Online Appendices

Appendix A: Email Data Collection Methods

Email data provides advantages over survey self reports in that direct observation of communications traffic avoids biases in respondent recall of social networks (Marsden 1990, Reagans & McEvily 2003). Securing access to complete email communications required us to develop original software tools, coded specifically for this project. A more complete description is provided in Van Alstyne & Zhang (2003) and Reynolds, Van Alstyne & Aral (2009). Here we summarize our methods, and describe the problems we confronted while collecting email data and our approaches for resolving each as a means to ensure data completeness and validity. Study participants who allowed use of their email data received $100 in Amazon gift certificates in compensation. Email coverage of eligible employees exceeded 87%.

The code developed for email capture runs on a Microsoft Exchange™ server, under a standard email configuration. It downloads all messages from the exchange server (storing To, From, CC, BCC, Subject, Body, and Attachments) and provides secure backup and nightly export of encrypted data to servers outside the research host site. Due to the potential intrusiveness of email capture, the research underwent three rounds and nine months of human subjects review by the University of Michigan IRB and two rounds and five months of human subjects review at MIT prior to launch. We addressed several data completeness and validity issues as follows.

1. *Privacy*: Potential invasions of privacy raised two concerns for data validity. First, the human subjects boards required voluntary participation of all study participants. Voluntary non-participation might therefore have reduced the representativeness of our data sample if too many employees had chosen to opt out. Second, among remaining employees, privacy concerns could cause them to change their communication behaviors. To address these issues, we developed a data masking and encryption algorithm that permits analysis of header information and content but prevents reading of any individual message. The hashing algorithm and the information theoretic and statistical properties for granting privacy protection are not part of this article but are described in Van Alstyne & Zhang (2003), Reynolds, Van Alstyne & Aral (2009), Van Alstyne & Zhang (US Patent 7503070), and in a master's thesis we supervised at MIT (Farrokhzadi 2007). Privacy assurances, however, are a primary reason why more than 87% of employees chose to remain in the study.

2. *Data Bias*: Different people exhibit different patterns of behavior with respect to retaining and deleting email. This risks introducing data bias in the volume and frequency of communication, and in the shape of egocentric social networks. To prevent deletion bias, administrative switches in the Exchange email server and email capture code recorded properties of deleted messages for a period of 24 hours before details were expunged.

3. *Load Balancing*: Significant data processing – necessary for capturing data, privacy sanitization, and secure data export – introduced risk of increased load on the email server and decreased user responsiveness. Poorer system performance, in turn, could also have influenced user behaviors. To ensure compliance with all normal functions on the Microsoft Exchange email server, we tested our software at a Microsoft research facility in Redmond, WA and received a clean bill of health. This also provided assurance that data intended for capture was correctly recorded. To minimize interference at the research site, we captured daily increments of email communications, and exported sanitized encrypted data to university servers at 03:00 am local time at the research host, utilizing a lull in normal activity immediately after normal backup operations.

References:

Farrokhzadi, M. 2007. "Entropy, Information Rate and Mutual Information Measures for the Email Content of Information Workers" MEng Thesis – Electrical Engineering and Computer Science, May 25th, MIT.
Marsden, P. 1990. "Network Data & Measurement." Annual Review of Sociology (16): 435-463.
Reagans, R. & McEvily, B. 2003. "Network Structure & Knowledge Transfer: The Effects of Cohesion & Range." Administrative Science Quarterly, (48): 240-67.
Reynolds, M., Van Alstyne, M., Aral, S. 2009. "Privacy Preservation of Measurement Functions on Hashed Text" 8th Annual Security Conference, April 15-16, Las Vegas, NV.
Van Alstyne, M. & Zhang, J. 2003. "EmailNet: A System for Automatically Mining Social Networks from Organizational Email Communication," NAACSOS.
Van Alstyne, M. & J. Zhang. 2009. US Patent 7503070 "Methods and systems for enabling analysis of communication content while preserving confidentiality."

Appendix B. Robustness Checks

Table B1. Heckman Selection Model of Multitasking and Output

Dependent Variable		**Output**	
Model	1	2	3
Multitasking	.36**	.36**	.36**
	(.02)	(.02)	(.02)
Multitasking Squared	-.08**	-.08**	-.08**
	(.01)	(.01)	(.01)
Average Duration	-.15**	-.14**	-.14**
	(.02)	(.02)	(.02)
Rho	.13	.05	.01
Sigma	.21	.19	.19
Lambda	.03	.01	.002
Observations	630	450	450

Notes: This table reports results of a Heckman two-step selection model in which the first stage regression predicts multitasking as a function of several observable characteristics of recruiters and the regional economic climate in which they operate. In model 1, we estimated the first stage of the model by predicting the level of multitasking using observable characteristics of recruiters including age, gender, industry experience, firm tenure, position within the firm, as well as the primary city in which recruiters work. In model 2, we predicted multitasking using data on exogenous shocks to recruiters multitasking collected from the Bureau of Labor Statistics on the levels and growth of statewide employment and GDP in the states where the firm operates (both their offices and the cities in which they have clients) for sectors of relevance to their work (which include executive recruiting, professional services, education, and health care as well as overall levels and growth of statewide employment and GDP in these states). We weighted these variables by the number of projects a recruiter had in each state. In model 3, we predicted multitasking using all of these variables. Significance levels are as follows: **p<.001; *p<.05.

Table B2: Hazard Rate Analysis of Multitasking and Project Completion Rate			
Dependent Variable	*Project Completion Rate*		
Variables	**Model 1**	**Model 2**	**Model 3**
	RSE-c	RSE-c	RSE-c
nMultitasking	.858** (.030)	.843** (.029)	.851** (.030)
IT Controls			
nFTF Contacts	1.002 (.027)		1.015 (.029)
nPhone Contacts	1.109 (.061)		1.073 (.062)
nEmail Contacts	.974 (.043)		.958 (.046)
nESS Use		1.118** (.035)	1.114** (.038)
nESS Skill		.947 (.051)	.949 (.060)
Team Controls			
Team Size	.854** (.067)	.820** (.067)	.842** (.071)
Industry Experience	.989** (.004)	.990** (.005)	.991** (.005)
Task Controls			
nRoutineness	1.003 (.042)	1.074 (.043)	1.042 (.047)
nInterdepend.	.980 (.038)	.957 (.034)	.979 (.041)
Job Class Controls?	YES	YES	YES
City Controls?	YES	YES	YES
Log Likelihood	-7080.3	-7077.6	-7076.03
X^2 (d.f)	185.07*** (19)	193.16*** (18)	196.79*** (21)
Obs.	1180	1180	1180

***p<.001; **p<.05; *p<.10. RSE-c = Robust Clustered SE (n = 505 Clusters)

Notes: Team size and industry experience are associated with longer project duration and slower completion rates. Teams with more members may take longer to execute projects due to the added complexity of coordination, or the firm may resort to 'throwing more labor at' difficult jobs or jobs that are taking longer to complete than expected. Controlling for team size therefore may also account for differences in project difficulty not picked up by controls for job type, task, and city characteristics. Industry experience also corresponds to longer project duration perhaps because less experienced employees receive less demanding work. Cost of living, crime rates, and greater commute times all reduce the project completion rate on average, meaning these characteristics may be less attractive to potential candidates, while good weather is associated with increased completion rate. Routine tasks consistently finish faster, and greater interdependence among team members is associated with slower completion rates. Significance levels are as follows: **p<.001; *p<.05.