Online Appendix

Socializing at Work: Evidence from a Field Experiment with Manufacturing Workers

Sangyoon Park*

1 Estimation of Worker Ability (Production Skill)

To estimate workers' abilities, or production skills, I build on the approach of Mas and Moretti (2009) by exploiting information on workers' positions and social ties to take into account of the fact that a worker's productivity may be affected by the composition of coworkers at contiguous positions and her social relationship with these coworkers. Thus, I estimate the θ_i terms using the following specification:

$$y_{ird} = \theta_i + \mathbb{V}'\mathscr{C}_{-ird} + \pi N_{ird} + \phi F_{ird} + \lambda_{rd} + \varepsilon_{ird} \tag{1}$$

where θ_i denotes worker fixed effects; N_{ird} denotes the number of workers that are spatially contiguous to worker i on a given day; F_{ird} denotes the vector of dummy variables that indicate the presence of the focal worker's friend at specific proximities (low, medium, and high proximity); λ_{rd} is a vector of all possible combinations of the room and day.

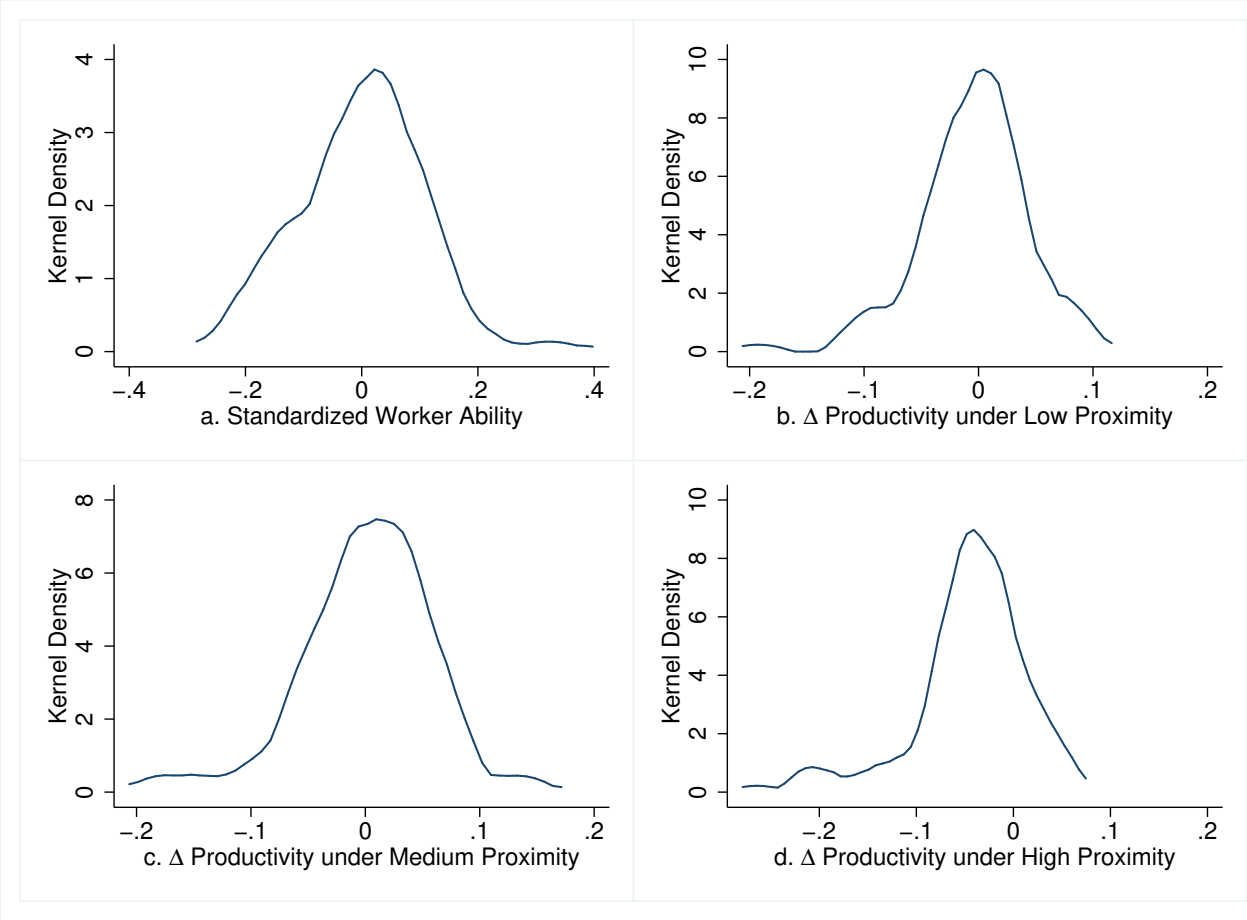
The term $\mathscr{C}_{-ird}$ accounts for worker-specific influences on coworker productivity. $\mathscr{C}_{-ird} = \{I_1, \ldots, I_{i-1}, I_{i+1}, \ldots, I_k\}$ is a set of dummy variables, one for each worker that shares the same processing room with worker i. Each $I_{j\neq i}$ is equal to one if worker j is working at the same table with worker i and zero, otherwise. The vector $\mathbb{V}$ contains k parameters, one for each worker $1, \ldots, k$. This allows consistency with equations (4) and (5) in the main text because $\mathscr{C}_{-ird}$ absorbs social interaction effects. Note that this approach is analogous to that used by Arcidiacono et al. (2017) where the authors estimate player specific effects on their teammates' scoring chances in a professional basketball context. In this setting, the idea is to account for productivity effects from the presence of friends that could differ across workers as well as across physical proximities between the worker and her friends.

Figure 1-1 plots estimates on worker's ability with respect to the presence of friends at each proximity. Panel a. presents kernel density estimates of worker ability, standardized at the room level, when there is no friend at a contiguous position. Panels b, c, and d each plots kernel density estimates of change in worker productivity when a friend is present in low, medium, and high

*Faculty of Business and Economics, University of Hong Kong, Pokfulam Road, Hong Kong. Email: sangyoon@hku.hk

proximity positions, respectively, relative to when no friend is present. Note that the density distribution shown in panel d, compared to other panels, supports the findings in the main analysis of this paper that workers are less productive when a friend is at high proximity.

Figure 1-1: Kernel Density Estimates of Worker Ability Depending on Proximity to Friends



Note: All density estimates use an Epanechnikov kernel with 'optimal' bandwidths. Panel a. uses ability estimates $(\hat{\theta}_i)$ standardized with respect to the room average. Panels b, c, and d each plots kernel density estimates of the change in individual worker productivity when working with a friend at the corresponding proximity relative to when working without a friend. For example, a worker who is equally productive when there is a friend at low proximity and when there is no friend will show up in Panel b as having a value of zero on the horizontal axis. A negative value indicates that the worker is less productive in the presence of a friend at the corresponding proximity relative to when there is no friend nearby.

2 Robustness Checks

While equation (5) in the main text provides a simple specification for characterizing how the presence of friends at different proximities affects worker productivity it may be worthwhile to explore how the presence of unconnected friend pairs (i.e. no direct friendship exists between the focal worker and workers in the friend pair) working in close proximity may affect the focal worker's productivity. A friend pair may be disruptive to others if they engage in chats during work. Goofing off with one another may also spawn negative attitudes in nearby workers. Failure to take into account potential spillovers from other friend pairs may result in misinterpretation of the effect of friends on productivity.

To allow for possible externalities from other friend pairs, I include variables that indicate whether any of the workers, excluding the focal worker, at contiguous positions have a friend working in specific proximities. This can be written as

$$\begin{aligned} y_{irt} &= \gamma_L \cdot \text{Low Prox}_{irt} + \gamma_M \cdot \text{Med Prox}_{irt} + \gamma_H \cdot \text{High Prox}_{irt} \\ &\quad + \sum_{j \in \mathscr{C}(i)} (\delta_L \cdot \text{Low Prox}_{jrt} + \delta_M \cdot \text{Med Prox}_{jrt} + \delta_H \cdot \text{High Prox}_{jrt}) \\ &\quad + X_{irt} + \theta_i + \lambda_{rt} + \varepsilon_{irt} \end{aligned} \tag{2}$$

where Low Prox_{jrt}, Med Prox_{jrt}, and High Prox_{jrt} are indicator variables equal to one if worker j in room r on day t has a friend working at low proximity, medium proximity, and high proximity, respectively. These externality variables are summed over the set of workers at spatially contiguous positions to worker i, denoted by $\mathscr{C}(i)$. Then, δ_L is the externality on productivity from an additional contiguous worker, which may not necessarily be in low proximity to the focal worker, working with a friend at low proximity.[1] Similarly, δ_M and δ_H are the externalities arising from an additional worker at a contiguous position working at medium proximity and high proximity to her friend, respectively. If both workers of a friend pair are working nearby, this is counted as two rather than one. In this way, a friend pair with both workers being spatially contiguous to the focal worker could be considered twice as influential as when only one worker from the friend pair is contiguous to the focal worker.

Column 1 of Table 2-1 provides IV estimates for own proximity variables ($\gamma_L, \gamma_M, \gamma_H$) and contiguous workers' proximity variables ($\delta_L, \delta_M, \delta_H$). Estimates on own proximity variables are almost unchanged.[2] However, the estimate on other workers' high-proximity variable suggests that working at positions near a high-proximity friend pair is associated with a 2 percent loss in productivity. This implies that the presence of high-proximity friends negatively affects other workers, which is possible if a friend pair distracts their surrounding peers. In contrast, friend pairs in low- or medium- proximity positions have no influence on their peer's productivity. The result is significant with the standard errors being clustered two ways by worker and room$\times$day level to allow for arbitrary correlations of the error terms across workers in the same room and day. Failure

[1] In general, it is possible to have one externality variable for each proximity of contiguous worker (to focal worker) and proximity of contiguous worker to friend (of contiguous worker) pair. This would result in nine externality variables. I only consider the proximity of the contiguous worker to her friend to simplify the equation.

[2] Taken together with the significant estimate on contiguous workers' high proximity variable, this indicates that the variation in a worker's proximity to own friends is orthogonal to the variation in contiguous coworker's proximity to friends.

to account for this is especially problematic when the regressors of interest ($\delta_L, \delta_M, \delta_H$) are also correlated across workers within the room×day level, which is known to lead to an over-rejection of the true null hypothesis (Wooldridge, 2003; Angrist and Pischke, 2009).[3]

Next I perform a series of robustness checks. Column 2 includes interaction terms between each proximity variable and an indicator variable that is equal to one if the friend's position is at the same table, and zero otherwise. At each level of proximity, the interaction term captures the difference in the productivity effects from the presence of a friend at the same table relative to the presence of a friend at a different table. If workers were to behave differently depending on whether the friend is at the same table or not — possibly due to the input allocation process — the coefficient on any of the interaction terms would be nonzero. In column 2, coefficient estimates are shown to be close to zero.

Because some work stations, in particular, stations at the corner of the rooms, only have one adjacent work station, if these work stations naturally provide higher productivity then this would create a spurious negative correlation between the presence of friend and productivity. The estimates in column 3 indicate that the main finding is not driven by correlations associated with working in a corner position. In columns 4-6, I include different sets of fixed effects to test the robustness of the main result to different specifications. Column 4 uses table×day fixed effects instead of room×day fixed effects and finds larger declines in productivity when a friend is in high proximity. Columns 5 and 6 show that including work station fixed effects in the main specification makes a minimal difference in the estimates. Column 7 assesses whether the effect on worker productivity from working with a friend changes over the five month experimental period. The estimate on the interaction term, High Prox×Month, indicates that the effect neither intensified nor abated during the experiment.

Next, I examine whether and to what extent the estimate of the effect of working with friends depends on how friendship is defined between two workers. The purpose of this exercise is to test the robustness of the self-reported friendship measure that is widely used in this study. In general, I rerun the IV regression used in Table 9 using different definitions of friendships.

The estimation results are reported below in Table 2-2. For comparison, column 1 shows the IV estimates presented in Table 9. Column 2 presents results using only reciprocal connections in the sense that workers i and j are friends if they both reported each other. The estimates are close to those in column 1. Column 3 uses only reported friendships and excludes those that are mutual to gauge how the effect differs when the report is not reciprocal. That is, if worker i reported j but j did not report i then j appears as i's friend but i would not appear as j's friend. Note that the median worker in the sample had one such relationship. The size of the estimate on High Prox is smaller than the previous two columns but still remains statistically significant. In contrast, as shown in column 4, when friendship is defined to exist only when a worker receives but not reports that same worker, the estimate loses significance.

In column 5, I estimate the proximity coefficients using only friendships that have formed prior to their current job. The effect size is slightly larger than when using all friendships. This suggests that workers are less productive when working next to a friend that they have known before their current job than when working next to a friend they have met on the job. It is also possible that

[3]To see why the regressors are correlated, if worker i has a friend pair working at a contiguous position, then that friend pair also appears in the regressors of all other workers that are contiguous to the friend pair. The room×day cluster correlation for other workers' proximity variables at low, medium, and high proximity is 0.12, 0.07, and 0.12, respectively.

the self-reports are just a list of coworkers who work nearby and happen to chat often. To partially address this concern, in the last column, I limit friendships to those that have been rarely observed working contiguously to each other before the experiment but was reported by at least one worker during the baseline survey on friendships. In this way, friends here are the ones that a worker has been less likely to work nearby and arguably less likely to have been reported just because they often chat with each other during work. The coefficient estimate for working in high proximity is negative and statistically significant, although the magnitude is about 9 percent smaller than the benchmark estimate in column 1. The result indicates that even among friendships that have been observed to work nearby less often prior to the experiment, there is a significant negative effect on productivity when they are assigned to work alongside each other.

The attrition rate during the experiment period was low: 101 of the 105 workers were still working at the end of the experiment. Nonetheless, the mean absence rate during the experiment was 21.4% which creates a missing data problem.[4] In order to examine the robustness of my main result, I use the bounding approach of Lee (2009) to construct lower and upper bounds for the effect of working alongside friends. Taking the weighted average of individual estimates based on number of observations, I obtain a lower bound of -0.0287 and an upper bound of -0.0944, compared to the IV estimate of -0.056 in Table 9. The lower bound is only relevant, however, if workers were able to anticipate their assignments and decided not to show up on days when none of their friends were assigned to work alongside them and they were expecting to be less productive than other days. However, this is unlikely, since, as mentioned above, assignment to friend is not a significant predictor of work attendance.

[4]The mean absence rate during the six weeks before the experiment was 21.8%. I found no significant difference in average worker attendance rates between the two periods (p-value = 0.65). Moreover, although there is a 2.1% difference in the absence rate during the experiment period between when a worker is assigned to work next to a friend (19.8%) and when there is no friend assigned next to the worker (21.9%) the difference is neither statistically significant nor is friend assignment a predictor of work attendance (results not reported here).

Table 2-1: Robustness Checks

	IV Estimates - Dependent Variable: log(productivity)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Low Prox	0.001	−0.008	0.001	0.001	0.000	0.001	0.026∗∗
	(0.006)	(0.012)	(0.006)	(0.010)	(0.006)	(0.010)	(0.011)
Med Prox	−0.003	−0.003	−0.006	0.023∗	−0.002	0.026∗∗	−0.002
	(0.010)	(0.010)	(0.011)	(0.012)	(0.010)	(0.012)	(0.019)
High Prox	−0.058∗∗∗	−0.051∗∗∗	−0.060∗∗∗	−0.069∗∗∗	−0.057∗∗∗	−0.069∗∗∗	−0.057∗∗∗
	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)	(0.012)	(0.015)
$\sum_{j\in\mathscr{C}(i)}$ Low Prox$_j$	−0.001	−0.001	−0.001	0.002	0.000	0.002	0.000
	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)
$\sum_{j\in\mathscr{C}(i)}$ Med Prox$_j$	−0.002	−0.001	−0.001	0.011∗∗	−0.001	0.012∗∗∗	−0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
$\sum_{j\in\mathscr{C}(i)}$ High Prox$_j$	−0.010∗∗	−0.010∗∗	−0.010∗∗	−0.017∗∗∗	−0.010∗∗	−0.017∗∗∗	−0.010∗∗
	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)	(0.006)	(0.004)
Low Prox × Same Table		0.011					
		(0.013)					
High Prox × Same Table		−0.010					
		(0.010)					
Low Prox × Corner Position			0.003				
			(0.013)				
Med Prox × Corner Position			0.012				
			(0.016)				
High Prox × Corner Position			0.021				
			(0.013)				
Low Prox × Month							−0.014∗∗
							(0.006)
Med Prox × Month							−0.001
							(0.007)
High Prox × Month							−0.001
							(0.007)
Include X_{irt}?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Worker FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Room × Day FE?	Yes	Yes	Yes	No	Yes	No	Yes
Table × Day FE?	No	No	No	Yes	No	Yes	No
Workstation FE?	No	No	No	No	Yes	Yes	No

Note: X_{irt} is a vector of control variables at the individual spatial level, which includes the number of workers working at contiguous positions and the mean of their average productivities. Same Table is an indicator variable equal to one if the friend at the respective proximity is at the same table with the reference worker, and zero otherwise. Corner position indicates whether the worker was assigned to a corner position. Month represents a linear monthly time trend. Standard errors are two-way clustered by worker and room×day. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 2-2: Does the estimate depend on how friendship is defined?

	IV Estimates - Dependent variable: log(productivity)					
Friendship exists if	Report or receive	Both Report	Report but not receive	Receive but not report	Formed before current job	Work apart (pre-exper.)
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Low Prox	0.000	0.005	−0.020	0.003	−0.004	0.002
	(0.006)	(0.008)	(0.015)	(0.013)	(0.013)	(0.009)
Med Prox	−0.003	−0.007	0.010	0.001	−0.012	0.003
	(0.010)	(0.014)	(0.012)	(0.012)	(0.018)	(0.009)
High Prox	−0.056***	−0.057***	−0.046***	−0.027	−0.060***	−0.051***
	(0.010)	(0.010)	(0.015)	(0.020)	(0.014)	(0.013)
Mean (# Friendships)	4.94	2.90	1.12	0.92	1.27	2.52
Median (# Friendships)	5	3	1	1	1	2
Observations	5,731	5,731	5,731	5,731	5,731	5,731

Note: Column 1 defines two workers as friends if either one reported the other. In column 2, a friendship exists only if two workers reported each other. Columns 3 and 4 are set differences between the set of friends reported by the worker (outgoing) and the set of reports received from other workers (incoming). In column 5, two workers are defined as friends if either one reported the other and if their social connection formed before starting work at their current job: duration of friendship $>$ current job tenure. Column 6 uses either reported or received friends but only those who rarely worked near each other during the pre-experiment period. All regressions include room$\times$day fixed effects, externality variables, and controls at the individual spatial level. Standard errors are two-way clustered at the worker and room$\times$day level. * $p < .10$, ** $p < .05$, *** $p < .01$.

3 Additional Tables & Figures

Table 3-1: Heterogeneous Effects - Production Skill

	IV - Dependent var.: log(productivity)			
Variable	(1)	(2)	(3)	(4)
Low Prox	0.001	0.001	0.000	0.001
	(0.006)	(0.006)	(0.012)	(0.013)
×Own Ability	0.064			
	(0.066)			
×Friend's Ability		0.251		
		(0.170)		
×Moreable			0.001	
			(0.017)	
×Moreable×|Own Ability - Friend's Ability|				−0.013
				(0.163)
×Lessable×|Own Ability - Friend's Ability|				0.016
				(0.160)
Med Prox	−0.003	−0.004	−0.002	−0.007
	(0.010)	(0.010)	(0.010)	(0.014)
×Own Ability	−0.035			
	(0.085)			
×Friend's Ability		−0.025		
		(0.068)		
×Moreable			−0.002	
			(0.010)	
×Moreable×|Own Ability - Friend's Ability|				0.026
				(0.104)
×Lessable×|Own Ability - Friend's Ability|				0.041
				(0.097)
Observations	5,731	5,731	5,731	5,731

Note: This table reports estimates not reported in Table 10. Dependent variable is the worker's log productivity (kilograms/hour). All regressions include worker and room×day fixed effects along with all externality variables at the individual spatial level. Standard errors are two-way clustered by worker and room×day level and corrected for the sampling variability of the estimated ability term using a Bayesian parametric bootstrap procedure. In columns 1 and 2, proximity variables are interacted with estimates on worker's own ability and friend's ability at each proximity, respectively. Column 3 uses an indicator variable equal to one if the reference worker has higher ability than her friends at each proximity, and zero otherwise. Column 4 uses a measure of absolute difference in ability with respect to that of her friends at each proximity. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 3-2: Heterogeneous Effects - Personality Skills

	IV - Dependent var.: log(productivity)				
Friendship based on:	All Reports			Mutual Reports	
Variable	(1)	(2)	(3)	(5)	(6)
Low Prox	0.000	−0.004	−0.001	0.006	−0.023
	(0.005)	(0.023)	(0.005)	(0.007)	(0.018)
× Extraversion	0.008	0.008	0.007	0.007	0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
× Agreeableness	−0.006	−0.006	−0.006	−0.001	0.002
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
× Conscientiousness	0.013	0.012	0.012	0.011	0.007
	(0.007)	(0.008)	(0.007)	(0.008)	(0.008)
× Neuroticism	0.008	0.009	0.009	0.009	0.011
	(0.006)	(0.006)	(0.007)	(0.008)	(0.008)
× Openness	0.005	0.005	0.005	−0.001	−0.001
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
× Age		0.000			0.001
		(0.001)			(0.001)
× Job Tenure		0.000			−0.001
		(0.001)			(0.001)
× Ability			0.026		0.083
			(0.072)		(0.072)
× Friend's Ability			0.226		0.064
			(0.185)		(0.211)
Med Prox	−0.002	0.000	−0.002	−0.004	−0.009
	(0.010)	(0.033)	(0.010)	(0.013)	(0.043)
× Extraversion	−0.014	−0.014	−0.014	−0.018	−0.014
	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)
× Agreeableness	−0.007	−0.008	−0.008	−0.003	−0.001
	(0.008)	(0.007)	(0.008)	(0.009)	(0.008)
× Conscientiousness	0.010	0.009	0.010	0.008	0.007
	(0.008)	(0.008)	(0.008)	(0.014)	(0.015)
× Neuroticism	−0.006	−0.007	−0.006	−0.001	−0.003
	(0.009)	(0.009)	(0.009)	(0.014)	(0.015)
× Openness	0.005	0.005	0.005	0.002	0.001
	(0.007)	(0.007)	(0.007)	(0.010)	(0.010)
× Age		0.000			0.000
		(0.001)			(0.001)
× Job Tenure		0.000			0.000
		(0.001)			(0.001)
× Ability			−0.008		−0.117
			(0.065)		(0.132)
× Friend's Ability			0.008		−0.041
			(0.051)		(0.105)
Worker FE?	Yes	Yes	Yes	Yes	Yes
Observations	5,731	5,731	5,731	5,731	5,731

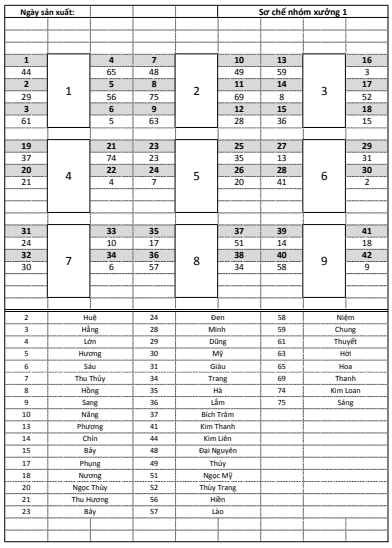
Note: This table reports estimates not reported in Table 11. Dependent variable is the worker's log productivity (kilograms/hour). The Big Five personality scores are self-reported and standardized with respect to the sample mean and standard deviation. All regressions include room×day fixed effects and and controls at the individual spatial level. Standard errors are two-way clustered by worker and room×day level. * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 3-3: Proximity to Friends during Pre-Experiment Period

Proximity (=1 if yes)	Mean (S.D.)		Difference (1) - (2)	P-value
	Observed (1)	Simulated (2)		
Contiguous	0.87	0.41	0.45	0.00
	(0.17)	(0.17)	(0.02)	
Low Prox	0.49	0.19	0.29	0.00
	(0.25)	(0.11)	(0.03)	
Med Prox	0.51	0.12	0.39	0.00
	(0.24)	(0.06)	(0.02)	
High Prox	0.51	0.19	0.32	0.00
	(0.27)	(0.09)	(0.03)	
Total workdays	32			

Note: Statistics reported in column 2 are obtained from 500 simulations based on a random assignment rule using the number of friends and coworkers present in the processing room for each worker and day. Columns 1 and 2 report standard deviations in parentheses. Column 3 reports standard errors from t-tests of difference in parentheses. One sided p-values are reported in the last column.

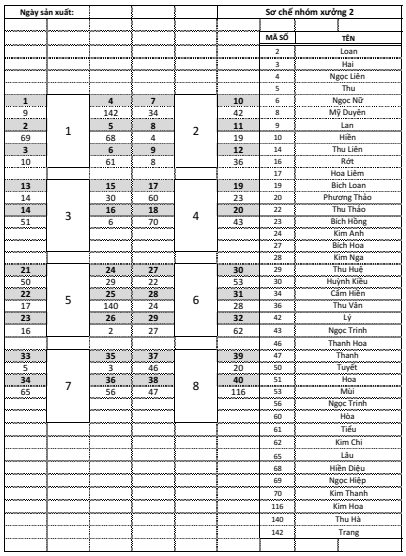
Figure 3-1: Worker Position Assignment Forms



Ngày sản xuất:					Sơ chế nhóm xưởng 1				
1	1	4	7	2	10	13	3	16
44		65	48		49	59		3
2		5	8		11	14		17
29		56	75		69	8		52
3		6	9		12	15		18
61		5	63		28	36		15
19	4	21	23	5	25	27	6	29
37		74	23		35	13		31
20		22	24		26	28		30
21		4	7		20	41		2
31	7	33	35	8	37	39	9	41
24		10	17		51	14		18
32		34	36		38	40		42
30		6	57		34	58		9

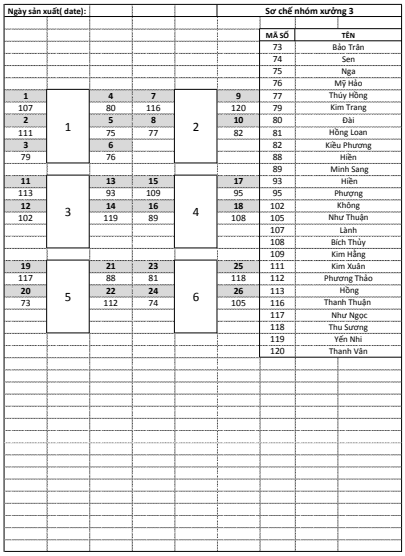
2	Huệ	24	Đen	58	Niêm
3	Hằng	28	Minh	59	Chung
4	Lớn	29	Dũng	61	Thuyết
5	Hương	30	Mỹ	63	Hời
6	Sáu	31	Giàu	65	Hoa
7	Thu Thủy	34	Trang	69	Thanh
8	Hồng	35	Hà	74	Kim Loan
9	Sang	36	Lắm	75	Sáng
10	Năng	37	Bích Trâm		
13	Phương	41	Kim Thanh		
14	Chín	44	Kim Liên		
15	Bảy	48	Đại Nguyên		
17	Phụng	49	Thúy		
18	Nương	51	Ngọc Mỹ		
20	Ngọc Thúy	52	Thúy Trang		
21	Thu Hương	56	Hiền		
23	Bảy	57	Lào		

(a) Processing Room 1



Ngày sản xuất:					Sơ chế nhóm xưởng 2		
					MÃ SỐ	TÊN	
					2	Loan	
					3	Hai	
					4	Ngọc Liên	
					5	Thu	
1	1	4	7	2	10	6	Ngọc Nữ
9		142	34		42	8	Mỹ Duyên
2		5	8		11	9	Lan
69		68	4		19	10	Hiền
3		6	9		12	14	Thu Liên
10		61	8		36	16	Rớt
						17	Hoa Liêm
13	3	15	17	4	19	19	Bích Loan
14		30	60		23	20	Phương Thảo
14		16	18		20	22	Thu Thảo
51		6	70		43	23	Bích Hồng
						24	Kim Anh
						27	Bích Hoa
						28	Kim Nga
21	5	24	27	6	30	29	Thu Huệ
50		29	22		53	30	Huỳnh Kiều
22		25	28		31	34	Cẩm Hiền
17		140	24		28	36	Thu Vân
23		26	29		32	42	Lý
16		2	27		62	43	Ngọc Trinh
						46	Thanh Hoa
33	7	35	37	8	39	47	Thanh
5		3	46		20	50	Tuyết
34		36	38		40	51	Hoa
65		56	47		116	53	Mùi
						56	Ngọc Trinh
						60	Hòa
						61	Tiếu
						62	Kim Chi
						65	Lâu
						68	Hiền Diệu
						69	Ngọc Hiệp
						70	Kim Thanh
						116	Kim Hoa
						140	Thu Hà
						142	Trang

(b) Processing Room 2

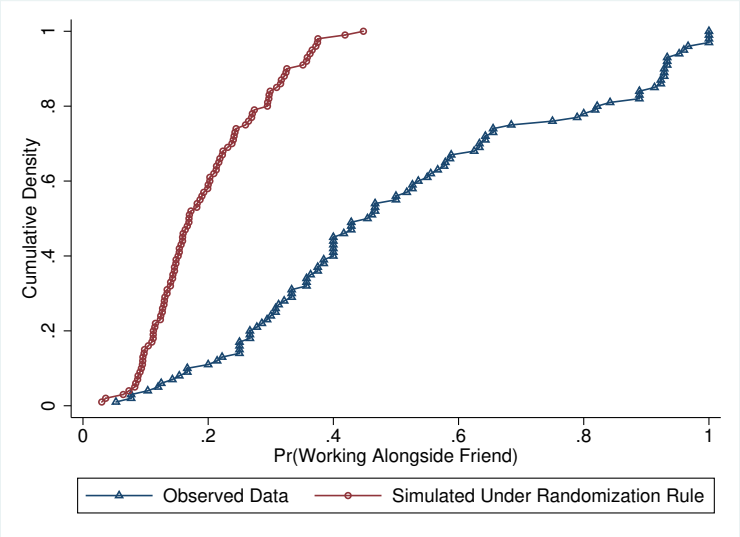


Ngày sản xuất(date):					Sơ chế nhóm xưởng 3		
					MÃ SỐ	TÊN	
					73	Bảo Trân	
					74	Sen	
					75	Nga	
					76	Mỹ Hảo	
1	1	4	7	2	9	77	Thúy Hồng
107		80	116		120	79	Kim Trang
2		5	8		10	80	Đài
111		75	77		82	81	Hồng Loan
3		6				82	Kiều Phương
79		76				88	Hiền
						89	Minh Sang
11	3	13	15	4	17	93	Hiền
113		93	109		95	95	Phượng
12		14	16		18	102	Không
102		119	89		108	105	Như Thuận
						107	Lành
						108	Bích Thúy
						109	Kim Hằng
19	5	21	23	6	25	111	Kim Xuân
117		88	81		118	112	Phương Thảo
20		22	24		26	113	Hồng
73		112	74		105	116	Thanh Thuận
						117	Như Ngọc
						118	Thu Sương
						119	Yến Nhi
						120	Thanh Vân

(c) Processing Room 3

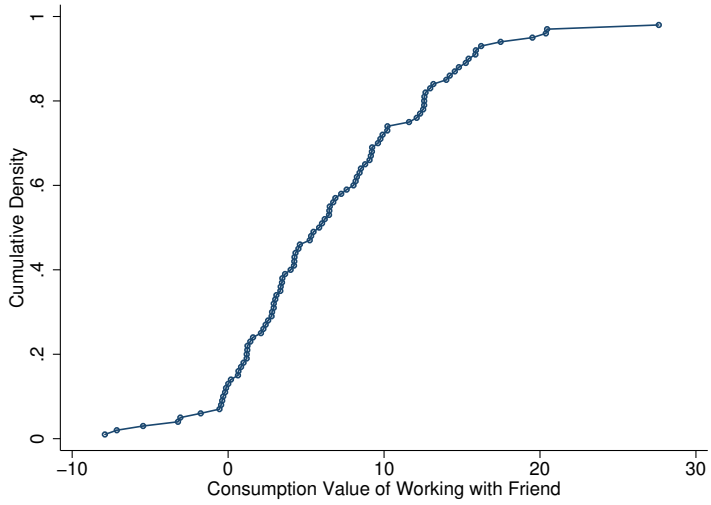
Note: The number of tables and worker positions are different across rooms due to differences in room sizes. Pseudonyms are used to generate these samples and not that of the actual workers.

Figure 3-2: Cumulative Distribution Functions of the Probability of a Friend at High Proximity Positions



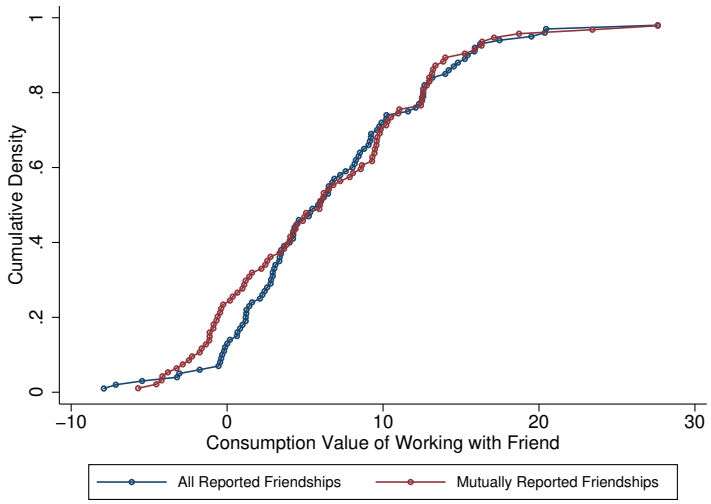
Note: Counterfactual data is from 500 simulations of worker positions with random assignment during the pre-experiment period. A comparison of the two distributions using the Kolmogorov-Smirnov equality of distribution test rejects that the two distributions are equal (p-*value* = 0.00).

Figure 3-3: Cumulative Distribution Function of Estimates on Consumption Value of Working with Friends



Note: This figure plots the empirical cumulative distribution function of worker-specific estimates of consumption value of working with friends, $\hat{s}_i$.

Figure 3-4: Comparison of CDFs of Workers' Consumption Values of Working with Friends



Note: Estimates on mutually reported friendships are derived only using friendships that are reported from both sides of the relationship. A Kolmogorov-Smirnov equality of distribution test fails to reject that the two distributions are equal ($p\text{-}value = 0.316$).

4 Derivation of Estimator for Consumption Value in Appendix

The probability of worker i working with a friend in high proximity is given as

$$\Pr(\mathrm{F}_i = 1) = \Lambda(s_i - x_i). \tag{3}$$

Then, the log-likelihood function for worker i is

$$\log \mathscr{L} \propto \left(\frac{1}{T} \sum_{t=1}^{T} \mathrm{I}_{\{F_{it}=1\}} \right) \cdot \log \Lambda(s_i - x_i) + \left(1 - \frac{1}{T} \sum_{t=1}^{T} \mathrm{I}_{\{F_{it}=1\}} \right) \cdot \log(1 - (\Lambda(s_i - x_i)) \tag{4}$$

and the first order condition is

$$\frac{d\ log\mathscr{L}}{d\ s_i} = \left(\frac{1}{T} \sum_{t=1}^{T} \mathrm{I}_{\{F_{it}=1\}} \right) \cdot \frac{\lambda(s_i - x_i)}{\Lambda(s_i - x_i)} - \left(1 - \frac{1}{T} \sum_{t=1}^{T} \mathrm{I}_{\{F_{it}=1\}} \right) \cdot \frac{\lambda(s_i - x_i)}{1 - \Lambda(s_i - x_i)} = 0. \tag{5}$$

Rearranging equation (5), I obtain

$$\frac{1}{T} \sum_{t=1}^{T} \mathrm{I}_{\{F_{it}=1\}} = \Lambda(s_i - x_i). \tag{6}$$

Under the assumption that the scale parameter of the logistic distribution is one, I can derive s_i as

$$s_i = x_i + \mathrm{logit} \left(\frac{1}{T} \sum_{t=1}^{T} \mathrm{I}_{\{F_{it}=1\}} \right) \tag{7}$$

since logit function is the inverse of the cdf of the logistic distribution.

References

Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Arcidiacono, P., Kinsler, J., and Price, J. (2017). Productivity Spillovers in Team Production: Evidence from Professional Basketball. *Journal of Labor Economics*, 35(1):191–225.

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies*, 76(3):1071–1102.

Mas, A. and Moretti, E. (2009). Peers at Work. *American Economic Review*, 99(1):112–145.

Wooldridge, J. (2003). Cluster-Sample Methods in Applied Econometrics. *American Economic Review*, 93:133–138.