Workplace Networks and the Dynamics of Labor Organizing

Hana Shepherd, Rebecca Roskill, Suresh Naidu, and Adam Reich

Preliminary: Please do not cite or circulate.

November 7, 2022

Abstract

How does workplace organizing – attending to, and intervening upon, relationships among workers – relate to organizing outcomes? We present detailed, establishment-level evidence of the importance of workplace networks to organizing outcomes among Walmart workers from 2010 to 2015. We process over 80000 unstructured organizer field notes from 118 store-level campaigns conducted at Walmart stores matched to an organizing outcome (signing a card indicating membership in a worker voice organization). We reconstruct workplace networks as perceived by the organizer, and then, motivated by a DeGroot model of labor organizing, construct a new measure of network-driven organizing (NDO): the correlation between organizer attention to a worker and that worker's network centrality. Using interview transcripts from 35 workers, combined with a qualitative analysis of organizer notes, we demonstrate what network-driven organizing looks like in practice, and corroborate our constructed measure. We then show that this measure is positively correlated with campaign success at the workplace level, and that this correlation is robust to a variety of controls. We also construct an instrument using the leave-one-out mean network-driven organizing of the organizer team assigned to a store, and show that this instrument is highly correlated with both store-level network-driven organizing as well as organizing success. A campaign going from an NDO correlation of 0 to a correlation of 1 is associated with at least a 100% increase in cards signed, over a baseline mean of 24 cards. These results indicate that organizing strategies that focus on workplace leaders are more likely to be successful, and provides a statistic for organizing efficacy that could be employed by unions and other worker organizations.

1 Introduction

Labor unions and other worker voice organizations have recently received renewed public, policy, and academic attention. In the U.S., these organizations have generally formed through a process of organizing at the level of the establishment, as staff organizers and worker leaders persuade other workers within a particular workplace to engage in collective action in order to establish a union or a group to improve working conditions. While the effects of unions on key outcomes like poverty, mobility, inequality, and racial disparities are are the subject of extensive literatures, there is little quantitative research on the dynamics of organizing new unions or worker organizations.

We use new network data constructed from unstructured organizer field notes from 118 store-level organizing drives at Walmart between 2010 and 2015. Our main outcome of interest is the number of membership cards signed within a store, a measure of collective action. Drawing on extensive interviews among Walmart workers involved in campaigns, in combination with a qualitative examination of the organizer notes, we show how successful and unsuccessful organizing strategies differentially interacted with workplace networks. Quantitatively, we show that organizers that targeted their effort and attention, as measured in the field notes, to workers perceived to be central in the workplace relationship network, are able to generate significantly higher levels of collective action, as measured by membership cards signed and monthly dues.

These results speak to longstanding comparative questions about the low level of union density in the United States (Naidu 2022, Eidlin 2020). Owing to both employer opposition and labor law, recognition of a new union or other worker voice organizations entails overcoming formidable barriers to worker collective action. In order to overcome such barriers, social capital and social networks among workers are key. But detailed examination of the role of relationships in the labor organizing process has remained largely confined to qualitative case-studies. Quantitative data on worker collective action and networks during an organizing drive remains rare, and our paper helps fill this gap.

In contrast to the literature on unions, the social movement literature has long recognized the role of relational organizing and social networks to collective action outcomes. This literature supports the idea that social networks are important to explaining movement participation and nonparticipation (for a succinct review, see Krinsky and Crossley 2014). Scholars have offered a range of explanations for why this might be the case, from processes of interpersonal influence, whereby people mobilize (or demobilize) their neighbors (Kim and Bearman 1997; McAdam and Paulsen 1993; Polletta 1998); to processes of information diffusion, whereby networks allow people to understand others' intentions for participation and thus to overcome collective action problems (Chwe 1999; Granovetter 1978; Kim and Bearman 1997; Macy 1991; Oliver et al. 1985); to processes of collective identity formation (Gould 1993; Gould 1995; Tilly 1978), whereby networks change how people understand themselves, and thus the collectivities on behalf of which they are willing to act.

For all its promise, however, this existing literature has empirical and theoretical limitations. Empirically, the vast majority of scholars who have explored the relationship between the formal structure of social networks and movement participation rely on simulation models rather than empirical evidence (Kim and Bearman 1997; Oliver et al. 1985; Chwe 1999; Macy 1991; Gould 1993; for an exception see Gould 1995). This is likely a result both of the difficulty of obtaining relational data regarding participants and non-participants within the context of a social movement, and of the small number of cases of insurgency within most social movement analyses, which make it difficult to generalize about the importance of network characteristics beyond the context of a particular campaign.

Theoretically, the literature is limited by the assumption that networks (and network positions) are stable characteristics of a setting that exist outside the control of movement actors themselves. Scholars have thus tended to overlook the ways in which movement actors might attend and respond to - as well as intervene upon- network structures and processes. Such a second-order understanding of, and strategic engagement with, social networks is likely a key characteristic of one of the most important yet undertheorized roles within a social movement or labor organizing campaign: that of the "organizer." Understanding whether and how organizers' mapping of, and engagement with, the networks of potential supporters – i.e. their learning about existing ties; their focusing on particular individuals; their rewiring of the network through the introduction of new ties – impacts the success of a campaign has implications both for our academic understanding of the role of networks in movements, as well as for social movement organizations themselves.

Below, we present new statistical evidence on the determinants of success in labor organizing at America's largest employer, Walmart. Motivated by the qualitative literature on labor organizing as well as simple models of network-driven social learning, we develop a new measure of network-driven organizing: the correlation between an organizer's attention to a potential participant and that potential participant's centrality in a wider network of potential participants (as this network is understood by the organizer.) We then test this measure using detailed organizing notes collected by labor organizers across 118 workplaces, in order to examine the relationship between organizers' network practices and their success at recruiting new members to a workplace organization. We find that workplaces in which organizers use network-driven organizing – i.e. in which organizers focus more on the workers they perceive to be central to the workplace network – are those in which organizers have more organizing success. While we cannot perfectly establish causality with OLS, we can rule out numerous confounds, including those selected by a double-LASSO procedure.

We further probe causality by constructing an instrumental variable based on the organizing team assigned to a given store. The labor organization we study divided Walmart stores into different regions, and each region received a different team of organizers, with different organizing strategies and cultures. We construct a leave-one-out average of network-driven organizing of the team assigned to a given store, and show that this variable is both strongly correlated with the level of network-driven organizing as well as the number of cards signed. Under a plausible assumption that the team assignment is independent (or independent conditional on observable variables) of other determinants of collective action at the store level, instrumental variables estimates show a large effect of network-driven organizing on cards signed.

The estimated coefficient magnitudes are large and meaningful in our context: going from a 0 correlation between network centrality and organizer effort to a perfect correlation of 1 results in at least a 100% increase in cards signed, with IV estimates increasing to over 200%. However, given the low baseline rate of card-signing, with only 13 cards signed at the median store and 24 cards on average, network-driven organizing is not on its own a magic bullet to sway an organizing drive. Nevertheless, it may be decisive in cases where a sizeable minority of workers are initially interested in organizing. We discuss the magnitudes and their intepretation in section 6.2.

2 Organizers as Networkers

The "organizer" has an almost mythical status within the study of social movements. On the one hand, social scientists generally appreciate the importance of organizers in explaining the emergence and success of social movements; on the other, they have struggled to define precisely what it is that these leaders do and how they do it. This is related to a more general struggle within the social sciences to theorize the practices by which actors successfully change institutional environments. While a wide variety of interesting theoretical constructs—from "institutional entrepreneurship" (DiMaggio 1988) to "robust action" (Padgett and Ansell 1993) to "social skill" (Fligstein and McAdam 2012)—have been proposed, such constructs tend to be defined in terms of the outcomes they produce (e.g. "Social skill can be defined as the ability to induce cooperation by appealing to and helping to create shared meanings and collective identities," Fligstein and McAdam 2012: 46), which limits their usefulness in terms of explaining when such collective outcomes succeed and when they fail.

Within the context of social movements, early efforts at theorizing the "organizer" recognized the relationship between the practices of an organizer and a network of social relationships with which the organizer would engage. In her account of the origins of the women's movement, for instance, Jo Freeman (1973) discussed "communication networks" as a necessary but insufficient condition for movement emergence. While occasionally a crisis could "galvanize[] the network into spontaneous action" (p. 794), she observed, rarely could a movement emerge or persist without the strategic action of organizers. However, while recognizing that the "role of the organizer in movement formation is... [a] neglected aspect of the theoretical literature" (p. 807), Freeman did not go much further in specifying either the network or the work that organizers do in relationship to it.

Within the historical literature on social movements, the most explicit treatment of organizing may be Charles Payne's (2007[1995]) *I've Got the Light of Freedom: The Organizing Tradition and the Mississippi Freedom Struggle*, in which he describes the work of the Mississippi Student Non-Violent Coordinating Committee field staff in 1963: a group of forty-one workers who were mostly young, "mostly Black, mostly southern, [and] mostly from working-class backgrounds" (p. 237). Payne describes in exacting detail the many ways that these young organizers went about the "slow work, respectful work" (p. 243) of organizing. In his account, the organizer played many different roles at once: "Organizers had to be morale boosters, teachers, welfare agents, transportation coordinators, canvassers, public speakers, negotiators, lawyers, all while communicating with people ranging from illiterate sharecroppers to well-off professionals and while enduring harassment from the agents of the law and listening with one ear for the threats of violence" (p. 246). Tying together these multiple roles was the goal of identifying, recruiting, and developing potential movement leaders and supporters. On the one hand, the organizer worked to identify and develop "informal leaders" (pp. 248-249) in a community, those people who were not necessarily endowed with any institutional authority but who were well liked and well respected by others; on the other hand, the organizer reached out to everyone they could contact through canvassing, "going door-to-door, trying to draw people in" (p. 250). Payne summarizes the role of an organizer by quoting the Civil Rights leader Bob Moses, who responded to a question about how to organize a town by saying,

> "By bouncing a ball," he answered quietly.
> "What?"
> "You stand on a street and bounce a ball. Soon all the children come around. You keep on bouncing the ball. Before long, it runs under someone's porch and then you meet the adults" (p. 243).

The organizer in Payne's account seems akin to a relationship worker or network architect, working within existing ties and developing new ones, with the goal of recruiting new people to a movement and deepening the commitment of those already involved.

Building on such historical case studies, sociologists and political scientists have sought to distinguish "organizing" from other types of movement activity. In unpublished course notes, which have nevertheless diffused widely among movement actors and academics alike, Marshall Ganz (2006) provides a succinct definition of organizers as those who "identify, recruit, and develop leadership; build community around leadership; and build power out of community."

Jane McAlevey (2016), focusing on the labor movement in particular, discusses how organizers "analyze the workers' preexisting social groups" (p. 34) using conversations with workers to learn which of their peers are considered "organic leaders" in the workplace. For McAlevey, leaders are not determined based on prior involvement in or enthusiasm toward the union but rather as a result of the influence they have with their coworkers. McAlevey quotes a labor organizer, Kristin Warner, who says, "[Organic leaders are] almost never the workers who most want to talk with us... They have a sense of their value and won't easily step forward, not unless and until there's a credible reason" (p. 34). Likewise, Hahrie Han (2014, p. 14), in her study of a range of NGOs, suggests that organizers "do not simply aggregate individuals but also create new relationships between them that generate new commitments and resources." In particular, Han continues, organizers do things like "make requests for action that bring people into contact with each other..." (p. 16); "focus on building relationships and community through interdependent (as opposed to individual) action" (p. 16); and develop people's leadership through "extensive training, coaching, and reflection" (p. 17).

Han and McAlevey both contrast organizing with what they call "mobilizing," which they each describe somewhat differently but each suggest is limited in its efficacy. Perhaps because "mobilizing" is the foil against which each author juxtaposes the "organizing" approach to which they are more sympathetic, mobilizing is not particularly well-defined in either account. Han gets closest to a definition of mobilizing when she writes that mobilizers "do not seek to transform people's interests as they recruit them for action" but rather are focused on "building their membership base" (p. 15). If organizers are interested in identifying, recruiting, and developing leaders, mobilizers are focused on finding the people who are already interested in acting and giving them opportunities for doing so, "allow[ing] people to self-select the level of activism they desire" (p. 15). If organizers are engaged in relational work, helping people to come together (and deepen their commitment) through interdependent action, mobilizers are more concerned with giving each individual supporter (or potential supporter) an opportunity to act in the way that they feel comfortable acting.

While this nascent literature on organizing uses the metaphor of networks to describe the work that organizers do, organizing has not been defined or evaluated in formal network terms. By bridging this literature with previous work that has examined social movements using more formal network methods, we can operationalize the concepts of "organizing" and "mobilizing," and compare these strategies against one another in the context of a multi-sited campaign.

One key difference between an organizing approach and a mobilizing approach is the organizer's or mobilizer's relationship to the social network of potential supporters. An organizer works to build support for a campaign by making use of a network of social influence: i.e. the organizer identifies supporters who are central in a network of potential supporters, and then works with them to generate support among those others. In contrast, a mobilizer is concerned with reaching out to the largest number of potential supporters as possible. Rather than working through central nodes in a network and encouraging these central actors to contact and recruit their peers, a mobilizer likely prioritizes finding new people to inform about a campaign or cause, in the hopes of encountering existing supporters.

We can operationalize this difference by attending to how an organizer or mobilizer allocates their attention across the universe of potential supporters. An organizer will likely focus their attention on the supporters or potential supporters who are most central in the network of potential supporters, even after those people have been successfully recruited. A mobilizer, in contrast, will likely focus their attention on reaching out to new potential supporters themselves. More formally, to the extent that a campaign is engaged in organizing, we expect there to be a correlation between the time devoted to supporters or potential supporters and their centrality within the full (perceived) network of potential supporters.

3 Data

In this paper, we use several types of information from a voluntary association of employees working at Walmart—OUR Walmart (henceforth, "OUR"). During the period in which these data were collected, OUR advocated for Walmart workers through collective actions such as attendance at the company's annual shareholders' meeting, media, and smaller-scale campaigns at specific stores; it did not and does not seek union recognition. We make use of an anonymized database maintained by OUR that includes information about workers with whom paid organizers for OUR were in contact between 2010 and 2015. During this period, the paid organizers (some but not all of whom were former Walmart workers themselves) sought to engender support from employees for actions and to recruit employees to be members of OUR. (The organization no longer has a similar membership structure.) The organizers would do so by making initial contact with workers in a store through brief and often surreptitious interactions on in their workplaces.

3.1 OUR Walmart Member Information

This database includes information about all members of OUR, including the stores for which they work or worked. The database also includes the date at which a worker was entered into the database and, for those employees who became OUR members, the date at which the member signed an OUR membership card. In our contexts, card signing is a meaningful outcome that indicates successful organizing, as signing is associated with a commitment to providing $5 in monthly dues.

3.2 OUR Walmart Field Organizer Notes

In addition to member information, the database includes a total of 91,019 "Notes" written by organizers in accordance with procedure established by OUR. Each note was logged by an organizer after they had a conversation with a worker and is associated with the date of the conversation and the unique identifier of the worker. We use these notes for two main purposes: first, we measure organizer attention to a worker as the number of notes indicating a conversation with that worker; second, we identify other workers' names in the text of the notes as indicators of relationships between workers that have been discovered or cultivated by the organizers (see subsection 3.5 on page 7). We provide some examples of the raw note data in Appendix B.

3.3 Organizing Teams (Unfinished)

We will see below that organizing teams are an important determinant of organizing strategy.

3.4 Campaign Data

Organizing activity, as reflected through card-signing activity and organizer notes, varies dramatically from store to store. Campaign length, measured as the number of weeks between the first and last organizer notes linked to the store, ranges from 63 weeks to 251 weeks (Figure 1).

3.5 Defining and Coding Network Edges

We construct the organizer's conception of the store-level network based on co-occurrence of individual employees' names within organizer notes about a store. For example, if an organizer indicated that they spoke with two employees about OUR within the same note, then we considered those two individuals to have a relationship with each other, denoted by a 1 in the store-level adjacency matrix. This method of network construction reflects an organizer's attention to relationships between employees within a store. Relationships that organizers recorded among employees could take two forms: relationships that pre-existed the organizer's attention and efforts (e.g., friendship or family relationships that the organizer recorded), *organizer-independent networks*, and relationships that the organizer facilitated among employees, such as asking two employees to come to a meeting together, *organizer-dependent networks*. We combined both of these types of relationships into an adjacency matrix representing any type of relationship between employees. We emphasize that the network does not represent the "true" underlying relationships between employees, but instead it represents the way the organizer conceives of the set of relationships among employees. Within each store adjacency matrix, we calculate the centrality of each individual using two centrality measures: degree centrality and Eigenvector centrality. The centrality measure of an employee reflects the number of coworkers that the organizers perceived to be socially connected to that employee through any type of relationship (degree centrality) or the calculation of the relative influence of an individual employee based on the Eigenvector score of his or her network connections.

3.6 Network-directed Organizing

Our key independent variable is our measure of network-directed organizing. The measure is based on field organizers a) recording relationships among employees in a store and b) recording that they spent more organizing time (assessed by the number of notes organizers associate with an employee) with those central employees, which we refer to below as organizer attention. In order to ensure that these two factors can vary independently of each other, our measure of organizer attention is calculated as the number of organizer notes for any individual employee minus notes that record their relationships (which were used to measure their centrality).

Given a network adjacency matrix in store j A^j, we define the network-directed organizing measure as the within-store rank-correlation between centrality (either degree or Eigenvector centrality) and organizer effort (notes that do not contain information about relationships with other employees, non-edge notes):

$$NDO_j = Corr(Rank(Cent_i^j), Rank(Notes_i^j))$$

Where $Cent_i^j$ is the degree or Eigenvector centrality of employee i at store j, $Notes_i^j$ is the number of non-edge organizer notes mentioning employee i, and $Corr()$ is the within-store correlation between employee centrality and the number of non-edge notes that include that employee taken over the $I(j)$ workers in the dataset for the duration of the campaign at store j. The main measure is the within-store (and, necessarily, within-campaign) correlation between organizer attention $Notes$ and employee centrality $Cent$.

We focus on rank correlations between organizer effort and worker centrality in our main specifications in order to minimize the influence of outliers and any non-normality in the underlying distributions. The distributions of the NDO measures are shown in Figure 8, and both versions have means slightly below 0, and vary from -0.5 to 0.56.

We illustrate the data underlying this measure in Figure 2, discussed above, which shows the networks corresponding to the Pico Rivera (one of the highest levels of NDO) and Federal Way (one of the lowest), with nodes (employees) scaled by the amount of organizer attention. Quantitatively, the NDO level of Pico Rivera is .56 while the NDO level of Federal Way is close to 0.

3.7 Outcome Variable: Campaign Card-Signing

We examine if campaigns marked by more network-driven organizing (a higher correlation between organizer attention and employee centrality) are more successful, measured by OUR membership cards signed.

We use signed membership cards as our primary indicator of store-level collective action, and we aggregate the total number of signed cards to the store level. As this is naturally determined by a variety of campaign-level characteristics beyond our network-driven organizing measure, we take care to either normalize by or control for mechanical determinants of card signing. For example, we control for the number of workers ever mentioned by the organizer, the number of workers ever contacted by the organizer, the length of the campaign, and the total number of notes made by the organizer in an attempt to isolate effects of network-driven organizer strategy that are independent of other determinants of card signing.

Card signing is not cheap talk. Signing a card was also a pledge to pay five dollars in monthly dues, as a signal of worker commitment to OUR Walmart's work at both the local and national levels. The organizer's objective was to gain signed cards, with the idea that the larger the share of workers who had signed cards, the more successful a variety of collective actions would be, ranging from specific policy demands like 'Respect the Bump' (pregnancy benefits) to actions like Black Friday strikes, in which workers walked off the job on the busiest shopping day of the year.

4 Qualitative Evidence From Two Walmart Campaigns

The quantitative data we have access to helps us to see that network-driven organizing is positively associated with more success in an organizing campaign. It does not, however, provide us with a picture of what, specifically, network-driven organizing might look like on the ground. We use two case studies, one of successful organizing and one of unsuccessful organizing, to help illustrate what types of processes might be at play in network-driven organizing.

Any store will have an existing, heterogeneous set of relationships between employees. Some of these relationships will be kin or friendship relationships, and others will be relationships based on sharing experiences, in the absence of a specific individual tie, as when workers share a shift or have children who attend the same school. The network we observe in these stores is one where the organizer is both learning about these existing ties and creating organizing-relevant ties by inviting coworkers to attend meetings together, suggesting that they talk with each other, and meeting with them together. To the extent that organizers continue to follow up with people who have more contacts, they are pursuing a network-driven strategy such that high-centrality people receive more energy and effort.

4.1 Organizing at Work

The organizing campaign at a store in Pico Rivera, a town in southeast Los Angeles County, began like it did in many other stores, with organizers from OUR Walmart trying to speak to as many workers as they could. This could be a somewhat long and painful process. For example, the very first worker that organizers recorded in their notes, in November of 2010, was Juan. He thought that organizing with OUR Walmart was a "good idea," he told the organizers, but wanted to wait until after the holidays to "see what happens." Another organizer met with him again in early December, but he was "still not ready to sign up" for the organization, and did not "want to be the first person in the store" to do so. Organizers contacted Juan, who ended up with 9 organizer notes by the end of the campaign, and a network centrality score of 0, the lowest possible, again in mid-January, when he was "still undecided," and again in March of 2011, when an organizer visited him at his home and spent at least ninety minutes in conversation, at which point he said he "still wants to wait and see if his situation changes." This was the last time that Juan appeared in the organizing notes, suggesting either that Juan left his job or that the organizers gave up on trying to bring him around.

At the beginning, as they were chatting up people like Juan, organizers were also experimenting with a range of tactics to get the names and contacts of potential recruits. For example, one strategy that appears in the notes from those early days was a holiday raffle—workers could enter the raffle to win a prize, and in return OUR Walmart would get the worker's name and contact information: an old-school data mining operation, which yielded several new names, but not many new members.

Around the same time that the organizers met Juan, though, they also met Dora Avila; according to Dora, one of the organizers was dating her ex-brother-in-law, and they started talking about OUR Walmart over a coffee at Starbucks. After the meeting, the organizer wrote that Dora had been working at Walmart for five-years and had a well-defined set of grievances: Walmart would cut people's hours and change their shifts arbitrarily; the managers would show favoritism to some workers over others, often in a way that reeked of sexism; and the company made her pay $160 a month out of pocket for health care. In Dora's account, she immediately saw the appeal: "I was like, 'Oh, sign me up.' [The organizer] was like, 'For reals?' So I was like, 'Yes. I got it, I understood it. Sign me up.'" But Dora remembers thinking to herself, "My number one thing was how do I get everybody else to sign up?" She was the fourth worker to sign up for the organization in the Pico store, but would quickly become one of the most essential. In our data, Dora accumulated 158 organizer notes and was the most central individual in the store network.

Having identified Dora as a potential leader, organizers began to meet consistently with her to support her as she reached out to her coworkers. Ten days after an organizer met with Dora for the first time, Dora brought a second worker (Lourdes) to a meeting with another organizer at a nearby shopping center. Lourdes signed up for the organization at that meeting, representing the first time that a sign-up occurred through the efforts of a worker leader. Lourdes is associated with 13 organizer notes, and is in the 67th percentile for network centrality.

Dora soon proved that she could recruit her coworkers to the organization in a way that organizers could not. This was not just because she knew who they were and how to reach them, but also because she was respected by her coworkers and thus able to influence them in ways that organizers were not. For instance, on December 7, an organizer had approached Lorena, who worked in the bakery department (25 organizer notes and the 67th percentile for network centrality). In that meeting Lorena had not been sure about the organization: on the one hand, she wanted "more help," and "more respect" from Walmart; on the other hand, she was "very scared" that she was "going to lose her job." A month later, though, on January 12th, Dora was able to convince Lorena to come to an organizing meeting. At this meeting, she "got her to sign." In early March, a worker who had recently joined the organization (Gabriel, 6 notes and 67th percentile in network centrality) told an organizer that "all contact to him should be through Dora," again suggesting the esteem in which her coworkers held her.

Dora was clearly central in the social network of potential supporters at Pico. Importantly, organizers from OUR Walmart seemed to recognize and nurture this centrality, providing Dora with support, advice, and encouragement as she took on and completed assignments. In early January of 2011, two organizers sat down with Dora to talk about how she might work to sign up her coworkers, and she left with five sign-up cards in hand. Rather than reach out to workers themselves, organizers took a step back and supported Dora as she became increasingly active in the nascent organization. On February 9th, Dora set up an organizing meeting with five of her coworkers. On March 18th, Dora led a meeting at the local Shakey's Pizza, where three more of her coworkers signed up for the organization. In the meantime, organizers were regularly in touch with Dora, strategizing with her about recruitment; putting her in touch with journalists who were beginning to cover the campaign; inviting her to meetings and trainings. As Dora organized among her coworkers, organizers increasingly invested time in supporting her. Excluding notes recording connections that Dora had made with others, organizers recorded 138 conversations or meetings with Dora individually over the course of the 2.5 year campaign.

We see a similar pattern in the way that OUR Walmart identified and developed another key leader at the Pico Rivera Store, Michelle Rogers. As documented in Reich and Bearman (2018:167-168), Michelle reports that she initially heard about OUR Walmart from "Crazy Dora Avila." As Michelle recalls, "She was always asking me, 'Hey, mama, how's things going?' And I would tell her, 'Not good,' you know... And she would say, 'You know, when we get a chance, let's talk.'" The two met at a Del Taco, a nearby fast-food joint, where Dora introduced Michelle to some of the OUR Walmart organizers. Michelle went home and looked up the organization online. She concluded that "if I was going to have to be here for a few more years," she would have to "either make changes or just take the beatings." She signed up for the organization on February 12, 2011. Michelle herself was well-connected in the store, in the 99th percentile for network centrality.

In those early months, Michelle attended a few meetings, but did not do much more than pay the organization's monthly dues. In July, however, an organizer sat her down to encourage her to be more active in the organization. Specifically, the organizer asked that Michelle take on two "assignments" to speak with her coworkers about OUR Walmart: as they did with Dora, organizers encouraged Michelle to recruit her coworkers to the organization. In this case, though, it is unclear whether Michelle ever completed these assignments, and her involvement diminished again.

As the campaign began to heat up in the early months of 2012, however, Michelle reached back out to the organizers saying she "felt bad that she ha[d] not been active." An organizer asked that she meet, and asked her to bring other coworkers who might be supportive. At this meeting, in March of 2012, the organizer and these three workers had a "heart to heart" about the importance of the organization, how important it was for their membership to grow at the store, and how critical it was for these three to take on more leadership. Michelle recommitted to the organizing process, this time along with her friends, and took on new assignments. Again, organizers sought to build support for the organization through the relationships that workers had.

To the extent that workers were willing and able to connect the organization with others, the organization invested more time and attention on them. When it became clear that Michelle was connected to others and able and willing to reach out to them, organizers invested more in her. Between late 2010 and February of 2012, organizers reported only six conversations with Michelle. Between February 2012 and mid-2014, organizers reported 46.

The strong ties that worker-leaders had with one another, in addition to their ties with organizers, also likely helped them take riskier action than they would otherwise have been willing to take (see Reich and Bearman 2018:168-169). For instance, both Dora and Michelle recall their first one-day strike in early October of 2012. Dora remembers being terrified: "I kept thinking, 'Oh my God, oh my God, oh my God, we just went on strike!'" But she had to hold it together for Michelle: "My friend was next to me. She was like, 'Are you okay?' And I snapped out of it." She continued, "I had to get into that mode of, like, you cannot show your fear, because you're going to show it to everyone else, and they see you as a leader and you can't." In other words, Dora explicitly linked her participation in the action to her desire to uphold the image that others had of her as a leader. Michelle, in turn, remembers how Dora helped her overcome her own fears: "I was getting sick. I was shaking and I kept telling myself, 'What am I doing?'" But Dora helped her persist: "I eventually did. I eventually walked with Dora."

Through the identification and development of leaders like Dora and Michele, OUR Walmart organizers at Pico were able to build an active committee of workers open to taking collective action. Of course, Walmart seemed to recognize the threat posed by workers at the Pico Rivera store as well. In April of 2015, the company announced that there was a plumbing problem at Pico, and that it would have to shut down and layoff its workforce. When it reopened six months later, those most active in OUR Walmart were not rehired.

4.2 Organizing without the Network

The campaign at the Federal Way Walmart outside of Tacoma, Washington, shared several features with the Pico campaign. Based on organizer effort (indicated by total notes logged by organizers), the campaigns seemed practically equivalent: organizers at Federal Way logged 1,350 notes over the course of the campaign, about seven percent less than the 1,448 notes logged at Pico Rivera. At Federal Way, organizers were in contact with more workers than they were at Pico: organizers log notes about 326 workers at Federal Way (94% of all store contacts), compared to just 170 at Pico Rivera (55% of store contacts). At Federal Way, though, organizers did not seem to make use of existing store networks the same way as they did at Pico: although they did use organizing conversations to map the shop's social network, they used this social network only to reach uncontacted workers, rather than to prioritize central workers with a greater capacity to influence their coworkers. Ultimately, organizers at Federal Way managed to sign up 83 workers for the organization, or 24% of contacts. At Pico, organizers managed to sign up 118 workers for the organization, or 38% of store contacts, a success rate nearly 60% higher. One reason for this difference in success rate seems to be the different strategies deployed by organizers in the two stores.

Early in the campaign at Federal Way, organizers seemed to identify several workers who were central in the workplace network, in that they were able to provide organizers with the names of other potential supporters. And yet organizers did not seem to follow up with these workers, or support them in their efforts to reach out to others. For instance, in June of 2011, a worker named Daniel convinced a coworker named Eleanor Bernard to sign up. Organizers had approached Eleanor earlier that year, in April, but she had demurred on participating in the organization. She had "answered the door with a baby in her arms," and had told organizers she was not interested in joining the organization because "she might be quitting soon and [Walmart wasn't] important to her." But Daniel had persuaded her, illustrating again the power of networks of influence.

Eleanor seemed to have the potential to be Federal Way's Dora Avila. She made efforts to introduce coworkers to organizers and she was in the 99th percentile of network centrality in the store. On June 30, 2011, as organizers waited outside her store, Eleanor convinced three coworkers to meet with them outside on their breaks. Just as some workers at Pico refused to talk to organizers but were willing to talk to Dora, several of Eleanor's coworkers refused to talk to anyone but Eleanor. In early August of 2011, she gathered a group of workers outside her store to sign a declaration of principles. And then Eleanor disappears from the notes. There are no logs of any conversations with Eleanor; not even a note indicating that Eleanor has cooled on the organization. The next note about Eleanor occurred on March 5, 2012, seven months after the declaration of principles, when her membership dues lapsed because her credit card was declined. She is recorded as having attended one final meeting a week later, on March 12th, and then disappears again. The final note about her (out of a total of 6 notes), in January of 2013, notes that she had been inactive for six months and was now opposed to the organization.

This seemed like something of a pattern at the Federal Way store. In early 2011, organizers stopped by the house of Erik Fraser, who expressed interest in the organization. Between July and October of that year, he provided organizers with information about fifty-six of his coworkers, reflecting his 100th percentile network centrality. And yet organizers did not seem to meet regularly with Erik, or support Erik in reaching out to his coworkers. While Erik is logged as having walked out on strike during a national action in November of 2012, there are few recorded meetings or conversations with him (a total of 8 notes). By September of 2013 he had left his job at Walmart. Marla Alexander was very active in October and November of 2011, identifying potential leaders and coming to meetings. Based on organizer records, she was in the 99th percentile of network centrality. Then, somewhat abruptly Marla seems to disappear from the log after 10 organizer notes.

We summarize the differences between the two campaigns in the network graphs in Figure 2, which show Pico Rivera on the top and Federal Way on the bottom. The Pico Rivera graph, besides showing a high number of card signing nodes, also clearly shows more organizer notes for the more central nodes. The Federal Way graph shows no differential investment by organizers in the most central nodes, and our hypothesis is that this explains the low rate of card signing in that store. The De Groot model of labor organizing in the Appendix formalizes this idea, and suggests the Network-Directed Organizing measure that we use in the regressions below.

To be fair, based solely on the organizing notes, it is difficult to establish definitively whether the differences between Pico and Federal Way are differences in organizer strategy or differences in organizing context. While it appears, for example, that organizers at Pico made much more of an effort to develop Dora Avila as a leader than organizers at Federal Way did to develop Eleanor Bernard, we cannot rule out the counterfactual that Dora was simply more open to being organized than Eleanor – more supportive of the organization, more willing to devote time and effort, more committed to voice over exit. Did organizers do something differently with Michelle at Pico than organizers did with Erik at Federal Way, or was Michelle simply a different kind of person than Erik? This is a limitation of our study more generally, in that we definitely demonstrate that the organizing metric we identify here – the association between a worker's centrality and the attention that organizers pay to them – is driving organizing outcomes, or whether there is some confounding, contextual variable that explains both. Nevertheless, the qualitative data does seem to provide support for our interpretation that organizers attended to central leaders at Pico in ways that was missing from Federal Way—and, more generally, that our organizing metric may be causally related to organizing outcomes, though future research ought to explore this further.

5 Empirical Specification

We begin by presenting simple bivariate scatterplots that show our main result. Recall that our primary outcome of interest is log number of cards signed. But this must be divided by some measure of campaign duration, number of workers, or organizer effort. In Panel A of Figure 7 we show the log number of cards per worker-week, as a measure of share of workers who sign cards per week, plotted against our network-directed organizing measure. In Panel B we normalize the number of signed cards by organizer effort, as measured by number of total notes, as an alternative measure. In both panels we see a statistically significant relationship between the measure of network-driven organizing and number of cards signed per worker-week or per unit of organizer effort.

While Figure 7 shows the basic pattern, we turn to OLS regressions to show robustness to various sets of possible confounds. We estimate ordinary least squares regressions of the form:

$$log(Cards_j) = \beta NDO_j + X_j'\gamma + \epsilon_j \tag{1}$$

Standard errors are heteroskedasticity-adjusted. While we have a large number of potential control variables and confounders, discussed in the next section, our limited sample size prevents us from saturating the specification with all the controls that we have. We instead control for a parsimonious set of controls in our baseline specification, and then use double LASSO (Chernozukov et al.) to select among the other controls.

5.1 Controls

The first set of controls (Group A) are factors related to organizer mobilization efforts and campaign characteristics that are relevant to how many cards were signed at a store: the number of total organizer notes (representing the amount of effort the organizer expended during the campaign; logged), the length of the campaign in weeks, the total number of employees at each store as represented in the organizer notes (logged), and the total number of employees the organizer contacted (logged).

The second set of controls (Group B) relate to features of the employee networks as represented by the organizer notes about relationships between employees. Here, we consider the mean and variance of centrality (for either degree and Eigenvector centrality, depending on the model) for the store-specific network, as well as the mean degree, the number of relationships (edges) in the network, and the average clustering coefficient of the network. By controlling for these characteristics of the network, we attempt to isolate the unique relationship between network-driven organizing and card signing apart from how the organizer understood and represented the characteristics of the store network.

The third set of controls (Group C) included zip code-level demographic characteristics of the store, including the percentage of Black residents, the percentage of Latino/a residents, and the percentage male. We use these controls to account for the possibility that card-signing rates are related to features of social organization or solidarity for which this demographic information can serve as a proxy.

Finally, we follow procedures established by for LASSO-selection of controls among the above variables. Those controls were average clustering coefficient of the store network, variance in store-network centrality, number of edges in the store network, length of the campaign, logged number of organizer notes, logged number of workers in the store, logged number of workers contacted by the organizer, and zip-code level characteristics of the percentage Black, percentage Latino/a, percentage male, and average Adjusted gross income.

6 Results

Results from estimating equation 1 are in Table 2 . Column 1 shows a minimal specification where only the (log of) store mean of degree and log number of organizer notes are controls. These are the two variables whose correlation defines network driven organizing, and the coefficient on all three are positive and significant. Of note is that the coefficient on the network-driven organizing variable is positive and significant despite controlling for both organizer effort and a measure of the mean centrality of workers, suggesting that the targeting of effort to the high-centrality workers has an independent effect on cards signed.

Column 2 includes a large set of controls. We include controls for store-level demographics, computed from zip-code of the store address. This includes fraction Black, fraction Hispanic, and fraction white. We further control for the log of workers contacted or mentioned by the organizers, to measure the free-rider problems that come with larger numbers of workers. Finally, we include a large set of additional network statistics include the log of the variance of worker centrality, log number of edges, and the average clustering coefficient. We display this relationship in a binned scatterplot visually in Figure A1, visually confirming little in the way of non-linearities.

In this saturated specification, the R-sq increases by about 20%, and the coefficient on the network-driven organizing variable falls by around 20% as well. This parameter instability may suggest some important omitted variables. To explore this further, in column 3 we use double-LASSO (Chernokuzov et al. 2014) as a device to select important sets of covariates. The double-LASSO first uses L1-penalization to select variables (besides the NDO measure) that significantly predict the outcome, and then uses a separate L1-penalized regression to select variables that predict the NDO measure. Any variable selected in either of these regressions gets included in a final regression, with standard errors adjusted as in Chernozukov et al. (2014).

The resulting R-sq remains virtually identical, at .68, but the coefficient on the NDO measure is now quite close to the simple specification in column 1. This exercise suggests that once a sparse set of covariates that predict the outcome (or the treatment) are included, there is little parameter instability while the explanatory power of the saturated specification is maintained.

Columns 3-6 replicate these specifications using the rank correlation between degree centrality and organizer effort, with similar patterns and magnitudes. Recent econometric work has suggested that when networks are sparse and subject to measurement error, degree may have better finite sample properties than eigenvector centrality (Choi 2022); the similarity of the coefficients on the two measures should thus be reassuring.

As a final check on our basic empirical setup, we conduct a randomization-inference based test. We construct 500 datasets of 115 stores each, but within each store we draw workers with replacement. We assume each worker is connected to themselves, and then calculate all our network statistics, including eigenvector centrality, degree, and the correlation of organizer effort and worker centrality. We then estimate equation 1 on each of these datasets (including all the controls), obtaining 500 placebo β. Figure 8 shows the resulting histogram of the placebos, with a vertical line indicating the true coefficient. As can be seen, the true coefficient is well beyond two standard deviations higher than the mean (or 0).

To summarize, regardless of whether centrality is measured using Eigenvector or degree, net of the LASSO-selected controls representing organizer mobilization, store network characteristics, and local area demographics, the network-driven organizing term is positively associated with the number of cards signed in that store over the course of the campaign. That is, when the organizer has more interactions with high centrality workers, the campaign is more successful. For every one unit (standard deviation) increase in the Eigenvector network-driven organizing term, 26 more cards are signed in the course of a campaign.

6.1 Instrumental Variables Estimates

While the OLS estimates are quite robust, there is one covariate that reduces the magnitude and significance of the coefficient on our network driven organizing measure: an indicator for the team assigned to organize that store.

The fact that significant variation in the NDO measure is absorbed by the team indicators suggests a candidate instrumental variable. We construct a leave-one-out team score, defined formally as:

$$NDO_j^{Team} = \frac{1}{|Team(j)| - 1} \sum_{k \in \{Team(j)/j\}} NDO_k \tag{2}$$

where $Team(j)$ is the team assigned to store j. Thus the team measure is the leave-one-out average of the NDO measure of all the other stores with the same team.

Panel A of Figure 6 shows the basic bivariate binned scatterplot corresponding to the first-stage of the instrument. Despite only 10 teams, there is considerable variation in the underlying measure, partly driven by the heterogeneity in team size (CHECK THAT), and the leave-one-out average is significantly correlated with the left-out store's NDO measure.

The discussion above about how the teams were assigned to regions should provide some confidence that teams are exogenous to other determinants of organizing success besides strategic organizing.

Panels B and C of Figure 6 show reduced form scatterplots of the transformed cards measures against the leave-one-out team average. Both show strong and significant associations.

6.1.1 Specification and Results

In order to examine robustness to a variety of controls, we estimate a series of 2SLS regressions. We obtain instrumental variable estimates starting with the first-stage given by:

$$NDO_j = \alpha NDO_j^{Team} + X_j'\gamma + \epsilon_j \tag{3}$$

and reduced form equation given by:

$$log(Cards_j) = \eta NDO_j^{Team} + X_j'\gamma + \epsilon_j \tag{4}$$

The IV estimate of β will be given by $\widehat{\beta^{IV}} = \frac{\widehat{\eta}}{\widehat{\alpha}}$. The X_j will be the same families of covariates from equation 1. We report robust standard errors, but also reported clustered standard errors at the level of the team below in square brackets.

Table 3 shows IV estimates of β from specifications parallel to those in 2. The F-statistic from the first-stage is reported at the bottom of the table, and shows that the instrument is generally strong across specifications.

The results in Table 3 again suggest strong and significant effects of the NDO measure on cards signed, roughly double the size of the OLS effects. An increase in the correlation between organizer effort and worker centrality from 0 to 1 increases cards signed by between 2 and 3 log points, or roughly 40 to 70 cards signed.

In columns 3 and 6 we also use a variant of the double-LASSO procedure, adjusted for the IV. Following Belloni and Chernozukov (2014), we estimate three L1-penalized regressions (with penalization parameter chosen by 5-fold cross-validation) with different outcomes, and our full set of controls. We look at LASSO regressions for NDO_j, NDO_j^{Team}, and $log(Cards_j)$ as outcome variables. Any control variable selected in any of these 3 regressions in then included in the 2SLS regression as a control.

The IV coefficients are significantly larger than the OLS coefficients. This could be due to measurement error in the NDO measure, but also could be due to a failure of the exclusion restriction. Teams may differ in many respects besides the degree of network-driven organizing, and the assignment of teams to stores may not be uncorrelated with other determinants of campaign success. Further, the complier stores with the team-based instrument may be ones with larger responsiveness to network-based organizing, as the organizers that do not comply with the team-culture have some different information about what would work in their particular store. While we think these instrumental variables estimates of network-driven organizing point towards a causal interpretation, we cannot rule out all possible confounds, and suggest that establishing definitively causal estimates should be a focus of future work (with many more stores and other sources of quasi-experimental variation).

6.2 Interpretation of magnitudes

We estimate quantitatively large coefficients, particularly in the IV specifications. Even taking the 95% lower confidence bar, estimates imply an increase in the correlation between organizer effort and worker centrality from 0 to 1 increase cards signed by at least 100%. The sample median cards signed is 13, less than 10% of the average Walmart store employment, and likely less if turnover is accounted for.

Further, the highest level of NDO is 0.56, suggesting that even the most network-driven organizing campaign has some difficulties in targeting perfectly.

To put these numbers in some general context, organizing folk wisdom is to not file for NLRB election without more than 65% of unit having signed cards. This is because the employer opposition mounted during the election period can effectively cut down support by 15%, on average. While OUR Walmart was not aiming to file for NLRB recognition, it also means the organizing was happening without the pressure of a concerted and well-organized employer opposition campaign.

So while using networks in the organizing process can drastically improve organizing outcomes, even the most perfect network-based targeting is unlikely to move enough workers to cross majority, let alone the 65% threshold. While strategic use of workplace networks important, unlikely to drive organizing success on its own, highlighting the structural disadvantage organizers face in the high-turnover, low-wage environment of U.S. retail.

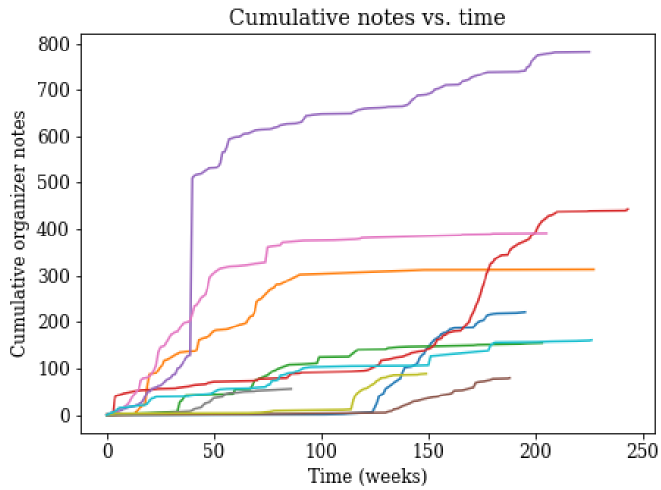
7 Conclusion

Our paper has implications not just for the study of networks in labor movements but also for the relationship between labor scholarship and the practice of labor organizing. We are able to analyze the relationship between organizer practice and organizing outcomes because of the uniquely rich relational data collected by organizers over the course of the OUR Walmart organizing campaign. This type of data is rarely preserved among labor organizations, and rarely shared with researchers.

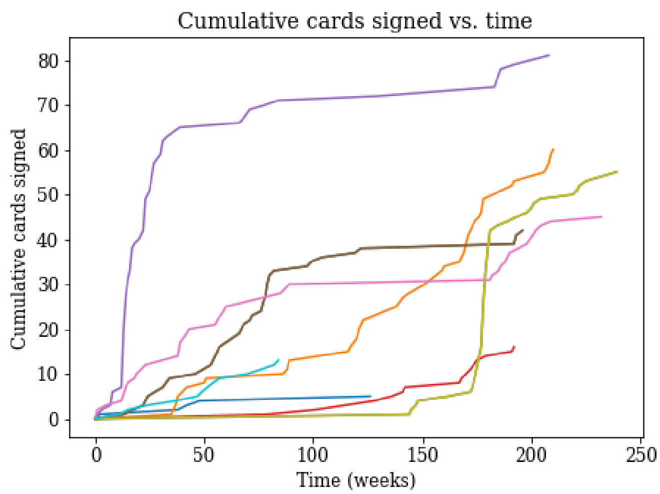
Conversely, the association we find between network-driven organizing and card signing, suggests a use for the data collected by labor organizations in the course of organizing. One immediate use of our results is that campaigns can be assessed on their degree of NDO in real-time, and organizer effort updated accordingly. Hopefully future work will modify and extend our results using both larger and more precise datasets, along with a variety of sources of quasi-experimental designs, to assess the efficacy of organizing strategies deployed by labor organizations.

8 Figures

Figure 1: Organizing Activity Over Time



Cumulative organizer notes over time for ten randomly selected stores in sample.

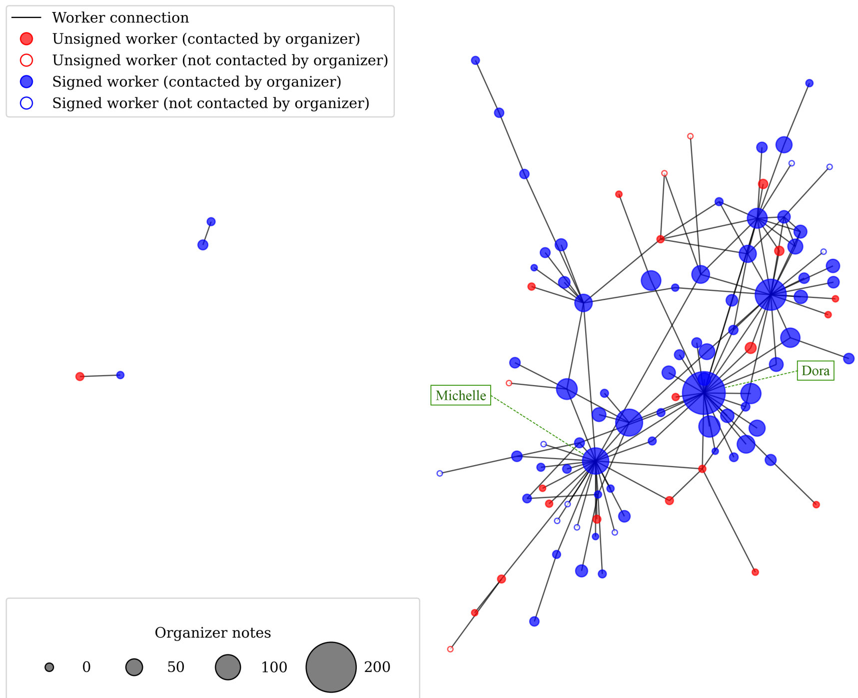


Cumulative membership cards signed over time for ten randomly selected stores in sample.

Organizing activity as measured through organizer notes and membership cards signed.

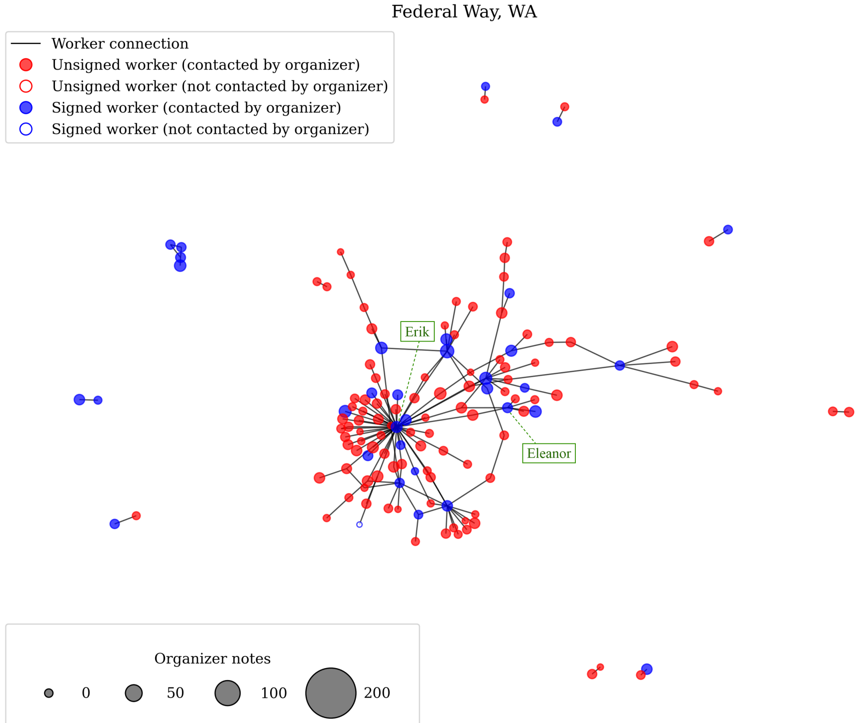
Figure 2: Pico and Federal Way Workplace Networks

Pico Rivera, CA



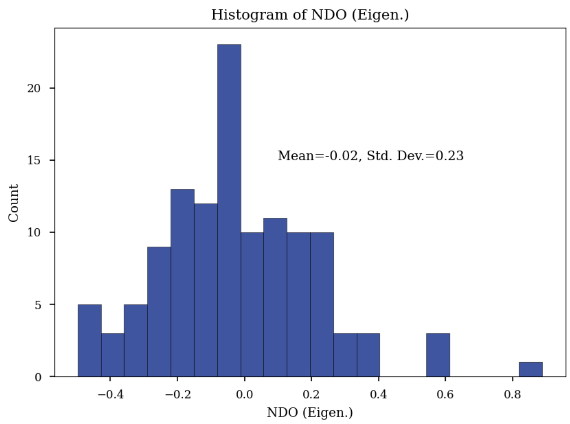
Workplace network from Pico Rivera, CA.

Federal Way, WA

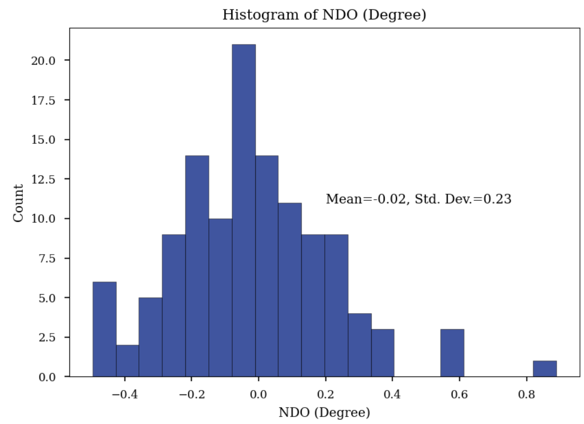


Workplace network from Federal Way, WA.

Figure 3: Network-Driven Organizing Measure



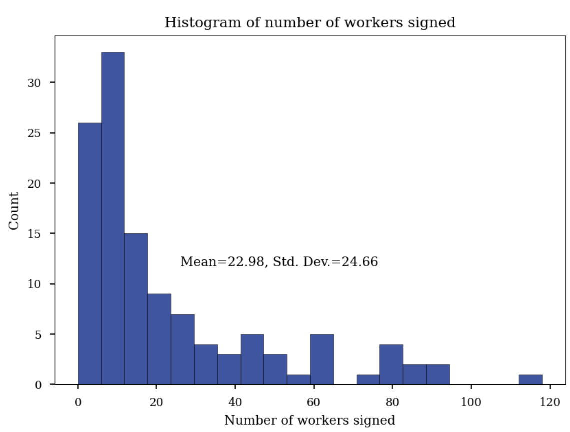
NDO scores using eigenvector centrality.



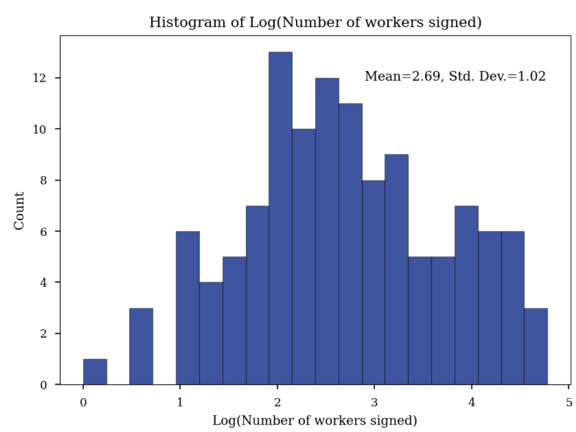
NDO scores using degree centrality.

Distributions of network-directed organizing (NDO) scores using the two selected centrality measures.

Figure 4: Signed Cards Measure



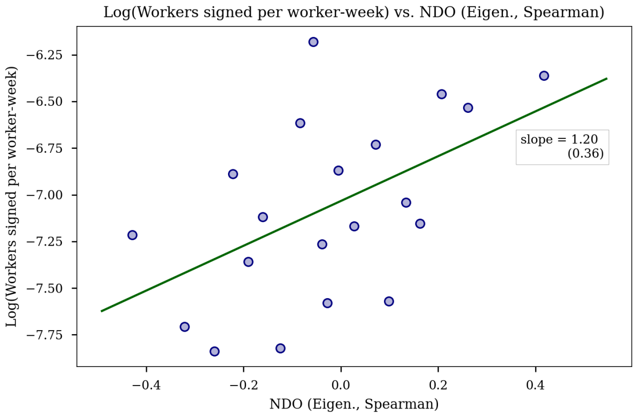
Histogram of number of workers signed.



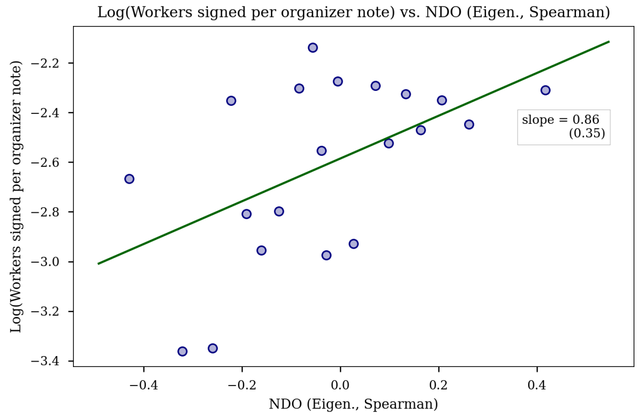
Histogram of number of workers signed (log-transformed).

Distributions of outcome variable, number of workers signed, which was log-transformed in the main regression to reduce skewness.

Figure 5: Cards Signed and Network-Driven Organizing, No Controls

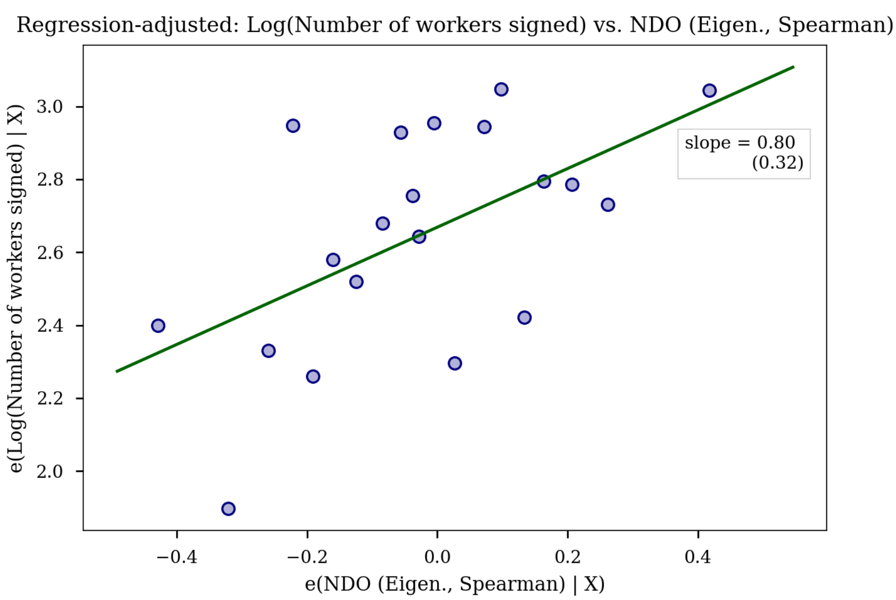


Number of cards signed per worker-week (log-transformed) vs. Eigenvector network-directed organizing measure.



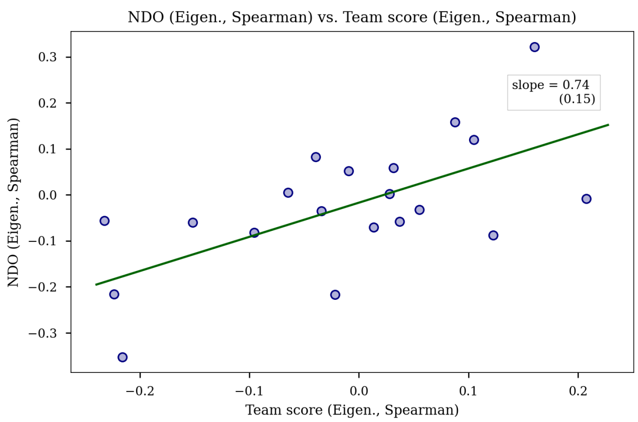
Number of cards signed per organizer note (log-transformed) vs. Eigenvector network-directed organizing measure.

Figure 6: Cards Signed and Network-Driven Organizing, Regression Adjusted

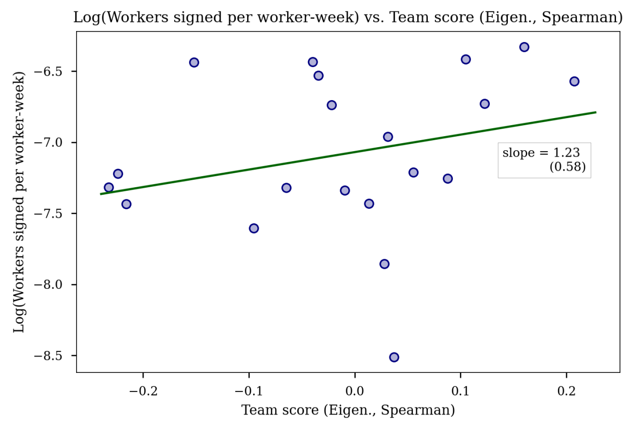


Scatter plot, showing network-directed organizing (NDO) vs. number of cards signed (log-transformed), and respective partial regression plot, conditional on all controls in Table 1.

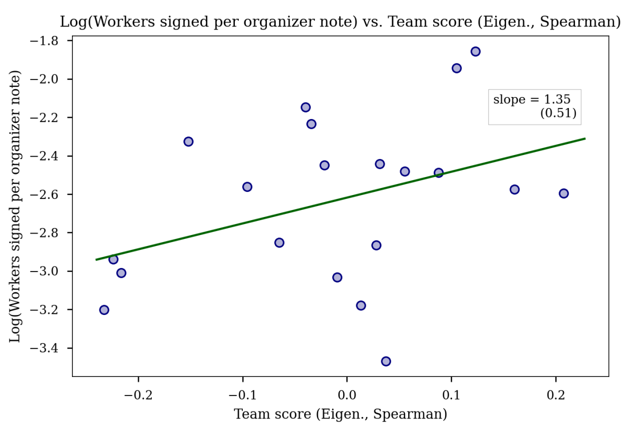
Figure 7: Leave-One-Out Team Score Instrument, No Controls



First stage: Eigenvector network-directed organizing measure vs. leave-one-out team score, based on Eigenvector network-directed organizing measure.

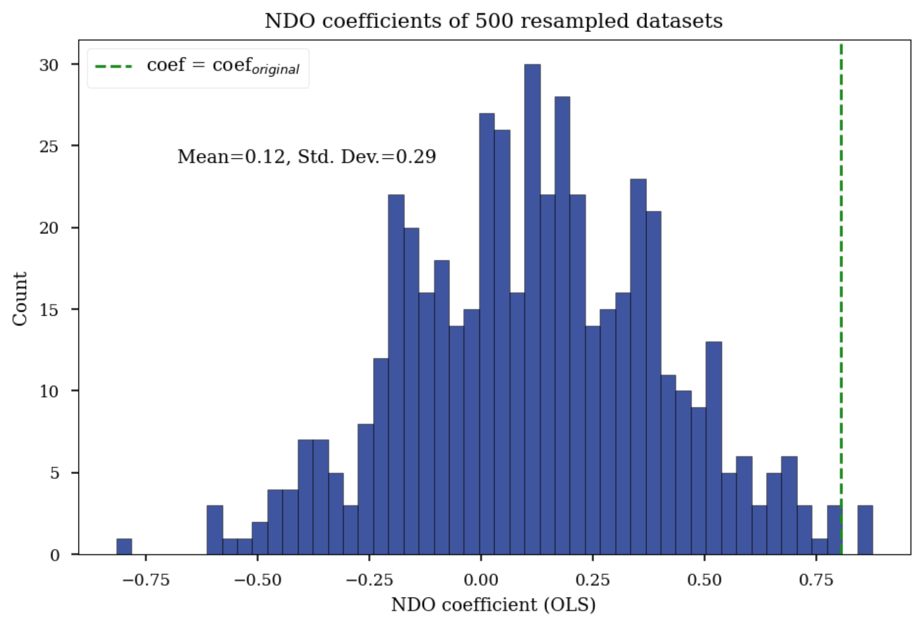


Reduced form: Number of cards signed per worker-week (log-transformed) vs. leave-one-out team score, based on Eigenvector network-directed organizing measure.



Reduced form: Number of cards signed per organizer note (log-transformed) vs. leave-one-out team score, based on Eigenvector network-directed organizing measure.

Figure 8: Distributions of coefficients in regressions on re-sampled networks



Distributions of OLS coefficients, regressing Log(No. workers signed) on Network-Driven Organizing, conditional on all controls in Table 1. 500 datasets were constructed from unique re-sampled networks, generated by randomly sampling (with replacement) nodes from the original networks to produce networks of the same number of workers, with permutated network structures.

9 Tables

Table 1: Summary statistics

	Mean	Median	Std. Dev.	Min	Max
Number of organizer notes	274.21	185.00	254.34	30.00	1416.00
Number of workers discovered	139.30	96.00	133.84	17.00	946.00
Number of workers contacted	104.52	68.00	116.85	10.00	793.00
Campaign length	179.55	202.29	57.68	22.14	251.14
Mean degree	0.16	0.11	0.16	0.00	0.78
Number of edges	18.53	9.00	26.49	1.00	141.00
Centrality variance	0.00	0.00	0.00	0.00	0.01
Average clustering	0.01	0.00	0.03	0.00	0.13
Percent black in ZIP	0.17	0.08	0.22	0.01	0.98
Percent Latino in ZIP	0.26	0.19	0.24	0.01	0.93
Percent white in ZIP	0.59	0.59	0.24	0.01	0.97
Percent male	0.40	0.41	0.10	0.00	0.64
Mean AGI in ZIP	41.62	37.42	19.02	21.58	141.08
NDO (Eigen., Spearman)	-0.03	-0.03	0.21	-0.50	0.56
NDO (Degree, Spearman)	-0.03	-0.03	0.21	-0.50	0.56
NDO (Eigen., Pearson)	0.05	0.05	0.36	-0.68	0.86
NDO (Degree, Pearson)	0.07	0.11	0.41	-0.83	0.92
Number of workers signed	23.83	13.00	24.71	1.00	118.00

N = 115. Summary statistics for all variables utilized in regressions.

Table 2: OLS Results

	1	2	3	4	5	6	7	8
NDO	**1.10**	**0.80**	**0.89**	0.47	**1.08**	**0.82**	**0.90**	0.47
	(0.34)	**(0.30)**	**(0.29)**	(0.30)	**(0.34)**	**(0.30)**	**(0.29)**	(0.31)
Log(No. organizer notes)	**0.88**	**0.65**	**0.63**	**0.70**	**0.88**	**0.64**	**0.62**	**0.69**
	(0.07)	**(0.18)**	**(0.19)**	**(0.20)**	**(0.07)**	**(0.18)**	**(0.19)**	**(0.20)**
Log(Mean degree)	**0.29**	**108.38**	**69.64**	**95.01**	**0.29**	**105.98**	**67.64**	**94.91**
	(0.05)	(73.45)	(61.43)	(84.28)	**(0.05)**	(73.41)	(61.03)	(84.23)
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 Controls	N	Y	Y	Y	N	Y	Y	Y
Team Fixed Effects	N	N	N	Y	N	N	N	Y
Adjusted R_{sq}	0.57	0.68	0.68	0.73	0.57	0.68	0.68	0.72
N_{obs}	117	117	117	117	117	117	117	117

Results of regressions showing the effects of network-directed organizing metrics 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 3: Instrumental Variable Results

	1	2	3	4	5	6
NDO	**1.97**	**2.90**	**2.94**	**1.75**	**2.44**	**2.45**
	(0.75)	**(1.26)**	**(1.31)**	**(0.66)**	**(0.98)**	**(1.01)**
	[1.17]	**[1.03]**	**[1.11]**	**[1.05]**	**[0.77]**	**[0.81]**
Log(No. organizer notes)	**0.90**	0.35	0.34	**0.90**	0.42	0.41
	(0.07)	(0.32)	(0.32)	**(0.07)**	(0.28)	(0.27)
	[0.06]	[0.32]	[0.34]	**[0.06]**	[0.30]	[0.30]
Log(Mean degree)	**0.30**	-2.87	11.57	**0.30**	15.28	23.03
	(0.06)	(106.21)	(87.05)	**(0.06)**	(95.39)	(77.69)
	[0.06]	[108.69]	[89.31]	**[0.06]**	[96.17]	[75.59]
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 Controls	N	Y	Y	N	Y	Y
First-stage F-Stat (Robust)	24.70	7.70	7.12	31.19	11.43	10.74
First-stage F-Stat (Clustered)	45.89	10.01	8.60	69.59	16.48	14.18
Adjusted R_{sq}	0.56	0.57	0.58	0.57	0.62	0.63
N_{obs}	115	115	115	115	115	115

Results of regressions using the leave-on-out team score as an instrument, showing the effects of network-directed organizing metrics on the log of cards signed, with various controls. Leave-one-out team score is calculated as the mean of a store's associated regional "team"'s scores at all other stores associated with that team. 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.

10 Appendix

Table A1: First Stage Results (Spearman Correlation)

	1	2	3	4	5	6
Team score	**0.72**	**0.50**	**0.47**	**0.76**	**0.55**	**0.53**
	(0.15)	**(0.18)**	**(0.18)**	**(0.14)**	**(0.16)**	**(0.16)**
	[0.11]	**[0.18]**	**[0.18]**	**[0.10]**	**[0.16]**	**[0.16]**
Log(No. notes)	-0.00	**0.18**	**0.18**	-0.00	**0.18**	**0.19**
	(0.02)	**(0.04)**	**(0.05)**	(0.02)	**(0.04)**	**(0.05)**
	[0.02]	**[0.04]**	**[0.05]**	[0.02]	**[0.04]**	**[0.05]**
Log(Mean degree)	-0.02	41.94	22.40	-0.02	42.22	21.96
	(0.01)	(26.01)	(23.94)	(0.01)	(25.59)	(23.46)
	[0.02]	[49.78]	[43.71]	[0.02]	[48.67]	[42.21]
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 Controls	N	Y	Y	N	Y	Y
Adjusted R_{sq}	0.19	0.36	0.33	0.23	0.40	0.37
N_{obs}	115	115	115	115	115	115

Results of first-stage regressions, showing the effects of leave-one-out team scores on network-directed organizing metrics, with various controls. Leave-on-out team score is calculated as the mean of a store's associated regional "team"'s scores at all other stores associated with that team. 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 A2: Reduced Form Results (Spearman Correlation)

	1	2	3	4	5	6
Team score	**1.43**	**1.44**	**1.39**	**1.33**	**1.35**	**1.31**
	(0.49)	**(0.54)**	**(0.54)**	**(0.46)**	**(0.51)**	**(0.51)**
	[0.88]	**[0.54]**	**[0.54]**	**[0.82]**	**[0.51]**	**[0.51]**
Log(No. organizer notes)	**0.90**	**0.86**	**0.87**	**0.90**	**0.86**	**0.87**
	(0.07)	**(0.18)**	**(0.17)**	**(0.07)**	**(0.18)**	**(0.17)**
	[0.08]	**[0.23]**	**[0.20]**	**[0.08]**	**[0.23]**	**[0.20]**
Log(Mean degree)	**0.26**	118.78	77.32	**0.26**	118.17	76.90
	(0.05)	(78.32)	(68.50)	**(0.05)**	(78.32)	(68.44)
	[0.07]	[110.37]	[88.17]	**[0.07]**	[109.61]	[87.35]
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 Controls	N	Y	Y	N	Y	Y
Adjusted R_{sq}	0.57	0.68	0.68	0.57	0.68	0.68
N_{obs}	115	115	115	115	115	115

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 A3: OLS Results (Pearson correlation)

	1	2	3	4	5	6	7	8
NDO	**0.59**	**0.40**	**0.41**	0.24	**0.52**	**0.41**	**0.45**	0.25
	(0.21)	**(0.18)**	**(0.18)**	(0.20)	**(0.19)**	**(0.17)**	**(0.17)**	(0.20)
Log(No. organizer notes)	**0.87**	**0.70**	**0.69**	**0.73**	**0.86**	**0.67**	**0.66**	**0.70**
	(0.07)	**(0.17)**	**(0.17)**	**(0.19)**	**(0.07)**	**(0.17)**	**(0.17)**	**(0.20)**
Log(Mean degree)	**0.31**	**116.60**	**120.05**	**103.50**	**0.31**	**126.86**	**88.37**	**114.18**
	(0.05)	(74.40)	(73.57)	(83.82)	**(0.05)**	(72.97)	(60.73)	(83.39)
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 Controls	N	Y	Y	Y	N	Y	Y	Y
Team Fixed Effects	N	N	N	Y	N	N	N	Y
Adjusted R_{sq}	0.57	0.67	0.67	0.72	0.57	0.68	0.68	0.72
N_{obs}	117	117	117	117	117	117	117	117

Results of regressions showing the effects of network-directed organizing metrics 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 A4: First Stage Results (Pearson correlation)

	1	2	3	4	5	6
Team score	**0.78**	**0.63**	**0.62**	**0.85**	**0.77**	**0.75**
	(0.14)	**(0.17)**	**(0.17)**	**(0.11)**	**(0.15)**	**(0.15)**
	[0.08]	**[0.16]**	**[0.17]**	**[0.05]**	**[0.16]**	**[0.17]**
Log(No. organizer notes)	0.00	**0.22**	**0.20**	0.03	**0.31**	**0.31**
	(0.04)	**(0.08)**	**(0.08)**	(0.04)	**(0.09)**	**(0.09)**
	[0.04]	**[0.09]**	**[0.09]**	[0.05]	**[0.10]**	**[0.12]**
Log(Mean degree)	**-0.06**	45.83	54.89	**-0.07**	-2.27	-37.92
	(0.03)	(44.70)	(44.31)	**(0.03)**	(45.98)	(44.15)
	[0.02]	[77.28]	[81.83]	**[0.03]**	[90.28]	[85.08]
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 Controls	N	Y	Y	N	Y	Y
Adjusted R_{sq}	0.25	0.42	0.41	0.35	0.48	0.45
N_{obs}	115	115	115	115	115	115

Results of first-stage regressions, showing the effects of leave-one-out team scores on network-directed organizing metrics, with various controls. Leave-on-out team score is calculated as the mean of a store's associated regional "team"'s scores at all other stores associated with that team. 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 A5: Reduced Form Results (Pearson correlation)

	1	2	3	4	5	6
Team score	**0.94**	**0.91**	**0.90**	**0.78**	**0.74**	**0.72**
	(0.27)	**(0.28)**	**(0.28)**	**(0.21)**	**(0.22)**	**(0.22)**
	[0.39]	**[0.28]**	**[0.25]**	**[0.29]**	**[0.22]**	**[0.22]**
Log(No. notes)	**0.89**	**0.89**	**0.88**	**0.89**	**0.91**	**0.91**
	(0.07)	**(0.18)**	**(0.18)**	**(0.07)**	**(0.18)**	**(0.17)**
	[0.07]	**[0.23]**	**[0.23]**	**[0.07]**	**[0.24]**	**[0.21]**
Log(Mean degree)	**0.26**	89.42	100.96	**0.27**	84.49	49.21
	(0.05)	(79.72)	(78.28)	**(0.05)**	(79.47)	(68.44)
	[0.07]	[103.12]	[101.47]	**[0.07]**	[99.37]	[75.76]
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 Controls	N	Y	Y	N	Y	Y
Adjusted R_{sq}	0.58	0.69	0.69	0.58	0.69	0.69
N_{obs}	115	115	115	115	115	115

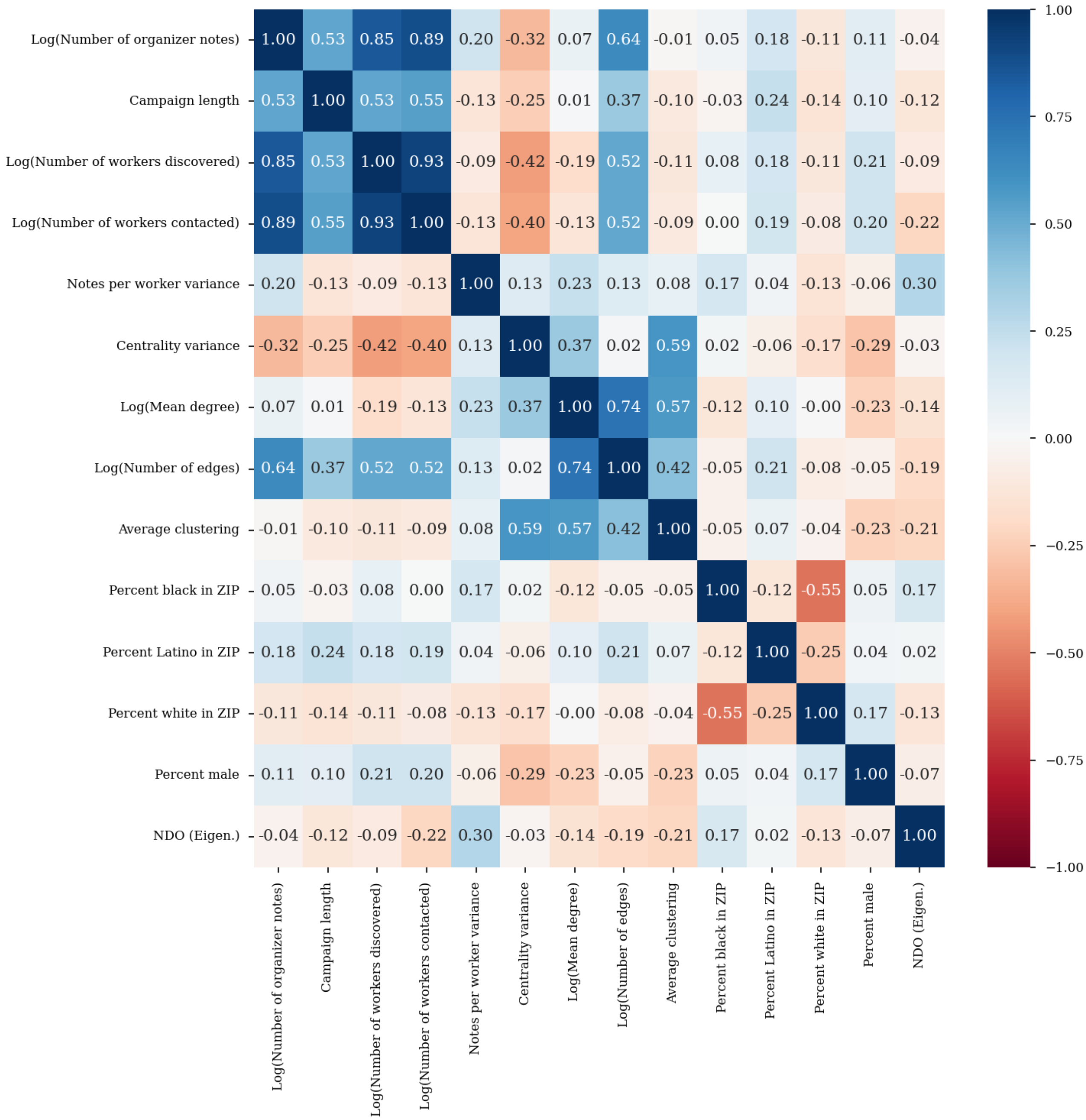
Results of reduced-form regressions showing the effects of leave-one-out team scores on card-signing, with various controls. No. workers includes log number of workers ever contacted and log number of workers mentioned in field notes. Campaign length controls are indicators for quintiles of campaign 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 A6: Instrumental Variable Results (Pearson correlation)

	1	2	3	4	5	6
NDO	**1.21**	**1.44**	**1.44**	**0.92**	**0.97**	**0.96**
	(0.38)	**(0.56)**	**(0.57)**	**(0.27)**	**(0.33)**	**(0.34)**
	[0.47]	**[0.51]**	**[0.54]**	**[0.32]**	**[0.32]**	**[0.33]**
Log(No. organizer notes)	**0.89**	**0.58**	**0.59**	**0.86**	**0.61**	**0.61**
	(0.07)	**(0.24)**	**(0.24)**	**(0.07)**	**(0.21)**	**(0.20)**
	[0.07]	**[0.35]**	**[0.34]**	**[0.07]**	**[0.32]**	**[0.32]**
Log(Mean degree)	**0.34**	23.61	21.98	**0.33**	86.68	85.67
	(0.06)	(97.64)	(99.28)	**(0.06)**	(78.76)	(64.62)
	[0.06]	[98.27]	[106.06]	**[0.07]**	[78.25]	[66.04]
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 Controls	N	Y	Y	N	Y	Y
First-stage F-Stat (Robust)	31.53	13.82	13.34	57.34	27.27	24.66
First-stage F-Stat (Clustered)	101.64	19.58	17.14	342.44	29.88	25.51
Adjusted R_{sq}	0.55	0.62	0.62	0.57	0.66	0.67
N_{obs}	115	115	115	115	115	115

Results of regressions using the leave-on-out team score as an instrument, showing the effects of network-directed organizing metrics on card-signing, with various controls. Leave-on-out team score is calculated as the mean of a store's associated regional "team"'s scores at all other stores associated with that team. 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.

Figure A1: Variable Correlation Heatmap



Pearson correlation between pairs of indicator variables used in regression analysis.

A A DeGroot Model of Labor Organizing

We motivate our measure with a canonical framework for network social learning, the DeGroot model (DeGroot 1974). This is by no means the only motivation, as foundations based on strategic interactions (e.g. Ballester et al. 2008) could also be given. But based on the qualitative evidence, it seems like social learning is an important mechanism, and we formalize the intuition here.

We suppose the subjective value to worker i of signing a card at time t (also worker $i's$ probability of signing a card) 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.

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 workers i update 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 of the matrix G corresponding to the eigenvalue of 1 (which is the highest eigenvalue as G is a stochastic matrix), denoted C. In steady-state $v(\infty)_i = C \cdot v(0)$ for all i, i.e. there will be consensus.

It is easy to see that $C_i = \frac{d_i}{\sum_j d_j}$ satisfies $\mathbf{C}'\mathbf{G} = \mathbf{C}$, so that the normalized degree of a worker measures the influence of a worker on co-workers. However also note that any rescaling of C by scalar will also satisfy the equation.

We modify the assumptions of a DeGroot model by including the existence of an organizer. Assume an organizer can choose an allocation of effort e_i, with $\sum_i e_i = 1$, corresponding to how much effort to spend influencing worker i. Effort can increase a worker's value of signing permanently, but at some time $t(i)$ by e_i, at a cost of $\frac{C}{2}e_i^2$, reflecting that there are increasing marginal costs to influencing any individual worker. While somewhat arbitrary, 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.

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

We further assume the organizer aims to maximize the steady-state sum of values, i.e. $\lim_{t\to\infty}\sum_i v_i(t) = \sum_i C_i(v_i + e_i)$. It is then easy to see that steady-state card signing will be maximized where $C \cdot e$ is maximized, and the organizer invests $e_i = \frac{C_i}{C}$ 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.

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.

Note that the organizer does not necessarily know the true influence vector C_i, 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.

$C \cdot e$ is proportional to the correlation of e_i and C_i within a store. However, since e and C are on arbitrary scales, and the quantitative prediction should hold with any positive rescaling of the vectors, in the empirical work we will use Spearman (rank) correlations, to ensure that our measures are not being driven by any particular normalization. We examine robustness to standard correlations (Pearson) in the Appendix.

B Examples of Notes

- ”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”
- ”Friend of Isabel Garcia - Schneider, D”
- ”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”
- ”CTW: 3/24/09 - Signed union card. Referred by Charlotte Jansen. Arthur Jones collected card. Loves Obama. Makes $12.39/hr. - Hoang, T”

C References (Unfinished)