Online Appendix for "The Influence of Peers in Worker Misconduct: Evidence From Restaurant Theft"

Appendix

A. Instruments Using Different Exclusion Windows

For robustness, we construct three sets of instruments by varying the length of time window, separately excluding coworker observations within one week, two weeks, and three weeks of the focal worker observation. We run the original IV regression using these three sets of instruments. The results are shown in Table A.1. We find very similar peer-effect estimates.

Table A.1 Average Theft Peer Effects With Different Exclusion Windows

	DV:Server Theft Count		DV:Server Theft Value	
Model	Estimate	Cluster Std.Error	Estimate	Cluster Std.Error
IV(1 Week)				
PeerTheft	0.046	0.010	0.027	0.0031
IV(2 Week)				
Peer Theft	0.042	0.015	0.027	0.0032
IV(3 Week)				
Peer Theft	0.045	0.017	0.025	0.0043

B. Peer Effects in Productivity

We also measure productivity peer effects in our setting using four measures: employee tip amount, sales revenue, drink sales, and add-on sales. Tip amount is a measure of service quality. Drink and add-on sales represent a significant portion of restaurant profits because of high margins. Add-on sales are also important for restaurants because of their high margins. Since not all restaurants in the data set recorded drink sales, tipping, and add-on sales, we must drop these restaurants in the regressions.

The OLS and IV regression results are shown in Tables B.4 and B.5, with standard errors clustered at the restaurant-shift level. In each case, estimated peer effects are positive, although only add-on and drink sales are statistically significant at the 5 percent level.

Table A.2 Peer Effects in Theft with Alternative Clustering

	DV:Server Theft Count		DV:Server Theft Value	
	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
Avg Peer Theft Count	-0.01***	0.04**		
	(0.0018,0.0021)	(0.015,0.019)		
Avg Peer Theft Value			0.007***	0.027***
			(0.001,0.0015)	(0.003,0.007)
Individual Fixed Effect	Yes	Yes	Yes	Yes
Manager Fixed Effect	Yes	Yes	Yes	Yes
Restaurant*Year Fixed Effect	Yes	Yes	Yes	Yes
Restaurant*Week Fixed Effect	Yes	Yes	Yes	Yes
Restaurant*Weekday Fixed Effect	Yes	Yes	Yes	Yes
Restaurant*IT-Monitoring Fixed Effect	Yes	Yes	Yes	Yes
Observations	4058783	4058783	4058783	4058783
R^2	0.182	0.125	0.241	0.240
1st Stage F Statistics		1396***		14260***
Difference in Estimate	0.05		0.02	
Hausman Test Statistics	10.14***		86.4***	

Note:Standard errors are presented in parentheses. The first number represents standard error clustered at restaurant shift level. The second number represents standard error clustered at restaurant level. Significance level:*p<0.1; **p<0.05; ***p<0.01.

Table B.3 Descriptive Statistics of Variables in Regression

Variable	Focal Server	Avg Peer
Theft Count Mean	0.06	0.07
Theft Count SD	0.60	0.35
Theft Value Mean	1.14	1.35
Theft Value SD	18.13	11.6
Tipping Mean	4.0	4.1
Tipping SD	17.8	9.3
Revenue Mean	43.2	43.1
Revenue SD	204.0	101.5
Drink Sales Mean	8.3	8.8
Drink Sales SD	10.6	8.3
Add-on Sales Mean	7.0	7.4
Add-on Sales SD	8.6	7.1

Note: Descriptive statistics for productivity measures for both dependent variable (Focal Server) and instrumental variable (Avg Peer).

Table B.4 Average Peer Effect in Tipping and Revenue

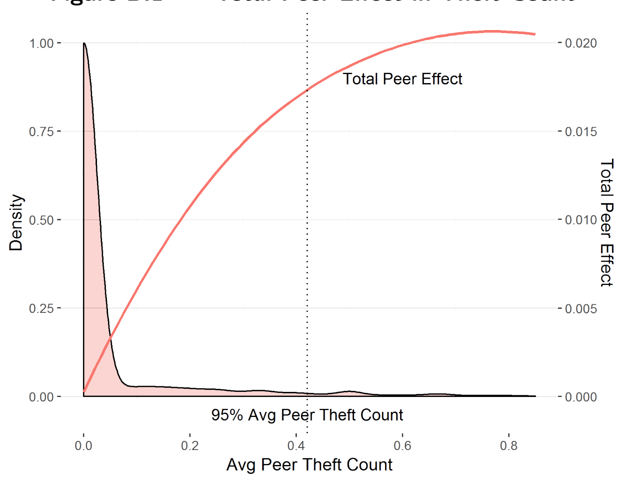
	DV:Server Tip		DV:Server Revenue	
	OLS	IV	OLS	IV
Avg Peer Amount	0.008	0.03	0.048	0.059
	(0.03)	(0.08)	(0.05)	(0.06)
Observations	4058783	4058783	4058783	4058783
R^2	0.190	0.188	0.207	0.206
1st Stage F Statistics		45.13***		59.62***
Difference in Estimate	0.022		0.011	
Hausman Test Statistics	2.2		0.14	

Table B.5 Average Peer Effect in Add-on and Drink Sales

	DV:Server Add-on Sales		DV:Server Drink Sales	
	OLS	IV	OLS	IV
Avg Peer Amount	0.20***	0.169***	0.20***	0.19***
	(0.005)	(0.029)	(0.02)	(0.04)
Observations	3769578	3769578	3769578	3769578
R^2	0.652	0.647	0.558	0.556
1st Stage F Statistics		151.3***		280.4***
Difference in Estimate	-0.031		-0.01	
Hausman Test Statistics	1.17		0.08	

Note:Standard errors, clustered at the restaurant-shift level, are presented in parentheses. Significance level:*p<0.1; **p<0.05; ***p<0.01.

Figure B.1 Total Peer Effect in Theft Count



This figure shows how the total peer effect in theft count changes across the data support range of average peer theft count, based on the non-linear model in Table 4. The total peer effect is increasing across the entire range of our data.

Table B.6 Peer Effects in Theft Count by Month for All New Employees

	(1) 1st Month	(2) 2nd Month	(3) 3rd Month	(4) 4th Month	(5) 5th Month	(6) 6+ Months
Avg Peer Theft Count	0.125***	0.10***	0.122**	0.121**	0.120*	0.032
	(0.02)	(0.03)	(0.05)	(0.06)	(0.07)	(0.05)
Individual Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Manager Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*Week Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*Weekday Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*IT-Monitoring Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	985,136	569,481	365,177	247,899	169,968	305,492
R^2	0.154	0.192	0.225	0.257	0.280	0.201

Note:Standard errors, clustered at the restaurant-shift level, are presented in parentheses. Significance level:*p<0.1; **p<0.05; ***p<0.01.

Table B.7 Peer Effects in Theft Value by Month for All New Employees

	(1) 1st Month	(2) 2nd Month	(3) 3rd Month	(4) 4th Month	(5) 5th Month	(6) 6+ Months
Avg Peer Theft Value	0.048***	0.042***	0.035***	0.043***	0.027	-0.004
	(0.006)	(0.008)	(0.01)	(0.015)	(0.02)	(0.01)
Individual Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Manager Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*Week Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*Weekday Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*IT-Monitoring Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	985,136	569,481	365,177	247,899	169,968	305,492
R^2	0.201	0.232	0.269	0.315	0.35	0.33

Note:Standard errors, clustered at the restaurant-shift level, are presented in parentheses. Significance level:*p<0.1; **p<0.05; ***p<0.01.

Figure B.2 Survival Probability for High- and Low-Theft Workers

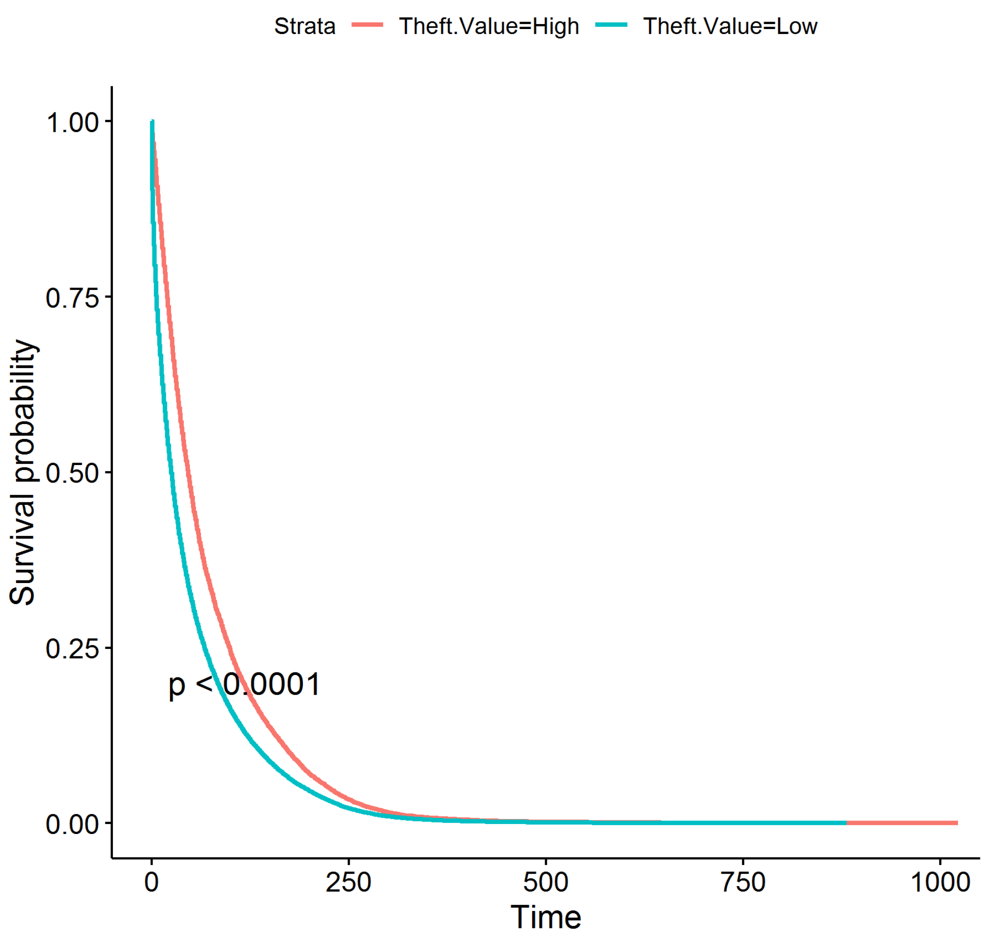
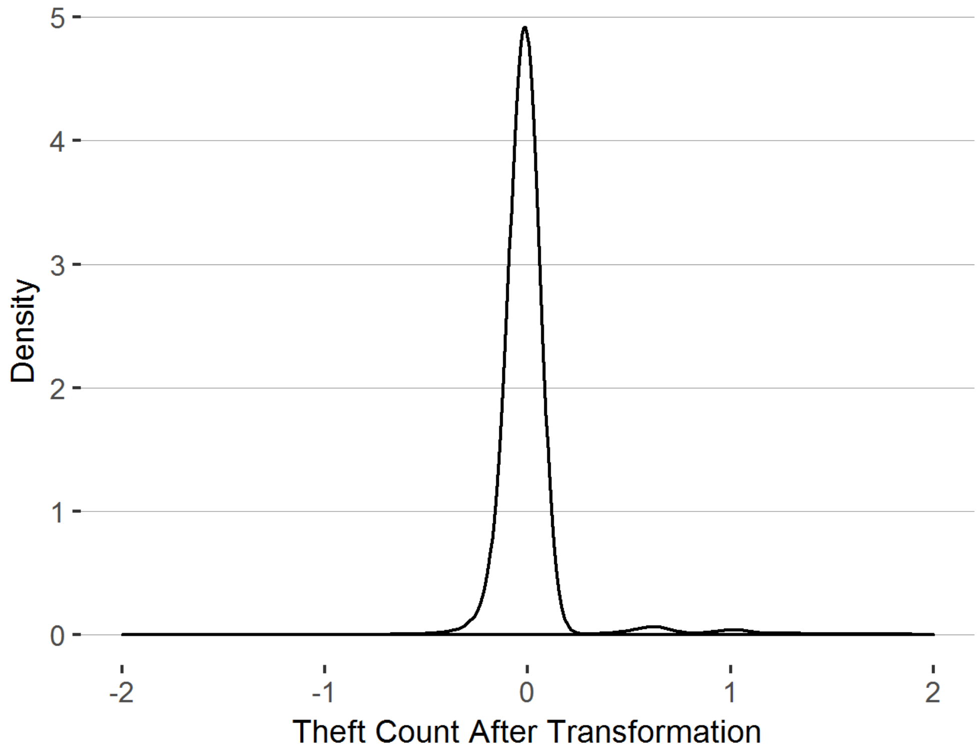


Figure shows raw survival data for all high- and low-theft workers, defined by whether their average daily theft value is below or above the median worker.

C. Count Data and Linear Model

In Figure C.3, we show the distribution of theft count after logarithmic and demeaning transformation. The distribution after transformation is close to normal, which could alleviate concerns of applying a linear model to discrete count data.

Figure C.3 **Distribution of Theft Count After Transformation**



D. Calculating Total Impact of Peer Effects

Here we use a toy numerical example to show that the endogenous peer-effect coefficient in our model has a multiplier effect on the overall outcome because of the reflection structure in peer effect equations. To simplify the analysis, we use a group of two workers i, j. Suppose that without peer effects, each worker's intrinsic theft level is α_i, α_j. With β as the peer-effect coefficient, the real theft level (Y_i, Y_j) should be the solution to the system of Equations (1).

$$\begin{cases} Y_i = \beta Y_j + \alpha_i \\ Y_j = \beta Y_i + \alpha_j \end{cases} \tag{1}$$

Solving the equation, we have

$$\begin{pmatrix} Y_i \\ Y_j \end{pmatrix} = \begin{pmatrix} \frac{\beta\alpha_j + \alpha_i}{1-\beta^2} \\ \frac{\beta\alpha_i + \alpha_j}{1-\beta^2} \end{pmatrix} \tag{2}$$

Now suppose the intrinsic theft level of worker i, α_i increases by an amount of Δ. Solving Equations (1) again, we have the new theft level $(Y_i^{'}, Y_j^{'})$ as

$$\begin{pmatrix} Y_i' \\ Y_j' \end{pmatrix} = \begin{pmatrix} \frac{\beta\alpha_j+\alpha_i+\Delta}{1-\beta^2} \\ \frac{\beta(\alpha_i+\Delta)+\alpha_j}{1-\beta^2} \end{pmatrix} \tag{3}$$

Comparing the theft levels as shown in Equations (2) and (3), the increase in theft level for worker i is: $Y_i' - Y_i = \frac{\Delta}{1-\beta^2}$. The increase in theft level for worker j is: $Y_j' - Y_j = \frac{\beta\Delta}{1-\beta^2}$. With a peer-effect coefficient $\beta < 1$, we could clearly see that the increase of worker $i's$ theft level is larger than Δ, and that the impact of $i's$ increase in intrinsic theft level on j is larger than $\beta\Delta$. This is caused by the reflection structure in peer-effect Equations (1); an increase in $i's$ theft value will have impact on j through the peer-effect coefficient β, and because $i's$ theft level is reflectively affected by $j's$ theft level, the increase in $j's$ theft level will be reflected back on $i's$ own theft. As a result, the overall increase in real theft level ($\frac{\beta+1}{1-\beta^2}\Delta$) is larger than the case without the endogenous peer effect (Δ).

The overall effect of such endogenous peer effect increases with the size of the peer group because the reflection exists in any pair of individuals in the peer group and the number of bilateral relationships, which is $\frac{n\times(n-1)}{2}$ for a group of n individuals, increases the size of the peer group. In summary, the interpretation of the endogenous peer-effect coefficient in our model should account for the reflection structure, which results in a multiplier effect on the observed outcome, and such multiplier effect increases the size of the peer group and the peer-effect coefficient.

E. Serial Correlation Test

We conduct serial correlation tests on residuals to understand whether our two-week window is sufficiently long to establish instrument exogeneity. We first estimate the autocorrelation coefficient between residuals ϵ_{it} and $\epsilon_{\mathrm{i(t\text{-}k)}}$, with lags $k = 2, 4, \cdots, 14$ from the IV regressions on theft count and theft value. We use the incremental value of 2 days because the average time between two consecutive shifts for a focal worker is two days in our data. Figure E.5 shows the decreasing trend of the autocorrelation coefficient over time. Within 10 lags, the auto correlation coefficient is 0.03 for theft count and 0.02 for theft value. From lag 10 to lag 14, the decreases to 0.02 for theft count and 0.005 for theft value before converging to zero. This suggests that even if a focal worker were aware of peer behavior during shifts she did not work, and this word-of-mouth changed her behavior that day, that behavioral change would be unlikely to carry forward two weeks into the future to violate the exclusion restriction.

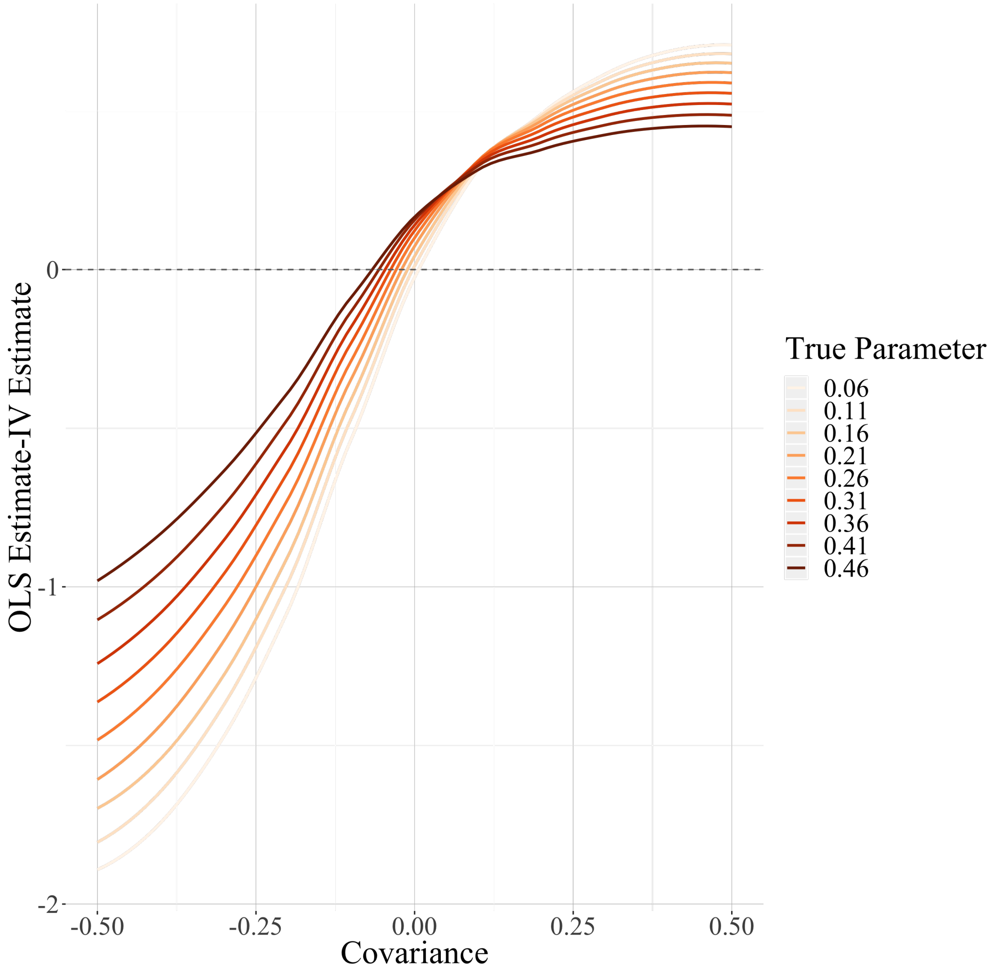
We also conducted a Ljung-Box test for each worker's time series ϵ_{it} with lags between 14 and 28 days, since these days represent data from which we build our instruments. We find that only 2.2% (theft count) and 1.6% (theft value) of all workers demonstrate any serial correlation with a p value less than 0.05. This further alleviates concerns that serial correlation from our instruments might violate the exclusion restriction.

F. Peer Effects Based on Restaurant Size

We also explored possible differences in peer influence between large and small restaurants, splitting our sample based on the number of median employees at that restaurant in a week. We run the average peer effect regressions separately for large and small restaurants, with results shown in Table F.8.

The IV estimates show much stronger peer effects in large restaurants than in small restaurants. Comparing the bias of OLS estimates with IV estimates, we find that the downward bias is also larger in large restaurants

Figure D.4 Correlation, True Parameter, & OLS Bias



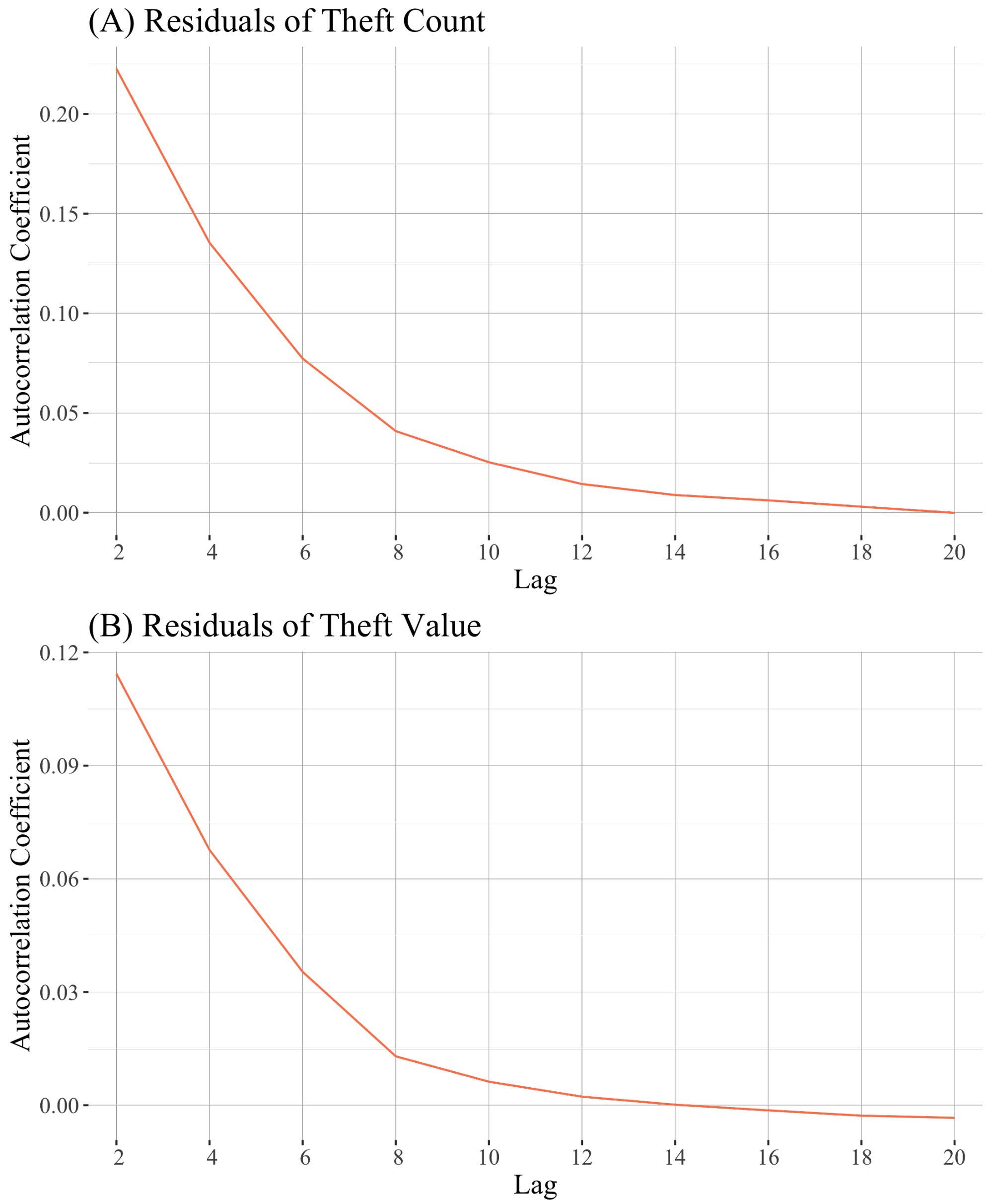
This figure shows how the OLS bias (OLS - IV estimates) varies with different covariance in residuals for true parameters ranging from 0.06 to 0.46. Each point in each of the nine curves includes 1,000 simulations

Table F.8 Average Peer Effect Small Restaurant VS Large Restaurant

Model	OLS	IV
Theft Count		
Small Restaurant	-0.012	0.02
	(0.0019)	(0.015)
Large Restaurant	-0.003	0.14
	(0.005)	(0.05)
Theft Value		
Small Restaurant	0.01	0.02
	(0.001)	(0.003)
Large Restaurant	0.005	0.06
	(0.002)	(0.009)

Note:Standard errors, block-bootstrapped at the restaurant-shift level, are presented in parentheses.

Figure E.5 Auto Correlation Coefficient Over Time



than in small restaurants. This indicates that the negative correlation in daily error terms is higher in large chains. This finding could be explained by the monitoring attention difference in large and small restaurants. In small restaurants, workers usually occupy multiple roles (bartenders, to-go server, etc.) due to small scale limiting specialization. As a result, managers in small restaurants could take other roles in addition to monitoring the servers and thus have less attention for monitoring. In our data, 65% of servers in small restaurants have worked as managers, while the percentage is 44% in large restaurants. Why then are peer effects larger in large restaurants?