Title: *Decision-making in a governance network: an empirical and counterfactual analysis*

Author: Carey Doberstein, PhD candidate, University of Toronto

Paper presented at the 2013 Canadian Political Science Association Meetings
June 4-6, 2013. Victoria, BC.

Introduction

Governments have acknowledged that with 'wicked' public problems, governance arrangements that bring diverse state and societal actors around the table in institutionalized governance networks can help address the complexity that one level of government or organization cannot do alone. Homelessness is such an example, with complex roots and intersecting personal and societal pathologies—as such, all levels of government have a stake in policy since the issues touch many portfolios, including health, housing, corrections, skills development, and employment. The challenge of the policy issue and the political, institutional, and financial dependence of local government on other public and private actors and organizations, makes it such that civil society actors—whether they are shelter service providers, mental health professionals, affordable housing providers—have always played an important role in policy development and implementation with regards to homelessness (O-Reilly-Fleming 1993; Pierre 1998). Yet there are those who argue that the lessons from governance networks that involve civil society actors is not so positive: there are threats of agency (and interest group) capture, fractured and unproductive debates, reinforced power differentials, and resource hoarding (Agranoff, 2006). These are empirical questions however, related to how a governance network is designed and managed, rather than representative of a broader theoretical debate (Head, 2008; Edelenbos and Klijn, 2006).

In response to early critiques by Dowding (1995), Thatcher (1998) and Borzel (1998) that scholars have not demonstrated that networks 'matter' to policy, the connections between the activity of networks and concrete policy dimensions like policy change and instrument type have since been made by scholars.[1] Less is known, however, about the mechanisms through which networks 'matter'. That is, *why* do networks matter? This paper seeks to contribute to this question by presenting the results and analysis of a natural experiment in network governance decision-making in Vancouver, interweaving a unique collection of quantitative and participant observation data.[2] The quantitative data gathered and analyzed in this study is unique because it answers an important counterfactual question in the scholarly debate over the extent to which (and why) networks 'matter': would decisions be made differently if bureaucrats (traditional decision-makers), rather than the network actors, made the policy and resource allocation decisions? This data was collected during a two-month period in which the researcher was embedded within the network and the network was tasked with setting policy priorities and resource allocations for the next two years under the Government of Canada's Homelessness Partnering Strategy (HPS) in Metro Vancouver.

This paper begins by briefly presenting the network literature to demonstrate the theoretical justification for the creation of governance networks for issues like homelessness, showing that we indeed may expect public servants to differ from civil

---

[1] See for example: Head, 2008; Adam and Kriesi, 2007; Klijn and Skelcher 2007; Sorenson and Torfing, 2007; Benz and Papadopoulos 2006; Leach 2006; Montpetit, 2005; Howlett, 2002; Bressers and O'Toole, 1998; Daugbjerg, 1998; Coleman, Skogstad and Atkinson, 1997.

[2] Kubik (2007) suggests that ethnographic or participant observation research is especially valuable in helping to identify and understand the mechanisms at work behind the various statistical relationships discovered through quantitative research (see also Allina-Pisano, 2007).

society actors in their interpretations of the policy problem and the potential solutions. The experimental data for the case is then presented, weaving together a quantitative and qualitative story to concretely answer the counterfactual question of whether networks 'matter', as well as *why* network activity produces distinct policy decisions and resource allocations, drawing on contemporary deliberative democracy theory. The final section offers tentative conclusions.

Creation of governance networks

Governments typically create or harness networks as vehicles to pool resources, resolve knowledge deficiencies, and improve the implementation of programs and services (Agranoff, 2006). Governance networks are seen as useful in some policy domains because they bring together a wide range of expertise, knowledge and resources that not only enables new thinking about complex issues, but also lends itself to more successful implementation (Head, 2008). The inclusion of civil society actors, who are generally more connected to the issues on the ground than public servants, offers a diversity of lived experience and therefore information, interpretations, priorities and perspectives about what works and is worthwhile in terms of policy (Head, 2008; Edelenbos and Klijn, 2006). Such actors thus not only possess critical resources to realize policy goals and outcomes, but may also conceive of the problem differently and have unique information and ideas on the most effective solutions (Edelenbos and Klijn, 2006).[3]

Given that it is precisely the rationale behind the creation of governance networks in complex files, we indeed *expect* public servants and civil society actors to view policy problems and solutions differently. Such differences are theorized to exist between public servants and civil society actors because of varying organizational incentives and experiences of the respective groups. Civil society actors tend to provide unique perspectives vis-à-vis government as a result of the tacit knowledge they possess from their closeness to the target population and experience on the ground (Agranoff, 2006). Archon Fung (2008) provides a rich empirical example of community participatory policing boards in Chicago, demonstrating that citizens often develop different priorities and approaches than professional police officers would have developed on their own, given their unique placement and experience, as well as their freedom from the sometimes stale arguments in the professional ranks. Government actors thus face rather different incentive structures than civil society actors, and also possess knowledge and skills generally not brought by civil society actors. With respect to incentive structures, it is well established that public servants, who are subject to greater public scrutiny, tend to operate in an institutional environment that breeds risk-aversion and rewards conventional interpretations of policy problems and solutions (Bozeman and Kingsley, 1998; Wilson, 1989).[4]

---

[3] Some scholars go further, arguing that go-alone strategies and hierarchical policy processes often lead to poor or narrow solutions because one rationality or perception dominates the formulation of the policy and others are excluded (Klijn and Koppenjan, 2000).

[4] Indeed, staying under the radar of elected officials (who may want to intervene) is a powerful incentive for public managers to pursue cautious and conformist behavior (Tirole, 1994). This is should not be interpreted as claiming that bureaucrats are cautious and conformist individuals, but "that government bureaucracies are caught up in a web of constraints so complex that any big changes are likely to rouse the ire of some important constituency", and thus tend to favor the most defensible and objective criteria when making decisions (Wilson, 1989: 69).

A natural experiment in network governance decision-making

The Metro Vancouver Regional Steering Committee on Homelessness (RSCH) was created in 2000, under the auspices of the Government of Canada's National Homelessness Initiative (NHI)—now called the Homelessness Partnering Strategy (HPS)—which provides financial incentives for cities to form local networks of diverse government and civil society actors to jointly craft 'Community Plans' (CP) to address homelessness. Civil society representation is diverse on the RSCH, with representatives from homeless shelters, transitional housing providers, mental health professionals, street nurses, private foundations like United Way, as well as those with lived experience of homelessness, including vulnerable populations like youth, Aboriginal and women. The purpose of the production of the CP is to develop a broad local consensus around the unique pressures and solutions with regard to homelessness in each city, and is to be used to inform the allocation of federal dollars towards service and housing programs and programs by the local governance network.

From December 2011 to February 2012, the researcher was embedded as a participant observer in the RSCH as they prepared for, deliberated, and allocated $11 million HPS dollars for the Metro Vancouver area for 2012-2014. The Manager of the Homelessness Secretariat at the Metro Vancouver regional government, under the direction of the RSCH network, selected and placed individuals from the policy community into four teams (each consisting of 5-6 members)—one of which was an exclusively Aboriginal network team—to divide up the 87 proposed homelessness programs for review and decision.[5] Parallel to all of this activity were a group of bureaucratic staff from the Homelessness Secretariat at Metro Vancouver and the Vancity Community Foundation who individually reviewed, scored, and deliberated each team's programs, in order to support the decision-making at the team level.[6] The researcher was present for all staff-level meetings and deliberations on the program proposals, as well as all team meetings and the final deliberations—where all teams came together—at which the final decisions were made. Extensive scoring data from network members and staff, before and after deliberations, was also made available to the researcher.

This decision-making context functions as a natural experiment in decision-making, offering a controlled environment in which to compare how bureaucrats (who represent 'normal/non-network decision-making') and network members conceptualize and evaluate proposed homelessness programs. In the language of the political experiments literature—the treatment group (network) and the control group (bureaucrats) are evaluating and scoring exactly the same homelessness program proposals (Gerber and Green, 2012). This is rarely achieved even in the political experiment world, as it is very

[5] Three teams were established to evaluate and recommend non-Aboriginal-specific programs and a fourth team was created to evaluate and recommend Aboriginal-specific program proposals. This is in recognition of the strikingly disproportionate share of Aboriginals among the homeless population and the demonstrated need for culturally sensitive policy and programs (Ward, 2008). Network members on the review teams included representatives from homelessness service providers, community-based philanthropic organizations, local government, provincial government, and prominent members of the community with a demonstrated interest in homelessness.

[6] The role of staff is to assist network members in their decision-making by understanding each program proposal, to analyze the proposed budget in detail, and to make sure the deliberations are structured to be effective, efficient, and under conditions of maximum information.

uncommon to get experimental and control groups to complete *precisely* the same tasks (Gerber and Green, 2012). As such, we can directly compare their scoring and decision-making without engaging in typical (speculative) counterfactual analysis—in this case, we have data for the counterfactual. This is a very unusual opportunity and offers the clearest way to make inroads to understanding to what extent, and why, governance networks 'matter' to policy and program choices.

There are 87 homelessness program proposals for which a group of network actors and bureaucrats separately evaluated and scored. The proposed programs included providing outreach services, shelter and supports, transitional housing, mental health and addiction treatment, life skills programs, and youth safe houses, among others, across the Metro Vancouver area. These program proposals are the 'subjects' in the experiment, in that they are subject to the treatment (network decision-making as opposed to bureaucratic decision-making). Thus the critical measure is the average treatment effect (ATE), which means for each program proposal scored, the difference between the average network actor score and average bureaucratic actor score. This provides a quantifiable estimate for how much difference network governance makes to homelessness policy and program choices.

Average treatment effect (ATE) on each subject = Treatment score (network) – Control score (bureaucrat)

Thus the central question to answer: are their systematic differences in the scoring and decisions between network actors and bureaucrats? If so, this lends credibility to theory that suggests that networks 'matter' to policy and program choices.

Hypotheses

To answer the broader question of whether governance networks 'matter' to policy development and implementation, three hypotheses are tested, informed by public administration and governance network theory, and tested by weaving together quantitative data analysis and qualitative narrative from participant observation.

Null hypothesis 1: The average treatment effect on the subjects is zero
Alternative Hypothesis 1: Bureaucrats and network members will score and deliberate programs differently, such that different policy and program choices would be made.

Null hypothesis 2: The average treatment effect on the Aboriginal subjects is zero
Hypothesis 2: The Aboriginal network team will score and deliberate differently from non-Aboriginal teams and bureaucrats, such that different policy and program choices would be made.

Null hypothesis 3: Deliberations will not systematically alter the preferences network actors
Hypothesis 3: Deliberations will result in different program selections than aggregated individual scores of network members would suggest.

Hypothesis 1 captures the counterfactual at the heart of the research question: would decisions be made differently if bureaucrats, rather than the network actors, made the policy and resource allocation decisions? One may expect staff and network members to score and deliberate policy and programs differently because of the experience they bring to the table and particular incentive structures that constrain them (Agranoff, 2006; Fung, 2008; Bozeman and Kingsley, 1998; Wilson, 1989). Hypothesis 2 is likewise a key claim

to test, given the increasingly accepted arguments by scholars and Aboriginal leaders that their unique ontological and epistemological systems inform how they conceive of public problems and their solutions. This is particularly appropriate to test in the context of homelessness since Aboriginals often constitute a dramatically disproportionate share of the homeless population and many argue that culturally sensitive policy and programs are key to their uptake of services (Ward, 2008).[7] The third hypothesis aims to test one of the key arguments in the governance network literature: that bringing diverse actors around a table results in decisions that are different than if government actors performed such tasks internally and in a more hierarchical context. That is, deliberations among government and civil society actors in a structured and authoritative setting has policy consequences, as alternative perspectives, arguments, and facts on the ground are presented to the collective decision-makers, which may not be initially valued or understood by each individual network member (Agranoff, 2006; Fung, 2008).

The analysis of the natural experiment proceeds in two steps: (i) the analysis of the nature and extent of differences between staff and network members scores of each program on the fifteen criteria established by the RSCH in advance of the review process; and (ii) the analysis of the differences between aggregated individual scores of each team and the programs recommended to fund after team deliberations, assisted by OLS regression. All presentations of quantitative and qualitative data are accompanied by further contextualization from participant observation.

Measuring the average treatment effect (ATE): network member and bureaucratic staff score differences for each proposed homelessness program

The first step to analyzing governance networks in this context is to establish that network members do indeed make different decisions than bureaucratic actors.[8] Recall that network actors formally make the decisions at the RSCH, but parallel to all of this activity is a group of bureaucratic staff who individually reviewed, scored, and deliberated proposed programs considered by each network team as a system of support for the decision-making.

Figure 1 below presents the difference in scores for each of the 87 programs evaluated by the network teams and bureaucratic teams, calculated by subtracting the average bureaucratic staff score from the average network member score, in percentage terms. These values are the average treatment effect (ATE) and are measured in absolute value terms because at this point we are most interested in identifying difference, not a direction of difference[9]. The average ATE across all proposed homelessness programs is

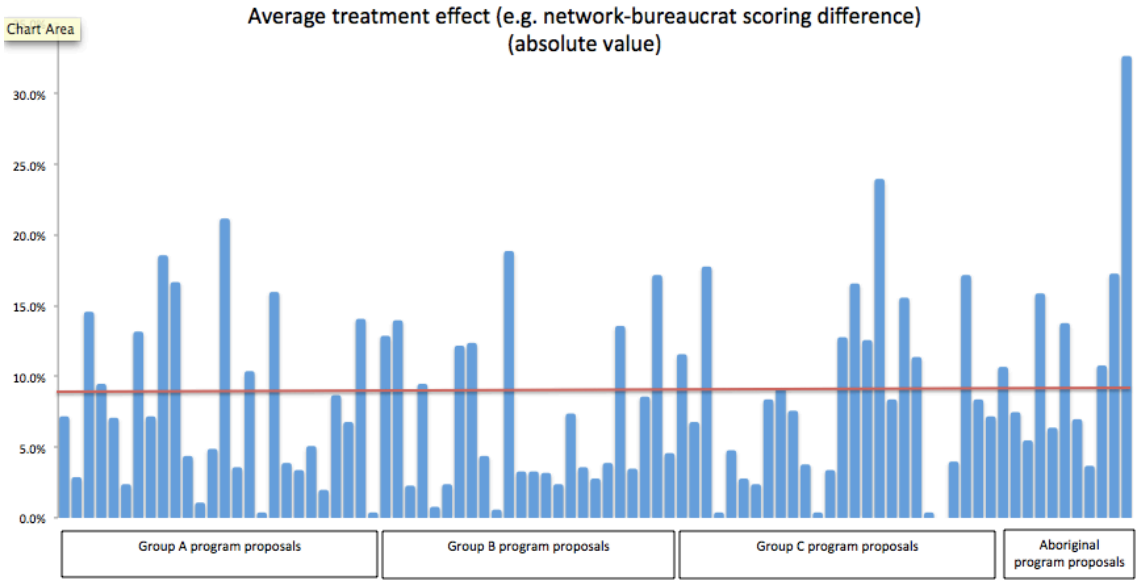
[7] The Government of Canada recognizes this argument, and has helped to create exclusively Aboriginal homelessness governance networks in many Canadian cities, but there has been no opportunity to systematically test such claims until now.

[8] Many make this claim in the literature (Bozeman and Kingsley, 1998; Wilson, 1989), though we rarely, if ever, have the opportunity to compare the actions and decisions by both groups in the same decision arena. However, the RSCH HPS 2012 decision-making context provides such an opportunity.

[9] Theory does not predict whether bureaucrats or network members would be more 'generous' in their scoring, but rather that they may conceive of problems and solutions differently. Also, averaging the score differences by using non-absolute values would dramatically underestimate the difference (e.g. average of -10% and +10% difference = 0; yet average of their absolute values is 10% difference).

8.1%. That is, across all 87 programs, network actors and bureaucrats differ in their total scores by 8.1%, with minimums of 0 and a maximum difference of 33%.

Figure 1: Average treatment effect (e.g. average score difference between network member and bureaucrats for each proposed homelessness program)



Whether an 8% scoring variation among network actors and bureaucrats on the precisely the same proposed homelessness programs constitutes systematic variation may be subject to debate, so there is value in differentiating the scoring differences according to the 15 criteria the evaluators used. Table 1 below summarizes the count and distribution of programs in which the staff and network member scores differed by greater than 25% on each criterion.[10] The first column shows the number of instances for each criterion in which the average network member score was at least 25% higher or lower than the average staff score. Differentiating by each criterion helps identify which criteria the staff and network members exhibit more common and substantial variation than others. Based on the data in Table 1, the criterion of 'community support'—that is, does this homelessness program leverage community support (financial or in-kind) to enhance the proposed work?—shows 23 (out of 87) instances of differences between staff and network members exceeding 25% (most of the time with the staff giving higher scores).

The next most common differences are whether the proposed program will deliver positive outcomes for the population, if it meets defined policy objectives, and whether

[10] The first question that arises is what constitutes a meaningfully 'different' score between the average bureaucratic staff and average team member score for each criterion? That is, in percentage terms, how much variation (in either direction) signals a real source of interpretation on the part of the reviewer, rather than noise? I contend that a difference of 25% or greater is worth investigating further, since the total value for each criteria ranges from 5 points to 15. This is selected as such to be conservative in making claims about the existence and nature of difference. Thus for a criterion with a maximum value of five, a meaningful difference would imply, for example, the average staff score of 2.7 and network score of 4. Conversely, counting as a 'difference' a score of 4.2 and 4.5 is very difficult to justify, so when evaluating categories with small total values, a higher percentage difference standard is needed than the difference in total scores (out of 125 points).

the program demonstrates a clear need and long-term impact. So not all criteria used in the evaluation resulted in systematic differences between network members and bureaucrats. But interestingly, the criteria that display the most variation between the treatment and control groups are fundamental features of the policy problem and their proposed solutions. Network actors and bureaucrats viewed these features very differently from one another.

Table 1: Number of programs which network and staff scores differed by > 25% and mean differences.

Evaluation criteria	count total >/< 25%	mean difference in scores ALL (non Aboriginal)	mean difference in scores Aboriginal only
1. Impact	14	4.2%	3.5%
2. Objectives	9	2.0%	-0.2%
3. Activities	11	1.7%	-10.4%
4. HPS objectives	15	-6.0%	-11.6%
5. HPS outcomes	19	-2.1%	-11.6%
6. Comm priorities	13	-1.9%	-5.3%
7. Outcomes	7	-1.7%	-14.1%
8. Evaluation	7	5.1%	-9.5%
9. Funding	10	4.8%	-5.6%
10. Sustainability	15	2.3%	-3.4%
11. Comm support	23	5.1%	0.0%
12. Timeframe	8	7.2%	-0.5%
13. Location	11	12.8%	10.2%
14. Sponsor	14	11.8%	-3.2%
15. Proposal	9	-1.9%	-14.7%
TOTAL	**1**	**+3.0%**	**-8.0%**

Even though I contend that the absolute value of the average treatment effect (ATE)—the difference between average network member and bureaucrat member scores for each program—is the most appropriate way to capture the variation between the two groups, there is value in comparing the real value (with directionality) to get a sense of whether patterns exist across the experimental and control groups. In particular, comparing the non-Aboriginal network Groups (A, B, C) and the Aboriginal Group against bureaucratic scoring. Ostensibly, carving out a network group of exclusively Aboriginal decision-makers implicitly suggests that they bring unique perspectives to the table.

Thus one particularly noteworthy finding when comparing the average staff and network member scores for each program proposal is revealed when one examines only Aboriginal-specific program proposals. The trend among all non-Aboriginal specific program proposals is that staff gave on average 3.0% *higher* total scores than network members, yet among the Aboriginal-specific programs, the trend is precisely the opposite, and captured in the last row of the last column of Table 1: staff scores are 8.0% *lower* than network members. That is, comparatively, staff found the Aboriginal programs not only to be weaker than the non-Aboriginal programs, but also weaker relative to the Aboriginal network members who also evaluated them. What explains this systematic difference?

The empirical observation that staff systematically scored Aboriginal programs lower than network members (who were Aboriginal)—which is precisely the opposite case for non-Aboriginal programs—is a surprising finding. Theory and previous empirical work suggests that evaluations by staff and Aboriginal network members may be different

because Aboriginals and non-Aboriginals may have unique interpretations of the policy problem and appropriate solutions (Ward, 2008), but one would not predict the systematic undervaluation as evidenced in Table 1. Without the benefit of participant observation, one might first hypothesize that this is because the Aboriginal network members are simply more generous in their scoring than non-Aboriginal members. Based on observations in the separate deliberations among both staff and Aboriginal network members, this is not the case. What is clear when witnessing Aboriginal network members articulate their reasons for scoring programs as they did, is that they interpreted the criteria and program proponent argument for funds much differently than staff did.

An example will help illustrate this. The framework in which program proponents were asked to articulate the purpose of their program, budget ask, and anticipated impact of activity was through a 'logic model'—essentially a linear flow chart that links activity inputs to program outputs, and then outcomes.[11] From observing Aboriginal network member deliberations, and confirmed in subsequent interviews, the logic model was considered totally inappropriate for the assessment of traditional Aboriginal policy development and program funding. In fact, in most cases, the Aboriginal members would simply ignore the logic model and justify their scoring and decisions based on knowledge of the organizations historical service record, the emerging issues on the ground, and the specific culturally sensitive services they proposed. Staff, by contrast, assessed the Aboriginal programs as they would for the non-Aboriginal programs, placing great emphasis on the logic model and other technical elements like the budget ask or eligibility for funding. Also, the staff review rarely spoke about the need for the service or what was situation was like on the ground, readily admitting that they lack the knowledge and expertise. Instead they focused on the budget and logic model, which no one explicitly acknowledged might be an inappropriate lens through which to assess Aboriginal programs.[12] Staff noted that there was an observable lack of capacity among all program proponents to produce coherent program proposals with respect to the logic model, and it was more pronounced among Aboriginal applicants.[13]

Counter-factual: would different programs be funded?

The average treatment effect (ATE) for the 87 propose homelessness programs that there are several important dimensions of difference, as demonstrated in Table 1. Figure 1 below shows the often-substantial difference in rankings of programs between network and staff.[14] For example, Program X (identifying information removed), highlighted in the figure, staff ranked this program 23rd out of the 87, while the network members

[11] It is an approach favored by the Government of Canada and accepted by Metro Vancouver staff, as it promotes a professional understanding of accountability and performance management.

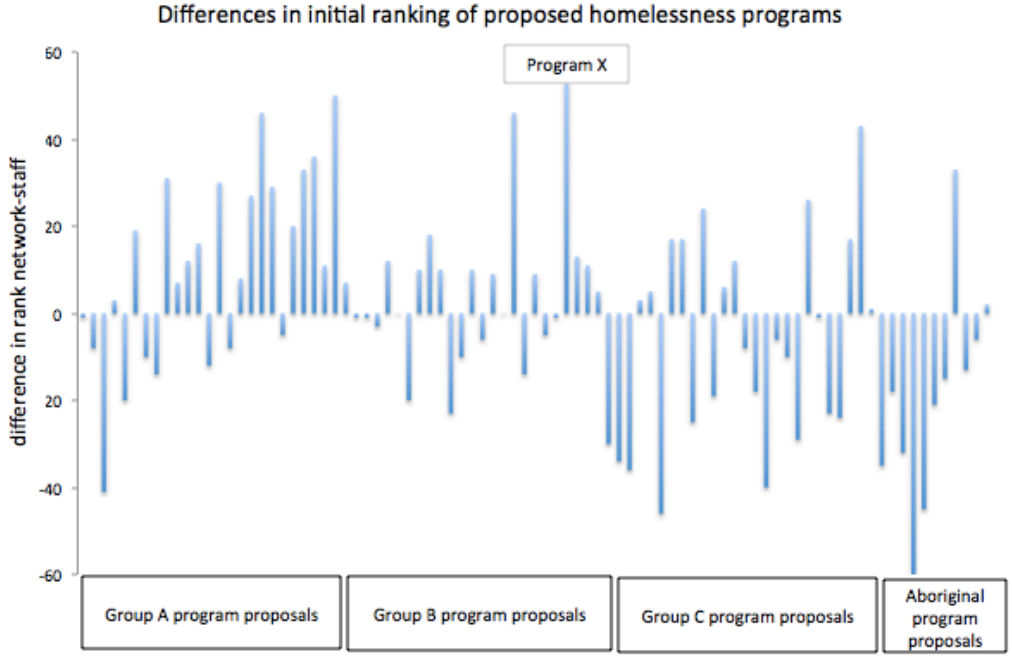
[12] Staff are not insensitive to Aboriginal issues, it is that they are in somewhat of a difficult position. They have certain expectations from the Government of Canada with regards to the use of the logic model, and are required to administer and monitor the contracts with organizations after the network makes their funding allocations, so it is not surprising that they focus on the tangible elements of proposals, like budgets, eligibility and logic model.

[13] In fact, one staff member claimed that several of the Aboriginal programs would not make it through initial evaluation in other government settings because they were so unclear and incomplete.

[14] Since staff scoring exhibits a bias such that on average they score programs 3.0% higher than network members, in order to confirm that such a difference is consequential, it is more appropriate to compare them on their relative rankings of the 87 programs.

ranked it 76th, resulting in a difference in ranking of 53. The negative values in the figure represent the cases in which the network rank is better than the staff.

Figure 2: Differences in initial ranking of 87 programs between network and staff.



Yet are these differences in ranking consequential from a policy perspective? Would different programs end up being funded if the bureaucrats, rather than the network, had full decision-making authority? We can estimate this by comparing the rankings of the programs after initial individual rankings and aligning them against each other, and then adding them up until we reach the total funding envelope of $11M.[15] This is prior to deliberations among network members, which included technical assistance and advice from staff—thus it captures an unadulterated average network member choice, free from staff influence.

Table 2 below captures the highest ranked programs from network members and staff after the initial evaluation, along with their budget ask, which are then totaled until the funding envelope has been allocated ($11,094,666)[16]. Table 2 demonstrates that there would be significant differences in the programs funded based on the initial evaluations. For example, of the 31 programs that the network members would have funded, 13 (or 42%) of those the staff would not. *Nearly half the programs selected by the respective groups would not have been selected by the other.*

[15] This also helps to address the fact that staff overall scored programs 8.6% higher than network members. Comparing the programs by their relative ranking captures which set of programs network and staff members felt were strongest.

[16] This, of course, does not imply that these are the programs that either group would ultimately fund, since it is prior to deliberations, but it does provide an opportunity to directly compare whether the network and staff differences in scoring would have real effects on allocations in a hypothetical sense.

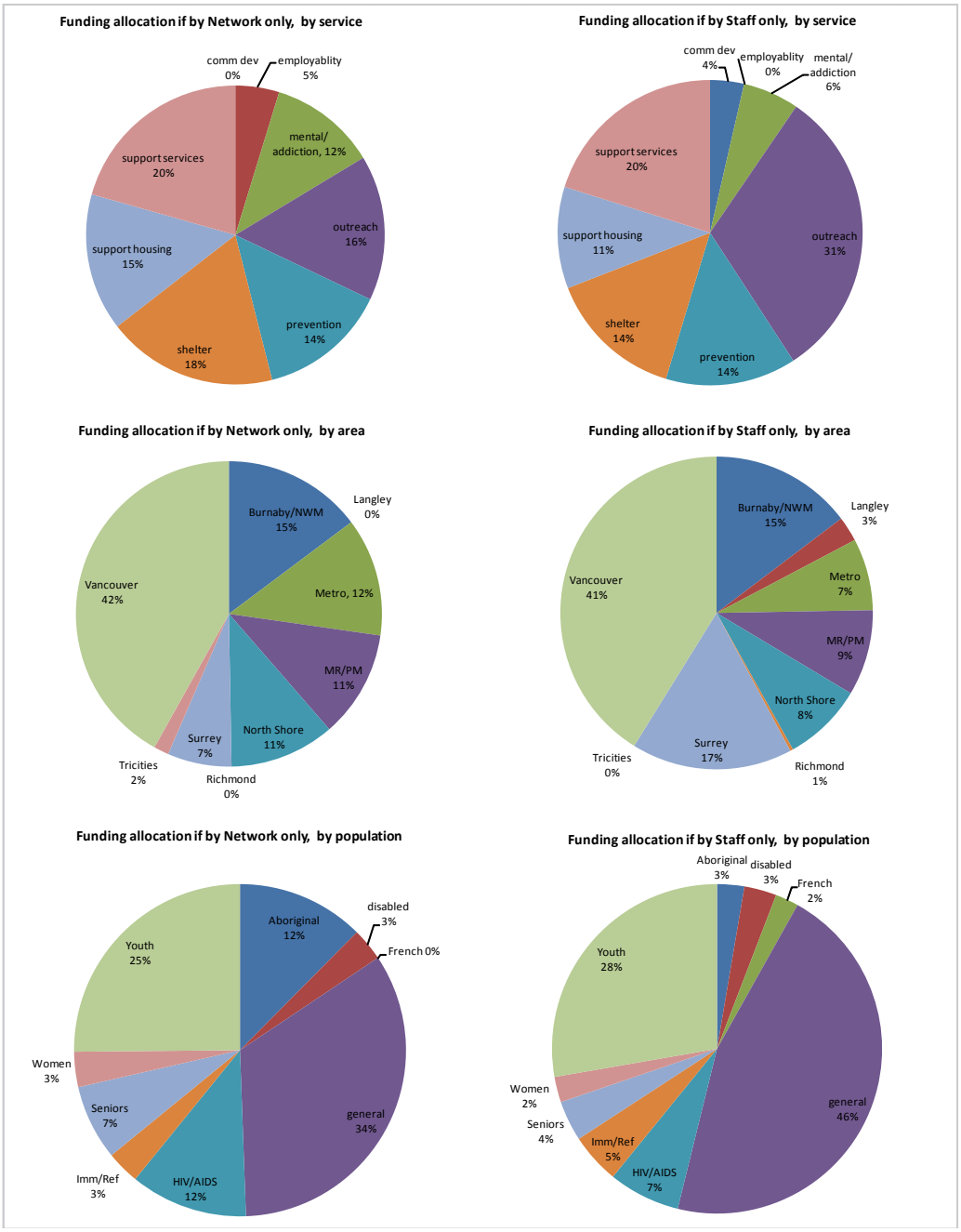
Table 2: Network and staff differences in programs funded, based on initial evaluations

FUNDS to allocate: $11,094,666			
IF NETWORK ONLY		IF STAFF ONLY	
Program	**Budget ask**	**Program**	**Budget ask**
Beilstein[17]	$629,782	Hellerman	$296,592
Maeno	$186,069	Calchera	$288,444
Lidder	$799,513	Maeno	$186,069
Cornelson	$277,318	Lidder	$799,513
Papranec	$315,800	Cornelson	$277,318
Imholte	$353,393	Papranec	$315,800
Nessmith	$379,240	Imholte	$353,393
Lettiere	$132,460	Nessmith	$379,240
Milkent	$877,369	Boccard	$265,990
Bellhouse	$186,878	Lettiere	$132,460
Kilstofte	$103,616	Baison	$34,000
Calleo	$221,290	Lindau	$219,506
Milz	$90,000	Milkent	$877,369
Metenosky	$434,848	Louria	$68,597
Morgret	$802,704	Milz	$90,000
Alrod	$177,470	Bandera	$134,027
Crawford	$390,000	Metenosky	$120,000
Bejerano	$259,914	Morgret	$802,704
Casoria	$453,942	Emberly	$585,772
Helley	$176,900	Bejerano	$259,914
Lesco	$489,700	Bartmes	$254,212
Cloman	$128,300	Bedaw	$64,000
Horneff	$332,392	Gunkelman	$189,885
Gettenberg	$530,320	Lordi	$912,114
Assalone	$275,185	Mettille	$304,294
Declark	$331,668	Horneff	$332,392
Bertman	$99,422	Gettenberg	$530,320
Corde	$568,480	Declark	$331,668
Drafall	$154,886	Bertman	$99,422
Mednis	$175,000	Feazelle	$324,640
Basone	$781,861	Corde	$568,480
		Basone	$781,861
TOTAL	**$11,115,721**	**TOTAL**	**$11,120,179**
count: program differences	**13**	**count: program difference**	**14**
percent difference	**42%**	**percent difference**	**44%**

So while we can be confident that network and staff differences in scoring based on initial reviews would result in different programs being funded, what is not clear yet is whether the funding differences would have policy implications. That is, would different types of services, in different areas of Metro Vancouver, targeting different sub-populations, be funded? By looking deeper into the policy and programmatic goals of each of the proposed homelessness programs hypothetically funded in this scenario, we can make such an assessment. Figure 3 below captures the differences in network and staff funding allocations, according to services funded, location, and target population.

[17] Organization names changed to protect the confidentiality of the deliberations.

Figure 3: Hypothetical differences in network and staff funding allocations.



The top two pie charts in Figure 3 show the differences, in percentage terms, of the funding envelope devoted to various homelessness services between network and staff hypothetical allocations. The most substantial differences among network and staff evaluations are with outreach, mental health and addiction, and employability services. Staff members would allocate twice the resources to outreach services than network members, which amounts to over $1.7 million in variation in allocations to this service. Network members, by contrast, would allocate twice as much in mental health and

addiction services, amounting to nearly $700,000 difference in investment, as well as $600,000 more in employability services (when staff would invest $0). In terms of the distribution of services in the Metro Vancouver area, staff would allocate 10% more ($1 million) to Surrey than network members. Finally, with respect to the various subpopulations among the homeless, staff show a tendency to fund services that target the general homeless population—that is, services which all can access—allocating 12% ($1.3 million) more than network members, whereas network members would allocate 12% ($1.3 million) to Aboriginal-specific programs, at the expense of general services.

Deliberative activity: does it matter?

The foregoing analysis has demonstrated that not only do network and staff members evaluate programs differently on an individual basis, but also that it has policy implications for the types of services funded, the location, and the sub-population targeted. Yet policy choices and funding allocations are not made on an aggregated individual basis. Rather, network actors come together after their initial individual evaluations to deliberate and make collective decisions on which programs to fund, with bureaucratic staff present to assist network decision-making. Thus whereas the first section sought to unambiguously demonstrate that there are meaningful differences in how policy problems and solutions are conceived in network governance contexts—specifically, including civil society actors in decision-making—the second section documents the final step of the real decision-making scenario: when network actors come together to deliberate over which homelessness programs should be funded, in the presence of bureaucratic staff to assist with decision-making. As such, there are new variables introduced in the second stage of decision-making: deliberation among network actors, pieces of information about the programs previously withheld are introduced (discussed below), and the influence of bureaucratic staff. Thus the first section captured the fundamental differences between network and bureaucratic evaluators, this section investigates the effect of deliberative activity in networks on decision-making.

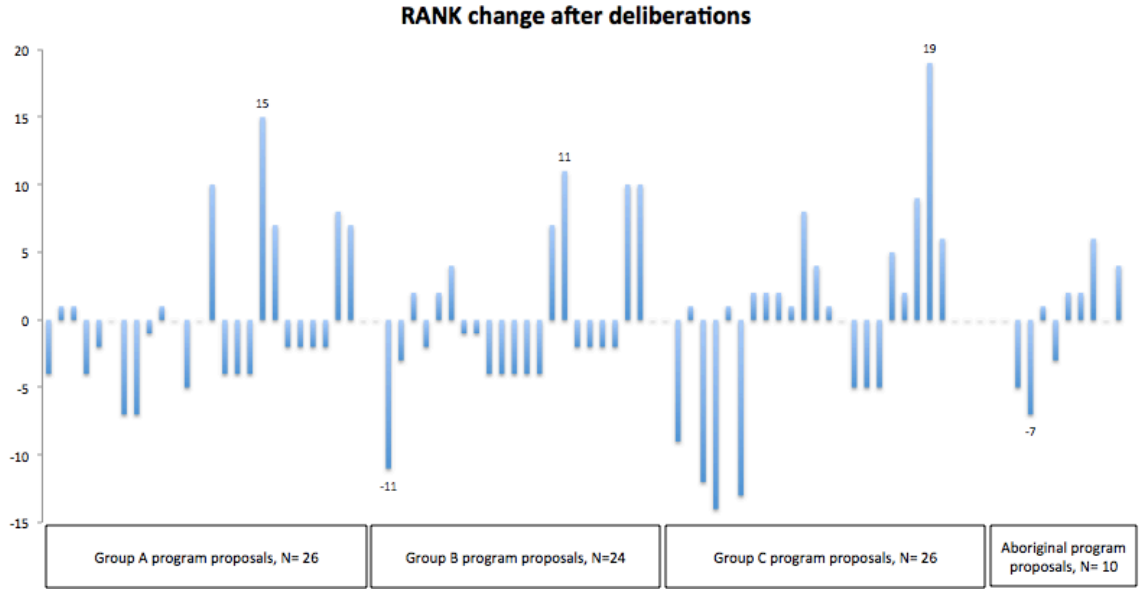
The deliberative context

After all members of each team completed their individual evaluations of the program proposals assigned to their group, they met as a group to discuss each program, deliberated over their strengths and weaknesses to create a ranked list of recommended programs for the whole RSCH to formally approve. Each network team review session had all of the same bureaucratic staff members present to help structure the decision-making, to manage all the information and data, as well as to provide technical advice when requested by network members. The demonstrated differences between network and staff members on preferred programs, analyzed in the previous section, thus become important to follow in this second stage of decision-making.[18]

[18] Advice requested from staff by network members ranged from eligibility of the program from the perspective of the Government of Canada, to the finer details of the budget request, as well as some value-for-money opinions. Formally, staff are present to help the process move along efficiently and effectively, and provide technical advice and support to network members as required—it was stated explicitly to all network members that they were the final decision-makers, not staff. This does not mean, however, that staff did not have any influence on decisions.

The deliberations among team members clearly matter, as evidenced by Figure 4 below, which shows how the initial rankings based on the aggregation of individual network scores can vary quite substantially from the network team rankings after deliberations. For example, in Figure 4, the highest column among the Group A programs increased from an initial network rank of 18 (out of 26) to become the 3rd most preferred program for network Group A, resulting in a rank shift of +15. Other substantial rank increases and decreases after network deliberations are identifiable for each network team. Not all rank changes were as dramatic, but Figure 4 reveals that a significant amount of reordering of preferred programs occurred as a result of team deliberations, and the longest bars denote the programs that experienced the largest shifts.

Figure 4: Change in rank of programs after deliberations, all network teams.



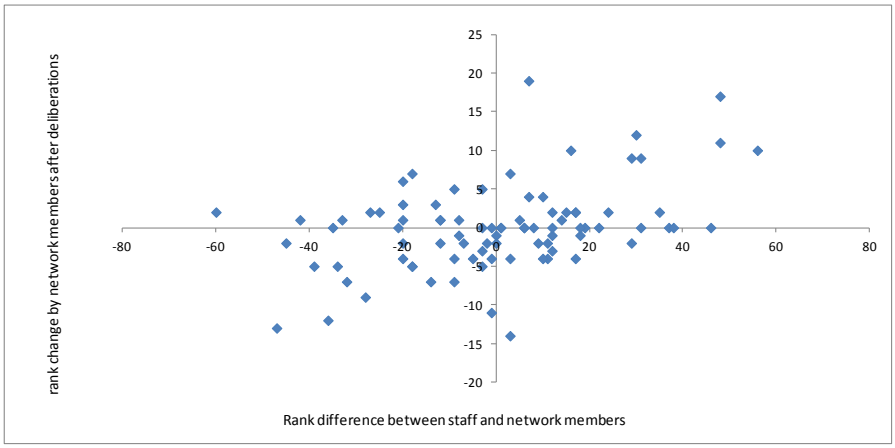
OLS Regression

Deliberations between network members on the merits of each program, which included bureaucratic staff offering technical advice, resulted in some significant shifts in the rankings of the programs from their aggregated scores based on the individual evaluations of network members. Less clear, however, is what is driving the change in rankings. OLS regression analysis can allow us to estimate the partial impact of factors that may influence the change in ranking of programs (pre-deliberations to post-deliberations), while holding the other variables constant. Public administration and governance theory inform the selection of variables, as do empirical observations from the researcher being embedded in the decision-making. In this last regard, the OLS regression analysis can help differentiate the anecdotal observations from the larger patterns that influence decisions in the network.

The dependent variable for the regression analysis is the *rank change of a proposed homelessness program after deliberations*.[19] That is, we know the rank of all programs based on the aggregation of scores from individual network members, as well as the rank of programs after the network met to deliberate their merits and collectively select the programs they wish to go forward. As such, a large (positive) value in change in rank after deliberations means that the rank improved, while a small (i.e. large negative) value means that the rank worsened after deliberations.

The independent variable of most theoretical interest is the difference in staff and network scores of each programs. That is, in cases in which there is large disagreement between these two groups of evaluators, we might expect that this could drive change in rankings of programs at the deliberation stage. And conversely, for the programs in which staff and networks generally agree on their placement relative to others, we would not expect much movement in rankings at the deliberations. In addition to anecdotal evidence obtained via participant observation, the bivariate correlation between the staff-network difference and the network rank after deliberations is r = 0.47, graphically depicted in Figure 5 below.[20]

Figure 5: Correlation between staff-network rank difference and network rank after deliberations.



There are several other independent variables likely to drive changes in the dependent variable (the change in rank of program after deliberations), all of which were identified from observing the decision-making: the 'budget ask' (e.g. how much the proposed program costs), whether the program was currently funded (under previous years of RSCH funding allocations), if eligibility concerns were raised, the type of service, its location, and the target population. These considerations were largely shielded from

[19] The change in rank after deliberation value is formed by subtracting the original network rank (from aggregating individual network scores) by the new network rank (after deliberations).

[20] The difference in staff and network scores can be measured in two ways: (i) the difference in average scores between staff and network members for a given program, which is measured by subtracting the average staff score from the average network score. With this measure, the larger the difference in average scores, the more the staff liked the program than the network members. It thus becomes a proxy measure for staff influence. The second way to measure staff-network differences is (ii) by the difference in rank of a program between staff and network members. This has an advantage over the score-based measure because staff, on average, scored programs 6.3% higher (Table 1) than network members—but comparing the differences in relative rankings corrects for this bias.

evaluators in the initial individual scoring, but featured prominently in the deliberations among network members and staff. With respect to the budget, it is expected that programs with larger budget asks, in the context of deliberations, will drive the rank down (i.e. negative rank change) after network deliberations. Another variable that appeared to be relevant in some cases during the decision-making was whether it was currently funded (by previous year's allocations of HPS funds).[21] This information was revealed at the deliberation stage to add context to the collective decision-making. For some programs, if network members agreed to not fund the program, its services would cease to exist (e.g. a homeless shelter would close), which is very different from agreeing not to fund a 'new' program.

Another new piece of information introduced at the deliberation stage were any concerns among staff that a given program was not technically eligible under the terms of the program defined by the Government of Canada. While all programs were linked to homelessness issues, there were a number of them—in the view of staff in particular—which were too tangentially related to homelessness, and thus could be disallowed by the Government of Canada under the terms of the HPS funding program.[22] This variable is also measured as categorical (dummy) variable and is applied to measure direct staff intervention in decision-making (though does not capture all staff influence).

Three additional variables are believed to be relevant to decision-making at the deliberation stage, identified in interviews with members and staff involved in previous deliberations at the RSCH: the type of service proposed, its location, and the target population. Since this is a regional homelessness governance network, members are keenly aware of the distribution of programs in the Metro Vancouver area, as well as the specific service needs and underserved populations. Again, the individual scoring by network members considered the programs in the abstract, according to their individual merit. But the deliberations force network members to think about the programs in relation to one another, and what would be the consequences for funding certain programs over others, particularly in terms of location, service, and target population.[23]

Regression function:

$$Y_i \text{ (change in rank)} = f \text{ (staff-network rank diff, budget, funded, eligibility , service, location, population)}$$

$$Y_i \text{ (change in rank)} = \beta_o + \beta_1(\text{staff-network rank diff}) + \beta_2 \ln(\text{budget}) + \beta_3(\text{funded}) + \beta_4(\text{eligibility}) + \beta_5(\text{service}) + \beta_6(\text{location}) + \beta_7(\text{population}) + \mu_i$$

[21] This variable is thus measured as a categorical variable, with a currently funded program = 1.

[22] Some examples included community development programs, research, employment programs and health services. It is formally up to the network membership to decide questions of eligibility, thus these concerns were raised by staff after initial evaluations so as to avoid biasing the initial perceptions of the programs among network members.

[23] Program locations were coded as: Vancouver, Inner suburbs (North Shore, Burnaby-New Westminster, Tricities, and Richmond), and Outer suburbs (Surrey, Maple Ridge-Pitt Meadows, and Langley). Program service types were coded: outreach, prevention (of homelessness), and employability. Program 'target populations' were coded as: general (services open to all groups) and vulnerable (women, youth and Aboriginal women and youth). There are additional types within each of the categories specified among the programs, but not in numerically sufficient numbers to make statistical inferences (i.e. target population of HIV/AIDS, n= 3; immigrant/refugee, n=4, leaving prison, n=3).

The results of a step-wise OLS regression are summarized in Table 3. The first estimation includes only the key independent variable of theoretical interest, staff-network rank difference, against the dependent variable, the change in rank of program after deliberations. The beta coefficient for staff-network rank difference is 0.47 and is significant at p=0.01. This can be interpreted as that the rank after deliberations will *increase* by 0.47 for every 1-point increase in the staff-network rank difference before deliberations. Recall that a higher staff-network rank difference means that staff favor a given program much more than network members. Therefore, this provides statistical evidence that when staff and network members disagree, the shift in rankings after deliberations *bends towards the original staff rank*. This variable alone explains 27% of the variation in the change of rank after deliberations, as per the adjusted R-square.

Table 3: OLS regression results for RSCH HPS deliberations.

Variable	Estimation						
	1	2	3	4	5	6	7
Scoring differences							
Rank difference	0.47*** (5.57)		0.38*** (5.46)	0.40*** (4.86)	0.36*** (4.34)	0.37*** (4.25)	0.37*** (4.27)
Staff-network % diff		0.16** (2.48)					
New Information							
Budget ask			-2.22*** (-3.39)	-2.04*** (-3.05)	-2.19*** (-3.30)	-2.24*** (-3.38)	-2.26*** (-3.44)
currently_funded			2.59*** (2.42)		2.29*** (2.06)	2.30*** (2.06)	2.46*** (2.21)
eligibility_concerns			-2.17 (-1.53)	-3.05**(-2.16)	-1.87 (-1.27)	-1.96 (-1.30)	-1.66 (-1.10)
Type of Service							
Service_outreach					0.40 (0.31)	0.39 (0.28)	0.70 (0.50)
Service_employability					-1.03 (-0.74)	-1.04 (-0.73)	-1.48 (-1.10)
Service_prevention					1.34 (0.85)	1.37 (0.87)	1.05 (0.66)
Program location							
Location_Vancouver						1.25 (0.78)	0.96 (0.60)
Location_Inner_suburbs						1.74 (0.34)	1.30 (0.72)
Location_Outer_suburbs						3.35* (1.76)	3.11* (1.72)
Target population							
Population_general							2.30* (1.69)
Population_vulnerable							2.06 (1.42)
Adjusted R-Square	0.27	0.07	0.39	0.35	0.38	0.39	0.39
N	87	87	87	87	87	87	87
Degrees of Freedom	85	85	82	83	80	77	75

t-values are given in parentheses. *** significant at p = 0.01, ** significant at p = 0.05, * significant at p = 0.10.

The subsequent estimations add the 'budget ask', currently funded and eligibility variables. The results indicate that 'budget ask' has a negative effect on the change in rank after deliberations, such that rank after deliberations will *decrease* by 0.22 for a 10% increase in budget ask, significant at p=0.01.[24] This is consistent with the expectation that a larger budget will negatively affect the rank of the program at the deliberation stage. The third estimation also introduces two categorical variables, whether the program is currently funded and if eligibility concerns were raised by staff. For the 'currently funded' variable, the results indicate that the rank after deliberations will *increase* by 2.42 if the program is currently funded, and for the 'eligibility' variable, the rank after deliberations will *decrease* by 2.51 if eligibility concerns are raised by staff, significant at p=0.01 and p=0.10 respectively. This model explains 39% of the variation in the dependent variable.[25]

[24] The budget variable was converted into natural log form, given that its bivariate correlation to the dependent variable is approximately s-shaped.

[25] A fourth estimation was conducted given that the 'eligibility concerns' variable was the expected sign, yet not significant, despite considerable anecdotal evidence of the phenomenon when the researcher was embedded in the network. The 'currently funded' variable and 'eligibility concerns' have an elevated bivariate correlation (0.30), so 'currently funded' was removed in the fourth estimation, which revealed a statistically significant relationship between

The final estimations focus on the specific policy and programmatic characteristics of the programs—the type of service, its location, and the target population—to gauge to what extent they influence the deliberations and the consequent change in rankings of programs. In terms of the type of service, none of the coded categories of dummy variables exhibited statistically significant relationships with the dependent variable. This could be the result of network members viewing the location or target population as more important than a particular service. Among the location categorical variables, the results indicate that the rank after deliberations for a program located in outer suburbs will increase by 3.35 (significant at p=0.10). This is consistent with expectations, since the RSCH is sensitive to the historical critique that funds were more likely to flow to the core of Vancouver. Among the target population variables, the coded category that captured services to be made available to the whole homeless or at-risk population (i.e. not restricted to a particular group, like women, youth or Aboriginals), the results indicate that the rank after deliberations will *increase* by 2.30 (significant at p=0.10). Serving the widest population in the context of very scarce resources, with network members trying to achieve economies of scale in service provision, is likewise an expected statistical finding.

Leveraging participant observation to interpret statistical analysis

The regression results allow claims of larger patterns to be made in statistical terms—and in fact show that staff-network differences have a statistically significant effect on the rankings of programs after deliberation—yet there are micro-phenomena and smaller, contingent patterns that are also part of the story on how network actors deliberate over policy and allocate funds. Only by observing the network in real-time can we isolate the process through which such networks and their deliberative activity matter. Observation of the deliberations yields one dominant explanation for the shifts in rankings: persuasion.

Persuasion as an emergent property of network activity provides the explanation for *why* networks matter for policy development and implementation in this decision setting, given that deliberative theory rests on a premise of the transformation of preferences among actors along the path to consensus (Young, 2000). Persuasion as a causal process operates not on the transmission of objective 'facts', but instead on the basis of 'mutual justification' (Mansbridge et al., 2010)— arguments that can be compelling and justified to those who reasonably disagree with them. Scholars suggest that network activity has policy consequence because various actors exposed to different experiences, new research and other jurisdictions will promote them within the network (Mintrom, 1997).[26] Persuasion involves the exchange of information among network actors (the premise

---

rank change and 'eligibility concerns', with the expected sign. Removing the 'currently funded' variable, however, had a negative effect on the R-square, and thus was reinserted in subsequent estimations in the interest of estimating the model that best explains the variation in the dependent variable.

[26] Some employ the concept of 'policy learning' as a causal process for understanding change in network or deliberative contexts, but the concept of persuasion, typically employed in the international relations literature, is more action-oriented and complex process capturing the influence of information, deliberation, and social context in changing actor's minds in decision-settings than typically applied concepts of 'policy learning' in the public policy literature (for example, May 1992, Sabatier and Jenkins-Smith, 1993).

behind 'learning' concepts), but adds the component of contestation and argument (Risse, 2003).

The concept of persuasion as a causal process helps understand how the change in the behavior of the participants in a social interaction is not limited to exogenous constraints on the individual—there are also endogenous processes, emergent from network activity (Johnston, 2001; Wendt, 1994). Johnston (2001) claims there are three ways in which an actor may be persuaded, all in the absence of overt material or mental coercion: (i) after engagement in a high-intensity process of cognition and reflection, (ii) by the relationship to the persuader or their position in the policy community, and (iii) by personal characteristics of the persuadee (experience, decision-making autonomy, strength of existing attitudes).

Deliberations among diverse actors are principally driven by the knowledge they bring. New information or reframed knowledge is thus always being inserted and floated in deliberative settings. Of course not at all contributions by network actors are equally persuasive or consequential. Based on participant observation, I argue that there are three key prisms through which knowledge claims flow through in deliberative settings: rules, credibility, and transmission prisms. The 'rules' prism is initially most consequential as a filter of arguments made by actors, and captures power relations in the deliberations, as those can set and enforce rules can use them to steer decision-making (Mansbridge et al., 2010). Especially relevant rules include the order of speech and the parameters of decision-making. Some arguments can be quickly dismissed or less persuasive on this basis. The institutional design of deliberations can thus focus deliberations or steer them from certain arguments. Any number of competing claims can emerge out of the 'rules' prism, which then must be assessed for their credibility, a second key prism. Features of credibility include how other actors interpret the position and experience of the actor, thus also capturing power relations among actors. The final prism through which contested knowledge claims are assessed is 'transmission'. Deliberations are a discursive process in which what matters is not just what one says, but how one says it, when and to whom. The transmission prism thus captures how arguments are presented, whether technical/scientific, politicized, or emotion-based means (see Triadafilopoulos, 1999 for theoretical bridging of rational and emotive persuasive rhetoric). An argument transmitted in a manner considered inappropriate by the collective norms of the participants may be dismissed from serious consideration. New or reframed information is constantly being introduced in deliberative settings, resulting in a cycling between credibility and transmission prisms until actors arrive at the convergence position (Mansbridge et al., 2010).

Figure 6: Persuasion, embodied three prisms, as a key process of the deliberative cycle

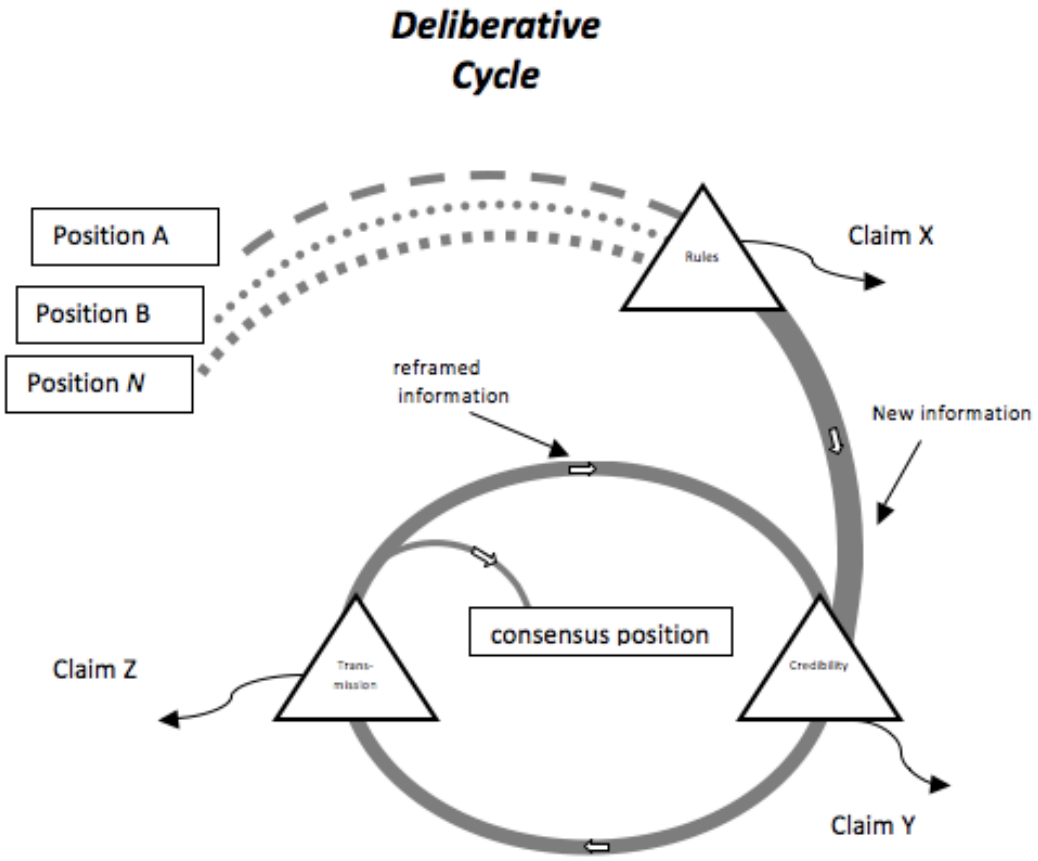


Figure 6 represents an inductively derived, abstracted deliberation cycle, where features of persuasion are the principle processes through which contested arguments are assessed and transformed in advance of reaching a convergence position. Examples from the RSCH decision setting can help demonstrate how arguments are accepted, transformed or ejected from deliberations.

Examples from the RSCH HPS decision setting can help demonstrate how arguments are accepted, transformed or ejected from deliberations. In the RSCH HPS deliberations, actors held various positions on the desirability of specific programs to fund. Yet the rules set in place governing the deliberations influenced the persuasiveness of certain arguments for or against program proposals. Perhaps the most consequential rule was that, typically the bureaucratic staff were the first to speak for every program under consideration. This had the effect of setting the tone for the deliberations around specific programs, depending on whether staff liked or disliked the proposal. There were a number of programs where team members individually ranked high, but when staff spoke out against them, they did not defend earlier scoring or challenge the staff argument. This institutional rule thus served to quell some dissent, as it is more difficult to shift the emerging tone than it is to build on it. Other rules in place that influenced the

persuasiveness of various arguments were with regard to the parameters of decision-making, specifically rules about the eligibility of program proposals under this funding program. Team member arguments on the desirability of rent subsidies, for example, in a program proposal were persuasive to many, yet were argued to be ineligible by staff in this funding program, and thus were dismissed on several occasions. Thus arguments can be persuasive to actors, yet suffer from institutional rules that marginalize them from consideration. Rules that set the parameters of decision-making exist in all deliberative settings, and have the effect of sidelining some arguments early in the process.

After the parameters of the deliberative setting filter out some arguments, the core process of deliberation begins, with actors offering arguments and introducing new information to support their claims or refute others. In the RSCH HPS deliberations, with some of the most significant shifts in ranking of preferred programs among team members, a network member introduced new data and the perception of the value of the proposed program changed. New data included knowledge about what was happening on the ground (ie. a shelter in the area is about to close), correcting a misconception about the organization or proposed program, or that the program was currently funded (and would thus close down if they did not fund it). In fact, almost all network members introduced at least one piece of information unknown or previously undervalued by other network members or staff during the deliberations and had an appreciable effect (good or bad) on the ranking of a program proposal.

When new information is introduced or arguments are articulated, its credibility must be assessed. The credibility of an argument rests not only on its logical connection to accepted facts, but also on the position and experience of the persuader. Some staff claims of the ineffectiveness of particular program proposals in reducing homelessness were strongly rebutted by team members with practical knowledge and experience of its potential. Staff certainly have credibility on the intergovernmental dimensions of policy, homelessness data, and the technical elements of budgeting, but less credibility on programmatic elements of proposals. Among team members, in certain scenarios one individual would be afforded substantial credibility given their experience and position on an aspect of homelessness (youth, Aboriginal, shelters, etc.). The credibility stemming from position and experience is not mechanically granted, but also depends on a demonstrated analytical approach. For example, by virtue of one network actor's professional position and expertise in the private sector, the consensus around an employability program— which was initially ranked very low by other network actors— shifted after this individual credibly articulated stories of programmatic successes. It was not merely the information presented by the actor that resulted in the ranking shift, but the legitimacy of the storyteller.[27]

After arguments and knowledge are filtered through the rules and credibility prisms, the persuasiveness is also assessed through how it has been transmitted to the group. The RSCH HPS decision setting is highly bureaucratized, driven by data and policy analysis. The vast majority of members were systematic in their approach, and visibly appalled

[27] A few team members who would seem to be credible specialists on paper were not granted this elevated status by team members because they exhibited a bias or non-analytical approach to their claims.

when network members launched overtly emotional appeals to support a program. Emotions were not absent, but only effective when complementary to a technical assessment of the merits of the proposal. Empty emotional appeals by individuals, and occasional threats of 'extreme disappointment' if the group did not decide in a certain direction, were universally unsuccessful in this deliberative setting. Arguments making links between generally accepted data and the objectives of the program proposal, transmitted in a passionate and bold manner, were most successful. In some cases, staff and network member disagreement was revealed and staff lobbied hard for (or against) the program, though it was always framed in terms of value-for-money, significant budget problems, or ineligibility (according to their interpretation of Government of Canada rules)—that is, in technical terms.

The process of persuasion, as conceptualized in Figure 6, thus allows for the relatively simple process of 'new facts' to be introduced and consequently change actor choices, but also for when more contested evidence and arguments are advanced, and why they are more likely to be accepted in certain constellations of actors and decision-settings. The deliberative cycle offered in Figure 6 is derived from extensive participant observation in a deliberative setting, but is abstracted as a generalized and abstracted model of how features of persuasion operates as a causal process in all deliberative contexts. The rules that govern the deliberations and set the parameters of decision-making, and the credibility and transmission of argument are the fundamental prisms through which arguments are deemed persuasive or not along the path of reaching a consensus position.

Conclusions

The data obtained from extended participant observation during the above-described process is both qualitative and quantitative, and only by weaving them together do we get a systematic and contextualized understanding of how and why governance networks matter to policy development and implementation. The central question introduced at the beginning focused on a counterfactual: would different decisions be made if bureaucrats, rather than network actors, made policy and allocation decisions with respect to homelessness? The data collected and presented in this paper functioned as a natural experiment and was uniquely positioned to answer this question, and the evidence indeed confirms that network members and staff show considerable variation (statistical and anecdotal) in their evaluation of programs and policy choices, and that deliberative activity in networks matters to the policy and programmatic outputs.

The first hypothesis predicted that bureaucratic staff and network members would score and deliberate programs differently, largely the result of different tacit knowledge, experience, and incentive structures associated with their positions. The quantitative data analysis confirms that along a number of dimensions, the staff and network members scoring differences were systematic, and the participant observation data suggests that this is largely explained by the on-the-ground knowledge and expertise network members possess that allow them to assess program proponent claims with more rigor. The implication of this finding is that there is value to extending this and other decision-making processes to one out of exclusive government control. Network actors contribute

meaningfully to the discussions, and not in a narrow self-interested manner, which is a frequent claim by critics of governance networks.

The second hypothesis predicted that bureaucratic staff and Aboriginal network members would score programs differently, though did not predict that staff would systematically under score Aboriginal programs, while systematically over score non-Aboriginal programs (vis-à-vis network members).[28] The implication of this finding is that institutionalized space for Aboriginal decision-making must be created and sustained for these urban issues. Otherwise they will not succeed in obtaining funding in such generic calls for proposals from government agencies. The third hypothesis predicted that deliberations would result in different program preferences than aggregated individual scores of network members may indicate. The OLS regression results demonstrate that deliberations matter greatly to decision-making, as there was significant shifting as a result of discussions among network members and staff technical advice. Evidence gathered from participant observation suggests that there were patterns for why deliberations mattered: persuasion is the dominant mechanism explaining how and why decisions were made, as the diverse actors shared tacit knowledge or information not known to others, which at times, radically changed their assessment of the program (usually around feasibility, local need, and organization's history). Persuasion, as articulated in the deliberative cycle in Figure 6, is the key causal process operating at the micro-level in deliberative settings. The deliberative cycle proposed does not tell us whether a specific decision will be made in a deliberative context—the process is too complex and contingent—yet it does specify a general model for why some arguments pass through key filter points (or prisms) in some contexts over others, with rules, credibility and transmission being the key determinants of persuasiveness of arguments on the path towards a consensus.

The broad narrative emerging out of this decision-making context is one of an interactive relationship between network members and staff that harnesses the expertise and knowledge of community and local government actors, while making decisions within the comfort level of bureaucrats responding to various political constraints. The decisions represent a blend of community priorities and government priorities. The implication of this finding is, of course, that staff exerted strong influence at times on the projects they did not wish to see go forward, which undermined the choices of network actors. Yet they also provide a complementary lens and perspective that network decision-makers relied upon and generally welcomed. For governance network critics, the findings should make clear that we should not be afraid that these networks are not sufficiently supervised or steered by government. And for network governance proponents, the evidence simultaneously demonstrates that the value of having diverse perspectives in a less-hierarchical decision-making structure—networks clearly matter to decision-making.

[28] This finding reveals what many Aboriginal leaders would already claim to know, but provides concrete evidence of its manifestation, even when bureaucratic actors are sensitive to Aboriginal issues and genuinely interested in the improving the life chances of homeless Aboriginals.

References

Adam, S. and Kriesi, H. (2007) "The Network Approach", in *Theories of the Policy Process*, ed., Paul Sabatier, Westview Press: Boulder, CO.

Agranoff, R. (2006) "Inside Collaborative Networks: Ten Lessons for Public Managers", *Public Administration Review*, 66 (s1), pp 56-65.

Allina-Pisano, J. (2007) "How to Tell an Axe Murderer: An Essay on Ethnography, Truth and Lies", in *Political Ethnography: What Immersion Contributes to the Study of Power*, Edward Schatz, ed., University of Chicago Press: Chicago. pp 25-52.

Benz, A. and Papadopoulos, Y. (2006) "Introduction. Governance and democracy: concepts and key issues". In *Governance and Democracy. Comparing National, European and International Experiences*, eds., Benz, A. and Papadopoulos, Y. London: Routledge, pp 1–16.

Borzel, T. (1998) "Organizing Babylon - on the Different Conceptions of Policy Networks," *Public Administration*, 76, pp. 263-73.

Bozeman, B. and Kingsley, G. (1998) "Risk Culture in Public and Private Organizations", *Public Administration Review*, 58 (2), pp 109-118.

Bressers, H. and O'Toole, L. (1998) "The Selection of Policy Instruments: A Network-Based Perspective", *Journal of Public Policy*, 18 (3), pp 213-239

Bridgman, R. (2003) "Bridging Public-Private Partnerships in a Case Study of Housing and Employment Training for Homeless Youth", *Canadian Journal of Urban Research*, 12 (2), pp 205-230.

Coleman, W., Skogstad, G., and Atkinson, M. (1997) "Paradigm Shifts and Policy Networks: Cumulative Change in Agriculture" *Journal of Public Policy*, 16 (3): pp 273-301.

Daugbjerg, C. (1998) "Linking Policy Networks and Environmental Policies: Nitrate Policy Making in Denmark and Sweden 1970-1995", *Public Administration*, 76 (2), pp 275-294.

Dowding, K. (1995) "Model or Metaphor? A Critical Review of the Policy Network Approach," *Political Studies*, 43, pp. 136-58.

Edelenbos, J. and Klijn, E.H. (2006) "Managing Stakeholder Involvement in Decision Making: A Comparative Analysis of Six Interactive Processes in the Netherlands", *Journal of Public Administration Research and Theory*, 16 (3), pp 416-446.

Fung, A (2008) "Citizen Participation in Government Innovations" in *Innovations in Government: Research, Recognition, and Replication*, ed. Sandford Borins, Brookings Institution Press: Washington, D.C., pp 52-70.

Gerber, A. and Green, D. (2012) *Field Experiments: Design, Analysis, and Interpretation*. W.W. Norton and Company: New York.

Head, B. (2008) "Assessing Network-based Collaborations: Effectiveness for whom?", *Public Management Review*, 10 (6), pp 733-749.

Howlett, M. (2002) "Do Networks Matter? Linking Policy Network Structure to Policy Outcomes: Evidence from Four Canadian Policy Sectors 1990-2000", *Canadian Journal of Political Science*, 35 (2), pp 235-267.

Klijn, E. and Skelcher, C. (2007) "Democracy and Governance Networks: Compatible or Not?", *Public Administration*, 85 (3), pp 587-608.

Klijn, E. and Koopenjan, J. (2000) "Public Management and Policy Networks: Foundations of a Network Approach to Governance", *Public Management*, 2, pp 135-58.

Kubik, J. (2009) "Ethnography of Politics: Foundations, Applications, Prospects", in *Political Ethnography: What Immersion Contributes to the Study of Politics*, ed., Edward Schatz. University of Chicago Press: Chicago. pp 25-52.

Leach, W. (2006) “Collaborative Public Management and Democracy: Evidence from Western Watershed Partnerships”, *Public Administration Review*, 66 (s.1), pp 100-110.

Mansbridge, J., Bohman, J., Chambers, S., Estlund, D., Føllesdal, A., Fung, A., Lafont, C., Manin, B., Martí, J. (2010). “The Place of Self-Interest and the Role of Power in Deliberative Democracy”, *Journal of Political Philosophy*, 18 (1), pp 64-100.

Marin, B. and Mayntz, R. (1991) *Policy Networks: Empirical Evidence and Theoretical Considerations,* Westview Prcss: Boudler, CO.

May, P. (1992) “Policy Learning and Failure”, *Journal of Public Policy*, 12 (4), pp 331-354.

Mandell, M. (2001) *Getting Results Through Collaboration: Networks and Network Structures for Public Policy and Management*. Greenwood Publishing: Westport, CT.

Mintrom, M. (1997) “Policy Entrepreneurs and the Diffusion of Innovation”, *American Journal of Political Science*, 41 (3), pp 738-770.

Mintrom, M. and Vergari, S. (1998) “Policy Networks and Innovation Diffusion: The Case of State Education Reforms”, *The Journal of Politics*, 60 (1), pp 126-148.

Montpetit, E. (2005) “A Policy Network Explanation of Biotechnology Policy Differences between the United States and Canada”, *Journal of Public Policy*, 25 (1), pp 339-366.

O’Reilly Fleming, T. (1993) *Down and Out in Canada: Homeless Canadians*. Canadian Scholars’ Press Inc: Toronto.

Pierre, J. (1998) *Partnerships in Urban Governance: European and American Experience*. MacMillan Press Ltd.: London.

Putnam, R. (1995) “Tuning in, Tuning Out: The Strange Disappearance of Social Scapital in America”, *PS: Political Science and Politics*, 28 (4), pp 664-683.

Rhodes, R.A.W. (2000) “Governance and Public Administration”, in *Debating Governance,* ed., Jon Pierre, Oxford University Press: Oxford.

Sorensen, E. and Torfing, J. (2007) “Introduction: Governance Network Research: Towards a Second Generations” in Sorensen, E. and Torfing, J., eds., *Theories of Democratic Network Governance*. Palgrave MacMillan: London.

Thatcher, M. (1998) “The Development of Policy Network Analyses”, *Journal of Theoretical Politics*, 10 (4), pp 389-426.

Tirole, J. (1994) “The Internal Organization of Government”, *Oxford Economic Papers*, 46 (1), pp 1-29.

Triadafilopoulos, T. (1999) "Politics, Speech, and the Art of Persuasion: Toward an Aristotelian Conception of the Public Sphere", *Journal of Politics*, 61 (3), pp 741-757.

Ward, J. (2008) “Dealing Effectively With Aboriginal Homelessness in Toronto”, Report for the Toronto Aboriginal Social Services Association. April 2008.

Whitehead, M. (2003) “In the Shadow of Hierarchy: Metagovernance, Policy Reform, and Urban Regeneration in the West Midlands”, *Area*, 35 (1), pp 6-14.

Wilson, J. (1989) Bureaucracy: *What Government Agencies Do and Why They Do It*. Basic Books Inc.: New York.

Young, I. M. (2000) *Inclusion and Democracy*. Oxford University Press: Oxford.