Reading between the Emails: Gendered Patterns of Communication in Local Government

Matthew Denny [*], **James ben-Aaron** [†], **Hanna Wallach** [† ‡], **and Bruce Desmarais** [*]

[*]Penn State University,[†]University of Massachusetts Amherst, and [‡]Microsoft Research NYC

We conduct an analysis of the relationship between a department manager's gender, and the emails they send and receive, in a sample of seventeen North Carolina county governments. First, we examine the aggregate patterns of email communication among department managers in these counties, the gender breakdown of managers for each department, and interdepartmental email flows by gender. Then, we apply a statistical model for communication network data to infer the topic-conditional structure of manager-to-manager email communication in these county governments. Our results indicate that while aggregate patterns of communication among department managers do not display significant gender bias, the topics of communication, and positions held by men and women differ significantly. From a policy perspective, our findings suggest that reaching parity in the gender composition of positions in an organization may not be enough to create a gender-equitable working environment by itself, as there is significant variation in email communication by gender, beyond department-level effects.

Gender in Organizations

Gender bias against female members of firms and government organizations is well documented in terms of lower pay and prestige, reduced opportunity for advancement, and gender-specific social interaction [5, 3, 11, 2, 7]. At the same time, providing equal treatment to male and female coworkers is a key objective for organizations that aspire to a just, efficient, and sustainable professional culture [9]. Scholars have sought to understand the roots, extent, and nature of gender bias in organizations, but have only had limited access to primary-source data in previous studies. The data used in these studies [e.g., 6, 1, 8] has traditionally been observational, ethnographic, or self-reported, which can be biased and is often limited in scope.

However, with the increasing use of electronic communication, and the rise of transparency initiatives within government, scholars are now increasingly able to use public records requests to gather primary source email communication data. Using public records requests as a method for data collection also provides a major benefit of making the data collection process replicable. Therefore, we choose to study the relationship between a department manager's gender, and the emails they send and receive, in a sample of local governments. To do this, we collect and analyze a large scale email data set from a sample of seventeen county governments in North Carolina.

We begin by examining the aggregate patterns of email communication among department managers in these counties by gender, and the gender breakdown of managers for each department. Then we analyze interdepartmental email flows by gender of the email sender and its recipients. Finally, we apply a statistical model for communication network data to infer the topic-conditional structure of manager-to-manager email communication in these county governments. Our results indicate that while aggregate patterns of communication among department managers do not display significant gender bias, the topics of communication, and positions held by men and women differ significantly.

Data

The data used in this study were collected in 2013 via a series of public records requests made to all one hundred North
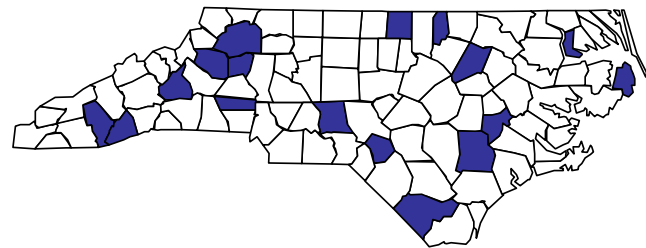

Fig. 1. North Carolina map with counties who provided useable data highlighted.

Table 1. Breakdown of department managers by gender, and the total number of emails sent by department managers, for each county in our sample.

| County | Manager Gender | | Emails Sent |
	Male	Female	
Alexander	12	9	907
Caldwell	12	8	121
Chowan	12	11	2,027
Columbus	14	10	920
Dare	15	12	2,247
Duplin	13	14	1,914
Hoke	13	11	1,106
Jackson	18	6	1,499
Lenoir	15	5	560
Lincoln	15	7	573
McDowell	12	5	326
Montgomery	8	10	680
Nash	11	8	1,147
Person	12	9	1,491
Transylvania	16	4	1,857
Vance	10	8	185
Wilkes	15	2	303
Totals:	362	139	17,863

Table 2. Email statistics by manager gender.

| | Manager Gender | |
	Male	Female
Average # of emails sent	48.3	51
Average # of recipients per email sent	1.45	1.43
Average # of emails received	70.8	71.6

[1]Some counties provided too little data (i.e. 30 emails for the entire county), or provided data in paper form, which was not amenable to our analyses.

[2]Requests to each county specified a random starting month for the request, between January and October, 2103.

Carolina county governments. We selected North Carolina for our data collection efforts because the state public records laws cover email data, and prevent counties from charging unreasonable fees for providing data. Twenty-three counties complied with our request, of which seventeen provided data was useable for our research[1]. Our records request to each county included all emails sent and received by all department managers (e.g. Health, Finance, Elections) over overlapping three-month periods in 2013[2].

We received approximately 500,000 emails from the seventeen county governments which provided useable data, of which approximately 17,000 were sent between county department managers, which we focus on in the analyses presented here. Figure 1 highlights the counties which responded to our requests for data. We also collected detailed metadata on the department and gender of the 362 department managers in our sample. We display some basic descriptive statistics for each county in table 1. Approximately 40% of department managers in our sample are women, although there is significant variation in the proportion of female department managers across counties. It is important to note that the counties in our sample are statistically indistinguishable from the rest of the counties in North Carolina on a number of demographic dimensions including population, per-capita income, and percent of the population that is white.

Descriptive Analysis

We begin our analysis of the relationship between gender and communication patterns in these county governments by look-ing for differences in aggregate email sending and receiving behavior, by gender, across our entire sample. Table 2 provides some basic descriptive statistics of the average number emails sent and received by male and female department managers in our sample. We can see that on average, male and female department managers send and receive a comparable number of emails, and that each email sent by male and female department managers has a similar number of recipients on average. These statistics tell us that if we observe gender differences in email communication, they are unlikely driven by some innate difference in the propensity for male and female department managers to send or receive emails.
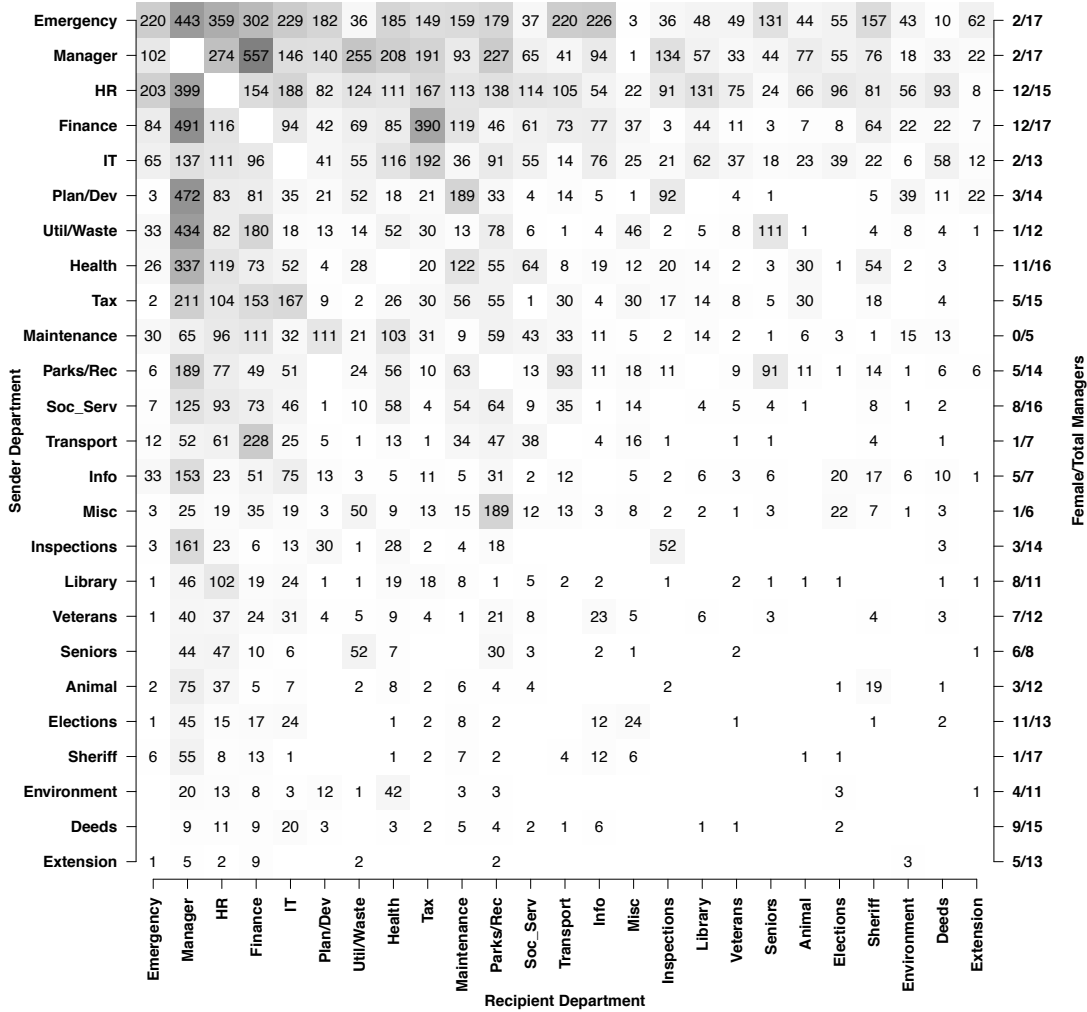
We can now disaggregate email sending and receiving by gender to see if there is a relationship between the gender of the sender of an email and the gender of the recipients of that email. If we do see a relationship, this would indicate that there is gender bias in communication within these organizations. To test whether the gender of an email sender is related to the gender of its recipients, in aggregate, we construct a contingency table of the gender of the sender of an email against the gender of the recipients of that email (see table 3). We then perform a χ^2 test for independence between the rows and columns.

We conduct a χ^2 test on this contingency table and the test statistic is $\chi^2 = 92.9$ with a p-value < 0.00001, indicating that the gender of an email sender and its recipients is not independent. While this result provides evidence for gender bias in communication in these county governments, our findings could potentially be confounded by gender differences in the departments men and women manage. Fortunately, our data includes the department managed by each manager, allowing us to disaggregate our analysis by department. We begin by looking for aggregate differences in the gender composition of departments (see table 4). Departments were hand coded into one of 25 different categories based on the title given the county directory, to group departments that perform a similar function. For example the

We can see that the managers of some department are mostly men: the county manager (the supervisor of all other department managers), the sheriff, and the emergency manager, for example. Other departments are mostly managed by women – the HR, finance, and health departments, for example. A a χ^2 test for the independence of department and gender confirms our qualitative finding that there is a strong bias in which departments tend to be managed by women ($\chi^2 = 130.8, p < 0.00001$). These results indicate a bias in the gender composition of different departments, but cannot tell us about differences in email communication by men and women who manage the same department. To answer this question, we want to know whether the gender of an email sender and the gender of its recipients is independent of the their respective departments.

Table 2 displays the aggregate number of times each department manager was the recipient of an email from the manager of each other department. We see pronounced differences in email sending an receiving by department, but cannot graphically disentangle the role of gender. To test for the independence of the gender of an email sender and the gender of its recipients from which department they direct, we construct a contingency table of department dyad-types (where each "dyad-type" is a unique combination of sender and recipient department, eg. finance $\longrightarrow$ HR) against gender dyad-types (where each "dyad-type" is a unique combination of sender and recipient gender, eg. male $\longrightarrow$ female). We then perform

Table 3. Each cell records the number of times a department manager of gender X was included as a recipient of an email sent by a department manager of gender Y. Statistics provided are calculated for all counties combined. Note that each email may have more than one recipient.

		Recipient		
		Male	Female	Total
Sender	Male	7,299	6,286	13,585
	Female	5,325	3,510	8,835
	Total	12,624	9,796	

Table 4. Number of male and female managers for each department. Note that not all departments are represented in each county.

	Emergency	Health	IT	Manager	HR	Library	Plan/Dev	Deeds	Parks/Rec	Finance	Soc_Serv	Veterans	Util/Waste
Male	15	5	11	15	3	3	11	6	9	5	8	5	11
Female	2	11	2	2	12	8	3	9	5	12	8	7	1

|	Elections	Sheriff	Info	Tax	Inspections	Animal	Maintenance	Seniors	Transport	Environment	Misc	Extension
Male	2	16	2	10	11	9	5	2	6	7	5	8
Female	11	1	5	5	3	3	0	6	1	4	1	5

Fig. 2. Heat map depicting the total number of emails sent from the managers of each department to the managers of each other department, aggregated across counties. The right-hand x-axis displays the number of counties that had that department, and the number of those managers who were women. Departments are ordered by the total number emails their managers sent.

Sender \ Recipient	Emerg	Mgr	HR	Fin	IT	Pl/Dev	Util/W	Health	Tax	Maint	Parks	SocSrv	Trans	Info	Misc	Insp	Lib	Vet	Sen	Anim	Elec	Sher	Env	Deeds	Ext	Female/Total Mgrs
Emergency	220	443	359	302	229	182	36	185	149	159	179	37	220	226	3	36	48	49	131	44	55	157	43	10	62	2/17
Manager	102		274	557	146	140	255	208	191	93	227	65	41	94	1	134	57	33	44	77	55	76	18	33	22	2/17
HR	203	399		154	188	82	124	111	167	113	138	114	105	54	22	91	131	75	24	66	96	81	56	93	8	12/15
Finance	84	491	116		94	42	69	85	390	119	46	61	73	77	37	3	44	11	3	7	8	64	22	22	7	12/17
IT	65	137	111	96		41	55	116	192	36	91	55	14	76	25	21	62	37	18	23	39	22	6	58	12	2/13
Plan/Dev	3	472	83	81	35		21	52	18	21	189	33	4	14	5	1	92		4	1		5	39	11	22	3/14
Util/Waste	33	434	82	180	18	13		14	52	30	13	78	6	1	4	46	2	5	8	111	1	4	8	4	1	1/12
Health	26	337	119	73	52	4	28		20	122	55	64	8	19	12	20	14	2	3	30	1	54	2	3		11/16
Tax	2	211	104	153	167	9	2	26		30	56	55	1	30	4	30	17	14	8	5	30		18		4	5/15
Maintenance	30	65	96	111	32	111	21	103	31		9	59	43	33	11	5	2	14	2	1	6	3	1	15	13	0/5
Parks/Rec	6	189	77	49	51		24	56	10	63		13	93	11	18	11		9	91	11	1	14	1	6	6	5/14
Soc_Serv	7	125	93	73	46	1	10	58	4	54	64		9	35	1	14		4	5	4	1		8	1	2	8/16
Transport	12	52	61	228	25	5	1	13	1	34	47	38		4	16	1		1		1		4		1		1/7
Info	33	153	23	51	75	13	3	5	11	5	31	2	12		5	2	6	3	6		20	17	6	10	1	5/7
Misc	3	25	19	35	19	3	50	9	13	15	189	12	13	3		8	2	2	1	3		22	7	1	3	1/6
Inspections	3	161	23	6	13	30	1	28	2	4	18									52			3			3/14
Library	1	46	102	19	24	1	1	19	18	8	1	5	2	2		1		2	1	1	1		1	1		8/11
Veterans	1	40	37	24	31	4	5	9	4	1	21	8		23	5		6		3			4		3		7/12
Seniors		44	47	10	6		52	7			30	3		2	1		2								1	6/8
Animal	2	75	37	5	7		2	8	2	6	4	4				2					1	19		1		3/12
Elections	1	45	15	17	24		1	2	8	2			12	24			1					1		2		11/13
Sheriff	6	55	8	13	1		1	2	7	2		4	12	6						1	1					1/17
Environment		20	13	8	3	12	1	42			3	3								3					1	4/11
Deeds		9	11	9	20	3		3	2	5	4	2	1	6		1	1				2					9/15
Extension	1	5	2	9		2					2											3				5/13

a χ^2 test for independence between the rows and columns. The test statistic $\chi^2 = 37,404$ with a simulated p-value[3] < 0.00002, indicates that the gender of an email sender and its recipients is not independent of the their respective departments.

This result indicates that there is gender bias in communication when we disaggregate to the department-dyad level, however, we are still stuck. This analysis does not allow us to disentangle the two potential sources of the pattern we find – bias in the positions women hold within these organizations, and bias in the way that male and female managers communicate. Furthermore, this analysis may be slicing our data too thin since not all counties have the same departments, and any particular department dyad may only exchange a handful of emails. Moreover, the department attribute of each manager is likely caused by several factors other than the topics of communication which may relate to gender bias.

One solution to the problems raised above is to model the email content, as general topics such as balancing budgets are commonly represented across counties. Practically, this also allows us to focus on a smaller number of communication topics that are shared across counties.
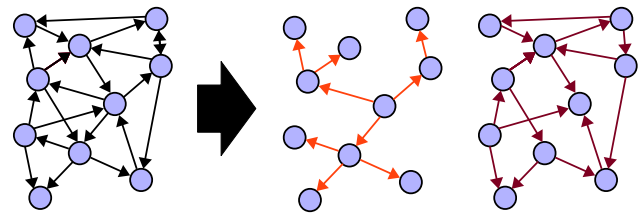
Fig. 3. Different patterns of communication across different domains.

To do this, we might want to use a statistical topic model to categorize emails, then model gender mixing using the LSM in each topic.

However, the gender mixing parameters we infer using this approach would be confounded by selection effects. As a toy example, we would not be able to tell whether men prefer to talk to male coworkers over female coworkers about football, or women prefer not to talk about football at work, and so they do not participate in those conversations.

To disentangle these effects, we need a joint model for email topical content, and the structure of communication.

A Model of Email Content

Here we provide a brief overview of our extension to the TPME model [12]. We focus on the ways in which our new model allows us to make inferences about the gender-specific patterns of communication in organizations, and how these patterns vary with the content of communication. Our model follows [12] by integrating the latent space network model [10] with a statistical topic model (latent Dirichlet allocation [4]) to jointly model the content and structure of a communication network.

At a high level, our model assumes the following generative process for a message sent across the communication network. First, we sample the content of the message following the generative process for latent Dirichlet allocation. Then, for a given message sender, we sample whether each other actor in the network is a receiver of that message following the latent space network model. However, different topics (or clusters of topics in our extension) are associated with different latent spaces, so how likely each actor is to be a receiver of a particular message is dependent on the topical content of that message. We review this process in greater detail below.

The message content is assumed to be generated via latent Dirichlet allocation. Under this model, we assume that all unique words in our vocabulary are associated to varying degrees with each of T latent "topics". Each topic is then a distribution over unique words (or "word types"), $\boldsymbol{\phi}^{(t)}$. For example "doctor", "virus", and "medicine" might all have high probability in a topic about hospitals, while "cat" and "music" might have low probability in that topic. We further assume that each topic distribution is drawn from a Dirichlet prior:

$$\boldsymbol{\phi}^{(t)} \sim \text{Dirichlet}(\beta, \boldsymbol{n}) \text{ for } t = \{1, ..., T\} \quad [\mathbf{1}]$$

The concentration parameter β controls how peaky this distribution is (how few word types will have high probability in a given topic), and the base measure $\boldsymbol{n}$ controls the bias in the degree to which word types will be assigned high probability in topics generally. For example, "patient" might generally have a higher probability than most word types in a corpus of documents about healthcare.

To generate each of the D documents (messages) in our corpus, we first draw a document-specific distribution over topics

$$\boldsymbol{\theta}^{(d)} \sim \text{Dirichlet}(\alpha, \boldsymbol{m}) \text{ for } d = \{1, ..., D\} \quad [\mathbf{2}]$$

Here, the concentration parameter α controls how topically specific the documents generated are, and $\boldsymbol{m}$ controls the bias in how likely certain topics are to appear. Once we have sampled our document-specific distribution over topics we can then generate each of N tokens (words) in the document via a two step process. First we draw the latent topic assignment for that token from the document specific distribution over topics.

$$z_n^{(d)} \sim \boldsymbol{\theta}^{(d)} \text{ for } n = \{1, ..., N\} \quad [\mathbf{3}]$$

Then we draw the word type of that token from the topic specific distribution over word types.

$$w_n^{(d)} \sim \boldsymbol{\phi}^{(z_n^{(d)})} \text{ for } n = \{1, ..., N\} \quad [\mathbf{4}]$$

We proceed in this manner for all tokens in all documents.

In our extension of TPME, we further assume each topic is uniquely associated with one of C clusters, sampled from a discrete uniform distribution.

$$C_t \sim \text{Discrete Uniform}(1, C) \text{ for } t = \{1, ..., T\} \quad [\mathbf{5}]$$

The intuition is that different broad content areas of communication (e.g. "planning", "sports","meeting planning") – each

of which is a collection of more specific topics – will imply a different pattern of communication. Importantly, under this model, a message can be about a number of topics, and these topics can be associated with different clusters. Each cluster is associated with a different pattern of communication, so the receivers for any particular message must be sampled following an add-mixture of several underlying communication patterns. For example, if a department manager in an organization sent an email message to schedule a budget meeting, they would likely include both staff whose job includes setting up meetings, and staff who needed to provide input on the budget as recipients.

Finally, for a given cluster c, the probability that an actor is selected as a recipient of a particular message is specified by the latent space network model [10]. Under this model, we assume there is some baseline propensity to include message recipients on any email, which is governed by an intercept parameter.

$$b^c \sim \text{Normal}(\mu, \tau^2) \quad [\mathbf{6}]$$

For example, messages about sensitive HR matters will probably include fewer recipients than messages announcing a department party. Second, some attributes of the sender and potential receiver (which we assume are observed) may also affect the probability of that actor being a recipient. The effect that each of these L different attributes have on the probability of an actor being included as a message recipient can then be parameterized by a vector:

$$\gamma_l^c \sim \text{M. V. Normal}(\lambda, \eta^2) \text{ for } l = \{1, ..., L\} \quad [\mathbf{7}]$$

For example, there is a great deal of social science literature showing that people tend to preferentially communicate with others of the same gender. Our model could generate this propensity by sampling positive parameters for Male-Male and Female-Female communication, and negative parameters for Female-Male and Male-Female communication.

The additional variation (not captured by covariate effects) in the probability that a message is sent between two actors is governed by how close they are in a k dimensional "latent social space". This latent space captures all un-modeled factors associated with the propensity for two actors to form a tie. In a typical social network this might include difficult or impossible to measure quantities like how nice a person is, or whether two people have "chemistry". Additionally, it may capture observable traits that the researcher was unable to collect data on (e.g. sexual orientation, or ethnicity). This makes the latent positions difficult to interpret when including covariates in the model, so great care should be taken when doing so. To capture these latent positions, each actor a is assumed to have some position in the k-dimensional latent space

$$\mathbf{s}_a^c \sim \text{M. V. Normal}(0, \sigma^2) \text{ for } a = \{1, ..., A\} \quad [\mathbf{8}]$$

and the probability that they send a message to an actor $r \neq a$ is decreasing in their latent distance from r. Thus the probability of actor r being a recipient of a message from actor a under this model is:

$$P(y_{a,r}^c = 1) = \text{logit}^{-1} \left(b^c + \sum_L \left[\mathbf{X}_{a,r}^l \gamma_l^c \right] - |\mathbf{s}_a^c - \mathbf{s}_r^c| \right) \quad [\mathbf{9}]$$

where $\mathbf{X}^l$ is a matrix recording the type of an edge for attribute l that would be sent between a and r (e.g. Male to Female or employee to supervisor). Edge values for an individual message are then sampled via Bernoulli trials. However, the probability that $y_{a,r} = 1$ for a given message d may be dependent on multiple latent spaces (because documents can be about multiple topics). The weight given to $P(y_{a,r}^{(c)})$ for

each cluster is determined by the proportion of tokens in the message that are assigned to topics associated with cluster c.

$$P(y_{a,r}^{(d)} = 1) = \sum_C P(y_{a,r}^c = 1) \times * \qquad [\mathbf{10}]$$

where $*$ is the proportion of tokens in document d assigned to topics associated with cluster c.

Given that we do not directly observe the generative process, the object of inference becomes the posterior distribution of our model parameters given the data. This problem is analytically intractable, so we must approximate the posterior distribution via Markov chain Monte Carlo methods. We perform inference for this model via block Metropolis within Gibbs sampling[4]. A further discussion of this model will be provided in a related paper, but this model provides a powerful and flexible framework to investigate the content-conditional gendered patterns of communication in an organization.

Inference. We have to invert the generative process to perform inference on the model parameters. We use block Metropolis Hastings within Gibbs sampling. A beta R package is available for those interested. Discuss our model specification and justify our hyper-parameter choices.

Gender Mixing

Since we have modeled content and structure together, we have broken up the confounding we were concerned about, and we can now interpret the gender mixing parameters that come out of our model with more confidence.

Lets return to the question we posed at the end of our descriptive analysis section: is there gender bias in the patterns of communication in these county governments.

Add in Bruce's MV KS test results for each county here.

Interpret the results.

Take as a case study, a county with the highest degree of gender bias, as identified by this method.

interpret the topic model output from two different clusters and discuss

References

1. Susan M Adams, Atul Gupta, Dominique M Haughton, and John D Leeth. Gender differences in ceo compensation: evidence from the usa. *Women in Management Review*, 22(3):208–224, 2007.
2. James Albrecht, Anders Björklund, and Susan Vroman. Is there a glass ceiling in Sweden? *Journal of Labor Economics*, 21(1):145–177, 2003.
3. William T. Bielby and James N. Baron. Men and Women at Work : Sex Segregation and Statistical Discrimination. *American Journal of Sociology*, 91(4):759–799, 1986.
4. DM Blei, AY Ng, and MI Jordan. Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3:993–1022, 2003.
5. Daniel J. Brass. Men's and Women's Networks: A Study of Interaction Patterns and Influence in an Organization. *Academy of Management Journal*, 28(2):327–343, 1985.
6. Emilio J Castilla. Gender, race, and meritocracy in organizational careers. In *Academy of Management Proceedings*, volume 2005, pages G1–G6. Academy of Management, 2005.
7. Colin Duncan and Wendy Loretto. Never the Right Age? Gender and Age-Based Discrimination in Employment. *Gender, Work & Organization*, 11(1):95–115, 2004.
8. Kim M Elsesser and Janet Lever. Does gender bias against female leaders persist? quantitative and qualitative data from a large-scale survey. *Human Relations*, 64(12):1555–1578, 2011.
9. Robin J Ely and Debra E Meyerson. Theories of gender in organizations: A new approach to organizational analysis and change. *Research in organizational behavior*, 22:103–151, 2000.
10. Peter D Hoff, Adrian E Raftery, and Mark S Handcock. Latent Space Approaches to Social Network Analysis. *Journal of the American Statistical Association*, 97(460):1090–1098, December 2002.
11. Herminia Ibarra. Homophily and Differential Returns : Sex Differences in Network Structure and Access in an Advertising Firm. *Administrative Sceince Quarterly*, 37:422–447, 1992.
12. Peter Krafft, Juston Moore, Bruce A Desmarais, and Hanna Wallach. Topic-partitioned multinetwork embeddings. In *Advances in Neural Information Processing Systems Twenty-Five*, 2012.

---

[4] The inference algorithm is currently implemented in a beta version as an R package, and is available here: github.com/matthewjdenny/ContentStructure