A Field Study on Matching with Network Externalities*

Mariagiovanna Baccara† Ayse Imrohoroglu‡ Alistair Wilson§
Leeat Yariv¶

August 4, 2009

Abstract

We study the effects of network externalities on a unique matching protocol for faculty in a large U.S. professional school to offices in a new building. We collected institutional, web, and survey data on faculty's attributes and choices. We first identify the different layers of the social network: institutional affiliation, coauthorships, and friendships. We demonstrate and quantify the effects of network externalities on choices and outcomes. Furthermore, we disentangle the different layers of the social network and quantify their relative impact. Finally, we assess the matching protocol from a welfare perspective. Our study suggests the importance and feasibility of accounting for network externalities in general assignment problems and evaluates a set of techniques that can be employed to this end.

JEL classification: D02, D61, D62, D85, C93.
Keywords: Matching, Social Networks, Externalities.

*We thank Dan Ackerberg, Allan Collard-Wexler, Jeremy Fox, Guillaume Fréchette, Maggie McConnell, Ariel Pakes, Debraj Ray, and Matt Shum for helpful conversations and feedback.

†Department of Economics, Stern School of Business, New York University, mbaccara@stern.nyu.edu

‡Department of Finance and Business Economics, Marshall School of Business, University of Southern California, aimrohor@marshall.usc.edu

§Department of Economics, New York University, alistair@nyu.edu

¶Division of Humanities and Social Sciences, California Institute of Technology, lyariv@hss.caltech.edu

1 Introduction

1.1 Overview

Externalities are commonplace within assignment processes: In the housing market, the value of a property depends on the demographics of neighboring homeowners. In an oligopolistic market, the returns from joining one firm depend on the composition of rivals. In universities, the desirability of a specific dorm room may depend on the peers in nearby rooms. In politics, the benefit from joining a particular party or coalition varies with the other political alliances formed. In team sports, the value in joining one team depends on the quality of other teams' players. And so on and so forth. Despite the wide range of applications featuring externalities, the matching literature, both positive and prescriptive, has largely ignored their presence.[1]

One of the significant challenges in assessing the role of externalities is that the underlying networks generating them are often unobservable or difficult to pin down. In particular, while attributes such as income, professional qualification, and education are frequently available, other important measurements of social connection—friendship, a shared professional history, etc.—are more difficult to obtain. Beyond the scarcity of data, the matching literature lacks a definitive framework that accounts for externalities, while still enabling empirical evaluations.

The current paper uses unique field data from *a centralized assignment process in which connections between individuals were mapped at both the professional and social level.* Specifically, our data originate from a matching process that assigned faculty members from a U.S. professional school to newly renovated offices. Using web and survey sources, we identify the institutional, coauthorship, and friendship networks of associations between the faculty involved.

Our study has three goals: First, to provide an empirical account of the effects of externalities resulting from agents' connections on behavior and outcomes. Second, to assess the differing networks' relative effects. Third, to evaluate the efficiency of matching protocols in terms of welfare, accounting for the identified externalities. As a by-product, our analysis sug-

[1] For several exceptions, see the literature review below.

gests and appraises several econometric techniques for estimating network externalities, and, in particular, assessing the relative importance of the different layers of agents' interactions.[2]

Our data describe the matching process and final assignment of 73 faculty into offices. The offices vary in their physical attributes—in particular, elevation, geographical exposure, and size, as well as their position and spatial relation to other offices. School officials designed a random serial-dictatorship matching protocol in which faculty members were coarsely ranked into three tiers according to career seniority, those with the greatest seniority choosing first; the order of choice was then randomized within each seniority level. Based on their rank number, each faculty member was called upon to select an office from those remaining, observing each of the selections made before them. After the selection process was completed, faculty were free to trade offices, and, additionally, were permitted to use transfers from their research budgets to facilitate trades.

In this environment, the externalities across agents can be easily mapped and separated into three layers of a social network. The first is institutional: the faculty members are divided into *departments*, and, within each department, specialization fields yield a further division into *research clusters*. The second social network is mapped using the past and current *coauthorship* links between faculty members. This network provides an alternative map of professional proximity, in which links between individuals were not incumbent on institutional affiliation or choice of research interest, but allowed to arise spontaneously across different clusters and departments through a bilateral choice. Finally, making use of a survey, we map a third social network, the *social interactions and friendships* between faculty members.

Our analysis follows several stages. As a preliminary step, we estimate an array of discrete-choice models in which, at each decision node, a particular faculty member faces a choice from a menu of offices, and makes their decision based on each office's physical attributes, as well as the network characteristics at the time of choice (for example, the number of coauthors who have located nearby). If network effects play no role in choice, the corresponding network

[2]To the best of our knowledge, all existing work estimating peer effects elicits one layer of interactions (be it social, professional, or geographical). In our study, we allow the data to speak as to which of the network layers matters and to what extent.

elements in our model should have no weight. However, all of our specifications generate significant network effects—in fact, the estimates suggest that network effects have a comparable impact to those of physical attributes. Nonetheless, this approach is tantamount to assuming that faculty are myopic, ignoring the implications of their choices on their peers' impending selections. In that respect, while we can reject the hypothesis that networks have no impact on choice, the magnitude of these effects should be interpreted carefully. This leads us to inspect more closely the dynamic and strategic aspects of the matching process.

In order to quantify the magnitudes of network sorting effects, while accounting for the strategic aspects present during the matching process, we compare the empirical assignment to a counterfactual in which faculty choose offices based only on physical attributes. Using the order in which faculty made decisions, we examine the outcomes that would have occurred had faculty made choices purely on the basis of offices' physical characteristics—for instance, preferring offices on higher floors to lower ones, large corner offices to all others, or western-exposure offices within a floor to eastern ones. Where faculty face a choice from a subset of offices with the same physical characteristics, we assume the particular office is chosen randomly from the subset. This allows us to simulate the resulting network clustering (for several preference specifications over offices' physical attributes) and compare it to that observed in the data.[3] The results from this comparison suggest that office proximity (both at the floor and office-neighbor level) among linked individuals occurs significantly more frequently in the observed assignment than in the simulated ones. Specifically, in the simulated assignments, members of the same department, coauthors, and friends are on the same floor 16%, 28%, and 17% less often than in our data, respectively. Similarly, proximity of office neighbors from every network layer were lower by 30% to 60%. From a general perspective, these results are illustrative (in both significance and magnitude) of the potential importance of network externalities on assignment outcomes.

Next, we disentangle the relative importance of each of the three network layers. In

[3] This is sometimes referred to as a dartboard approach in the context of spatial econometrics, see Ellison and Glaeser (1997), Guimarães, Figueiredo, and Woodward (2009), and Head and Reis (2005).

particular, we are interested in separating the effects of the institutional network, generated by department affiliation, from the idiosyncratic choice networks, described by coauthorships and friendships.

As mentioned before, following the sequential choice process, faculty were allowed to exchange their allocated offices using transfers between their research budgets. This allows us to define a simple notion of stability pertaining to the final assignment (after all swaps were carried out). We say that an assignment is *pairwise stable with transfers* if there is no switch in office assignments between two faculty members that results, with a transfer, in an improvement for both faculty, keeping all other office assignments fixed. We show that pairwise-stable assignments exist when utilities are such that: (i) the effects of offices' physical attributes are common across faculty and separable from network effects; and (ii) network effects are consistent in sign (for example, all positive) and separably additive (for example, utilities depend linearly on the number of peers that are within the relevant neighborhood).[4]

Pairwise stability (with transfers) imposes a sequence of constraints corresponding to all faculty pairs in our data. Using techniques developed recently for matching games without externalities (Bajari and Fox, 2009, and Fox, 2009), we estimate utility parameters for each of our network layers. We find that the coauthorship network has a greater impact than the institutional network, which is followed by the friendship network, which we find to have a negligible effect. Beyond the relevance to the matching process per se, this observation highlights the importance of studying the appropriate network of connections when examining peer effects. From a methodological point of view, these estimations underscore the importance of accounting for strategic behavior in dynamic matching markets. Indeed, the relative magnitudes of our estimates are different than those we achieve using standard discrete-choice models, which, as stressed above, omit the forward-looking strategic aspects.

[4]This is interesting from a theoretical point of view. Indeed, the literature on matching with externalities has mostly concentrated on existence of stable outcomes. As discussed in the literature review below, the difficulty arises due to the freedom one has in specifying beliefs regarding the reaction of agents to deviations (since, with externalities, a swap may affect the well-being of all agents, including those not involved in the swap). Our notion essentially entails myopic beliefs about the swaps that ensue. This assures existence.

Given the significance of externalities in individuals' utilities, it is interesting to contemplate the design of high-welfare assignments. In principle, designing the most efficient assignment is a complex problem due to the vast number ($73! > 10^{105}$ in our data set) of possible assignments.[5] As it turns out, designing the most efficient assignment for a class of preferences taking into account network externalities (that encompasses those we estimate) is a special case of *quadratic assignment problems* (see Koopmans and Beckman, 1957). While generally difficult computationally, and subject of an active line of investigation in operation research, we show how new techniques, still unexploited in the economics literature, can be used to estimate upper and lower bounds on the efficiency of the optimal assignment.

Under our assumptions that the effects of offices' physical attributes are shared across faculty and are separable from network characteristics, utilitarian efficiency is influenced only through the network effects present in our population. In fact, given our utility specification and the estimation results, any assignment that would increase the clustering of members from the different network layers would increase efficiency. Using our estimates of the relative weights different network variables carry in faculty's preferences, we can evaluate the efficiency of the matching protocol at hand. We assess a lower bound on the welfare loss generated by the outcome of the field experiment by identifying an assignment that achieves a 181% improvement over it. From a general institutional-design point of view, this analysis suggests the importance of recognizing and accounting for the underlying networks of relevant connections when constructing assignment mechanisms, and illustrates a computational technique for doing so in practice.

1.2 Related Literature

The idea that externalities may play a crucial role in group formation appears in some of the recent theoretical work on cooperative games. The general setup of games that are often referred to as "partition function games" (Lucas and Thrall, 1963, and Myerson, 1977)

[5] Furthermore, the presence of network externalities makes the problem significantly more intricate than the one pertaining to well-known problems of assigning goods exhibiting complementarities, for example, those relevant for spectrum auctions.

or "global games" (Gilboa and Lehrer, 1991) presumes that players' payoffs depend on the *partition* of the population. There are two general approaches that the literature takes. One strand focuses on core-like or Shapley value notions in which a particular belief structure (pertaining to the entire population's reaction to a coalitional deviation) is imposed (for example, Gilboa and Lehrer, 1991, De Clippel and Serrano, 2008, and Hafalir, 2008). The goal of this literature is to provide conditions under which the relevant solution concept exists. The other line of work is more explicitly dynamic in that it proposes a particular "bargaining protocol" by which coalitions are formed and analyzes the resulting set of equilibria in terms of efficiency and the pattern of emerging coalitions (see Bloch, 1996, Maskin, 2003, Ray and Vohra, 1999, and Yi, 1997).

In the context of matching, Sasaki and Toda (1996) illustrate the large freedom in beliefs regarding deviations from any market match that assure the existence of stable matches *for any* prevailing preferences.[6]

Without externalities, there is a large body of theoretical work that studies housing matching environments similar to ours (starting from Shapley and Scarf, 1974 and more recently explored in, for example, Che and Gale, 2009, Ehlers, 2002, Pycia and Ünver, 2007, and references therein).[7]

Empirically, while we are not aware of any studies quantifying the effects of network externalities in cooperative setups (matching environments in particular), the idea that peers may affect behavior and outcomes has been explored in many realms (see, for instance, Jackson, 2008 and Wasserman and Faust, 1994 for references).[8] Also, there is a recent literature

[6] The matching literature has also considered different types of externalities in many-to-one matching environments in which agents (say, workers in a firm, or students in a school) care about the peers who are assigned with them (but *not the entire assignment in the population*). That literature focuses mostly on conditions under which particular notions of stability generate non-empty predictions. See Echenique and Yenmez, 2007, and Pycia, 2009.

[7] In the no externalities world, there is also a budding literature studying decentralized dynamic matching games in which, similarly to our setting, agents may consider other agents' actions when deciding to commit to an irreversible match, see Niederle and Yariv (2009).

[8] In fact, several recent papers have mapped friendship networks in order to test for their effects on behavior in experimental games (for example, Leider, Möbius, Rosenblat, and Do, 2009 and Goeree, McConnell, Mitchell, Tromp, and Yariv, 2009).

on field performance of assignment mechanism, which does not account for externalities (see Abdulkadiroğlu, Pathak, Roth, and Sönmez, 2006, and references therein).

Methodologically, the dartboard approach used in Section 4 to estimate the impact of network externalities on the observed huddling of connected faculty has been used in other empirical studies on geographical clustering (for instance, Ellison and Glaeser, 1997, who use a similar approach to estimate geographic concentration of U.S. manufacturing industries). The estimations we perform in order to assess the relative magnitudes of the effects of the different layers of the underlying networks utilize identification techniques developed by Bajari and Fox (2009) and Fox (2009).

Finally, our welfare analysis involves finding the optimal solution for a quadratic assignment problem, which dates back to the specification of location assignments with externalities in Koopmans and Beckmann (1957). Solving this problem, which has been shown to be NP-hard, is a continuing area of research within the operations-research and combinatorics literatures.[9]

Several recent papers contain welfare assessments of assignments via random serial-dictatorship, *without externalities*. Manea (2007) characterizes subgame-perfect equilibrium outcomes of serial-dictatorship procedures for multiple objects, and finds that outcomes are not generically efficient, in contrast to the single-object case. Budish and Cantillon (2009) analyze data from a university's course-assignment process and find that the university's manipulable mechanism provides ex-ante welfare improvement over the strategy-proof and ex-post efficient random serial dictatorship.

Another assignment mechanism extensively studied in the literature is auctions. Bajari and Fox (2009) analyze the welfare loss in the sale of FCC spectrum licenses via auctions after constructing estimates of license complementarities. Again, externalities across different bidders' license assignments are assumed not to be present. Sönmez and Ünver (2009) discuss the welfare losses caused through auction mechanisms with endowment of fiat currency,

[9] Liola, Nair, de Abreu, Boaventura-Netto, Hahn, and Querido (2007) provides an extensive list of references in the operations research literature and Brandeau and Chiu (1989) provides a general taxonomy for a planner's location/assignment problems.

demonstrating the failure of these markets over straightforward statements of ordinal preference. Krishna and Ünver (2008) empirically analyze the results from course assignments with bidding, finding the auction mechanism inferior to a standard Gale-Shapley approach.

2 The Allocation Process

In this section, we describe the details of the matching protocol that was utilized in the field experiment, as well as the components of our data set.

2.1 The Matching Protocol

In 2006, plans to renovate one building of a large U.S. professional school were revealed to the faculty. The renovation would result in 74 vacant offices. Dean-level negotiations produced an initial list of 74 faculty members from 4 departments to occupy the new building.[10]

The assignment process used was a random serial-dictatorship procedure. As a first step, the school officials produced a coarse ranking of the 74 faculty members according to career seniority: priority was given first to chaired professors and department chairs, then full and associate professors, and, finally, assistant professors. The ordering of faculty within each group was determined by a *random draw* administered by the dean's office, under the supervision of department representatives.

Once the ranking was determined, the faculty members bound for the new building received a memo that provided the complete ordering, as well as instructions on how the process would evolve. These instructions indicated that all the office choices were to be conducted in one day. Each faculty member was able to see all the choices made up to the time of his/her own choice. New faculty members and others who could not be present on the day of the draft were asked to fill out a proxy form, and asked to give their proxy to someone else or provide a detailed list of their preferences so that a choice could be made on their behalf.[11] On the

[10] Before the renovation, three different buildings housed the offices of the school's faculty members. Departments were assigned to different floors within these buildings.

[11] Also, three offices were reserved to potential new hires, who were given a number in the ranking similarly to the other faculty members. Department representatives were in charge of selecting those offices.

Table 1: Summary Statistics on Office and Population Characteristics

Variable	Mean	Std. Dev	Min	Max	N
Office Variables:					
Large, corner office	0.10	0.30	0	1	73
Western exposure	.38	0.49	0	1	
Floor	6.37	1.24	4	8	
Faculty Variables:					
Years since PhD	13.64	11.09	0	37	72
Years in Department	9.51	8.78	0	35	
Female	0.21	0.41	0	1	73
Married	0.05	0.22	0	1	
Coauthors	1.56	1.63	0	6	
Colleagues	2.67	1.86	0	7	56
Social Events	0.89	1.15	0	4	
Lunches	2.00	2.04	0	9	
Friendship Network	2.39	2.18	0	9	
Institutional Variables:					
Department Size	19.60	5.21	13	26	4
Research Cluster Size	6.32	2.31	1	9	15

day of the office selections, a list of the faculty choices was posted publicly.

Conversations and discussions among the faculty took place prior to the selections. Faculty members were encouraged to make pre- and post-draft exchanges (prior to the draft, exchanges of rank numbers, and after the draft exchanges of offices). Further, *faculty were allowed to use funds from their research accounts to facilitate both types of exchange.* Indeed, while no ex-ante draft-number trades took place, ex-post trades involving 5 offices occurred immediately following the draft. We have detailed data regarding 73 of these faculty, which are the subject of our study.

2.2 The Assignment Data

We collected three types of data: data on office characteristics, population characteristics, and the matching process, which we now describe in turn.

2.2.1 Office Characteristics

The building had housed one of the departments for many years prior to the renovation. Therefore, faculty members from that department had detailed information on the desirability of different offices. Moreover, before the office selections were made, the dean's office provided detailed descriptions regarding office attributes to all faculty.

The top panel of Table 1 summarizes the characteristics for the available offices. The offices are located on the top five floors of an eight-floor building. Half of floor 4 and floors 5, 6, 7, and 8 had undergone renovation.[12] Each floor has offices that face east, west, and south. There are differences in size, floor, and view of the offices.

In terms of size, there are three office types. The majority of offices are identically sized (at about 213 square feet). These are the 56 offices facing either west or east, aligned on the two sides of a main corridor, on floors 5, 6, 7, and 8 (accounting for 76% of all offices). Then there are 8 large offices in the corners of the south sides of floors 4 through 8 (corresponding to 10% of all offices). These have an area of approximately 261 square feet and include an additional 20 square feet of closets Finally, there are 10 smaller offices in the south sides of floors 4 through 8 that have an area of approximately 200 square feet. Since the offices are very similar in terms of size, the view was considered an important distinguishing characteristic. Faculty were told that the preferred views tend to be on the higher floors, and on the sides facing west and south (in high floors, the west and south-exposed offices have open city views, while the east ones look onto a high-traffic artery).[13]

In what follows, we consider two alternative measures of proximity between two offices. In particular, we illustrate the impact of both *floor proximity* (that is, the two offices are on the same floor), and *neighborhood proximity* (the distance between two offices' doors is less than *30 feet*) on outcomes.[14]

[12] The lower floors and the other half of the 4th floor contained classrooms and were not changed.

[13] In fact, faculty members were encouraged to tour the building with its current tenants and examine the rooms regarding these attributes.

[14] The 30ft figure is chosen to include offices that are *less than two doors away* on either side of the main corridor. Since there are no offices directly in front of them, offices facing south have neighbors only on their

2.2.2 Population characteristics

The bottom two panels of Table 1 contain the summary statistics of our population, and Section 3 provides a detailed description of its characteristics. Faculty members' attributes were collected using two sources:

1. Web-harvested individual data on department, research cluster, arrival at the school, Ph.D. cohort, coauthorship, education background, and gender.[15]

2. Survey results. Faculty members were surveyed after the draft and the ex-post trades took place. There were 37 survey respondents (50% of the total number of faculty members). The survey elicited information on the faculty's social network as well as their preferences over the offices' physical (floor, view, and size) and non-physical (colleagues' proximity) attributes. The respondents were also asked to assess the importance of each office attribute keeping the other attributes constant. The complete list of questions asked in the survey and the summary of responses are available in Appendix A.

2.2.3 The Matching Process Data

Our data contain the complete results from the matching process. In particular, besides the final assignment of offices after swaps, for every choice made we know the set of faculty who had already chosen an office, how the partial assignment looked at the time of choice, and the remaining faculty who still had to make a choice.

3 The Underlying Networks

Individuals interact in different spheres. Since the location of an office may affect one's quality of work-life on both purely social and purely intellectual levels, we elicited peer connections on three dimensions: institutional links, determined according to the department or research

own side of the corridor. See Figure 2 for a graphical illustration of the offices' spatial arrangement.

[15] Additionally, the variable "Married" in Table 1 indicates whether an individual is married to another faculty member of the school.

cluster each faculty belongs to, coauthorship links, and friendship links. Below, we describe each of these layers of the network of connections and the correlations between them.

3.1 Institutional Network

The first network we consider addresses the research interests of faculty members, dividing them according to their specific research fields. The 73 faculty members are divided into 4 departments according to main research fields. Each department is further divided into different research clusters according to research sub-fields, resulting in a total of 15 clusters. The average department size is 19.6 individuals (ranging from 13 to 26). The average cluster size is 6.3 individuals, ranging from 1 to 9. The research network appears in Figure 1 below. In the figure, each node's shape corresponds to a particular research department, where each of the four departments is located in a different quadrant. Nodes are encircled in a set if they belong to the same research cluster. Since in this network individuals within a cluster (or department) are all connected to one another, each connected component in this network is complete (in particular, the average distance[16] within a connected component is 1).[17]

3.2 Coauthorship Network

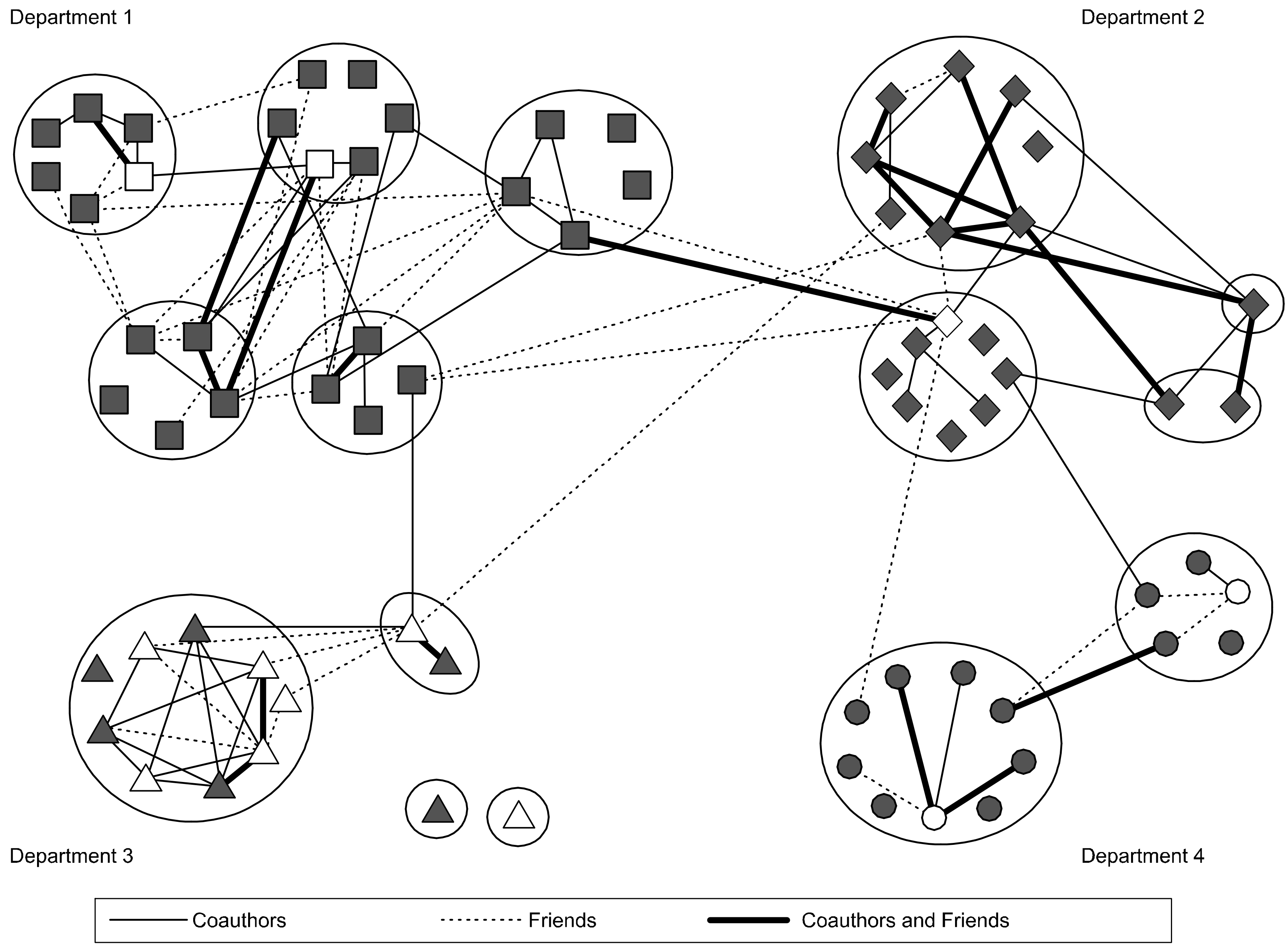
The second network encapsulates professional interactions among faculty, as captured by the existence of coauthored work. This network has been built combining web-harvested and survey data. In particular, in this network we consider two faculty connected if they coauthored at least one paper together in the past or are currently collaborating on a project (the latter element declared in survey responses). The coauthorship network is described by the solid lines between nodes (both bold and faint) in Figure 1.

At the time of choice, the average number of coauthors each faculty member had among other faculty members was 1.56, ranging from 0 to 6. The average distance between connected

[16] The distance between two nodes is defined as the number of links on the shortest path in the network connecting the two nodes.

[17] Moreover, we distinguish the faculty gender in Figure 1 using the shading of nodes (white for female faculty and gray for male faculty).

Figure 1: Network Diagram



Differing shapes represent department affiliation; ellipses research clusters; node shading represents gender with white female, gray male.

individuals in this network is 3.28, ranging from 1 to 10.

3.3 Friendship Network

The friendship network captures the interactions among the 37 faculty members who answered the survey, as well as the individuals socially connected to them (that is, individuals who were not survey respondents themselves, but were declared as a friend by at least one survey respondent).

In detail, Question 6 in the survey asked participants to name up to 5 fellow faculty members with whom they had lunch on a regular basis. Question 8 asked participants to name up to 5 personal friends (people with whom the participant interacts socially with outside school at least once a month) from within the school (see Appendix A). To build a friendship network, we combined the answers to these two questions. Two faculty are connected if they were mentioned one by the other in either Question 6 or 8. In particular, not all faculty included in this network are necessarily survey respondents—they are either survey respondents or individuals connected to a survey respondent, totalling 56 such *extended survey respondents* who compose the network. Twenty-one faculty members specified at least one individual on either one of these questions. We stress that survey respondents who did not specify colleagues' names answering Questions 6 and 8 could either (i) not have any social interactions with other members of the faculty, or (ii) have social interactions they prefer not to disclose in the survey.

This generates the network given by the dotted and bold line connections in Figure 1 (bold signifying links between nodes that are connected in both the coauthorship and friendship networks). We assume links are bidirectional. Indeed, Questions 6 and 8 were phrased so that individuals were specifically asked to report the frequency of interactions (lunches or social events outside the school), that are inherently symmetric.[18]

As reported in Table 1, the extended survey respondents are reported to have an average

[18] In our data, if we restrict attention to survey respondents alone, the probability that f considers f' a friend, conditional on f' considering f a friend, is 52%.

Table 2: Network Correlations

	Same Department	Same Cluster	Coauthors	Friends
Same Department	1	0.521	0.260	0.086
Same Cluster	0.501	1	0.322	0.093
Coauthors	0.223	0.275	1	0.160
Friends	0.188	0.172	0.294	1

Correlations for dummy variable indicating relationship between all faculty pairs (f, f'). Numbers below the diagonal correspond to the full sample, $N = 2628$, numbers above the diagonal correspond to the extended survey sample, $N = 1540$

of 0.89 individuals (ranging from 0 to 4) whom they interact with socially outside the school, and 2 colleagues with whom they regularly have lunch (ranging from 0 to 9). The average degree in the friendship network is 2.39, ranging from 0 to 9, and the average distance of individuals is 4.91, ranging from 1 to 12.

3.4 Overlap of Networks

Figure 1 demonstrates the complexity of the social network under examination, and the difference between the coauthor and friendship relations. Table 2 provides the exact correlations between the different layers of the social network. The correlations below the main diagonal are computed with observations associated with all faculty members. The correlations above the main diagonal are computed restricting the data to survey respondents and faculty members who are connected to them (in the friendship network). As clusters are subsets of departments, the two networks are highly correlated. However, the other network layers are not highly correlated with one another, and, notably, friendship ties seem fairly uncorrelated with department, cluster and, to some degree, coauthorship links.[19]

3.5 Existence of Network Effects in Office Selection

The choice made by each faculty member during the matching process could be influenced both by the office's physical attributes (floor, size, exposure), as well as the choices made (or

[19] The research network is rather different from the friendship and coauthorship networks in that links are arguably institutionally specified (whereas friendships and coauthorships are flexible individual choices).

expected to be made) by others. Figure 2 describes the outcome of the matching process (after ex-post trades took place) with dotted lines representing links within a floor in the friendship network, faint solid lines the coauthorship network, and bold solid lines the intersection of both networks. In particular, the figure represents the final spatial assignment by floor, with the shape of the nodes corresponding to different departments. Figure 2 is partially suggestive of the role department affiliation, coauthorship, and friendship may have played during the assignment process.

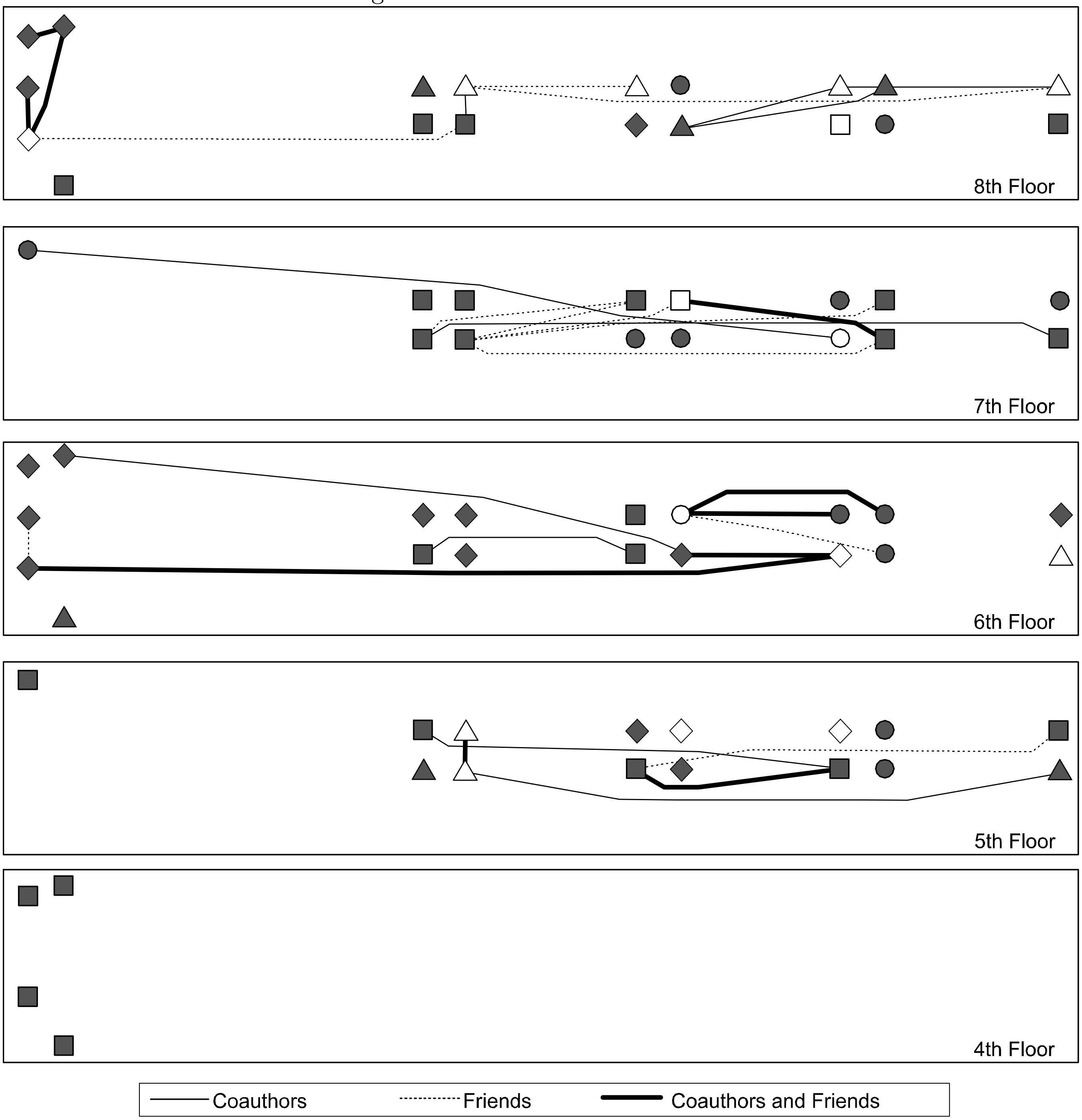
We start our investigation by assessing the *null hypothesis that network externalities are not taken into consideration during office selection.* As a first take, we consider the discrete choice each faculty is facing. Each observation in our sample corresponds to a pair (f, o), where f is a faculty member and o is an office *available to this faculty member at her time of choice.* We specify an array of models in which choices are explained by variables that correspond to both physical and network characteristics. Such an approach is inherently non-strategic in that we do not take into account forward-looking strategic aspects that are potentially present if network effects are at play (in particular, the approach does not allow us to quantify the effects of externalities). However, note that if the null hypothesis that network externalities were irrelevant to choice holds, any coefficient pertaining to network variables should not be significantly different from 0.

Table 3 contains the results of four conditional-logit regression specifications (where errors are always clustered by faculty[20]). The variables associated with offices' physical attributes are *Large Corner Office*, *Western Exposure*, and *Highest Available* (which are the corresponding dummy variables, that is, they take the value 1 if the office under consideration is a large corner office, has western exposure, and is on the highest floor available at the time of choice, respectively, and 0 otherwise).

In the first specification, denoted $CL(i)$, we include only the physical attributes of offices.

[20]Note that the set of available offices decreases by one unit after each choice is made. Thus, a faculty at position $k = 1, .., N$ in the ranking has $N - k + 1$ possible offices to choose from. Each observation in our data corresponds to a faculty and their menu of offices (excluding the last faculty who was left with no choice). Thus, for our 73 faculty, we have a total of $\frac{73 \times 74}{2} - 1 = 2700$ possible choices.

Figure 2: Networks on Floors



Differing shapes represent department affiliation; shading represents gender with white female, gray male.

Table 3: Conditional Logit Regressions

Variable	**CL**(i)	**CL**(ii)	**CL**(iii)	**CL**(iv)
Large Corner Office	0.155	0.284	0.237	0.277
	(0.097)	(0.078)	(0.088)	(0.087)
Western Exposure	0.312	0.394	0.377	0.388
	(0.045)	(0.033)	(0.036)	(0.039)
Highest Available	0.368	0.221	0.291	0.240
	(0.034)	(0.071)	(0.053)	(0.073)
Department Neighbor		0.174	0.229	0.243
		(0.043)	(0.038)	(0.047)
Department on Floor		0.082		
		(0.027)		
Coauthor Neighbor			0.186	0.221
			(0.103)	(0.113)
Friend Neighbor				0.291
				(0.344)
N	73	73	73	55

Coefficients represent marginal effects at zero; standard errors in parentheses.

Western exposure and high floor seem to have significantly affected faculty's choices. In the following specifications, denoted $CL(ii)$–$CL(iv)$, we introduce variables associated with the department, coauthorship, and friendship networks.[21] Consider an observation pertaining to a faculty and available office pair (f, o). *Department Neighbor* is a count of the number of offices neighboring o that are already taken by another member of f's department at the time of f's choice;[22] *Department on Floor* indicates the number of people in f's department that were present on the floor corresponding to o at the time of choice. Similarly, *Coauthor Neighbor* and *Friend Neighbor* are indicator variables representing the number of neighboring offices close to o that have been taken by faculty members with coauthor or social network ties, respectively, to f at the time of choice.[23]

The results of these specifications provide two main insights. First, network variables'

[21] Specifications CL(*ii*–*iv*) were chosen to correspond to our ensuing specifications in Section 5.

[22] Recall that in Section 2.2.1 we defined two offices as *neighbors* if the distance between office doors is less than 30 feet.

[23] The last specification, $CL(iv)$, is restricted to faculty who have links in our social network and as such pertains to fewer observations.

coefficients are positive and at the micro-neighborhood level significantly different from zero at any reasonable confidence level.[24,25] In particular, *we reject the null hypothesis that network externalities (at all three levels) did not influence faculty's office choices.* Second, the regressions suggest the importance of accounting for network effects when estimating such matching processes. Indeed, the coefficients corresponding to offices' physical attributes change significantly when we include network variables. Note that these coefficients respond in different ways to the omission of network variables: the effects of large corner offices are underestimated in $CL(i)$ relative to $CL(iii)$ since faculty choose offices close to colleagues when large ones are available; the effects of highest floor are overestimated in $CL(i)$ relative to $CL(iii)$, suggesting that faculty may be choosing higher floors to be in proximity to particular colleagues, rather than out of a preference for higher floors per se.[26]

4 A Dartboard Approach to Estimating Network Effects

The exploratory regressions discussed in Section 3 suggest the existence of non-trivial effects of network externalities on match outcomes. In order to assess the magnitude of these effects, we start by considering counterfactual assignment procedures that do not account for any observable externalities and compare the degree to which such procedures generate network proximity relative to that observed in our data. Put differently, we assess to what extent a random assignment based on purely physical office attributes can explain the observed patterns of social connection, accounting for the mechanism in place (namely, the order in which faculty chose offices).

[24] The joint hypothesis that the network variables are all zero in specifications CL(ii–iv) is rejected at the 99.9% level.

[25] When adding research cluster variables to any of these specifications, corresponding coefficients are not significant.

[26] In principle, there could be instances in which physical characteristics of an office are correlated with network attributes (for instance, corner offices are by geography more isolated than others). When controlling for offices' spatial arrangement by including offices' number of neighbors and a southern-corner office dummy, we find no significant effects associated with these variables.

We consider three types of procedures that differ in the prevailing faculty preferences for offices. For simplicity, we analyze lexicographic preferences that allow faculty to value certain physical attributes over others. In our first specification, denoted (I), we make all offices equivalent. In specification (II), we assume that each faculty has a lexicographic preference in which large corner offices precede western-exposure offices. In the third specification, denoted (III), we suppose faculty have lexicographic preferences such that higher floor precedes large corner offices, which precede western-exposure offices.

These specifications are consistent with survey results. Indeed, in Question 11 in the survey, 86% of respondents declared the top floors, floors 6–8, as their most preferred, and 83% declared the bottom floors, floors 4–5, as their least preferred.[27] In response to questions asking about the importance of office floor, size, and exposure (on a scale of 1–10), there were no significant differences in responses, though the distribution of responses of office floor did stochastically dominate those corresponding to size and exposure, suggesting that our counterfactual (III) may be of more relevance. Nonetheless, specifications (I) and (II) allow for greater freedom in the assignment process and therefore, in principle, make the observed assignment easier to achieve.[28]

Under both specifications, upon indifference, agents randomize among their top offices. For each of the two specifications, we used the order by which faculty chose to simulate the random-matching procedure $100,000$ times.

For every set of simulations, we considered the three network layers: institutional affiliation (captured through department), coauthorship, and friendship (encapsulating the connections determined through social interaction or lunch companionship as described in Section 2). We calculated the resulting average volume of faculty from each network in a participating faculty's macro-neighborhood (the floor of their office), and their micro-neighborhood (the set of close office neighbors), which are reported in Table 4. For example, the number of neighboring

[27] The assumption that faculty prefer larger offices to smaller ones and offices with city views to large road views seem natural first steps.

[28] Permuting the order of office characteristics that enter the lexicographic preferences does not alter results significantly.

faculty members who share a department affiliation is 89 in our data, but approximately 63 in each of our three specifications.

In fact, Table 4 illustrates how, across the network layers, simulated faculty numbers are consistently lower than those observed in the data. Floor-level proximity is lower by at least 16% for members of the same department and friends, and by at least 27% for coauthors.

At the neighborhood level our results are more striking, with proximity of office neighbors under the coauthor and friendship layers lower than the observed number by at least 40%. Department level sorting was lower by approximately 30%. The differences between simulated sorting and that observed were significant at any reasonable confidence level, with the probability of reaching or exceeding the observed levels of clustering in our simulations never surpassing 1% at the floor level, and 0.2% at the neighborhood level. We therefore conclude that the order in which faculty chose their offices does not seem, in itself, to explain a big fraction of the network huddling that occurred during the matching process.

While highlighting the effect of network externalities as a whole, this approach does not allow us to disentangle the differential effect from each of the different networks on agents' choices for two reasons: First, since the different layers of social networks are correlated, cross-layer comparisons of the dartboard approach results cannot be directly associated to agents' utility. Second, because of the strategic nature of the matching process, significant differences in how agents perceive the different layers of the social networks may result in small differences in the counter-factual estimates, and vice-versa.

5 Estimating the Relative Effects of Different Networks

The previous two sections motivated the importance of networks to the office selection mechanism. In particular, Section 3 made the case that externalities matter, while the dartboard approach presented in Section 4 allowed us to quantify the extent to which network externalities affected the final assignment, but without the ability to separate the effects of the different network layers. In this section, we seek to disentangle the relative importance of each of the

Table 4: Conterfactual Fills

Variable	**Match on**	**Specification**	**Observed**	**Mean Fill**	**Std. Dev**	$\tilde{\mathrm{F}}\{x \geq x^*\}$
Department	**Same Floor**	**(I)**	215	180.727	9.267	0.004
		(II)	215	180.560	9.155	0.003
		(III)	215	177.441	4.721	0.000
	Neighbor	**(I)**	89	62.804	6.505	0.001
		(II)	89	62.464	6.498	0.001
		(III)	89	62.597	5.098	0.000
Coauthor	**Same Floor**	**(I)**	21	12.102	3.066	0.007
		(II)	21	11.914	3.019	0.005
		(III)	21	15.201	1.710	0.000
	Neighbor	**(I)**	15	5.267	2.169	0.000
		(II)	15	4.865	2.056	0.000
		(III)	15	5.926	1.819	0.000
Friend	**Same Floor**	**(I)**	26	14.218	3.336	0.002
		(II)	26	14.200	3.277	0.001
		(III)	26	21.623	1.674	0.000
	Neighbor	**(I)**	18	6.193	2.365	0.000
		(II)	18	6.080	2.289	0.000
		(III)	18	8.070	2.039	0.000

Specification (I) fills randomly; specification (II) fills by large corner offices first and then western exposure; specification (III) fills by corner office, western exposure and floor from the top. $\tilde{\mathrm{F}}(x \geq x^*)$ gives the empirical probability of being equal to or greater than the observed value. Each procedure was simulated 100,000 times.

three layers on office choice: the department affiliation, the coauthorship network, and the friendship network. In particular, we are interested in separating the effects of the institutional network generated by membership in a department from the ones of the spontaneous networks based on coauthorship and friendship.

In principle, the mechanism instated for allocating offices defines an extensive-form game. One empirical strategy allowing to quantify the relative effects of network connections and physical characteristics of offices entails the assumption that faculty's assignments are irreversible and that they are in equilibrium.[29] That is, assume that agents' preferences take some functional form, allowing for the weight placed on network and physical characteristics of offices to be parametrized. Then, for any parameter value, there would be a corresponding set of equilibria of the assignment mechanism. In principle, this approach would allow us to select the parameters that best match our data. Note, however, that strategies in this game are contingent plans that specify, for each agent, a choice to be made after *any* conceivable selections by their predecessors. Since our data set contains 72 faculty who have a non-trivial choice, the set of strategies is vast and finding parametric equilibria profiles is not computationally feasible.

In order to overcome this computational difficulty and still exploit the strategic elements inherent in the matching process, we focus on natural restrictions on the emergent assignment. Namely, recall that faculty members were allowed to swap offices after the draft was completed, and that monetary transfers across research accounts were allowed to facilitate such swaps. In what follows, we assume that the transfers were not subject to budget contraints. Indeed, faculty were allowed to borrow against future years' research budget allocations. It follows that once the assignment has been determined (after all ex-post office swaps), we can assume that there are no remaining beneficial swaps, that is, the assignment is *stable*. Furthermore, we use the fact that faculty could make monetary transfers to assume transferable utility. In the spirit of Bajari and Fox (2009), and Fox (2009), we add structure to faculty's preferences

[29]The irreversibility of assignments is important in order to avoid the specification of the structure of the effective game that takes place after the initial allocation is made. In practice, we saw few immediate office exchanges.

and require that no two faculty would benefit from exchanging offices (accounting for network effects derived from such an exchange) regardless of the monetary transfers between them. This requirement provides us with a manageable set of restrictions that allows for preference estimation.

5.1 Stability with Externalities

Consider a finite set of faculty $\mathcal{F}$ and a finite set of offices $\mathcal{O}$. We ultimately observe an assignment $\mu : \mathcal{F} \to \mathcal{O}$, a bijection assigning each faculty member to a particular office. The utility of faculty member f can be generically represented by the utility function $u_f(\mu)$.[30] For any assignment μ, we denote by $\mu_f^{f'}$ the assignment derived from μ by exchanging the office assignments of f and f':

$$\mu_f^{f'}(x) := \begin{cases} \mu(f') & \text{if } x = f \\ \mu(f) & \text{if } x = f' \\ \mu(x) & \text{otherwise} \end{cases}$$

The notion of stability (with transfers) we use requires that for any faculty pair (f, f') there does not exist a transfer t from f to f' such that the bilateral exchange of offices specified by μ improves both their outcomes. That is,

$$u_f(\mu_f^{f'}) - t \geq u_f(\mu)$$

and

$$u_{f'}(\mu_f^{f'}) + t \geq u_f(\mu)$$

and at least one of these inequalities is strict. Or, equivalently:

Definition 1 (Pairwise Stability)**.** *An assignment μ is* pairwise stable *if for every pair $(f, f') \in \mathcal{F}$*

$$u_f(\mu) + u_{f'}(\mu) \geq u_f(\mu_f^{f'}) + u_{f'}(\mu_f^{f'}).$$

[30] Unlike matching settings without externalities, in which an agent's utility depends solely on their own match, externalities imply that utilities may depend on the entire assignment. Note that the set of feasible assignments has cardinality $|\mathcal{F}|! < \infty$, so a utility representation for each agent over all assignemnts is well defined.

We remain agnostic as to the exact nature of any bargaining or distribution of any pairwise surplus from a switch, but maintain the condition that should a pairwise reassignment be improving, that it be carried out.[31]

Two comments regarding the assumptions underlying this definition are necessary. First, for technical tractability, our stability notion essentially assumes that faculty have myopic or boundedly rational beliefs over the process that ensues following a deviation. Indeed, in the presence of externalities, a switch by any pair of faculty affects others uninvolved in the swap. In general, one could contemplate beliefs specifying the reactions of all participants to such a deviation (in which case even existence can be problematic to obtain, see Sasaki and Toda, 1996 and Hafalir, 2008). Second, our notion considers only bilateral swaps, rather than larger groups' exchanges. We choose to focus on pairwise stability for simplicity and to match the behaviorally founded idea that it would be harder for larger coalitions to solve this assignment problem within their coalition.[32]

Pairwise stability generates $\frac{(|\mathcal{F}|-1)\times|\mathcal{F}|}{2}$ necessary inequalities. In our data, one faculty moved to a different building after the initial assignment. Therefore, with 72 faculty left after the post-draft moves, we generate 2556 inequalities. We will assume that $u_f(\mu)$ takes the form:

$$u_f(\mu) := P_{\mu(f)} + \beta \cdot R(f;\mu) \tag{1}$$

where P_o represents the physical desirability of office o (its view, size, exposure, etc.) and $R(f;\mu)$ is a vector of network effects specific to f induced by the assignment μ (proximity to coauthors, friends, departmental colleagues, etc.). In fact, throughout our analysis, we will assume that $R(f;\mu)$ depends (linearly) on the number of faculty from each network under consideration that end up on their floor or in their immediate neighborhood. That is, for any

[31] It is useful to contrast the notion of stability we use, exchange of an assigned object between two faculty f and $f^{'}$, and the blocking-pair notion of stability in the two-sided matching, where a faculty-office pair (f, o) would block an assignment. Due to a lack of agency or preference on the side of the offices, the blocking coalition is of the same size, two agents, but on one side of the market.

[32] Indeed, of the observed switches after the assignment process, the first involved a pair of faculty exchanging their offices, and the second involved a triple, where one of the three moved to an office outside of the building, and the other two organized a switch over the three offices (their own two and the one freed up by the move). So the maximal size of the switching groups was in fact two.

faculty f, let $k(f, \mu, l)$ be the number of faculty from network layer l, $l = 1, ..., L$ (research, coauthorship, or friendship) that are in f's neighborhood (say, floor) under the assignment μ. Then,

$$u_f(\mu) := P_{\mu(f)} + \sum_{l=1}^{L} \beta_l k(f, \mu, l).$$

This formulation allows for the volume of peers in someone's proximity to affect their well-being. For simplicity, we assume that the volume of faculty members unconnected to the individual have no effect on well-being at either of the proximity levels. This formulation is general in that: (i) Networks could be thought of as bilateral, with each pair of agents constituting a particular layer l, so it is additive separability, not linearity, that places the main constraint on utilities' functional form. (ii) The coefficients $\{\beta_l\}$ are not restricted in sign so that peer effects can be either positive or negative.[33]

Proposition 1, whose proof is given in Appendix B, shows that the market structure we impose allows for the existence of pairwise-stable assignments.[34]

Proposition 1 (Existence)**.** *There exists a pairwise-stable assignment.*

We now add a stochastic term to represent an idiosyncratic component for faculty f's preferences for a match μ so that preferences are represented by:

$$U(f, \mu) := P_{\mu(f)} + \sum_{l=1}^{L} \beta_l k(f, \mu, l) + \epsilon^f_{\mu(f)}$$

where ϵ is the match-specific unobserved idiosyncratic error. Given this specification, consider the pairwise stability-condition corresponding to two faculty members. The physical attractiveness of the office essentially serves as a fixed effect when contemplating a swap, which can be directly compensated for with a transfer. Consequently, pairwise-stability constraints put

[33] In particular, a layer must be symmetric: should faculty member f value the proximity of $f^{'}$ at x utiles, then it must be the case that $f^{'}$ values the proximity of f at x too. This symmetry rules out cycles where one faculty member desires close proximity to another who desires distance, and is key to the existence result in Proposition 1.

[34] We note that existence of stable assignments in the presence of externalities has been a major hurdle in the theoretical literature on the topic. Our existence result suggests that in environments such as those we study, stability is a managable concept.

restrictions only on the network components of faculty's utility. Formally, pairwise stability of a match μ translates into the following. For any two faculty f, f', noting that $\mu(f) = \mu_f^{f'}(f')$ and $\mu(f') = \mu_f^{f'}(f)$,

$$\beta \cdot (R_f(\mu) + R_{f'}(\mu)) + \epsilon^f_{\mu(f)} + \epsilon^{f'}_{\mu(f')} \geq \beta \cdot \left(R_f(\mu_f^{f'}) + R_{f'}(\mu_f^{f'})\right) + \epsilon^f_{\mu(f')} + \epsilon^{f'}_{\mu(f)}. \tag{2}$$

The inequalities captured in (2) allow us to estimate the underlying parameter vector β.

5.2 Implementation

The set of inequalities defined by (2) serve as the basis for maximizing a score function (see Manski, 1975) defined as:[35]

$$Q(\beta) := \sum_{f \neq f'} 1\{\ \beta \cdot \left(R_f(\mu) + R_{f'}(\mu) - R_f(\mu_f^{f'}) - R_{f'}(\mu_f^{f'})\right) > 0\} \tag{3}$$

Three remarks about this objective function are in order. First, note that the satisfaction of each inequality in (3) is defined in terms of strong inequalities. While inconsequential for the estimated parameters themselves, this allows us to get slightly more meaningful optimal score values. Indeed, in many cases our network measures are sparse, that is, two faculty are not likely to be connected across a particular measure. When individuals are not connected, the corresponding summand in (3) would always be satisfied if the inequality were weak. In particular, the values of the score would be shifted up by the number of faculty pairs who are not connected in any of the network layers relevant for the specification.[36]

Second, we use this objective rather than a smoothed version a-là Horowitz (1992) for simplicity and freedom from the need to choose tuning parameters for smoothing. This comes at the natural cost of the objective being discontinuous and not amenable to differential methods of optimization.

[35] Estimations were performed using *Mathematica*'s differential evolution algorithm, which has good properties when used to find global extrema of optimization problems (see Fox, 2009 and Santiago and Fox, 2008).

[36] We stress that, since our scores are lower than the ones obtained with a weak inequality version of (3), there is a difference in magnitude between our scores and the ones found in the literature that employs maximum scores with weak inequalities (see, for instance, Bajari and Fox, 2009).

Table 5: Single-variable Explanatory Power

Normalization	$\beta = 1$		$\beta = -1$	
Department Neighbors	1388	881	348	169
Department Floor	1045	621	921	569
Coauthor Neighbors	1092	820	193	90
Coauthor Floor	984	673	307	165
Friend Neighbors	-	903	-	151
Friend Floor	-	751	-	188
N	2556	1504	2556	1504

The column headed $\beta = 1$ ($\beta = -1$) corresponds to a positive (negative) normalization for variable. The scores correspond to number of inequalities predicted with objective (3).

Finally, the value of the score is invariant to scaling of the parameter β (that is, for any $a > 0$, $Q(a\beta) = Q(\beta)$). Identification therefore requires a normalization for one of the coefficients, which must have a non-zero contribution to preferences. As previously demonstrated through our discrete-choice estimations (see Table 3) and illustrated in Figure 2, locating near department colleagues plays an important role in location choice. In addition, since the average degree corresponding to the departmental affiliation network is high (relative to the other network layers we consider), many of the inequalities in (3) have non-trivial elements pertaining to departmental network effects. We therefore normalize the coefficient for the proximity of a departmental neighbor to 1, denominating the remaining variables in terms of foregone departmental neighbors. In order to further justify this normalization Table 5 provides the score Q when accounting for only one layer of the network. Since the magnitudes of the relevant coefficient β cannot be calibrated, we look at the scores for $\beta = 1$ and $\beta = -1$. Table 5 reports the score Q for both the entire data set and the subset of observations corresponding to participants of the friendship network. The department variables are consistently the ones generating the highest score levels, thereby explaining most of the restrictions.[37]

[37] Note that, as stressed above, we would have obtained higher scores using a weak inequality version of (3). In particular, the only inequalities that we would not have been able to predict are the ones explained by a negative coefficient ($\beta = -1$) when inequalities are strict. For instance, Department Neighbors would have predicted 2208=2556-348 with a positive normalization.

5.3 Results

Holding constant the department-neighbor normalization discussed above, we now estimate the intensities of each network layer relative to *Department Neighbors*. Our results are given in Table 6. Coefficients are reported as an identified interval $\left(\underline{\beta}^i, \overline{\beta}^i\right)$ for the specific variable i, where $\underline{\beta}^i$ is the minimal coefficient *that maximizes the objective (3),* and $\overline{\beta}^i$ is the maximal coefficient.[38] That is:

$$\begin{aligned} \underline{\beta}^i &= \arg\min \left\{ \tilde{\beta}^i \mid \tilde{\beta} \in \arg\max_{\beta \in \mathbb{R}^L} Q(\beta) \right\} \quad \text{and} \\ \overline{\beta}^i &= \arg\max \left\{ \tilde{\beta}^i \mid \tilde{\beta} \in \arg\max_{\beta \in \mathbb{R}^L} Q(\beta) \right\}. \end{aligned}$$

This approach is required by the lack of point identification for β, that we discuss below, caused by two factors: (i) a discontinuous objective; and (ii) integer measures for our network layers.[39]

The first column in Table 6, titled PS(i), gauges the relative importance of the micro and macro neighborhoods for department members. *Department Floor* is a count of the number of department colleagues currently located on the same floor for the particular faculty-office pair under consideration. As such, the floor variable can only aid in explaining floor choice, since the difference in floor counts under an intra-floor exchange is zero for both faculty members involved in the swap. The specification produces no significant effect for floor organization. However, even though the floor coefficient is small in size, the negative sign still explains approximately 500 inequalities over the single variable *Department Neighbors*. The negative sign could be interpreted as a decreasing-returns effect to *Department Neighbors* with which it is (unsurprisingly) correlated. Given the small estimated size of the coefficient, and the insignificance of the effect, we conclude that local proximity is much more important than

[38] Similarly, confidence intervals reported in Table 6 are constructed using the minimal coefficients for the lower limits and the maximal coefficients for the upper limit, a particularly conservative approach to finding the 95% confidence interval.

[39] Note that when the specification entails more than one coefficient in addition to the normalized one, $\underline{\beta}$ and $\overline{\beta}$ need not maximize the objective. In particular, focusing on a set of coefficients with minimal absolute values provides a conservative assessment of the effect of the corresponding variables.

Table 6: Discrete Variable Pairwise-Stability Estimates

Variable	**PS**(i)	**PS**(ii)	**PS**(iii)
Department neighbors	1	1	1
Department floor	(-0.07, 0)		
	(-0.11,0.07)		
Coauthor neighbors		(4.00,5.00)	(4.00,5.00)
		(1.00,5.00)	(0.80,5.00)
Friend neighbors			(0.00,0.50)
			(0.00,1.45)
Score	1954	1734	1215
	(2134)	(2165)	(1338)
N	2556	2556	1540

95% confidence interval in parentheses, derived from bootstrap $B = 1,000$, where the upper and lower limit correspond to the 2.5% and 97.5% percentiles of the minimal and maximal coefficient values. The first variable's coefficient in each column, β_1, is normalized to magnitude 1 for scale identification. Score in parentheses is the number of inequalities predicted with a weak inequality version of (3).

floor proximity.

In PS(ii), we evaluate the importance of the coauthorship network relative to the department network.[40] The first point to note from PS(ii) is the large and significantly positive effect coming from *Coauthor Neighbors*. In particular, looking at the assessed interval we find that having a single coauthor nearby is enough, ceteris paribus, to compensate for four to five department colleagues in neighboring offices. The constructed inference region for this coefficient is large, and is arbitrarily close to 1—however, we point out that this confidence interval is a conservative estimate of the true 95% confidence interval.

It is worth noting here that the score in PS(ii) is substantially lower than in PS(i), with the same number of explanatory variables when focusing on strict inequalities (this conclusion is reversed when looking at the weak inequality version of the objective in (3)). This is due to the larger number of degenerate inequalities generated by PS(ii). Restricting the calculation

[40] Specifications including floor-level variables for both coauthors and friendships found results similar to PS(i). While the inclusion increases the score slightly, the magnitude and significance of the coefficients remain similar. This leads us to believe that floor-level variables do not have powerful explanatory value. Similarly, the inclusion of variables related to research clusters do not add explanatory power to the specifications PS(ii)–PS(iii).

to inequalities with non-trivial content—that is, those swaps that can be *positively* explained by a vector β—leads to similar explanatory power for the two specifications: PS(i) predicts 82.2% of them correctly, against 81.6% for PS(ii).

Finally, we introduce data from the friendship network in PS(iii). In this specification we only include those inequalities corresponding to swaps between the 56 members of the extended-survey network.[41] Consequently, the sample size decreases by approximately 40%. In our previous empirical approaches, in Sections 3 and 4, the friendship and coauthorship networks appear to have a comparable impact on the final outcome. Specifically, inspection of Table 4 points to a similar level of additional sorting in each network, and the coefficients in Table 3 represent similar effects from each on the probability of choosing a particular office. Surprisingly, however, we find that the friendship network does not have a strong effect on office choice with respect to the other network layers. This result is due to an important difference between this approach and the previous techniques: indeed, the stability estimation incorporates the information derived from the lack of ex-post swaps that could have produced greater proximity among friends. Finally, note that of those pairs with a non-trivial inequality, PS(iii) predicts 85.7% of the inequalities correctly.

In order to measure the tradeoffs between the addition of variables and the number of inequalities explained we construct the following criterion. If regressors were i.i.d. random variables (with median 0) uncorrelated with outcomes, a single regressor would explain approximately one half of the non-trivial inequalities in a large sample. The addition of another independent regressor would explain half of the remaining inequalities, and so on. Thus, with two regressors our baseline is 75% of non-trivial inequalities explained, with three, 87.5%, etc.[42] Given the positive correlation between our network measures, this is a harsh criterion for evaluating the predictive power of an additional regressor. Both PS(i) and PS(ii)

[41] However, the vector of characteristics for each of the remaining faculty are calculated using data on the entire population. For example, a faculty member outside of the extended-survey sample will still be counted as a coauthor/colleague neighbor when considering the observed assignment or a prospective swap.

[42] Note that this rule does not specify anything about the intensity of the coefficients derived, just that a signed variable can explain this many of the remaining inequalities.

exceed this level of sorting but PS(iii) is slightly below the 87.5% threshold. We view the failure to meet this threshold as further evidence for the under-organization contributed by the friendship network when both department and coauthor networks are controlled for.

Two final notes are in order regarding the econometric techniques used. First, we utilize the bootstrap method in order to derive the confidence regions illustrated in this section. Subbotin (2008) provides asymptotic arguments for bootstrap inference over the class of algorithms containing the maximum-correlation estimator. One of the operating assumptions for this result is continuity of the independent variables. Our environment does not meet this particular requirement because of the binary nature of both the network relations and the physical proximity measure for offices. However, with finite samples, we believe the usefulness of the bootstrap method is in testing the sensitivity of the estimated coefficients to the over- and under-sampling of specific inequalities within the data, while its freedom from any tuning parameter is an additional benefit.[43]

A further problem relates to the asymptotic identification assumptions required for this estimation method. The assumptions in Bajari and Fox (2009) require that, given enough data, for any parameter vector different from the true one, more inequalities would be violated. In a setting in which all of the underlying explanatory variables are related to social-network degree, one may expect problems since: (i) network degrees (and, in fact, many other network characteristics) are discrete by nature; and (ii) it is unlikely that individuals have unboundedly many social connections, even as we consider bigger and bigger samples, so the support is bounded. Point identification is therefore inherently problematic. It is for this reason that we focus on estimating the interval of coefficients that maximize our objective function. This is well-defined, computationally feasible, and pessimistic in terms of coefficients' significance.

[43] An alternate approach would be subsampling, as outlined in Politis and Romano (1994), with the tradeoff being the necessity of specifying a subsample size parameter. In a strong form—using all the inequalities generated between random samples of 42 faculty members, a number chosen so as to produce roughly 1/3 the number of inequalities as the full sample of 72—subsampling generates similar though less significant results. For instance, for the variable *Coauthor Neighbors* in PS(ii), the lower bound of the confidence interval (corresponding to $\underline{\beta}^i$) is 0 at the 2.5% level and 0.5 at the 5% level. The upper bound of $\overline{\beta}^i$ is similar to that found with the bootstrap inference at the 97.5% level.

From a methodological point of view, we stress that econometric theory is still in flux with respect to set identification and our analysis is based on techniques engineered for similar environments.[44]

6 Welfare

Having established the importance of network externalities in individual preferences, a natural next step is identifying the socially optimal assignment. In this section, we illustrate techniques for doing so.

The analysis in the previous section allowed us to identify the network layers and the proximity notion that impact individuals' utilities the most. We identify a pairwise-stable assignment that, under the estimated preferences, would increase overall efficiency relative to the one instated and, thus, we provide a lower bound on the welfare loss generated by the observed assignment.

Given the utility specification introduced in the previous section, an efficient assignment μ^* must satisfy (using our previous notation)

$$\begin{aligned} \mu^* \in \arg\max_{\mu\in\Theta} \textstyle\sum_{f\in\mathcal{F}} u_f(\mu) &= \arg\max_{\mu\in\Theta} \textstyle\sum_{f\in\mathcal{F}} \left[P_{\mu(f)} + \sum_{l=1}^{L} \beta_l k(f,\mu,l)\right] \\ &= \arg\max_{\mu\in\Theta} \textstyle\sum_{f\in\mathcal{F}} \sum_{l=1}^{L} \beta_l k(f,\mu,l), \end{aligned}$$

where Θ is the set of all possible assignments, and the second equality follows from the fact that the utility derived from offices' physical characteristics is homogeneous across agents.

In light of the results of the preference estimation in Section 5, we now assume that the only links that matter to agents are those between departmental colleagues and coauthors in local neighborhoods. Following the results of specification PS(ii), we let the relative coefficient for coauthor with respect to department vary between 1 and 5, the estimated confidence interval.

Even with these simplifying assumptions, the problem of finding the most efficient assignment is still not trivial: there are $|\Theta| = 73! > 10^{105}$ possible assignments and the problem

[44] For current econometric results on partial identification see Imbens and Manski (2004), Chernozhukov, Hong, and Tamer (2007), and Stoye (2009), as well as references therein.

is inherently combinatorial, so differential techniques cannot be employed.[45] The problem fits into the class of *Quadratic Assignments Problems* (QAP henceforth), first described in Koopmans and Beckmann (1957) in the context of locating industrial plants with spillovers and transportation costs. Specifically, let $b_{ff'}$ (a generic entry in a symmetric $(N \times N)$-matrix B) be the overall intensity of network externalities between any two agents f and f' (obtained summing up the links present in the different networks weighted by the estimated relative importance of each network).[46] Also, let $h_{oo'}$ (a generic entry in a symmetric $(N \times N)$-matrix H) be a variable that takes value 1 if an office o neighbors an office o', and 0 otherwise. Finally, an *assignment matrix* X is an $(N \times N)$-matrix where each row and column is made up of $N-1$ zeros and a single 1, with element x_{fo} corresponding to whether agent f is assigned office o; the row and column restrictions guarantee a single assignment for each faculty member, and one faculty member to each office. Let Π be the set of assignment matrices:

$$\Pi = \left\{X \in \{0,1\}^{N \times N} : X\iota = X^T\iota = \iota\right\}$$

with ι the $N \times 1$ vector of ones. The problem of finding the most efficient assignment μ^* can therefore be specified in the Koopmans-Beckmann formulation as:[47]

$$\max_X \sum_{f=1}^{N}\sum_{f'=1}^{N}\sum_{o=1}^{N}\sum_{o'=1}^{N} b_{ff'}h_{oo'}x_{fo}x_{f'o'}$$

$$s.t. \text{ for all } f,o \in \{1,..,N\},\ x_{fo} \in \{0,1\}, \sum_{f'=1}^{N} x_{f'o} = 1, \text{ and } \sum_{o'=1}^{N} x_{fo'} = 1,$$

where the constraints simply assure that X is an assignment matrix, that is $X \in \Pi$.

As detailed in Çela (1998), the formulation above can be respecified as follows:

Problem 1 (Quadratic Assignment Problem)**.** *Find the matrix* $X \in \Pi$ *that maximizes* $\mathrm{tr}\left\{BXHX^T\right\}$*, where* $\mathrm{tr}\{\cdot\}$ *is the trace operator.* [48]

[45] We note that the subtlety of the network architectures make this a more intricate problem than others pertaining to efficient design in the presence of complementarities, such as, say, FCC spectrum allocations.

[46] Note that this formulation takes bilateral links as the de-facto networks. The values appearing in the matrix B essentially identify the coefficients $\{\beta_l\}$ in our utility specification, the utility flows between members in the (bilateral) network.

[47] The generalized formulation of QAP allows for the arbitrary term c_{ijlm} in place of $b_{ij}h_{lm}$.

[48] Note that the matrix $Z = BXHX^T$ has generic element $z_{f\tilde{f}} = \sum_{f'=1}^{N}\sum_{o'=1}^{N}\sum_{o=1}^{N} b_{ff'}x_{f'o'}h_{o'o}x_{f'o}$.

The QAP problem has been shown to be NP-hard; in fact, even the problem of finding an ε-approximation is computationally complex.[49] Full solutions to this class are still considered numerically intractable for $N > 30$. Thus, we begin our welfare analysis by specifying a slack upper bound for the problem using the properties of the trace and the eigenvalues of B and H. Given the ordered eigenvalues of H and B—($\upsilon_1 \leq \cdots \leq \upsilon_N$) and ($\rho_1 \leq \cdots \leq \rho_N$), respectively—we can give a simple upper bound on welfare since we have

$$\operatorname{tr}\left\{BXHX^T\right\} \leq \sum_{i=1}^{N} \upsilon_i \rho_i \tag{4}$$

for any $X \in \Pi$.[50] We build the matrix B from the relevant network structures (with preference weights derived from our pairwise-stability estimates as discussed above), and the matrix H from the office building's layout, considering two offices neighbors consistently with the neighborhood definition specified in our previous analysis. We find that as the weight given to a coauthor neighbor (relative to a departmental neighbor) varies from 1 to 5, *the upper bound defined in (4) exhibits a gain of 278% to 466% over the observed assignment's efficiency.*

Next, we assess a lower bound for the potential efficiency gain. We do so by identifying an alternative pairwise-stable assignment using an ant-colony algorithm (see Dorigo, 1992 and Dorigo, Di Caro, and Gambardella, 1999), which has been demonstrated to be effective in finding global optima within the QAP.[51] This algorithm, which is similar in many ways to simulated annealing, imitates the way ants lay pheromone trails when finding food, guiding the colony as a whole to food sources. The probability of an ant locating a faculty member in a particular office is determined by the pheromone levels for that assignment—that is, the

---

Thus, using the symmetry of H and changing summation orders, the trace is equal to $\sum_{f=1}^{N}\sum_{f'=1}^{N}\sum_{o=1}^{N}\sum_{o'=1}^{N} b_{ff'}h_{oo'}x_{fo}x_{f'o'}$, so the objectives in the two formulations are identical. In fact, the equivalence with Problem 1 is more generally true (see references in Çela, 1998).

[49] The proof that the problem is NP-Hard can be seen by reinterpreting the locations as a time sequence of visits to differing cities, with b_{ij} representing the distance between a city pair, and h_{lm} assuming value 1 if l and m are sequential time-periods. This reinterpretation gives the fairly well-known NP-hard Travelling Salesman's Problem. The complexity of the approximation is demonstrated in Sahni and Gonzalez (1976). They show that if the ε-approximation can be found in polynomial time, then $P = NP$.

[50] For a proof of the inequality in (4), see Hadley, Rendl, and Wolkowicz (1992).

[51] The efficiency gain we identify represents a lower bound on the potential efficiency gain only because, as is common in this literature, the most efficient assignment we identify represents a local maximum, but we are unable to tell whether it is also a global one.

number of times other ants have placed the faculty member in this location weighted by the global outcome from the other ants' assignments. We use a colony of 20 ants, and simulate the process for 50 periods, combining at every point the probabilistic location through the ants with direct local-search methods.

Recall that the coefficient corresponding to coauthorship links with respect to departmental ones was estimated between 1 and 5,with a midpoint of 3. Figure 3 illustrates the most efficient assignment found by the algorithm when coauthorship is valued 3 times as much as departmental affiliation.[52] The comparison with the observed assignment in Figure 2 suggests the potential for more network clustering in this matching process. In fact, given an equal effect of coauthor and departmental links, *the algorithm finds a lower bound on the potential efficiency gain of* 181%. *When coauthor links have five times the impact of departmental links, this lower bound is evaluated at* 213%.[53]

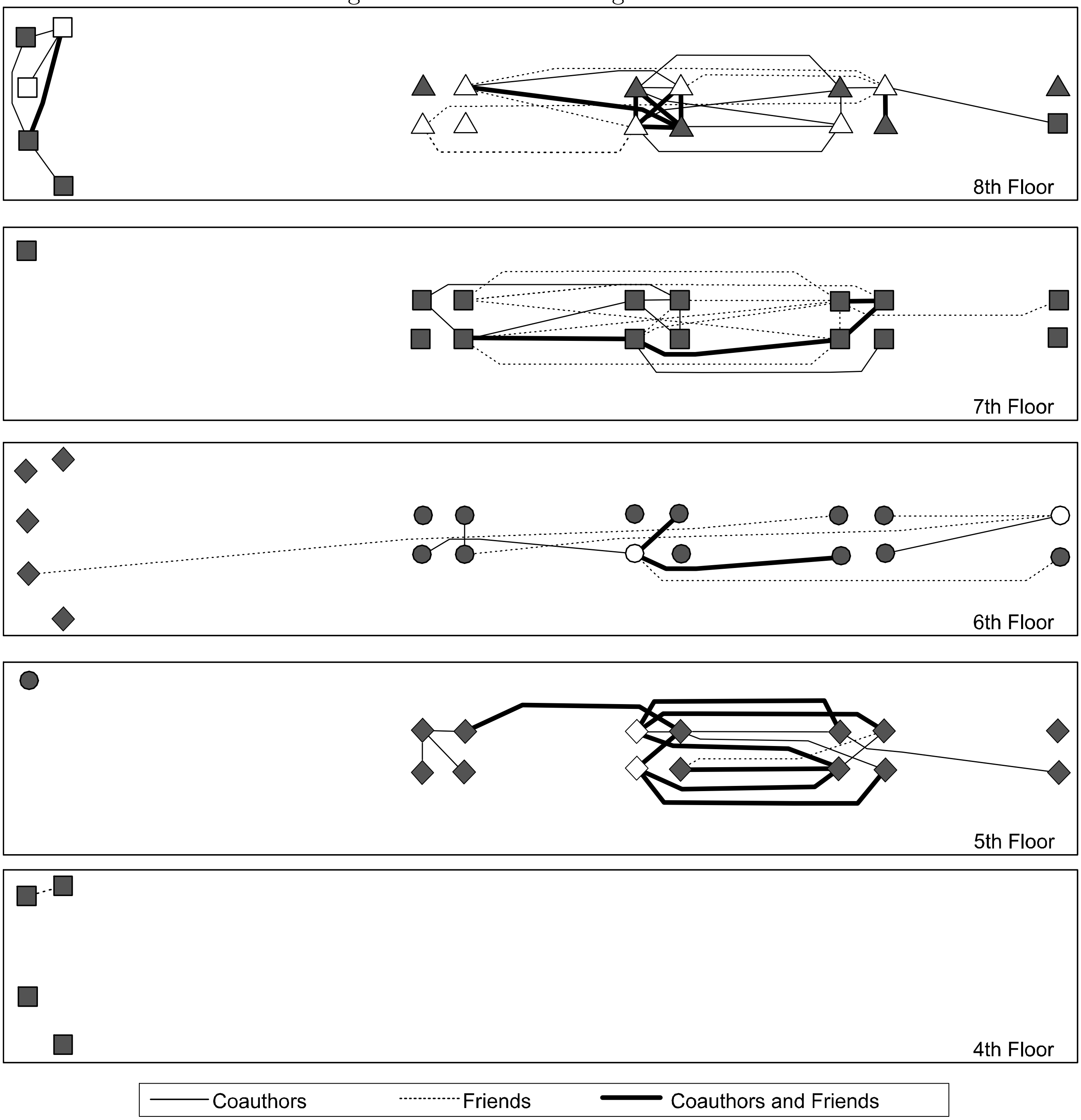
Given these results, we conclude that the equilibrium selected by the mechanism in place appears inefficient. However, we stress that the limitations of our data set forced us to assume homogeneous preferences across faculty. The efficiency of the observed assignment could in principle improve with respect to our estimates if we allowed different individuals to put different weights on network externalities (e.g., if faculty members that care less about externalities tend to be more senior, they may sort themselves toward better offices at the cost of less network links, allowing others to generate more high-efficiency links among themselves). Nonetheless, in light of the significant distance between the welfare performance of the observed assignment and the simulated one, one needs to conjecture a dramatic heterogeneity in preferences in order to commend the performance of the procedure in place.

The pragmatic message of this section is that ex-ante knowledge of the underlying network architecture appears to be crucial in generating a high efficiency match. Furthermore, computational methods from the QAP literature may be useful for identifying efficient assignments.

[52] The most efficient assignment is not unique. Indeed, notice that floors 5 and 7 are interchangeable as are floors 6 and 8.

[53] We note that despite friendships not appearing in the utility specification here, the number of friendship links in local neighborhoods improves from 18 in the observed assignment to 29 in the simulated one.

Figure 3: Best Found Assignment



Differing shapes represent department affiliation; containing rectangles floors; shading represents sex with white female, gray male.

7 Conclusion

We document a unique assignment protocol of faculty to offices in which locations (offices) varied in physical characteristics. We elicited three layers of network connections: institutional and choice-based (coauthorship and friendship). Our data allow us to study the role of network externalities.

Three main insights stand out. First, network externalities have a crucial impact on behavior and final outcomes in the assignment process. Second, the different network layers have unequal impacts on outcomes. Third, from a normative perspective, identifying the relevant networks is important for the design of efficient assignments.

From a methodological point of view, our study suggests the usefulness of a modified notion of stability for the estimation of network externalities in assignment processes. The paper also contributes to the empirical literature regarding social networks. Namely, we show how to account for the relative impact of different layers of peer connections. We also point out techniques that can be employed to evaluate the welfare performance of assignments in the presence of externalities.

Ultimately, this paper highlights the conceptual significance and empirical feasibility of considering network externalities in matching setups.

8 References

Abdulakadiroğlu A., P. Pathak, A.E. Roth, T. Sönmez, 2006, "Changing the Boston School Choice Mechanism: Strategy-proofness as Equal Access," NBER Working Paper 11965.

Bajari, P. and J. T. Fox, 2009, "Measuring the Efficiency of an FCC Spectrum Auction," Mimeo.

Bloch F., 1996, "Sequential Formation of Coalitions in Games with Externalities and Fixed Payoff Division," *Games and Economic Behavior*, 14, 90–123.

Brandeau M. L. and S. S. Chio, 1989, "An Overview of Representative Problems in Location Research," *Management Science*, 35(6), 645–674.

Budish E. and E. Cantillon, 2009, "Strategic Behavior in Multi-Unit Assignment Problems: Lessons for Market Design," Mimeo.

Çela E., 1998, *The Quadratic Assignment Problem*, Kluwer Academic Publishers.

Che, Y. K. and I. Gale, 2009, "Market versus Non-Market Assignment of Ownership," Mimeo.

Chernozhukov, V. and H. Hong and E. Tamer, 2007, "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75(5), 1243–1284.

De Clippel G. and R. Serrano, 2008, "Marginal Contributions and Externalities in the Value," *Econometrica*, 76(6), 1413–1436.

Dorigo M., 1992, "Optimization, Learning and Natural Algorithms," PhD thesis, Politecnico di Milano, Italy.

Dorigo M., **G. Di Caro and L. M. Gambardella, 1999,** "Ant Algorithms for Discrete Optimization," *Artificial Life*, 5 (2): 137–172.

Echenique, F. and M. Bumin Yenmez, 2007, "A Solution to Matching with Preferences over Colleagues," *Games and Economic Behavior*, 59(1), 46–71.

Ehlers, Lars, 2002, "Coalitional Strategy-Proof House Allocation," *Journal of Economic Theory*, 105, 298–317.

Ellison, G. and E. L. Glaeser, 1997, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 105(5), 889–927.

Fox, J. T., 2009, "Estimating Matching Games with Transfers," Mimeo.

Gilboa, I. and E, Lehrer, 1991, "Global Games," *International Journal of Game Theory*, 20, 129–147.

Goeree, J. K., M. A. McConnell, T. Mitchell, T. Tromp, and L. Yariv, 2009, "The 1/d Law of Giving," *American Economic Journal: Microeconomics*, forthcoming.

Guimarães P., O. Figueiredo, and D. Woodward, 2009, "Dartboard Tests for the Location Quotient," *Regional Science and Urban Economics*, 39(3), 360–364.

Hadley, S. W. , F. Rendl and H. Wolkowicz, 1992, "A New Lower Bound via Projection for the Quadratic Assignment Problem," *Mathematics of Operations Research*, 17(3), 727–739.

Hafalir, I. E., 2008, "Stability of Marriage with Externalities," *International Journal of Game Theory*, 37, 353–370.

Head, K. and J. Ries, 2005, "Judging Japan's FDI: The Verdict from a Dartboard Model," *Journal of the Japanese and International Economies*, 19(2), 215–232.

Horowitz, J. L., 1992, "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica*, 60(3), 505–531.

Imbens, G.W. and C.F. Manski, 2004, "Confidence Intervals for Partially Identified Parameters," *Econometrica*, 72(6), 1845–1857.

Jackson, M. O., 2008, *Social and Economic Networks*, Princeton University Press.

Koopmans, T. C. and M. Beckmann, 1957, "Assignment Problems and the Location of Economic Activities," *Econometrica*, 25(1), 53–76.

Krishna, A. and M. U. Ünver, 2008, "Improving the Efficiency of Course Bidding at Business Schools: Field and Laboratory Studies," *Marketing Science*, 27(2), 262–282.

Leider, S., M. Möbius, T. Rosenblat, and Q. A. Do., 2009, "What Do We Expect From Our Friends?," *Journal of the European Economic Association*, forthcoming.

Liola E. M, M. Nair, M. de Abreu, P. O. Boaventura-Netto, P. Hahn, and T. Querido, 2007, "A Survey for the Quadratic Assignment Problem," *European Journal of Operational Research*, 176(2), 657–690.

Lucas, W. F. and R. M. Thrall, 1963, "n-Person Games in Partition Function Form," *Naval Research Logistics Quarterly*, 10, 281–298.

Manea M., 2007, "Serial Dictatorship and Pareto Optimality," *Games and Economic Behavior*, 61(2), 316–330.

Manski C. F., 1975, “Maximum Score Estimation of the Stochastic Utility Model of Choice,” *Journal of Econometrics*, 3, 205–228.

Maskin E., 2003, “Bargaining, Coalitions, and Externalities,” *Presidential Address to the Econometric Society.*

Myerson R., 1977, “Values for Games in Partition Function Form,” *International Journal of Game Theory*, 6, 23–31.

Niederle M. and L. Yariv, 2009, “Decentralized Matching with Aligned Preferences,” NBER Working Paper Number 14840.

Politis D. N. and J. P. Romano, 1994, “Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions,” *The Annals of Statistics*, 22(4), 2031–2050.

Pycia M, 2009, “Payoff Rules in Coalition Formation and a Tractable Model of Matching with Complementarities and Peer Effects,” Mimeo.

Pycia M. and M. U. Ünver, 2007, “A Theory of House Allocation and Exchange Mechanisms,” Mimeo.

Ray D., and R. Vohra, 1999, “A Theory of Endogenous Coalition Structures,” *Games and Economic Behavior*, 26, 286–336.

Sahni S. and T. Gonzalez, 1976, “P-complete Approximation Problems,” *Journal of the Association for Computing Machinery,* 23(3), 555–565.

Santiago, D. and J. T. Fox, 2008, “A Toolkit for Matching Maximum Score Estimation and Point and Set Identified Subsampling Inference: Software Documentation,” Mimeo.

Sasaki, H. and M. Toda, 1996, “Two-Sided Matching Problems with Externalities,” *Journal of Economic Theory*, 70(1), 93–108.

Shapley L. S. and H. Scarf, 1974, “On Cores and Indivisibility,” *Journal of Mathematical Economics*, 1(1), 23–28.

Sönmez, T. and M. U. Ünver, 2009, “Course Bidding at Business Schools,” *International Economic Review*, forthcoming.

Stoye, J, 2009, “More on Confidence Intervals for Partially Identified Parameters,” *Econometrica*, 77(4), 1299–1315.

Subbotin, V., 2008, "Asymptotic and Bootstrap Properties of Rank Regressions," Mimeo.

Wasserman, S. and K. Faust, 1994, *Social Network Analysis: Methods and Applications*, Cambridge University Press.

Yi, S., 1997, "Stable Coalition Structures with Externalities," *Games and Economic Behavior*, 20(2), 201–237.

Appendix A: Faculty Survey

Hello and thank you for responding!

Your survey responses will be strictly confidential and data from this research will be reported only in the aggregate. Your information will be coded and will remain confidential. Please start with the survey now by clicking on the Continue button below.

1. Your name:

   Department:

2. How many days a week do usually come into your office?

Teaching period	1(2.63%)	2(2.63%)	3(13.16%)	4(39.47%)	5(23.68%)	>5(18.42%)
Non-teaching period	1(2.70%)	2(18.92%)	3(10.81%)	4(32.43%)	5(27.03%)	>5(8.11%)

3. How many hours (on average) do you spend at the office?

Teaching period	<2(0%)	$2-5$(5.26%)	$5-8$(39.47%)	>8(55.26%)
Non-teaching period	<2(0%)	$2-5$(2.63%)	$5-8$(60.53%)	>8(36.84%)

4. In a typical week, which days of the week to do you come into your office?

Monday	(18.24%)
Tuesday	(19.50%)
Wednesday	(17.61%)
Thursday	(18.87%)
Friday	(22.01%)
Weekend	(3.77%)

5. Do you try to come to the office when your office neighbors are around?

Yes, I try to coordinate	(31.58%)
I do not think about it	(68.42%)
No, I try to arrive when they are not there	(0%)

6. Please name up to 5 people you have lunch with on a regular basis and specify the number of times in a typical week that you have lunch with each of these.

7. Please name up to 5 of your most recent coauthors within the business school and the year in which you have last worked together.

8. Please name up to 5 personal friends (people with whom you interact socially with outside school at least once a month) from within the business school.

9. Please name up to 5 colleagues that would be valuable for you to have on your floor.

10. On a scale of 1-10, how important to you are the floor (4-8), exposure (east, west, or south), and size (corner office or standard office) for the quality of an office (where 1 is least important and 10 is most important)?

Floor	1(10.81%)	2(5.41%)	3(8.11%)	4(5.41%)	5(10.81%)	6(5.41%)	7(13.51%)	8(8.11%)	9(10.81%)	10(21.62%)
Exposure	1(8.11%)	2(8.11%)	3(5.41%)	4(8.11%)	5(18.92%)	6(8.11%)	7(8.11%)	8(16.22%)	9(8.11%)	10(10.81%)
Size	1(10.81%)	2(5.41%)	3(2.70%)	4(8.11%)	5(18.92%)	6(10.81%)	7(5.41%)	8(13.51%)	9(10.81%)	10(13.51%)

11. For a particular exposure and size of office, please rank the floors from 1-5 (where 1 would be your most preferred floor and 5 would be your least preferred floor).

Floor 4	1(14.29%)	2(0.00%)	3(5.71%)	4(0.00%)	5(80.56%)
Floor 5	1(0.00%)	2(14.29%)	3(11.43%)	4(71.43%)	5(2.78%)
Floor 6	1(8.57%)	2(11.43%)	3(74.29%)	4(5.71%)	5(0.00%)
Floor 7	1(8.57%)	2(68.57%)	3(2.86%)	4(17.14%)	5(2.78%)
Floor 8	1(68.57%)	2(5.71%)	3(5.71%)	4(5.71%)	5(13.89%)

12. On a scale of 1-10, what was the importance of your office neighbors to you prior to moving (where 1 is least important and 10 is most important)?

1(10.81%) 2(0.00%) 3(5.41%) 4(8.11%) 5(2.70%) 6(10.81%) 7(5.41%) 8(8.11%) 9(24.32%) 10(24.32%)

13. If you are part of a particular research cluster within your department, please identify it.

14. On a scale of 1-10, how important is it for you to be on the same floor with members of your own department and research cluster (where 1 is least important and 10 is most important)?

Department	1(5.41%)	2(2.70%)	3(2.70%)	4(2.70%)	5(5.41%)	6(10.81%)	7(13.51%)	8(16.22%)	9(16.22%)	10(24.32%)
Research Cluster	1(8.57%)	2(0.00%)	3(5.71%)	4(0.00%)	5(2.86%)	6(5.71%)	7(8.57%)	8(8.57%)	9(14.29%)	10(45.71%)

15. On a scale of 1-10, how important is it for you to be a direct neighbor, that is, sit in an adjacent office to, or across the hallway from members of your own department and research cluster (where 1 is least important and 10 is most important)?

Department	1(16.22%)	2(2.70%)	3(10.81%)	4(5.41%)	5(13.51%)	6(5.41%)	7(10.81%)	8(10.81%)	9(16.22%)	10(8.11%)
Research Cluster	1(8.57%)	2(2.86%)	3(14.29%)	4(2.86%)	5(8.57%)	6(2.86%)	7(8.57%)	8(17.14%)	9(17.14%)	10(17.14%)

16. At the time of your selection, how likely did you think you were to switch offices (where 1 corresponds to no switch and 10 corresponds to sure switch)?

    1(22.22%) 2(11.11%) 3(22.22%) 4(8.33%) 5(5.56%) 6(8.33%) 7(2.78%) 8(8.33%) 9(2.78%) 10(8.33%)

17. To what extent was your initial selection of a new office influenced by the possibility of ex-post trade, that is, by how desirable the office would be for others (where 1 corresponds to unimportant and 10 corresponds to very important)?

    1(36.11%) 2(5.56%) 3(13.89%) 4(11.11%) 5(5.56%) 6(11.11%) 7(2.78%) 8(5.56%) 9(5.56%) 10(2.78%)

18. Did you exhaust your research account this past year?

    Yes (61.76%)
    No (38.24%)

19. Did you exchange offices with anyone using your research account?

Yes	(11.43%)
No	(77.14%)
Tried but failed	(11.43%)

20. Did you exchange offices with anyone without using your research account?

Yes	(6.06%)
No	(81.82%)
Tried but failed	(12.12%)

21. We would appreciate it greatly if you could describe to us how you would made your decision of office in the space below.

22. Suppose that an additional office were made available and auctioned off in the business school. Please specify your 3 top choices for the location of that office (in terms of floor - 4 through 8 and exposure - east, west, or south) and the maximal bid you would be willing to pay out of your research account in order to move from your current allocated office to the new available one. Thus, if you specify an

amount X for any particular office, and all other bids fall below that, you would move to that office and pay X out of your research account. If any other bid surpasses X, you would stay in your current office. If several other colleagues would specify precisely the same X, we would randomly select one of you and exchange their office for X out of their research account.

Appendix B

Proof of Proposition 1

Since the number of possible assignments is finite, there exists a most efficient one. Let μ be the most utilitarian efficient assignment. We now show that μ is pairwise stable. Indeed, suppose that faculty f and f' form a blocking pair. That is, there exists some transfer t such that:

$$\begin{aligned} P_{\mu_f^{f'}(f)} + \sum_{l=1}^{L} \beta_l k(f, \mu_f^{f'}, l) + t &\geq P_{\mu(f)} + \sum_{l=1}^{L} \beta_l k(f, \mu, l); \\ P_{\mu_f^{f'}(f')} + \sum_{l=1}^{L} \beta_l k(f', \mu_f^{f'}, l) - t &\geq P_{\mu(f')} + \sum_{l=1}^{L} \beta_l k(f', \mu, l) \end{aligned}$$

and at least one of inequalities is strict. Since $\mu_f^{f'}(f) = \mu(f')$ and $\mu_f^{f'}(f') = \mu(f)$, summing up these inequalities leads to:

$$\sum_{l=1}^{L} \beta_l k(f, \mu_f^{f'}, l) + \sum_{l=1}^{L} \beta_l k(f', \mu_f^{f'}, l) > \sum_{l=1}^{L} \beta_l k(f, \mu, l) + \sum_{l=1}^{L} \beta_l k(f', \mu, l).$$

In particular,

$$\delta \equiv \sum_{l=1}^{L} \beta_l k(f, \mu_f^{f'}, l) + \sum_{l=1}^{L} \beta_l k(f', \mu_f^{f'}, l) - \sum_{l=1}^{L} \beta_l k(f, \mu, l) - \sum_{l=1}^{L} \beta_l k(f', \mu, l) > 0. \tag{5}$$

Note that δ is affected through network layers that encompass different numbers of faculty on each location following the switch. That is, suppose that under μ, in office $\mu(f)$, faculty f has $k_1, ...k_L$ connected faculty from layers $1, ..., L$ (including themselves), and, in office $\mu(f')$, faculty f has $r_1, ..., r_L$ faculty from layers $1, ..., L$. Then a shift to office $\mu(f')$ would correspond to an increase by 1 of all the relevant networks: $r_1 + 1, ..., r_L + 1$. This would render office $\mu(f)$ with $k_1 - 1, ..., k_L - 1$ faculty from each layer. Similarly, we will assume that under μ in office $\mu(f')$, faculty f' has $k'_1, ...k'_L$ faculty from layers $1, ..., L$ (including themselves) and in office $\mu(f)$, faculty f' has $r'_1, ..., r'_L$ faculty from layers $1, ..., L$. Condition (5) can then be written as:

$$\delta = \sum_{l=1}^{L} \beta_l (r_l + 1) + \sum_{l=1}^{L} \beta_l (r'_l + 1) - \sum_{l=1}^{L} \beta_l k_l - \sum_{l=1}^{L} \beta_l k'_l > 0, \text{ or}$$

$$\delta = \sum_{l=1}^{L} \beta_l \left(r_l + r'_l - k_l - k'_l + 2 \right) > 0 \tag{6}$$

Consider now the effects on utilitarian efficiency a shift from μ to $\mu_f^{f'}$ has. δ captures the sum of utilities gained for faculty f and f'. For each layer l, we need to take into account the changes (an addition or reduction of one peer for the shift of f and the shift of f') for all members of that network layer other than f and f'. For instance, r_l agents gain β_l utiles from the shift of f to $\mu(f')$, whereas $k_l - 1$ agents lose β_l utiles from this shift. Using (6), the overall efficiency gain is then

$$\Delta = \sum_{l=1}^{L} \beta_l \left(r_l + r'_l - (k_l - 1) - (k'_l - 1) \right) + \delta = 2\delta > 0$$

in contradiction to μ being the most efficient assignment. ■