on number of cards signed, with various controls. No. workers includes log number of workers ever contacted and log number of workers mentioned in field notes. Campaign length controls are indicators for quintiles of campaign length. Other network statistics include the log of the variance of worker centrality, logged number of edges, and average clustering coefficient. Standard errors are robust, and coefficients significant at 95% are in bold text.

Table A2: Instrumental Variable Controls Test

	A		B		C
Log(No. organizer convers.)	-0.01	Log(Centrality Var)	-0.01	Percent Black in ZIP	**0.16**
	(0.01)		(0.01)		(0.03)
	[0.02]		[0.02]		[0.08]
Campaign length (Quint. 1)	0.03	Log(Mean degree)	-0.01	Percent Latino in ZIP	0.01
	(0.03)		(0.01)		(0.04)
	[0.06]		[0.01]		[0.13]
Campaign length (Quint. 2)	0.05	Log(No. edges)	-0.00	Percent male	0.07
	(0.03)		(0.01)		(0.10)
	[0.02]		[0.01]		[0.10]
Campaign length (Quint. 3)	-0.01	Average clustering	0.12	Log(Mean AGI in ZIP)	**-0.06**
	(0.03)		(0.38)		(0.02)
	[0.02]		[0.20]		[0.02]
Campaign length (Quint. 4)	-0.00			Union density in ZIP	**-0.54**
	(0.02)				(0.18)
	[0.03]				[0.40]
Log(No. workers discovered)	0.00				
	(0.01)				
	[0.02]				
Log(No. workers contacted)	-0.02				
	(0.01)				
	[0.03]				

IV bivariate coefficients on controls (robust and clustered errors in round and square brackets, respectively). Spearman correlation. Coefficients significant at 95% are in bold text.

Table A3: Instrumental Variable First Stage Results (Rank Correlation)

	1	2	3	4	5	6
Team score	**0.76**	0.36	**0.40**	**0.79**	**0.42**	**0.46**
	(0.15)	(0.19)	(0.19)	(0.14)	(0.17)	(0.17)
	[0.09]	[0.15]	[0.14]	[0.08]	[0.14]	[0.13]
Log(No. organizer convers.)	-0.01	**0.17**	**0.18**	-0.01	**0.18**	**0.18**
	(0.02)	(0.04)	(0.05)	(0.02)	(0.04)	(0.04)
	[0.02]	[0.04]	[0.04]	[0.02]	[0.04]	[0.05]
Log(Mean degree)	-0.02	33.80	**-0.05**	-0.02	35.29	**-0.05**
	(0.01)	(28.93)	(0.02)	(0.01)	(28.62)	(0.02)
	[0.02]	[47.71]	[0.03]	[0.02]	[48.41]	[0.03]
LASSO-selected	N	N	Y	N	N	Y
Centrality metric	Eigen.	Eigen.	Eigen.	Degree	Degree	Degree
No. Worker Controls	N	Y	Y	N	Y	Y
Campaign Length Controls	N	Y	Y	N	Y	Y
Other Network Statistics	N	Y	Y	N	Y	Y
Demographic-Union Controls	N	Y	Y	N	Y	Y
Adjusted R_{sq}	0.21	0.38	0.36	0.25	0.42	0.40
N_{obs}	118	118	118	118	118	118

Results of first-stage regressions, showing the effects of leave-one-out team scores on network-driven organizing metrics, with various controls. The leave-one-out network-driven organizing team score is calculated as the mean of a store's associated regional stores' (based on teams of organizers) scores, excluding the score of that store. No. workers includes logged number of workers ever contacted and logged number of workers referenced in field notes. Campaign length controls are indicators for quintiles of campaign length. Other network statistics include the log of the variance of worker centrality, log number of edges, and average clustering coefficient. Standard errors are robust, and coefficients significant at 95% are in bold text.

Table A4: Instrumental Variable Reduced Form Results (Rank Correlation)

	1	2	3	4	5	6
Team NDO score	**1.39**	**1.11**	**1.19**	**1.29**	**1.04**	**1.11**
	(0.51)	(0.53)	(0.51)	(0.48)	(0.49)	(0.48)
	[0.84]	[0.30]	[0.24]	[0.79]	[0.27]	[0.21]
Log(No. organizer convers.)	**0.82**	**0.59**	**0.62**	**0.82**	**0.59**	**0.62**
	(0.09)	(0.19)	(0.20)	(0.09)	(0.19)	(0.20)
	[0.08]	[0.26]	[0.21]	[0.08]	[0.26]	[0.21]
Log(Mean degree)	**0.26**	73.57	-0.29	**0.26**	73.41	-0.29
	(0.05)	(96.72)	(0.73)	(0.05)	(96.53)	(0.73)
	[0.07]	[82.02]	[0.84]	[0.07]	[81.70]	[0.84]
LASSO-selected	N	N	Y	N	N	Y
Centrality metric	Eigen.	Eigen.	Eigen.	Degree	Degree	Degree
No. Worker Controls	N	Y	Y	N	Y	Y
Campaign Length Controls	N	Y	Y	N	Y	Y
Other Network Statistics	N	Y	Y	N	Y	Y
Demographic-Union Controls	N	Y	Y	N	Y	Y
Adjusted R_{sq}	0.55	0.64	0.64	0.55	0.64	0.64
N_{obs}	118	118	118	118	118	118

Results of reduced-form regressions showing the effects of leave-one-out network-driven organizing team scores on number of cards, with various controls. No. workers includes logged number of workers ever contacted and logged number of workers referenced in field notes. Campaign length controls are indicators for quintiles of campaign length. Other network statistics include the log of the variance of worker centrality, log number of edges, and average clustering coefficient. Standard errors are robust, and coefficients significant at 95% are in bold text.

Table A5: Instrumental Variable First Stage Results (Pearson correlation)

	1	2	3	4	5	6
Team NDO score	**0.77**	**0.35**	**0.35**	**0.83**	**0.48**	**0.53**
	(0.13)	(0.16)	(0.16)	(0.11)	(0.14)	(0.14)
	[0.09]	[0.09]	[0.09]	[0.06]	[0.11]	[0.11]
Log(No. organizer convers.)	-0.01	**0.22**	**0.22**	0.01	**0.31**	**0.32**
	(0.03)	(0.07)	(0.07)	(0.04)	(0.08)	(0.08)
	[0.04]	[0.12]	[0.12]	[0.05]	[0.12]	[0.11]
Log(Mean degree)	**-0.06**	46.01	46.01	**-0.07**	7.76	-22.56
	(0.03)	(43.60)	(43.60)	(0.03)	(47.52)	(46.40)
	[0.02]	[79.70]	[79.70]	[0.03]	[100.68]	[90.94]
LASSO-selected	N	N	Y	N	N	Y
Centrality metric	Eigen.	Eigen.	Eigen.	Degree	Degree	Degree
No. Worker Controls	N	Y	Y	N	Y	Y
Campaign Length Controls	N	Y	Y	N	Y	Y
Other Network Statistics	N	Y	Y	N	Y	Y
Demographic-Union Controls	N	N	Y	N	N	Y
Adjusted R_{sq}	0.26	0.43	0.43	0.36	0.48	0.46
N_{obs}	118	118	118	118	118	118

Results of first-stage regressions, showing the effects of leave-one-out team scores on network-driven organizing metrics, with various controls. The leave-on-out team NDO score is calculated as the mean of the associated regional team's scores at all stores associated with that team, excluding the focal store. No. workers includes log number of workers ever contacted and log number of workers mentioned in field notes. Campaign length controls are indicators for quintiles of campaign length. Other network statistics include the log of the variance of worker centrality, log number of edges, and average clustering coefficient. Standard errors are robust, and coefficients significant at 95% are in bold text.

Table A6: Instrumental Variable Reduced Form Results (Pearson correlation)

	1	2	3	4	5	6
Team NDO score	**0.79**	**0.64**	**0.64**	**0.68**	**0.55**	**0.62**
	(0.29)	(0.32)	(0.32)	(0.22)	(0.25)	(0.25)
	[0.36]	[0.22]	[0.22]	[0.27]	[0.20]	[0.18]
Log(No. organizer convers.)	**0.81**	**0.61**	**0.61**	**0.81**	**0.61**	**0.65**
	(0.08)	(0.19)	(0.19)	(0.08)	(0.19)	(0.20)
	[0.08]	[0.26]	[0.26]	[0.08]	[0.26]	[0.22]
Log(Mean degree)	**0.26**	58.94	58.94	**0.27**	54.45	6.34
	(0.05)	(99.81)	(99.81)	(0.05)	(99.07)	(98.43)
	[0.07]	[84.81]	[84.81]	[0.07]	[83.94]	[90.15]
LASSO-selected	N	N	Y	N	N	Y
Centrality metric	Eigen.	Eigen.	Eigen.	Degree	Degree	Degree
No. Worker Controls	N	Y	Y	N	Y	Y
Campaign Length Controls	N	Y	Y	N	Y	Y
Other Network Statistics	N	Y	Y	N	Y	Y
Demographic-Union Controls	N	N	Y	N	N	Y
Adjusted R_{sq}	0.56	0.64	0.64	0.56	0.64	0.64
N_{obs}	118	118	118	118	118	118

Results of reduced-form regressions showing the effects of leave-one-out team scores on card-signing, with various controls. No. workers includes log number of workers ever contacted and log number of workers mentioned in field notes. Campaign length controls are indicators for quintiles of campaign length. Other network statistics include the log of the variance of worker centrality, log number of edges, and average clustering coefficient. Standard errors are robust, and coefficients significant at 95% are in bold text.

Table A7: Instrumental Variable Results (Pearson correlation)

	1	2	3	4	5	6
Network-driven Organizing	**1.03**	1.83	1.83	**0.82**	**1.14**	**1.17**
	(0.38)	(1.11)	(1.11)	(0.27)	(0.57)	(0.51)
	[0.41]	[0.78]	[0.78]	[0.28]	[0.46]	[0.44]
Log(No. organizer convers.)	**0.82**	0.21	0.21	**0.80**	0.27	0.26
	(0.08)	(0.33)	(0.33)	(0.08)	(0.25)	(0.24)
	[0.07]	[0.42]	[0.42]	[0.07]	[0.37]	[0.36]
Log(Mean degree)	**0.33**	-25.05	-25.05	**0.33**	45.62	32.83
	(0.06)	(137.04)	(137.04)	(0.06)	(98.76)	(88.75)
	[0.07]	[155.00]	[155.00]	[0.07]	[108.80]	[110.50]
LASSO-selected	N	N	Y	N	N	Y
Centrality metric	Eigen.	Eigen.	Eigen.	Degree	Degree	Degree
No. Worker Controls	N	Y	Y	N	Y	Y
Campaign Length Controls	N	Y	Y	N	Y	Y
Other Network Statistics	N	Y	Y	N	Y	Y
Demographic-Union Controls	N	N	Y	N	N	Y
First-stage F-Stat (Robust)	32.34	4.66	4.66	56.71	11.76	13.82
First-stage F-Stat (Clustered)	88.28	17.98	17.98	191.17	23.88	26.85
Adjusted R_{sq}	0.55	0.54	0.54	0.56	0.62	0.63
N_{obs}	118	118	118	118	118	118

Results of regressions using the leave-on-out team score as an instrument, showing the effects of network-driven organizing metrics on card-signing, with various controls. The leave-on-out team NDO score is calculated as the mean of the associated regional team's scores at all stores associated with that team, excluding the focal store. No. workers includes log number of workers ever contacted and log number of workers mentioned in field notes. Campaign length controls are indicators for quintiles of campaign length. Other network statistics include the log of the variance of worker centrality, log number of edges, and average clustering coefficient. Standard errors are robust, and coefficients significant at 95% are in bold text.

B   Social Learning and Workplace Organizing

In this Appendix, we connect different measures of centrality to well-known models of social persuasion on networks (DeGroot, 1974) to motivate our measure of network-driven organizing (NDO). Motivations might also be conceived as strategic (e.g. Ballester et al., 2006) but based on the qualitative evidence, social learning seems like an important mechanism, and we formalize that intuition here.

We suppose the subjective value to worker i of signing a card at time t is given by $v_i(t)$. We further suppose that co-workers in a workplace are strongly connected by a symmetric unweighted graph given by adjacency matrix A, with $A_{ii} = 0$. This network can reflect relationships of trust, Bayesian priors on whose signals to put more weight on, or simply the number of interactions that result in learning at work.

We modify the assumptions of a social learning model by including the existence of an organizer. Assume an organizer can choose an allocation of conversation effort e_i, corresponding to how much effort to spend influencing worker i. Conversation effort is a one-time event, say at the beginning of the process, that can increase a worker's subjective value of signing permanently by e_i, at a cost to the organizer of $\frac{C}{2}e_i^2$, reflecting that there are increasing marginal costs to influencing any individual worker. The convex costs to investing effort in a single worker seems to accord with both qualitative evidence on the difficulties of locating the same worker over time as well as the increasing costs for organizers to induce large changes in worker beliefs.

Why might an organizer care about the sum of subjective values? One reason is that, particularly in the labor organizing context, having as many people as committed to signing possible is the best antidote against the anti-union campaign that begins once the employer learns of the organizing effort, for example when cards are filed with the NLRB. In the OUR Walmart context, more cards means more paying, committed members, and more participation and higher probability of success in workplace collective actions.

Degree Centrality: The simplest updating would have each worker engage in only one period of

social learning from their neighbors. Then the level of support is given by:

$$v(1) \cdot 1 = Av(0) \cdot 1 = \sum_i \sum_j A_{ij}(v_j(0) + e_j) = \sum_j d_j(v_j(0) + e_j)$$

An organizer trying to maximize $v(1) \cdot 1$ with quadratic costs would thus choose $e_j = d_j/C$, and allocate effort proportional to degree. Organizers with a higher correlation between degree centrality and effort (i.e. NDO with degree centrality) will have a higher level of support.

Eigenvector Centrality: We can motivate the eigenvector centrality NDO by changing the structure of social learning, so that individuals update every period, as in a De Groot model. Define the influence matrix induced by A as G, where $G_{ij} = \frac{A_{ij}}{d_i}$ if $i \neq j$ and $G_{ii} = 1$ otherwise (this ensures aperiodicity of the influence matrix), with $d_i = \sum_j A_{ij}$ being the degree of worker i.

In this model, individual worker i updates their beliefs about the value of signing a card based on the beliefs of the people they are connected to in the network G. After T periods of learning, with no organizer effort, the subjective value of worker i will be

$$v_i(T) = (G^T v(0))_i$$

. In an undirected, connected, and aperiodic network, the steady-state beliefs of everyone in the workplace converge to the same value, given by the dot product of initial beliefs $v(0)$ with the eigenvector C^E of the matrix G corresponding to the eigenvalue of 1 (which is the highest eigenvalue as G is a stochastic matrix). In steady-state $v(\infty)_i = C^E \cdot v(0)$ for all i, i.e. there will be consensus.

Empirically, our networks are disconnected due to limitations of measurement, so the classic result showing convergence to a consensus distribution Jackson, 2010 may not be applicable. But if we think the connected component is large and aperiodic, the objective function can be approximated by the steady-state vector of beliefs.

We assume that social learning occurs quickly, and the organizer aims to maximize the steady-state sum of values, i.e. $\lim_{t \to \infty} \sum_i v_i(t) = \sum_i C_i^E(v_i(0) + e_i)$. One can then see that steady-state card

signing will be maximized where $C^E \cdot e$ is maximized, and the organizer invests $e_i = \frac{C_i^E}{C}$ in each worker i. Where organizers invest the most in the most influential workers, the average long-run probability of card signing will be highest.

Note that the organizer does not necessarily know the true influence vector C_i^E, but merely sees a signal of it from a partially observed network. Because we only measure the network as observed by the organizer, we can still see that the expected long-run consensus number of cards signed will be maximized by the organizing investing the most effort in the workers they *perceive* to be most influential.

Katz-Bonacich Centrality: An alternative model, based on Friedkin and Johnsen (1999), would have organizers continuously exercising effort vector e in each period, persuading workers continuously rather than a one-time effort. Organizer effort pushes a belief by a factor α relative to learning from peers, so that the updating rule is:

$$v(t+1) = (1-\alpha)Gv(t) + \alpha e$$

. This gives rise to a steady-state belief vector

$$v(\infty, \alpha, e) = (I - (1-\alpha)G)^{-1}\alpha e$$

With quadratic costs as above, the optimal allocation of e that would maximize the steady-state sum of values is $e_i = \alpha C((I - (1-\alpha)G)^{-1})_i$, which means that it is proportional to Bonacich-Katz centrality with parameter α, as presented in Appendix Figure A3.

For all of these centrality measures the resulting level of support is proportional to the correlation of e_i and centrality within a store. However, since organizer effort is on an arbitrary scale, and the quantitative prediction should hold with any positive rescaling of the vectors, in the empirical work we will use rank correlations, to ensure that our measures are not being driven by any particular normalization. We examine robustness to standard correlations (Pearson) in Appendix Tables A1-A7.

C   Examples of Notes

Date	Worker name	Text
11/09/2010	Sofia Torres	10/25/10; Maria house called; card signer; Bakery Dept. Issues: Too much to do with not enough people. Not enough support from supervisors, and they want her to do more. Would cut her hours because she doesn't have open availability, but since it is high demand in the bakery, they haven't cut her hours. - Lopez, R
05/18/2012	Johnny Hughes	Friend of Isabel Garcia - Schneider, D
10/10/2011	Elena Herrera	Liz has worked at wmt 2280 for 4 years ICT Receiving and Inventory. She makes $12.10/hr works FT 7a-4p M-F. Issues: did not receive full raise in July Review ($0.40). They told her it is because she is not a role model. Her husband (Jerry Alberto Gomez, also a member, former) helped her write a letter to her manager after her review in July 2011, but they didn't give her any answer. She is very upset because she said she does too much work and deserves to make $12.30. Store manager Ben asked her to write a list of all of her qualifications and reasons why she deserves her raise. We helped translate list into English and are waiting to hear the results (9/30/11, Emma and Devika) - Farrow, L
11/28/2010	Ian Kimutai	CTW: 3/24/09 - Signed union card. Referred by Charlotte Jansen. Arthur Jones collected card. Loves Obama. Makes $12.39/hr. - Hoang, T

Examples of organizer notes. "Worker name" reflects the worker (pseudonym) with whom the organizer had a conversation, and "Date" reflects the date of that conversation. Network edges were drawn to all workers whose names appear in the note's "Text".

Reference

Friedkin, Noah E and Eugene C Johnsen. 1999. "Inuence networks and opinion change."
Advances in Group Processes 16:1–29.