The Influence of Peers in Worker Misconduct: Evidence From Restaurant Theft

Tat Y. Chan*
Olin Business School, Washington University in St. Louis, chan@wustl.edu

Yijun Chen
Olin Business School, Washington University in St. Louis, yijun.chen@wustl.edu

Lamar Pierce
Olin Business School, Washington University in St. Louis, pierce@wustl.edu

Daniel Snow
Saïd Business School, Oxford University, daniel.snow@sbs.ox.ac.uk

Forthcoming in *Manufacturing & Service Operations Management*

* The authors are listed alphabetically and contributed equally to the paper.

Problem Description: Misconduct, such as theft, is a major problem in operational settings, and staffing decisions can either amplify or mitigate this problem as workers influence their peers' behavior. Peers are known to influence coworker productivity, and likely also affect counterproductive behaviors.

Academic/Practical Relevance: Many studies have shown how such mechanisms as helping, knowledge transfer, teaching, and social pressure generate productivity peer effects in service and other settings. Yet few papers empirically examine these effects in counterproductive behaviors. We argue that while the same mechanisms driving productivity spillovers also generate peer effects in misconduct, an additional effect—strategic peer response—reflects how coworkers, under managerial monitoring, adjust misconduct in response to peers' daily behavior. An additional contribution of this paper is to identify the effect of peers on operational performance in a firm setting.

Methodology: We use transaction and theft data from 83,153 servers at 1,049 restaurants across 46 states in the US. We employ instrumental variables (IV) models to account for both reflection problems and correlated error terms in same-day peer theft. We use Monte Carlo simulations to present how biases identified by a combination of ordinary least squares (OLS) and IV models suggest that managerial oversight might generate negative correlation in the same-day error terms of peers that reflects strategic peer responses.

Results: Our results show that although servers are more likely to steal when working with high-theft peers, they steal less as peers steal more *on a given day*. We also show that this negative correlation in daily peer theft is higher under an IT system that increases managerial oversight by reporting likely theft to managers. Importantly, we demonstrate how reflection effects can significantly amplify even small peer effect coefficients to have large organizational implications. Our parameter estimates indicate that doubling a single worker's average theft amount will increase total theft in an average restaurant by 76%. Doubling all workers' theft amounts increases totals by 550%. Finally, we show that the positive peer effect from high-theft coworkers only exists for new workers in their first three to five months on the job, consistent with imprinting mechanisms that include knowledge transfer and norms.

Managerial Implications: The results show that the costs of employing unethical workers is higher than the direct cost of those workers' misconduct because their behavior spills over into coworkers' actions and amplifies through reflection effects. Yet our results also suggest that this contagion can be mitigated by managerial oversight. So long as there is sufficient monitoring of misconduct, workers will strategically limit such behavior in response to peers.

Key words: Peer Effects, Productivity, Misconduct, Service Operations, Theft, Shrinkage

1. Introduction

Performance spillovers from one worker to another are fundamental to operational performance in service, sales, and production environments. Existing work almost uniformly shows that top employees positively influence peer productivity in settings that include supermarkets (Mas and Moretti 2009), health care (Song et al. 2017), and science (Azoulay et al. 2010, Oettl 2012).[1]

[1] Tan and Netessine (2019) is an exception. Also see related work on peer effects among managers and entrepreneurs (Hasan and Koning 2017, Chatterji et al. 2019).

These positive peer effects are driven by mechanisms that include direct assistance, production complementarities, knowledge transfer, and peer pressure and behavioral norms.[2]

Although each mechanism is cited in studies involving a productive behavior and outcome, these same peer effect mechanisms likely also apply to costly behaviors such as theft, fraud, sabotage, and shirking. Indeed, evidence from non-firm settings suggests that the mechanisms behind productivity peer effects also generate peer effects in misconduct. Athletes cheat more when connected with unpunished cheaters (Palmer and Yenkey 2015). Similarly, crime, tax evasion, and underage drinking are amplified by peer behavior (Bayer et al. 2009, Alm et al. 2017). The limited evidence from firms is consistent with these results.[3] Ichino and Maggi (2000) show positive peer effects in shirking among automotive workers in one firm, while Dimmock et al. (2018) find similar results in financial-service career networks. What these few studies do not address, however, is a key mechanism that is likely to moderate productivity and misconduct peer effects—managerial oversight. Although it seems quite intuitive that working with a generally high-theft peer might normalize and enable misconduct on average, it is important to recognize that this interaction occurs within a larger organizational context. Indeed, a worker's bad behavior *on a given day* might produce the opposite effect. If higher daily levels of misconduct in the organization increase the likelihood of managerial detection, then high peer misconduct might incentivize *lower* misconduct in the focal worker in a given time frame. This presents two theoretical predictions about how peers may affect coworker misconduct in operational settings. On one hand, being staffed on a given day with a peer who steals more *on average* might justify, normalize, and enable higher individual misconduct—a positive "peer effect" as in Mas and Moretti (2009). On the other hand, a peer's higher misconduct levels *on a given day* might incentivize a worker to reduce cheating to avoid managerial detection and resulting punishment, what we term "strategic peer response."

In this paper, we disentangle the positive peer effects and negative strategic peer responses in misconduct. We do so in an important operational setting—US casual dining restaurants. These restaurants are characterized by table service and moderate prices and account for $269 billion of the $766 billion US restaurant market.[4] Theft is a common problem in restaurants, and although there are not precise figures available on the magnitude of theft, estimates suggest annual employee theft costs of $3 to $6 billion (Garber and Walkup 2004). While theft and illicit behavior are, by their very natures, difficult to measure, our unique setting allows us to directly observe daily worker

[2] Exceptions include cases where employees are incentivized to compete with or sabotage their peers (Siemsen et al. 2007, Chan et al. 2014a).

[3] Mohliver (2019) shows misconduct peer effects at the firm level that imply employee-level effects. Related laboratory experiments show peer influence in cheating and other misconduct (Gino et al. 2009, Gino and Pierce 2009).

[4] Maze, J. 2017. "Restaurant sales to hit $799B in 2017" Feb 28, 2017. https://www.nrn.com/sales-trends/nra-restaurant-sales-hit-799b-2017

staffing, sales productivity, and revenue theft estimates for 83,153 servers at 1,049 restaurants from 34 chains over seven years.[5] Variation in server staffing, and thus peers, allows us to examine whether working with high-theft peers (compared to low-theft peers) on a given day indeed increases a focal worker's theft in the same way that the productivity literature suggests.

We implement a specific identification strategy that derives peer-effect estimates from both ordinary least squares (OLS) and instrumental variables (IV) methods. This strategy serves two purposes. First, the IV method, which uses the average daily theft of the peer on days in which the two workers are not co-staffed as an IV for daily peer theft, addresses Manski (1993)'s two classic peer-effect estimate biases: (a) the reflection problem and (b) correlation in peers' regression residual terms. Second, as we show, using simulations, the difference between OLS and IV models can shed light on the strategic interaction between workers on a given day. These simulations show that for positive peer-effect models, OLS can only produce smaller parameter estimates than IV models if the residual terms are strongly negatively correlated. This implies that workers endogenously reduce theft on days in which peers choose to steal over their normal level. Our combined results therefore address two key statistical biases and also use the magnitude of the estimated bias to identify negative strategic peer response.

Our unbiased IV estimates show that working with high-theft peers increases a server's theft likelihood and magnitude. Our coefficient magnitudes imply peer effects of 4 percent in theft count and 2.7 percent in theft value from average theft levels of peers. Although these effects appear, the implication for firms is significant because of the reflection effect of the endogenous peer effect coefficient. An increase in one individual's theft level will have a multiplicative effect over group theft levels through the endogenous coefficient across multiple peers. We show through simulation that reflection effects in an average restaurant substantially increase the impact of a high-theft worker. Doubling a single worker's average theft amount implies a 76 percent increase in an average restaurant's total theft, while doubling all workers' theft produces an increase of 550 percent. Our significantly smaller OLS estimates indicate a negative correlation in the daily residual terms of peers in the regression models, however, which implies that workers reduce theft in response to higher than normal peer theft on that day. Although we cannot identify if these correlations are due to collusion or to independent adaptation to observing peers, they are consistent with workers responding to the increased threat of detection of higher daily restaurant losses.

We support this monitoring mechanism argument by comparing model estimates in a subset of restaurants before and after they adopted an IT-based theft-monitoring system. The system, which notifies managers of suspiciously high clusters of possible theft-related transactions, raises the risk

[5] Our data do not allow us to measure inventory losses.

that management will investigate high-theft days. We find that the biased OLS estimates become smaller and the negative correlation in the residual terms becomes larger in magnitude following monitoring adoption. This suggests that under a stronger monitoring regime, workers are more likely to reduce theft in response to higher peer theft on a given day.

Finally, we provide evidence for the information and social norms mechanisms by examining new employees, showing that positive theft peer effects on new employees are strongest in the first months of employment and disappear after the fifth month. Although these results cannot precisely separate each mechanism, they are consistent with an imprinting mechanism in which new employees are most vulnerable to environmental influences (Marquis and Tilcsik 2013).

Our paper contributes to four main streams of literature. First, it adds to the growing literature on people-centric operations. Relatively few papers have examined peer effects (Chan et al. 2014a, Tan and Netessine 2019) or other coworker interactions (Pierce et al. 2019, Moon et al. 2018) in operational settings, with related work studying learning from peers (Song et al. 2017, KC et al. 2013, Chan et al. 2014b, Siemsen et al. 2008, Yin et al. 2018, Valentine et al. 2018, Tucker et al. 2007) and the value of staffing heterogeneity (Aksin et al. 2015, Kesavan et al. 2014, Huckman and Staats 2011) and team experience (Huckman et al. 2009). Ours is the first to study peer effects in theft. This research also contributes to work on operational losses, which include not only theft (Pierce et al. 2015) but also broader categories of shrinkage (DeHoratius and Raman 2008, 2007) and other operational risks (Xu et al. 2017).

Second, this paper contributes to a growing literature on behavioral misconduct and unethical behavior in field settings. As Pierce and Snyder (2015) note, only a few papers objectively measure individual employee misconduct in operational settings (e.g., Nagin et al. 2002, Pierce and Snyder 2008, Bennett et al. 2013, Balafoutas et al. 2013, Derfler-Rozin et al. 2016). Our paper takes a behavioral effect commonly studied in the lab (Gino et al. 2009) and demonstrates its generalizability to an operational setting. In this way, our paper parallels recent field work on peer effects in other behaviors such as tax compliance (Alm et al. 2017) and fraud (Edelman and Larkin 2014)

Third, we add to a growing debate on the role and efficacy of monitoring in operational settings. Although substantial evidence shows how monitoring and other managerial oversight can improve compliance (Staats et al. 2016) and reduce misconduct (Nagin et al. 2002, Pierce et al. 2015), other work explains that such monitoring can have negative effects on motivation and performance (Bernstein 2012, 2017, Anteby and Chan 2018, Ranganathan and Benson 2019). Our work suggests that the effects of managerial monitoring can extend beyond the monitored worker to coworkers. Monitoring that reduces theft in the average worker will have an amplified effect as that reduction spills over to coworker behavior through peer effects.

[5] See related work on social networks and misconduct (Yenkey 2015, 2018, Aven 2015, Cohen et al. 2010).

Finally, we add a methodological contribution to a large empirical literature on peer effects in firms (Herbst and Mas 2015) by demonstrating that the use of both biased OLS and unbiased IV methods can do more than correct for reflection problems; it can also inform on collusion and other strategic interactions between peers on a given day.

2. Empirical Setting

The industry context for this study is a labor-intensive service operation——US casual dining restaurants. Some restaurant examples in this segment (not necessarily in our sample) include Applebee's, Buca di Beppo, and Johnny Rockets. Customers of restaurants in this segment receive table service from waitstaff, who typically take orders, deliver food, and process payment.

Staffing in our setting almost always involves more than one server in a given shift. Ninety-four percent of shifts in our sample involve two or more workers. In informal interviews at casual dining restaurants, waitstaff reported a heavy interaction with other servers and service staff. Each server with whom we spoke expressed the inevitability of interacting socially, financially, and sometimes romantically with other staff. These interactions often occur at choke points, such as terminals and order windows, and also occur in kitchen areas, in employee meetings, and after closing.

Theft by servers and others in this segment is a significant problem, although its precise magnitude is unknown. A significant portion of theft likely comes from unapproved "comped" meals given for free to customers and through employee consumption of food and beverages. A small number of studies have examined restaurant employee theft. Victor et al. (1993) use survey-based data on peer theft reporting, while Detert et al. (2007) find employee food theft to be associated with the number of and abusiveness of managers. Pierce et al. (2015) utilize a subsample of the data in the present paper with the first large-scale study using objective individual theft data. In that (and this) paper, the identification of employee theft is based on servers not reporting an item's sale or removing an item from the restaurant's IT system after customers have paid.

Although restaurant employees steal in many ways, we focus on three types that are observable in the data generated by point-of-sale (POS) systems. Restaurants in our sample are managed using a common POS system that tracks each employee's orders, sales, and job category. When a customer places an order, or a "ticket," with a server, that server enters the information into a touch-screen terminal. Order information is stored in the system's database and is passed to a display in the kitchen. After the customer pays and leaves, the server closes out the ticket.

The three POS-based "scams" in our data are common in the industry, even appearing in how-to tell-all books about theft (Francis and DeGlinkta 2004), and are observed using theft detection algorithms provided by the POS system provider. The first type is called the "wagon-wheel scam" in which, following customer payment, the server transfers an item from that bill in the POS system to the bill of another customer who ordered the same item. That bill is reprinted after the customer leaves and the server keeps the difference by taking cash from the terminal. The wagon wheel can be applied to cash and credit card transactions, with the latter achieved by increasing the tip amount by the transferred amount to maintain the total credit card bill. The second theft technique involves “comping,” or refunding a customer’s meal in the system after they have already paid but before the ticket has been closed. The third involves voiding a transaction as erroneous after the customer has already paid. When cash is paid, the server keeps all or part of the payment rather than depositing it in the terminal. For credit card transactions, the server takes cash from the terminal as a fraudulent tip. The level of flexibility in the POS system that allows for these types of theft is common in the industry and is meant to allow servers and managers to adjust for entry errors and changes in customer orders. The detection of this type of theft under normal circumstances falls to the manager, acting as the restaurant owner’s agent. Considerable effort would need to be applied in order to detect, investigate, and take action over this type of theft.

The data are from an IT firm that sold POS systems to 34 restaurant chains with 1,049 locations. The IT system stored information about, among other things, menu items ordered, times of events, payment types, tip amounts, server identifiers, and an indicator for a likely theft having occurred in a given transaction. Theft monitoring was an add-on feature sold to restaurant chains, costing less than $100 per month per location. Theft was detected using proprietary algorithms, with subscribers provided theft alerts identifying associated employees, times, and values. Although detected theft was only available to subscribers, it is visible to the researchers for all restaurant-days in the sample. The algorithms were constructed with a strong bias against finding a theft event because false positives (accusations) in this context are costly. Consequently, our theft measurements are substantially smaller than actual theft levels, although we have no reason to believe this under-measurement is biased. Such a bias is unlikely because the algorithm applies equally to all restaurants and servers and because its existence is never known by servers or managers in restaurants that don’t implement the monitoring system. Even in restaurants where it is implemented, the algorithm is unknown to servers and restaurant managers and would thus be difficult to game. Importantly, we can observe these theft events in the data regardless of whether or not the restaurant has adopted the monitoring system by applying the provider’s algorithms. Interviews with restaurant managers indicated that their use of and response to the monitoring system varied, although most indicated that they intervened when theft was repeated or substantial.

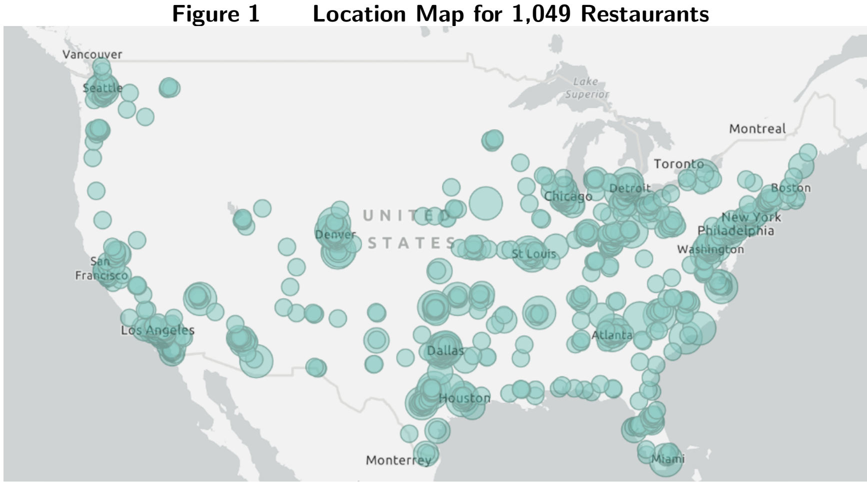
The particular structure of worker assignment to customers is helpful in identifying the managerial oversight mechanism because servers are quasi-randomly assigned to customers in ways that reduce concerns that certain workers choose customers who facilitate theft. As Tan and Netessine (2014) and Tan and Staats (2016) detail, servers in this segment are assigned an area of tables, with customers then matched to those tables either through specific rules or algorithms or by hosts at the front door. Similarly, the fact that customers do not directly suffer from server theft reduces concerns that customer monitoring might explain any peer effects in theft. In the wagon-wheel scam, where servers transfer drinks in the IT system, customers still receive the ordered drink and pay for it. Similarly, in falsely comped or voided tickets, the customer is unaware of the action because they still pay for and receive their food.

As Tan and Netessine (2019) detail, servers interact constantly before, during, and after shifts, and the small workforce and rotating schedules ensure worker awareness of their peers' behavior and personalities. Similar to productivity, peer theft can be directly observable by peers, either through queuing at the POS terminal or through comped or voided orders. In addition, workers carry reputations for both productive and unproductive behaviors, such that directly observing theft is unnecessary to know that one is working with a high-theft peer (Brass et al. 1998).

3. Data and Measures

The data set for this study contains transaction and theft data aggregated to the worker-day level. It contains 5,731,806 unique worker-day observations, covering 83,153 servers in 1,049 restaurants that belong to 34 chains over approximately seven years. Figure 1 presents all locations for the lower 48 states, with larger circles representing larger location counts in a given city. Typically, each restaurant will have a daily staffing schedule in which servers are assigned to different shifts throughout the day. For each server, we observe the shift's start and end times. For each shift, we also know the sales revenue associated with the server's transactions. Table 1 describes the data set.

Figure 1 Location Map for 1,049 Restaurants



Note: Each circle reflects all restaurants in a specific city, with larger circle sizes reflecting more unique restaurants. Alaska and Hawaii are not shown.

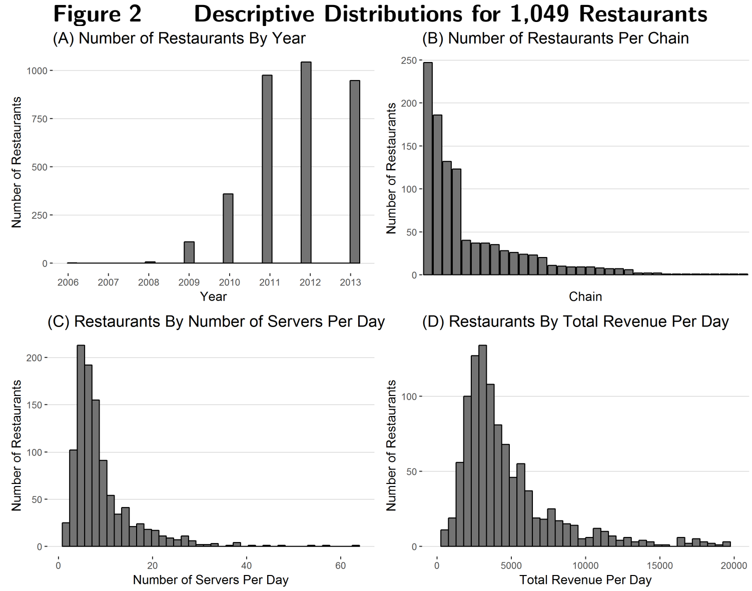
Table 1 Data Set Description

Number of Restaurants	1,049
Number of Chains	34
Number of Employees	83,153
Unique Restaurant-Days	806,541
Number of Employee-Day Observation	5,731,806
Years of Observations	2006-2013

Panel (B) of Figure 2 shows the distribution of locations across the 34 chains, ranging from single locations to four large national chains that constitute the majority of our observations. Panels (C) and (D) show the distribution of servers and revenues across restaurants on a given day.

Our worker-day panel is unbalanced at both the worker and restaurant levels for several reasons. First, workers rarely stay at a given restaurant for the entire time that location appears in the data. Second, restaurants appear in our data only for the period in which they are using the POS system, such that later adopters enter our data set at different points. We note that all models include dummy variables for restaurant-specific week, year, and day of week. In addition, we control for the time of day with four dummies indicating when the worker started their shift: 3am–9am, 9am–3pm, 3pm–9pm, and 9pm–3am. We also include unique identifiers for each individual worker (and thus for each restaurant). These dummies mitigate most potential concerns around an unbalanced panel. Panel (A) of Figure 2 presents the number of restaurants by year in our data.

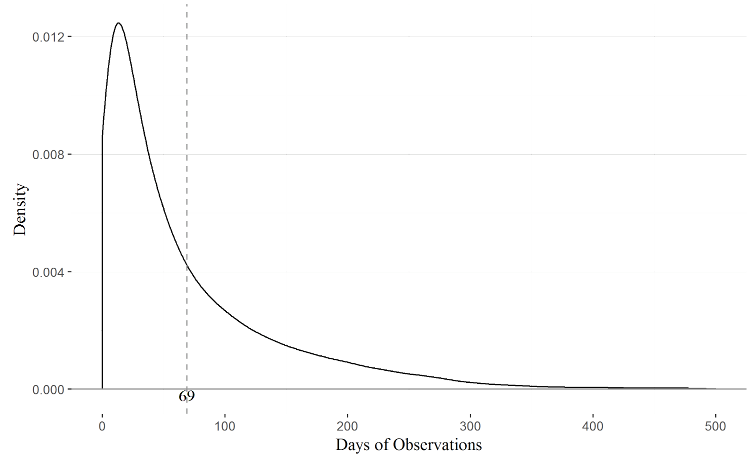
Figure 2 Descriptive Distributions for 1,049 Restaurants



Note: Figures show descriptive data on the restaurants in our sample.

Turnover is high in our restaurant setting, with the mean worker observable on only 69 days. Figure 3 shows the distribution of total tenure (in days) for the 83,153 servers in our data.

Figure 3 Tenure in Days for 83,153 Servers



Note: Figure presents kernel density plots for the distribution of total days worked for all servers in the data set.

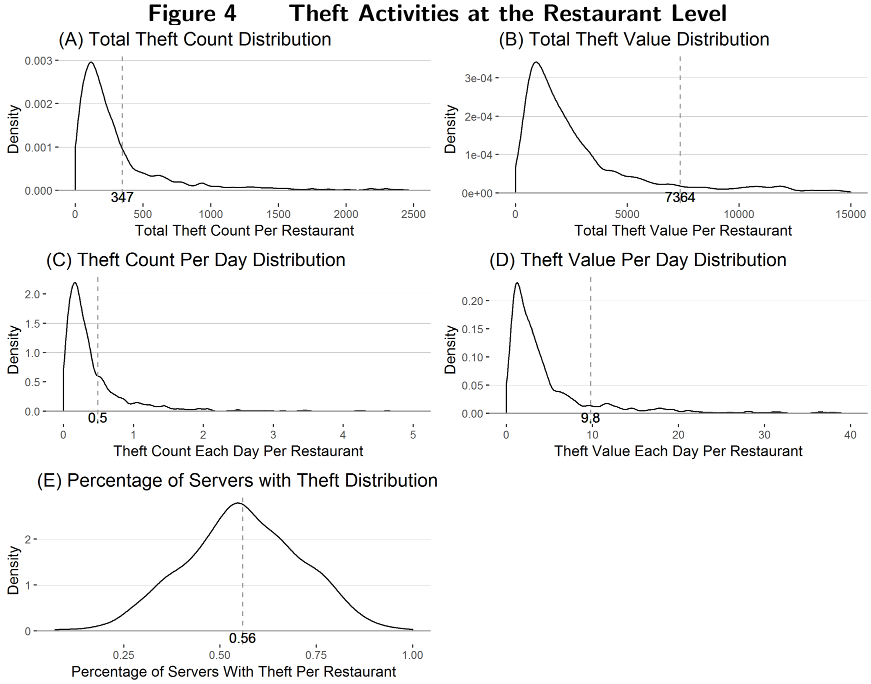
3.1. Theft Activities

We use two measures of server theft, both of which are calculated at the day level. *Theft count* is the number of unique theft occurrences identified for that worker on a given day. If three transactions on a given day are tagged by the IT system as involving theft, for example, then theft count will take a value of three. *Theft value* is the total monetary value stolen by a worker on a given day. We also observe the manager, who is likely to monitor the server as well as whether the theft-notification alerts are being sent to that restaurant. We know the exact date for each observation, which allows us to construct time variables, including year, week, and day-of-week indicators.

In Figure 4, we show the distribution of theft activities at the restaurant level. Panels (A) and (B) represent the distributions of theft count and value across restaurants, while panels (C) and (D) show the distributions in each day. There is a high degree of variation across restaurants, and the distributions are heavily right skewed. The figures suggest that for most restaurants, identified theft is rare on any given day (the average daily theft count is 0.5) and of low value ($9.80 per day). As described above, these figures significantly understate true levels of theft, as they are constructed to conservatively generate theft alerts. Panel (E), however, presents the distribution of the percentage of servers identified to have stolen at least once in the data. On average, 56 percent of servers in a restaurant commit identifiable theft at least once in the sample.

Figure 5 further presents the distributions of theft activities at the server level that are consistent with the observations in Figure 4. Panels (A) and (B) show that over their employment tenure, the average number of *observed* thefts by an individual server is about 10 and the total theft value is about $160, implying that the average value involved in each observed theft is about $16. On a daily basis, however, there is only a 16 percent chance that a server steals and the average value is just $1.67 per server (see Panels (C) and (D)).

Figure 4 Theft Activities at the Restaurant Level



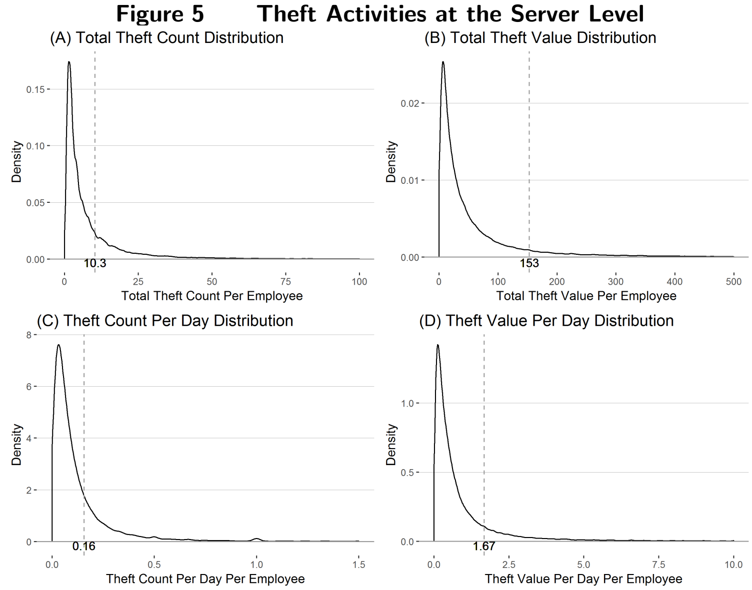
Note: Figure presents kernel density plots for the distribution of theft by individual restaurant location. Dotted lines represent the mean for each variable.

The right skewness in Figure 5 also shows that some servers steal with high frequency. For example, about 1.1 percent of servers stole more than 50 times in the data, and 1.8 percent have stolen more than $400. Different factors could contribute to the total theft behavior of those servers. First, there may be variation in underlying willingness to steal based on individual differences in moral or ethical codes, cognitive ability, or skill at theft (Pierce and Snyder 2015, Bandura 2014). Second, the individual details of particular restaurant chains and locations may influence the opportunity for theft. For instance, restaurants that have more drink items on the menu may be more prone to theft. Also, shift timing may affect theft behavior. Customers may be more likely to order drinks on weekends, leading to a situation in which servers working on weekend shifts have more opportunities to steal. For each of these reasons, our identification strategy will require controlling for observable differences in staffing schedules. Finally, and of primary interest in this study, theft may also be influenced by the identity of peers.

3.2. Shift Overlaps with Peer Employees

To identify peer effects in theft, we exploit variation in daily staffing schedules, a crucial element in manufacturing and service operations (Green et al. 2013, Batt et al. 2019). If a server always were to work with the same peers over time, we could not identify peer effects because of the impossibility of separating each server's permanent theft level through a fixed effect. As most restaurants in our data are open for long hours and have large numbers of employees, each server is assigned to

Figure 5 Theft Activities at the Server Level



Note: Figure presents kernel density plots for the distribution of theft by individual servers.

shifts that vary in work hours and days over the sample period. This, in fact, is what generates the variation necessary for our empirical analysis.

We first construct for each server the percentage of working days as their total number of working days divided by the total number of open days at that restaurant. Figure 6 shows that on average the percentage of working days per server is only 8 percent. Very few servers are observed working for the same restaurant for the whole sample period,[6] which reflects the high employee turnover rate in the industry. The low percentage suggests that servers commonly share shifts with new workers. This regularity is responsible for the (useful for measurement purposes) temporal variation in the pool of peers across working days.

Even though the pool of peers varies over time, a stronger identification criterion is that each server works with different pools of peers currently employed by the restaurant. The logic that if variation in peers for a given server were to come only from the restaurant's new hires, then the server would otherwise work with the same set of peers and we would be concerned about a significant selection issue (i.e., who is hired/fired and who is assigned to work with the server on a fixed basis) that could affect inference of a peer effect. To test whether the stronger criterion applies to the data, we further construct a measure to quantify the degree of overlap in working schedules between each pair of servers, on the condition that both are currently working for the same restaurant, using the following procedure: We first calculate the schedule set, $Schedule_i$, for

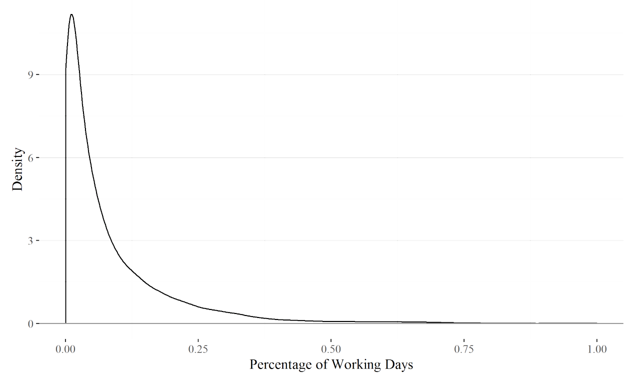
[6] Although it is likely that some servers work in two restaurants in the sample, we cannot identify those employees; server identifiers are specific to restaurants.

each server i. This includes all the shifts in which the server is observed to work in the data. We then select all other servers who are also working for the restaurant during the same time period. For another server j, we then construct the variable $Overlap_{ij}$ which is calculated as

$$Overlap_{ij} = \frac{|Schedule_i \cap Schedule_j|}{|Schedule_i|} \tag{1}$$

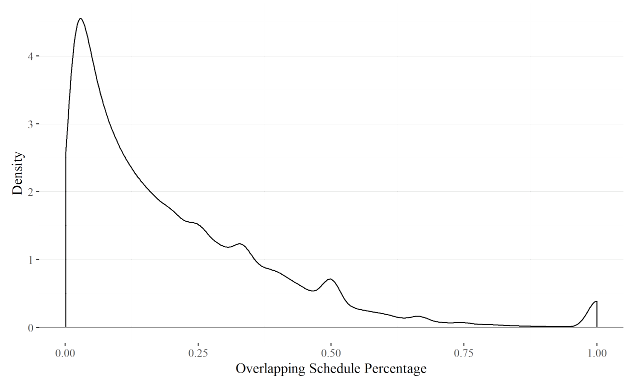
where $|\cdot|$ denotes the number of elements in a set, and the numerator represents the number of overlapping shifts. If there is no variation in the daily staffing schedule, then the overlapping measure will be 1 for servers always working in the same shift and 0 for other servers. The distribution of the overlapping shifts is shown in Figure 7. The average is only 20 percent across all worker pairs. In other words, each server on average only overlaps with other current servers in the restaurant in 20 percent of shifts, indicating a high degree of variation in shift assignment.

Figure 6 Percentage of Working Days per Server Distribution



Note: Figure presents kernel density plot of the percentage of total days that a given server works at a given restaurant. This indicates there is substantial variation in employees across time.

Figure 7 Percentage of Overlapping Shift Distribution



Note: Figure presents kernel density plot of the percentage of overlapping shifts for all pairs of coworkers in the data. This indicates substantial variation in coworkers for a given server.

4. Empirical Model and Its Estimation

We build a peer-effect model to investigate how a server's theft is affected by the theft behavior of peers who work in the same shift. There is, of course, risk of endogeneity in model estimation here. To address the endogeneity problem, we use an instrumental variables (IV) approach in the estimation. Later we discuss in detail the validity of the instruments for peer theft and use a simulation study to demonstrate the sources of estimation bias when ignoring the endogeneity problem and how the problem can be solved by the proposed method.

4.1. The Base Peer-Effects Model

We model how an individual server's theft is influenced by peer theft type in the following way:

$$Y_{it} = \beta \overline{Y_{-it}} + \alpha_i + \gamma_{it} + \epsilon_{it} \tag{2}$$

The dependent variable Y_{it} is the theft by server i on day t. We use the number of times the server steals (theft count) and the value of money the server steals (theft value) to measure the extent of the theft on that day. To deal with the highly right-skewed distribution of both measures observed in the data, we employ a logarithmic transformation. That is, $Y_{it} = log(TheftCount_{it} + 1)$ or $Y_{it} = log(TheftValue_{it} + 1)$. We add 1 inside the parentheses because there are many 0s for both theft count and theft value in the sample.[7]

The independent variables include the average theft by peers and a set of control variables. Included is α_i, a fixed effect for individual servers (there are 83,153 fixed effects altogether), which controls for the server's intrinsic preference or ability for theft. In addition, we use a vector of fixed effects, γ_{it}, to control for factors that can have macro impacts on employee behaviors. These include a fixed effect for each week, each day of the week (e.g., Monday, Tuesday, Wednesday), each year, each manager who oversees the restaurant during the shift, and the presence of the theft-monitoring technology.

The variable $\overline{Y_{-it}}$ represents the collective theft activity of peers. We use the theft count or theft value averaged across peers as the measure. That is, let P_{it} denote the set of peers for server i on day t, such that the peer theft activities are

$$\overline{Y_{-it}} = \frac{\Sigma_{j \in P_{it}} Y_{jt}}{|P_{it}|} \tag{3}$$

where $|P_{it}|$ is the count of the peer set. Note that $\overline{Y_{-it}}$ captures the actual theft behavior of peers on day t, instead of whether peers have a higher theft type (e.g., higher preference or ability to

[7] We use linear instead of nonlinear (e.g., ordered probit or logit) models because of the many fixed effects. Estimating a nonlinear model (while controlling for the endogeneity issue explained below) with close to 100K parameters is impractical for our computing resources. As Figure C.3 in the Appendix demonstrates, the transformed variables are nearly normally distributed, which alleviates concerns about implementing a linear model.

steal). We use this specification because each individual server's theft patterns fluctuate greatly across shifts. The coefficient for $\overline{Y_{-it}}$, β, represents the effect of peer workers' actual theft behaviors on the focal server's theft. Finally, ϵ_{it} is the unobserved residual term.

4.2. Endogeneity Issues

In estimating the peer-effect coefficient β, OLS suffers from the classic endogeneity problem that the unobserved residual term can be correlated with the actual theft behavior of peers, $\overline{Y_{-it}}$. In such cases, the OLS estimate will be biased and inconsistent. This endogeneity problem can come from two sources. The first is the reflection problem (Manski 1993). Suppose the true β is positive. If ϵ_{it} is positive, then other workers' theft behaviors will increase through a peer effect, thus causing an increase in $\overline{Y_{-it}}$ over time. Therefore ϵ_{it} and $\overline{Y_{-it}}$ will be positively correlated.

The second source of endogeneity is that the unobservable ϵs can be correlated across peers. For example, there may exist some common unobserved factors that will simultaneously influence the theft behavior of all servers (e.g., there are unexpectedly large numbers of customers coming into the restaurant, thus making it easier for all servers to steal). Suppose the correlation in ϵs is positive. The residual term ϵ_{it} will also be positively correlated with $\overline{Y_{-it}}$.

Finally, another potential source of endogeneity, well known in the peer-effects literature, comes from the potential nonrandom shift assignment for servers, which is similar to the challenge of endogenous network formation. As an example, if a potentially high-theft server always works with high-theft peers, OLS regressions will infer a positive peer effect. This should be less of a concern in our analysis, however, because we have included fixed effects α_i for every server in Equation (2). If the server continuously has a high likelihood of stealing, it should be picked up by the high estimate of α_i. The residual ϵ_{it} represents the deviation from the average theft level of the server. Even though the shift assignments are not random, α_i may correlate with $\overline{Y_{-it}}$, but the residual should not.

We investigate how and in what direction OLS peer effect estimates are biased with a Monte Carlo study that examines the first two potential endogeneity sources. We simplify the simulation with the following model that only includes the actual theft behavior of peers in the shift:

$$Y_{it} = \alpha_i + \beta\overline{Y_{-it}} + \epsilon_{it} \tag{4}$$

In the simulation, we set the true value of β to be 0.3 and assume the residual term ϵ_{it} to be normally distributed, with variance equal to 1. We simulate 100 servers, with each one's fixed effect α_i drawn from a log-normal distribution, with variance equal to 0.25. We then simulate 1,000 shifts, each shift with 3 to 10 individual workers randomly drawn from the server pool. We then simulate ϵ_{it} for each individual in every shift. To investigate how the correlation of ϵs will bias the OLS

Table 2 Monte Carlo Study Results

Model	(1) Scenario 1	(2) Scenario 2	(3) Scenario 3
OLS	0.47(0.02)	0.66(0.01)	-0.23(0.03)
IV	0.30(0.03)	0.306(0.03)	0.298(0.04)

Note: Table presents peer effect coefficients for regressions from 10,000 simulated datasets, where the true parameter is 0.3 and the covariance is either 0.15 (Scenario 1), 0 (Scenario 2), or -0.15 (Scenario 3). Standard errors are presented in parentheses. IV estimates are unbiased, while OLS bias depends on covariance.

estimates, we assume the following scenarios: In Scenario 1, ϵs are independent from one another. In Scenario 2, ϵs are positively correlated, with covariance equal to 0.15. Finally, in Scenario 3, ϵs are negatively correlated, with covariance equal to -0.15. With the simulated data, we then run OLS regressions. Table 2 reports the OLS estimates in the first row under the three scenarios.

Under the assumption of independent ϵs in Scenario 1, the second source of endogeneity does not exist and the bias should come only from the reflection problem. As the true peer effect β is positive, Table 2 shows that OLS will overestimate the peer effect by 56 percent, consistent with our discussion above.

When the ϵs are positively correlated in Scenario 2, OLS will overestimate the peer effect even more. Table 2 shows that the estimate is about 1.1 times higher than the true peer effect. This is because, compared to the case with independent ϵs, the residual term has a higher positive correlation with $\overline{Y_{-it}}$ in this scenario. However, when the ϵs are negatively correlated in Scenario 3, Table 2 shows that OLS actually underestimates the true peer effect. Given that the reflection problem in this scenario will bias the peer effect upward, the result implies that the negative correlation in the ϵs will lead to a downward bias for the OLS estimates and that the extent of the downward bias can dominate the upward bias due to the reflection problem.

To test the robustness of these results, we simulate the three scenarios 10,000 times. Figure 8 graphically illustrates the kernel density of the OLS estimates corresponding to each scenario (the red region in the diagrams). Compared to the true peer effect, represented by the dotted line at 0.3, the figure shows that the distribution of simulated OLS estimates in Scenario 1 (Panel A) and Scenario 2 (Panel B) never overlap with the true peer effect value, demonstrating the extent of the upward bias. In Scenario 3 (Panel C), however, the negative correlation in residuals more than offsets the reflection bias and the distribution of OLS estimates is far below the true β, demonstrating a downward bias. These figures show that negatively correlated residuals are necessary to achieve negative OLS bias with true positive peer effects.

We further investigate how the bias of the OLS estimates relates to the correlation of the ϵs. We conduct a series of simulations by varying the magnitude of the covariance. We fix the true value

Figure 8 **Monte Carlo Study Bootstrapping Results**

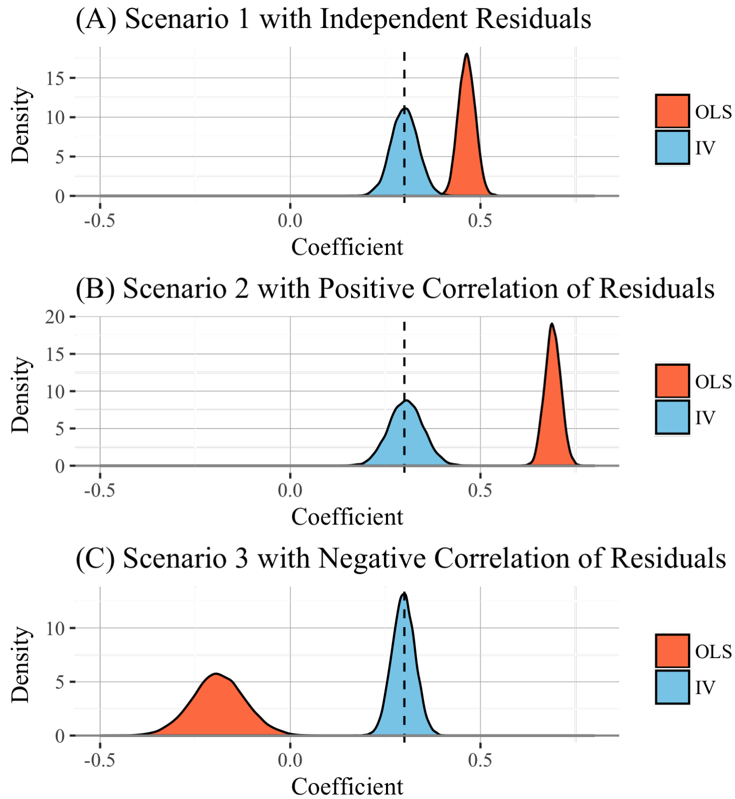


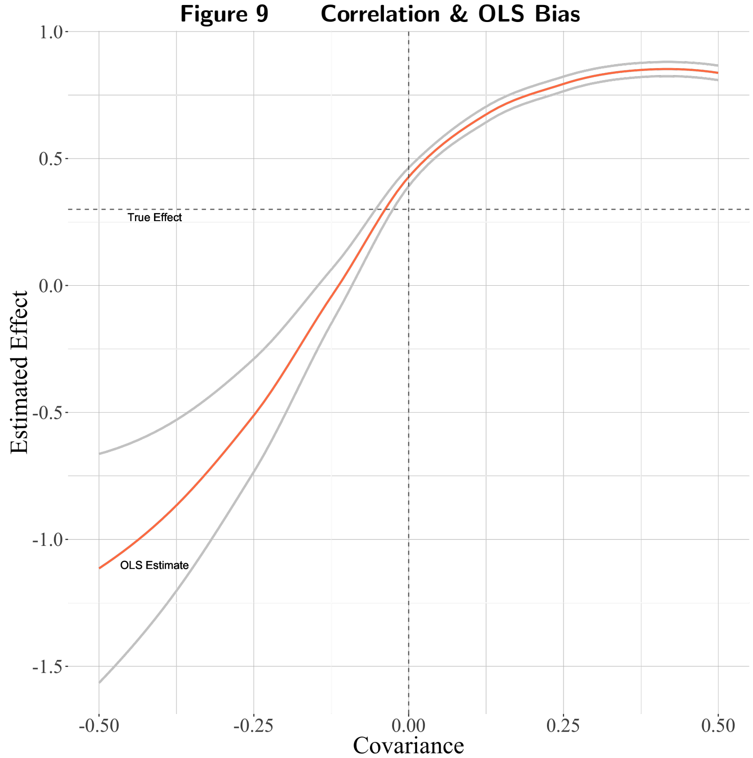
Figure shows results from 10,000 simulations for each of three types of residual correlations among coworkers: none, positive, and negative. Figure indicates that OLS coefficients can only be smaller than IV if residuals are negatively correlated.

of β at 0.3 and gradually increase the covariance of residuals from -0.5 to 0.5. For each covariance value, we run the simulation 1,000 times.

Figure 9 reports the fitted curve of the OLS estimate as the covariance in residuals varies from -0.5 to 0.5 (the gray lines represent 95 percent confidence intervals). The dashed horizontal line represents the true value of β. The OLS estimate is positively biased when the covariance is 0, which indicates that the reflection problem will cause a positive bias in OLS estimates when the true peer-effect coefficient is positive. With positive covariance, the OLS estimate is further upwardly biased. When the negative covariance lies in the range of -0.06 to -0.05, the OLS estimate is close to the true value of 0.3, indicating that weak negative correlation in residuals will offset the upward bias caused by the reflection problem. As negative covariance grows, however, the OLS estimate is biased downward. The series of simulations shows that the only way OLS can be negatively biased from a positive true peer effect is with negatively correlated residuals, and that this negative bias monotonically grows with the strength of the negative correlation.

We also investigate how the relationship between the OLS bias and covariance in residuals varies with different values of the true parameters β. We repeat the above simulation for nine values of

Figure 9 Correlation & OLS Bias



This figure shows how the OLS estimate becomes more downwardly biased as the negative covariance in peer residuals increases, while achieving positive bias as the covariance approaches 0. Estimates are generated from 1,000 simulations for each covariance level, where the true parameter is 0.3.

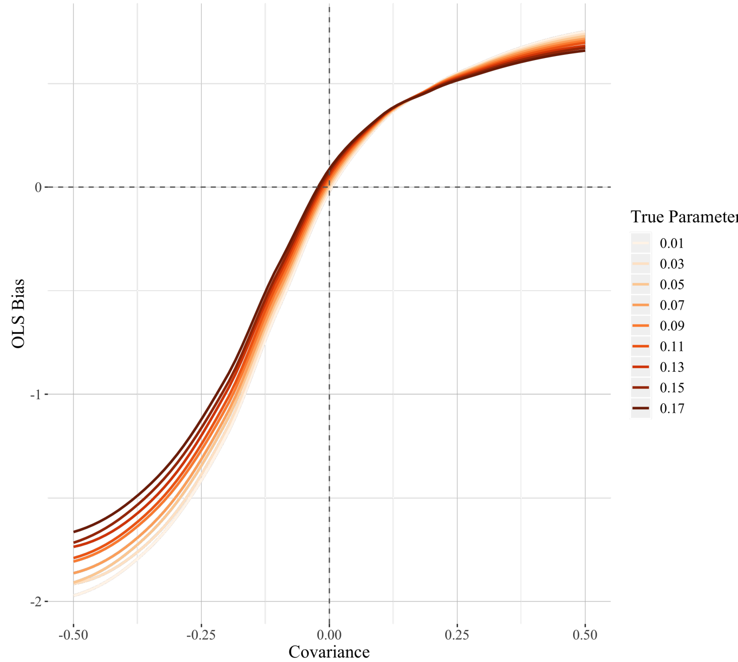
β, from 0.01 to 0.17. Figure 10 reports the fitted curves of OLS estimate bias as the covariance in residuals varies from -0.5 to 0.5 across the different values of βs. We see a nearly identical relationship between the covariance in residuals and the OLS bias for each true parameter β; negative covariance will result in downward bias of OLS estimates and that downward bias grows with the strength of negative correlation. This shows that for the range of these β values, the OLS bias can be used to compare the covariance in residuals across different unbiased estimates. Comparing the curves for different values of β, we find that the effect of negative covariance on the magnitude of downward OLS bias is slightly larger for smaller values of β. [8]

4.3. IV Estimation

To address the potential OLS endogeneity bias shown in Figures 8 through 10, we construct valid instruments for $\overline{Y_{-it}}$ that exploit variation in restaurant shift assignments across time. We start by constructing instruments for the theft of peer j, Y_{jt}. To satisfy the conditions for valid instruments, we need a variable not correlated with ϵ_{it} (i.e., the exclusion restriction) but that has a nontrivial correlation with Y_{jt} (i.e., the relevance restriction). For each peer j in worker i's shift t, we find all

[8] Figure D.4 in the Appendix increases the true peer-effect coefficient β to 0.46.

Figure 10 **Correlation, True Parameter, & OLS Bias**



This figure shows how the OLS bias varies with different covariances in residuals for true parameters ranging from 0.01 to 0.17. Each point in each of the nine curves includes 1,000 simulations.

of j's shifts that satisfy the following two criteria: First, the shift must not include worker i. This is because if i and j work together in another shift s, the reflection problem implies that ϵ_{is} will correlate with Y_{js}, and, if ϵ_{is} and ϵ_{it} are serially correlated, Y_{js} will be correlated with ϵ_{it}. Second, we further require that shift s be at least two weeks before or after shift t. This is because, assuming ϵ_{it} and ϵ_{jt} are correlated and that ϵ_{jt} and ϵ_{js} are serially correlated, Y_{js} will still be correlated with ϵ_{it} even though j and i are not in the same shift s. By requiring shift s to be at least two weeks away from t, the serial correlation will be minimized and thus can satisfy the exclusion restriction.[9]

Next, we take the average of j's theft in all of the shifts that meet the above criteria. That is,

$$\hat{Y_j^{it}} = \frac{\Sigma_{s \in Obs_j^{it}} Y_{js}}{|Obs_j^{it}|} \tag{5}$$

where $|Obs_j^{it}|$ is the number of shifts that meet the criteria and Y_{js} represents the theft in each shift of j that meets the criteria.

Finally, we take the average of $\hat{Y_j^{it}}$ for all of the peers working with the focal server as the instrument for $\overline{Y_{-it}}$ in Equation (2).

$$\widehat{\overline{Y_{-it}}} = \frac{\Sigma_{j \in P_{it}} \hat{Y_j^{it}}}{|P_{it}|} \tag{6}$$

[9] To further test the validity of the instruments, we alternatively use one- and three week windows to construct different IVs in the empirical estimation. We find very similar estimated peer effects, which we present in the Appendix.

The large variation in the daily staffing schedule provides an instrument for most of the observations (70 percent). For those observations for which we cannot find IVs satisfying both of the above criteria, we drop them in the regression analysis. To be valid, the instrument must be correlated with the endogenous variable (i.e., relevance restriction). Conceptually, as $\hat{Y_j^{it}}$ in Equation (5) is a proxy for the theft propensity of j, it is correlated with Y_{jt}. Consequently, the instrument $\widehat{\overline{Y_{-it}}}$ in Equation (6) is also correlated with $\overline{Y_{-it}}$ in Equation (2), thus satisfying the condition. As discussed above, the two selection criteria ensure that $\hat{Y_j^{it}}$ is uncorrelated with ϵ_{is}.

We estimate the IV model using our simulated data from the previous section, employing the instrument constructed from Equations (5) and (6). Given the simulated data construction, our model should satisfy both the relevance and exclusion restrictions. An examination of these data shows that it does. We find that in Scenario 1, (ϵs are independent), $cor(\overline{Y_{-it}}, \widehat{\overline{Y_{-it}}})$ is 0.61 and $cor(\widehat{\overline{Y_{-it}}}, \epsilon_{it})$ is merely 0.001. In Scenario 2 (ϵs are positively correlated), $cor(\overline{Y_{-it}}, \widehat{\overline{Y_{-it}}})$ is 0.48 and $cor(\widehat{\overline{Y_{-it}}}, \epsilon_{it})$ is 0.002. In Scenario 3 (ϵs are negatively correlated), $cor(\overline{Y_{-it}}, \widehat{\overline{Y_{-it}}})$ is 0.65 and $cor(\widehat{\overline{Y_{-it}}}, \epsilon_{it})$ is -0.001. We see a similar pattern in our restaurant data, with a correlation between $\widehat{\overline{Y_{-it}}}$ and $\overline{Y_{-it}}$ of 0.25 when we use theft count as the dependent variable and 0.59 for theft value.

We implement a two-stage least squares (2SLS) method to estimate peer effects for all scenarios and report them in Table 2 (see the second row). In each of the scenarios, the estimated β is very close to 0.3, demonstrating that the proposed IV method can recover the true parameter value if it satisfies the relevance and exclusion restrictions. Finally, we test the robustness of the results by bootstrapping the above three scenarios 1,000 times. The kernel density of the IV estimates corresponding to each scenario is shown in Figure 8 (see the blue region in the diagrams). All of the estimates are very close to 0.3, with the median exactly equal to 0.3, in each of the scenarios.

5. Regression Results and Analysis

In this section, we provide summary tables of the regression models. We provide both clustered and block-bootstrapped standard errors at the restaurant-shift level. We report the F-test statistic of the excluded instrument in the first stage of all IV models. We run Hausman tests to compare the OLS and IV regression for the average peer-effect model, where the null hypothesis is that both OLS and IV estimators are consistent. The Hausman test statistics and P value are reported in each regression table.[10]

5.1. Average Peer Effects

We run peer-effect regressions on theft count and theft value, with the results in Table 3. As discussed in the IV estimation section, we drop all observations for which we cannot find an IV

[10] Significance level $***p < 0.001, **p < 0.01, *p < 0.05, .p < 0.1$ for all statistics in this paper.

Table 3 Average Peer Effect in Theft

	(1)	(2)	(3)	(4)
	DV:Server Theft Count		DV:Server Theft Value	
	OLS	IV	OLS	IV
Avg Peer Theft Count	-0.01***	0.04**		
Clustered Std. Error	(0.002)	(0.015)		
Bootstrapped Std. Error	(0.003)	(0.020)		
Avg Peer Theft Value			0.007***	0.027***
Clustered Std. Error			(0.001)	(0.003)
Bootstrapped Std. Error			(0.002)	(0.006)
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	4,058,783	4,058,783	4,058,783	4,058,783
R^2	0.182	0.125	0.241	0.240
1st Stage F Statistics		1,396***		14,260***
Difference in Estimate	0.05		0.02	
Hausman Test Statistics	10.14***		86.4***	

Note: Standard errors, both clustered and block-bootstrapped at the restaurant-shift level, are presented in parentheses. Significance level:*p<0.1; **p<0.05; ***p<0.01.. Block-bootstrapped standard errors with 100,000 iterations for IV models, 10,000 for OLS models.

satisfying the exclusion restriction, restricting our regression sample to 71 percent of all observations in our data. We find positive and significant peer-effect coefficients in both theft count and value in our unbiased IV models in columns (2) and (4), with the coefficient in column (2) implying that doubling a coworker's theft count increases the likelihood of theft for each peer by 4 percent. Similarly, column (4) implies a 2.7 percent spillover. The OLS regression coefficients in columns (1) and (3), however, are downwardly biased from the IV regression coefficients. If the only source of endogeneity was the reflection problem, under positive peer effects OLS estimates would be larger in magnitude than IV coefficients (as demonstrated in our simulation). The reverse direction of the bias reflects the second source of endogeneity through a negative correlation in daily error terms. From the Hausman test presented, we can reject the null hypothesis for peer effects in theft count and theft value that both OLS and IV estimators are consistent. This difference in coefficients across the two models implies that although working with a high-theft peer increases a worker's theft, that worker's theft on days where the peer steals high values *decreases*.

The validity of our instrument depends on the exclusion restriction not being violated, such that $\hat{Y}_j^{it}$ is uncorrelated with ϵ_{is}. Given the construction of our instrument, an exclusion restriction violation would require two conditions to be true. First, the residuals would need to be serially correlated within workers for longer than our 14-day window. We test this two ways. We estimate serial correlation in the residuals for a series of two-day lags from two to 14 days,[11] which we

[11] We use two-day intervals because most workers don't work on consecutive days.

present in Appendix Figure E.5. Although serial correlation exists with short lags, it is below 0.02 for theft count and 0.005 for theft value at Day 14—the minimum lag used in our instrument construction—before quickly converging to zero. We also conducted a Box-Ljung test for each worker's time series ϵ_{it} with lags between 14 and 28 days. We find identifiable serial correlation (at p=0.05) in only 2.2% of all workers for theft count and 1.6% for theft amount. Collectively, these tests limit concerns of serial correlation violating the exclusion restriction. We also note that even with serial correlation of over two weeks, this process would still require word-of-mouth through which a worker would hear about the actions of peers on shifts she did not work. Although this is certainly possible, in combination with our low serial correlation, we are confident in the validity of our instrument.

5.2. Economic Significance of Estimates

The magnitude of our peer-effect estimates depends on the average peer theft count and value. With an average peer theft count of 0.065, for example, doubling the average theft count will result in a 6 percent increase in the focal worker's theft count. For an average peer theft value of 1.26, doubling this will increase the focal worker's theft value by 2 percent. The magnitude of these peer effects may seem minor, but they are substantially larger when accounting for reflection, where the spillover from one worker to the other will be reflected back and forth over all peers. With reflection, the total impact on theft count and theft value on all workers is far greater than the magnitude of the peer-effect coefficient itself. In fact, the total impact of spillover over all workers depends on several factors, including the number of workers in a shift and the base level of theft of each individual worker free of peer impact.

To demonstrate this overall impact, we run a simulation for a typical shift of seven workers in a restaurant. We set the base theft count and value at 0.16 and 1.67 for each individual worker, which are the average theft count and theft value in our data. For each worker, we then draw an idiosyncratic random shock from the empirical distribution of residuals in our regression results. Given the base theft count (value), random shock, and the peer-effect coefficient, the actual theft count (value) is determined by the system of peer-effect equations, where each worker's theft is affected by the average peer theft, while at the same time each worker also affects theft by peers. Thus to calculate the actual theft count (value) with endogenous peer effects, we need to solve the fixed point of the system of peer-effect equations. We run the above simulation 10,000 times. On average, comparing to the case where there is no peer effect, total theft count across all workers increases by 34 percent and total theft value increases by 200 percent.

We run two other simulations to show the impact of an increase in an individual worker's theft has across all peers. In the first, we double one of the seven worker's base theft count (value) from the first simulation while keeping the other workers, variables, and coefficient constant. We solve for the actual theft count (value) for all workers, with 10,000 simulations. On average, doubling the base theft count of one worker will increase total restaurant theft count by 20 percent and theft amount by 76 percent, compared to 14 percent increases without peer effects. Our second simulation doubles all workers' theft levels, which produces increases of 150 percent in total theft count and 550 percent in total theft amount.[12]

These simulations demonstrate that the overall impact of theft peer effects is significantly larger than the marginal effect implied by the peer-effect coefficient in estimation results and has substantial implications for overall operational performance.

5.3. Testing for Nonlinear Peer Effects

To address the possibility of nonlinear peer effects that Tan and Netessine (2019) raise, we adjust our models to include quadratic terms for peer theft. We rerun both the IV and OLS models and present results in Table 4 with standard errors clustered or block-bootstrapped at restaurant-shift level with 100,000 iterations.

Table 4 Average Peer Effect in Theft

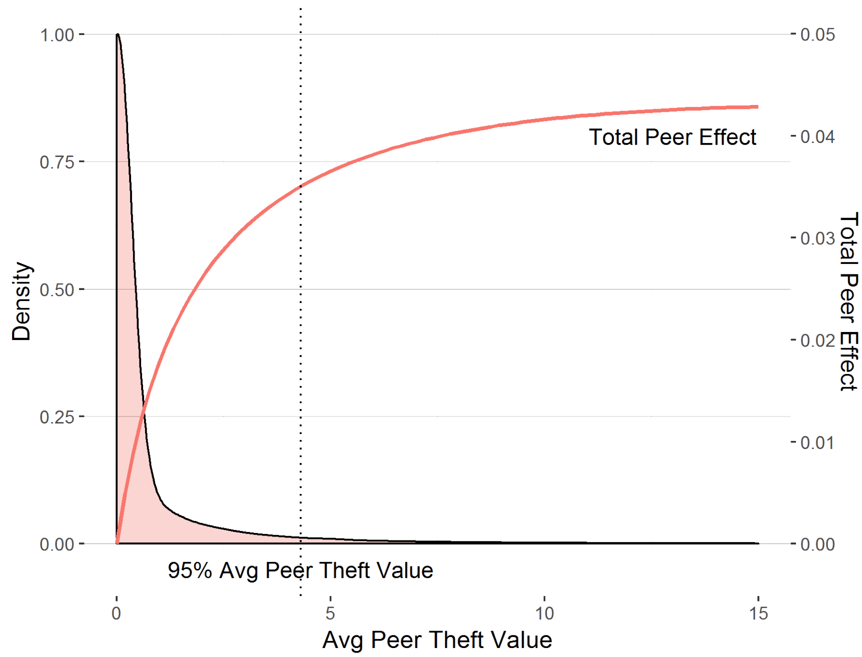
	DV:Server Theft Count			DV:Server Theft Value		
	(1)	(2)	(3)	(4)	(5)	(6)
Model	Estimate	Clustered Std. Error	Bootstrapped Std. Error	Estimate	Clustered Std. Error	Bootstrapped Std. Error
OLS						
$PeerTheft$	-0.01***	0.003	0.005	0.0028*	0.001	0.002
$PeerTheft^2$	-0.0017	0.003	0.006	−0.002***	0.001	0.001
IV						
$PeerTheft$	0.07***	0.016	0.021	0.029***	0.004	0.006
$PeerTheft^2$	-0.06***	0.010	0.026	−0.005***	0.001	0.002

Note: Standard errors, clustered or block-bootstrapped at the restaurant-shift level, are presented in parentheses. Significance level:*p<0.1; **p<0.05; ***p<0.01. Block-bootstrapped standard errors with 100,000 iterations for IV models, 10,000 for OLS models.

We consistently find positive linear coefficients and negative quadratic coefficients, indicating that the magnitude of the total peer effect increases as the peers' average count or value of theft increases. Although these results indicate some concavity over a broad range of theft, we show in Figure 11 that, unlike in Tan and Netessine (2019), the total peer effect for theft value is always increasing over our data support, with the dotted line representing the 95th percentile of average peers' theft count and value. Figure B.1 in the Appendix shows a very similar relationship for theft count. For both theft count and theft value, even though the derivative of a focal worker's theft to average peer theft decreases as average peer theft increases, the total peer effect increases.

[12] See Appendix, Section D, for an example of this compounding.

Figure 11 Total Peer Effect in Theft Value



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

5.4. Peer Effect Based on Employee Tenure

Evidence from studies on productivity peer effects shows that new workers are the most likely to be influenced by peers (e.g., Chan et al. 2014b). To investigate whether peer effects differ in magnitude for peers with different experience, we look at new employees by separately estimating peer effects in 30-day time windows using our IV model. We identify start dates based on the first observation of each employee at the restaurant in the data set. If the first observation for an employee is at least 30 days later than the first of any observation at the restaurant, we define the employee as new so as to avoid left truncation. We examine the number of days between any two consecutive observations for a particular employee and find that over 99% of the observations are smaller than 14 days (two weeks), indicating employees seldom take a break from work for more than two weeks before returning. In our data set, 84 percent of employees are defined as new employees. We restrict our regression to the 9.4 percent of new employees who stay for at least six full months in order to compare monthly peer effects across a consistent sample.[13] The number of observations we use for estimating peer effects conditional on tenure is 24 percent of the observations in our main regressions.

The regression results for the 30-day interval models are given in Tables 5 and 6. In models predicting theft count, peer-effect estimates are stable and statistically significant for the first

[13] See Figure B.2 in the Appendix for survival rates for high- and low-theft employees. High-theft employees stay longer.

Table 5 Peer Effects in Theft Count by Month for New Employees Staying Six Months or More

	(1) 1st Month	(2) 2nd Month	(3) 3rd Month	(4) 4th Month	(5) 5th Month	(6) 6+ Months
Avg Peer Theft Count	0.168***	0.128**	0.155**	0.093	0.041	0.032
Clustered Std. Error	(0.058)	(0.062)	(0.063)	(0.062)	(0.054)	(0.058)
Bootstrapped Std. Error	(0.082)	(0.086)	(0.092)	(0.081)	(0.075)	(0.056)
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	127,706	133,458	136,148	137,953	138,644	305,492
R^2	0.248	0.266	0.274	0.288	0.284	0.201
1st Stage F Statistics	425***	378***	402***	459***	585***	1,323***

Note: Estimates are from IV models. Standard errors, block-bootstrapped with 10,000 iterations at the restaurant-shift level, are in parentheses. Significance level:*p<0.1; **p<0.05; ***p<0.01.

Table 6 Peer Effects in Theft Value by Month for New Employees Staying Six Months or More

	(1) 1st Month	(2) 2nd Month	(3) 3rd Month	(4) 4th Month	(5) 5th Month	(6) 6+ Months
Avg Peer Theft Value	0.029*	0.031*	0.032*	0.044***	0.033*	-0.004
Clustered Std. Error	(0.018)	(0.019)	(0.018)	(0.016)	(0.017)	(0.01)
Bootstrapped Std. Error	(0.022)	(0.24)	(0.023)	(0.024)	(0.024)	(0.014)
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	127,706	133,458	136,148	137,953	138,644	305,492
R^2	0.319	0.316	0.332	0.344	0.363	0.332
1st Stage F Statistics	2,437***	2,327***	2,550***	3,373***	3,118***	13,010***

Note: Estimates are from IV models. Standard errors, clustered or block-bootstrapped with 10,000 iterations at the restaurant-shift level, are in parentheses. Significance level:*p<0.1; **p<0.05; ***p<0.01.

three months of employment before decreasing and losing precision. Theft value regressions show similar results, with peer effects consistent through the first five months then nonexistent after that. Collectively, these models support the idea that new employees are the most vulnerable to the norms and behavior of their peers. We caution, however, that experience at a given restaurant does not necessarily reflect overall experience as a server, which likely biases against us finding differences in our models. We also note that since attrition is endogenous, meaning that those who stay six months or more are different in many ways from those who leave earlier, we cannot easily generalize to all workers.[14]

[14] See the Appendix for regressions including new employees who left before six months.

5.5. Peer Effects Before/After IT Theft-Monitoring Implementation

We examine whether managerial oversight explains the negative correlation in same-day peer theft by estimating peer effects before and after restaurants implement an IT-monitoring system that increases the risk of detection (and likely punishment) for theft. If managers are more likely to intervene when observing high theft levels in a restaurant on a given day, this could explain the strategic peer response where servers reduce theft when their peers steal more. To test this, we separately run regressions before and after the IT theft-monitoring system implementation and compare the bias of OLS from IV regression. We drop all restaurants that never implemented the monitoring system (7 percent of the observations) because adoption is endogenous and the non-adopters consequently form a problematic control group. Furthermore, we can exploit the quasi-random staggered adoption rates to differentiate adoption from simple time trends.[15] The results, shown in Table 7, provide two sets of results. First, the true peer effect in the IV regressions is much larger following monitoring adoption. We note that this does not indicate a higher theft level (which is lower following monitoring adoption), but rather that peer theft is more dependent on the identity of peers after IT-monitoring adoption. This could be because in the presence of monitoring, theft decreases more strongly in the absence of high-theft peers. Second, the downward bias in the OLS coefficient estimates is larger after the monitoring system is introduced, which suggests that increased managerial oversight increases the same-day strategic peer responses. This supports the idea that monitoring and oversight constrain total theft in the organization, forcing employees to strategically respond to the misconduct of peers.[16]

6. Discussion and Conclusion

This paper provides the first evidence that the influence of peers in worker misconduct is more complex than the productivity spillovers identified in studies from operations, economics, and management. Our service operations setting shows that although high-theft peers indeed encourage more theft in coworkers, a second peer influence exists through strategic peer responses. These strategic peer responses, in the presence of managerial oversight and monitoring, produce negative same-day correlations in peer theft as workers reduce theft *on a given day* when peers steal more. These unique dynamics are identifiable because our simulations show that the difference between OLS and IV models reflects the same-day correlation in worker theft. Our monitoring explanation

[15] Adoption timing is quasi-random because of the way in which the chains introduced the monitoring system. See Pierce et al. (2015).

[16] We present in the Appendix additional analysis with restaurants split by median workforce size. Similar to the IT monitoring result, small restaurants show less bias in OLS estimates, which is consistent with lower monitoring. One possible explanation is the frequent absence of a specialized manager in small restaurants, such that the "manager" is frequently a deputized server. But this explanation is speculative, and others are reasonable as well.

Table 7 Peer Effect Before/After Theft-Monitoring System Implementation

	DV:Server Theft Count				DV:Server Theft Value			
	Before		After		Before		After	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Avg Peer Theft Count	-0.014**	0.0028	-0.018***	0.074***				
Clustered Std. Error	(0.005)	(0.02)	(0.002)	(0.02)				
Bootstrapped Std. Error	(0.006)	(0.03)	(0.003)	(0.034)				
Avg Peer Theft Value					-0.003***	0.03***	-0.006***	0.029***
Clustered Std. Error					(0.001)	(0.005)	(0.002)	(0.005)
Bootstrapped Std. Error					(0.002)	(0.006)	(0.002)	(0.008)
Individual Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Manager Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*Week Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Restaurant*Weekday Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	718,409	718,409	3,082,061	3,082,061	718,409	718,409	3,082,061	3,082,061
R^2	0.164	0.215	0.095	0.105	0.284	0.371	0.274	0.193
1st Stage F Statistics		398***		1,202***		7,158***		7,808***
Difference in Estimates	0.017		0.092		0.033		0.035	
Hausman Test Statistics	0.75		23.7***		45.3***		61.2***	

Note: Standard errors, clustered or block-bootstrapped at the restaurant-shift level, are in parentheses. Significance level:*p<0.1; **p<0.05; ***p<0.01. Block-bootstrapped standard errors with 100,000 iterations for IV models, 10,000 for OLS models.

is supported by the increased size of this difference following adoption of an IT-monitoring system that raises the risk of theft detection.

What do our findings imply? First, we confirm that peer effects in misconduct are important in the same ways as the productivity peer effects in prior work. Few studies have shown this in employees (Ichino and Maggi 2000, Dimmock et al. 2018), with none that we are aware of at the scale or level of detail as the present study. This field evidence in a service operations setting is important because it validates long-held beliefs in the behavioral ethics literature that “bad apples” have a broad influence on an organization beyond the direct outcomes of their own unethical actions (Gino et al. 2009, Kish-Gephart et al. 2010, Treviño and Youngblood 1990, Moore and Gino 2013, Pinto et al. 2008, Brass et al. 1998). We note that the organizational costs of bad apples may be even greater if peer effects dynamically increase coworker misconduct traits through learning processes or changes in norms as in Chan et al. (2014b) or Hasan and Koning (2017).

Second, our identification of the importance of reflection effects highlights how even small peer effects can multiply with large organizational implications. An unethical employee may affect many peers simultaneously, who in turn will affect others. It is this contagion or normalization of corruption that can generate astounding levels of misconduct within organizations (Pinto et al. 2008).

Third, our results suggest that managerial oversight not only reduces average misconduct levels (e.g., Pierce et al. 2015, Nagin et al. 2002), but also can constrain daily escalation in bad behavior among peers. Oversight, such as monitoring, changes the calculus for employees because it positively ties the risk of detection to the misconduct of peers. It is this strategic peer response that we observe in our biased OLS estimates. In settings where low levels of misconduct are acceptable but higher ones expose the firm to substantial risk of loss, the effects we identify may shorten periods of peak misconduct, instead smoothing these costly behaviors across time.

Fourth, we provide further evidence that new employees are particularly vulnerable to coworker norms and behavior. Firms must consider carefully with whom to staff new workers, given the potential of peers to shape worker behavior. These results also support the importance of successfully onboarding new workers through both formal training and mentoring processes (Cable et al. 2013).

Finally, we provide an important methodological contribution by demonstrating how the biases in standard OLS peer effect models can be used to reveal strategic interactions among peers in a given shift or day. Although scholars have long understood the importance of correcting for these biases (Manski 1993), our simulations show that the magnitude and direction of the biases can also reveal mechanisms that shape peer interactions.

We acknowledge that we are unable to precisely separate the many mechanisms through which workers' misconduct influences peer behavior. These mechanisms, which include social pressure, shame, helping, coordination, knowledge transfer, and others, are well established in prior work on productivity, but we are not able to differentiate them in our particular setting with the data available. Our evidence on new workers provides perhaps the best attempt at this, but there are reasons to believe new workers would be most strongly influenced by all of these. The peer effects identified in our models may also persist across multiple days, but such persistence will not bias our estimates because of the two-week window in our instrument, which we demonstrated was sufficiently long to effectively eliminate any serial correlation.

One challenge of our large-scale operational setting is that it is not a randomized experiment. Consequently, although the setting is important and generalizable, we are unable to control for either the hiring process or the staffing decisions, and, as we show in the Appendix, attrition is correlated with theft levels. As we note earlier, because our regressions control for the identities of focal and peer workers, these are unlikely to present major biases in our estimation. The one concern would be if managers were strategically staffing workers based on both their average theft levels and their vulnerability to peer influence. Unlike in productivity studies, however, this bias would work against our results. Endogenous staffing presents a problem for productivity peer-effect models because managers might strategically staff their best workers with those most likely to benefit from these star coworkers, thereby biasing estimates upward. If such endogenous staffing is present in our study, managers would instead likely staff the highest theft employees with those who are *least likely* to be influenced, biasing our estimates downward. So, if anything, our estimates would be smaller than the true effect.

Finally, we emphasize that the magnitude of our effects has significant implications for firms seeking to reduce misconduct among teams of workers. Our coefficients imply spillovers of between 2.5 percent and 4 percent, which, although smaller than prior estimates of productivity peer effects (Herbst and Mas 2015), are still important in low-margin industries, such as our restaurant setting. As we demonstrate, even small-magnitude coefficients have major implications for organizations due to each worker having multiple peers and the reflection effects between them. One weakness in interpreting our estimate magnitudes is that our theft measures represent only a small portion of the likely theft occurring in these restaurants because the data provider's forensic algorithms are necessarily conservative. Furthermore, our data cannot detect an equally, if not more, important type of theft occurring in parallel—inventory theft. Yet if our peer-effect estimates apply equally to these hidden types of theft, as well as to other important types of misconduct in service settings, such as sexual harassment and abuse, then even small increases in peer misconduct can have considerable implications for organizations.

How can managers apply our results to operations management? The first clear and important implication is that "bad apples" with high levels of misconduct are even more costly than their individual behavior. They also propagate a culture of misconduct, increasing the bad behavior of those around them. Managers must recognize that removing such high-misconduct workers from their organization is of paramount importance when that misconduct is contagious to peers, and that the continued employment of these workers cannot be justified by their contributions being greater than their individual bad behavior. This is an important implication because managers commonly must evaluate this trad-eoff on a variety of antisocial or illegal behaviors. Second, our results emphasize the importance of matching new or other easily influenced workers with those whose contributing and prosocial behavior will generate productive peer effects rather than those of misconduct. If high-misconduct employees cannot be separated from the organization, isolating them can potentially mitigate the peer effects in misconduct observed in this paper. Our paper also implies that although monitoring can't stop theft, it can constrain it on a given day. But even with effective monitoring, the worst employees will still have a strong influence on peers.

References

Aksin, O Zeynep, Sarang Deo, Jónas Jónasson, Kamalini Ramdas. 2015. Learning from many: Partner diversity and team familiarity in fluid teams. *Unpublished Working Paper* .

Alm, James, Kim M Bloomquist, Michael McKee. 2017. When you know your neighbour pays taxes: Information, peer effects and tax compliance. *Fiscal Studies* **38**(4) 587–613.

Anteby, Michel, Curtis K Chan. 2018. A self-fulfilling cycle of coercive surveillance: Workers' invisibility practices and managerial justification. *Organization Science* **29**(2) 247–263.

Aven, Brandy L. 2015. The paradox of corrupt networks: An analysis of organizational crime at Enron. *Organization Science* **26**(4) 980–996.

Azoulay, Pierre, Joshua S Graff Zivin, Jialan Wang. 2010. Superstar extinction. *The Quarterly Journal of Economics* **125**(2) 549–589.

Balafoutas, Loukas, Adrian Beck, Rudolf Kerschbamer, Matthias Sutter. 2013. What drives taxi drivers? a field experiment on fraud in a market for credence goods. *Review of Economic Studies* **80**(3) 876–891.

Bandura, Albert. 2014. Social cognitive theory of moral thought and action. *Handbook of Moral Behavior and Development*. Psychology Press, 69–128.

Batt, Robert J, Diwas S Kc, Bradley R Staats, Brian W Patterson. 2019. The effects of discrete work shifts on a nonterminating service system. *Production and Operations Management* **28**(6) 1528–1544.

Bayer, Patrick, Randi Hjalmarsson, David Pozen. 2009. Building criminal capital behind bars: Peer effects in juvenile corrections. *The Quarterly Journal of Economics* **124**(1) 105–147.

Bennett, Victor Manuel, Lamar Pierce, Jason A Snyder, Michael W Toffel. 2013. Customer-driven misconduct: How competition corrupts business practices. *Management Science* **59**(8) 1725–1742.

Bernstein, Ethan S. 2012. The transparency paradox: A role for privacy in organizational learning and operational control. *Administrative Science Quarterly* **57**(2) 181–216.

Bernstein, Ethan S. 2017. Making transparency transparent: The evolution of observation in management theory. *Academy of Management Annals* **11**(1) 217–266.

Brass, Daniel J, Kenneth D Butterfield, Bruce C Skaggs. 1998. Relationships and unethical behavior: A social network perspective. *Academy of Management Review* **23**(1) 14–31.

Cable, Daniel M, Francesca Gino, Bradley R Staats. 2013. Reinventing employee onboarding. *MIT Sloan Management Review* **54**(3) 23.

Chan, Tat Y, Jia Li, Lamar Pierce. 2014a. Compensation and peer effects in competing sales teams. *Management Science* **60**(8) 1965–1984.

Chan, Tat Y, Jia Li, Lamar Pierce. 2014b. Learning from peers: Knowledge transfer and sales force productivity growth. *Marketing Science* **33**(4) 463–484.

Chatterji, Aaron, Solène Delecourt, Sharique Hasan, Rembrand Koning. 2019. When does advice impact startup performance? *Strategic Management Journal* **40**(3) 331–356.

Cohen, Lauren, Andrea Frazzini, Christopher Malloy. 2010. Sell-side school ties. *The Journal of Finance* **65**(4) 1409–1437.

DeHoratius, Nicole, Ananth Raman. 2007. Store manager incentive design and retail performance: An exploratory investigation. *Manufacturing & Service Operations Management* **9**(4) 518–534.

DeHoratius, Nicole, Ananth Raman. 2008. Inventory record inaccuracy: an empirical analysis. *Management Science* **54**(4) 627–641.

Derfler-Rozin, Rellie, Celia Moore, Bradley R Staats. 2016. Reducing organizational rule breaking through task variety: How task design supports deliberative thinking. *Organization Science* **27**(6) 1361–1379.

Detert, James R, Linda K Treviño, Ethan R Burris, Meena Andiappan. 2007. Managerial modes of influence and counterproductivity in organizations: A longitudinal business-unit-level investigation. *Journal of Applied Psychology* **92**(4) 993.

Dimmock, Stephen G, William C Gerken, Nathaniel P Graham. 2018. Is fraud contagious? coworker influence on misconduct by financial advisors. *The Journal of Finance* **73**(3) 1417–1450.

Edelman, Benjamin, Ian Larkin. 2014. Social comparisons and deception across workplace hierarchies: Field and experimental evidence. *Organization Science* **26**(1) 78–98.

Francis, P, RC DeGlinkta. 2004. *How to Burn Down the House: The Infamous Waiter and Bartender's Scam Bible by Two Bourbon Street Waiters*. Promethean Books, New Orleans, LA.

Garber, A, C Walkup. 2004. Theft protection plan key to locking in profits. *Nation's Restaurant News* **38**(21) 92–98.

Gino, Francesca, Shahar Ayal, Dan Ariely. 2009. Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological science* **20**(3) 393–398.

Gino, Francesca, Lamar Pierce. 2009. Dishonesty in the name of equity. *Psychological Science* **20**(9) 1153–1160.

Green, Linda V, Sergei Savin, Nicos Savva. 2013. "nursevendor problem": Personnel staffing in the presence of endogenous absenteeism. *Management Science* **59**(10) 2237–2256.

Hasan, Sharique, Rembrand Koning. 2017. Prior ties and the limits of peer effects on startup team performance. *Strategic Management Journal* **40**(9) 1394—-1416.

Herbst, Daniel, Alexandre Mas. 2015. Peer effects on worker output in the laboratory generalize to the field. *Science* **350**(6260) 545–549.

Huckman, Robert S, Bradley R Staats. 2011. Fluid tasks and fluid teams: The impact of diversity in experience and team familiarity on team performance. *Manufacturing & Service Operations Management* **13**(3) 310–328.

Huckman, Robert S, Bradley R Staats, David M Upton. 2009. Team familiarity, role experience, and performance: Evidence from indian software services. *Management Science* **55**(1) 85–100.

Ichino, Andrea, Giovanni Maggi. 2000. Work environment and individual background: Explaining regional shirking differentials in a large italian firm. *The Quarterly Journal of Economics* **115**(3) 1057–1090.

KC, Diwas, Bradley R Staats, Francesca Gino. 2013. Learning from my success and from others' failure: Evidence from minimally invasive cardiac surgery. *Management Science* **59**(11) 2435–2449.

Kesavan, Saravanan, Bradley R Staats, Wendell Gilland. 2014. Volume flexibility in services: The costs and benefits of flexible labor resources. *Management Science* **60**(8) 1884–1906.

Kish-Gephart, Jennifer J, David A Harrison, Linda Klebe Treviño. 2010. Bad apples, bad cases, and bad barrels: Meta-analytic evidence about sources of unethical decisions at work. *Journal of Applied Psychology* **95**(1) 1–31.

Manski, Charles F. 1993. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies* **60**(3) 531–542.

Marquis, Christopher, András Tilcsik. 2013. Imprinting: Toward a multilevel theory. *Academy of Management Annals* **7**(1) 195–245.

Mas, Alexandre, Enrico Moretti. 2009. Peers at work. *American Economic Review* **99**(1) 112–145.

Mohliver, Aharon. 2019. How misconduct spreads: auditors' role in the diffusion of stock-option backdating. *Administrative Science Quarterly* **64**(2) 310–336.

Moon, Ken, Patrick Bergemann, Daniel Brown, Andrew Chen, James Chu, Ellen Eisen, Gregory Fischer, Prashant Kumar Loyalka, Sungmin Rho, Joshua Cohen. 2018. Manufacturing productivity with worker turnover. *Available at SSRN 3248075* .

Moore, Celia, Francesca Gino. 2013. Ethically adrift: How others pull our moral compass from true north, and how we can fix it. *Research in Organizational Behavior* **33** 53–77.

Nagin, Daniel S, James B Rebitzer, Seth Sanders, Lowell J Taylor. 2002. Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *American Economic Review* **92**(4) 850–873.

Oettl, Alexander. 2012. Reconceptualizing stars: Scientist helpfulness and peer performance. *Management Science* **58**(6) 1122–1140.

Palmer, Donald, Christopher B Yenkey. 2015. Drugs, sweat, and gears: An organizational analysis of performance-enhancing drug use in the 2010 Tour de France. *Social Forces* **94**(2) 891–922.

Pierce, Lamar, Daniel C Snow, Andrew McAfee. 2015. Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Science* **61**(10) 2299–2319.

Pierce, Lamar, Jason Snyder. 2008. Ethical spillovers in firms: Evidence from vehicle emissions testing. *Management Science* **54**(11) 1891–1903.

Pierce, Lamar, Jason A Snyder. 2015. Unethical demand and employee turnover. *Journal of Business Ethics* **131**(4) 853–869.

Pierce, Lamar, Laura W Wang, Dennis Zhang. 2019. Peer bargaining and productivity in teams: Gender and the inequitable division of pay. *Manufacturing & Service Operations Management* Forthcoming.

Pinto, Jonathan, Carrie R Leana, Frits K Pil. 2008. Corrupt organizations or organizations of corrupt individuals? two types of organization-level corruption. *Academy of Management Review* **33**(3) 685–709.

Ranganathan, Aruna, Alan Benson. 2019. A numbers game: Quantification of work, accidental gamification, and worker productivity. *Unpublished Working Paper* .

Siemsen, Enno, Sridhar Balasubramanian, Aleda V Roth. 2007. Incentives that induce task-related effort, helping, and knowledge sharing in workgroups. *Management Science* **53**(10) 1533–1550.

Siemsen, Enno, Aleda V Roth, Sridhar Balasubramanian. 2008. How motivation, opportunity, and ability drive knowledge sharing: The constraining-factor model. *Journal of Operations Management* **26**(3) 426–445.

Song, Hummy, Anita L Tucker, Karen L Murrell, David R Vinson. 2017. Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices. *Management Science* **64**(6) 2628–2649.

Staats, Bradley R, Hengchen Dai, David Hofmann, Katherine L Milkman. 2016. Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Science* **63**(5) 1563–1585.

Tan, Tom, Serguei Netessine. 2019. When you work with a superman, will you also fly? an empirical study of the impact of coworkers on performance. *Management Science* Forthcoming.

Tan, Tom, Bradley R Staats. 2016. Behavioral drivers of routing decisions: Evidence from restaurant table assignment .

Tan, Tom Fangyun, Serguei Netessine. 2014. When does the devil make work? an empirical study of the impact of workload on worker productivity. *Management Science* **60**(6) 1574–1593.

Treviño, Linda K, Stuart A Youngblood. 1990. Bad apples in bad barrels: A causal analysis of ethical decision-making behavior. *Journal of Applied Psychology* **75**(4) 378.

Tucker, Anita L, Ingrid M Nembhard, Amy C Edmondson. 2007. Implementing new practices: An empirical study of organizational learning in hospital intensive care units. *Management Science* **53**(6) 894–907.

Valentine, Melissa A, Tom Fangyun Tan, Bradley R Staats, Amy C Edmondson. 2018. Fluid teams and knowledge retrieval: Scaling service operations. *Manufacturing & Service Operations Management* **21** 346–360.

Victor, Bart, Linda Klebe Treviño, Debra L Shapiro. 1993. Peer reporting of unethical behavior: The influence of justice evaluations and social context factors. *Journal of Business Ethics* **12**(4) 253–263.

Xu, Yuqian, Tom Tan, Serguei Netessine. 2017. When is the root of all evil not money? the impact of load on operational risk at a commercial bank. *SMU Cox School of Business Research Paper* (18-13).

Yenkey, Christopher B. 2015. Mobilizing a market: Ethnic segmentation and investor recruitment into the Nairobi securities exchange. *Administrative Science Quarterly* **60**(4) 561–595.

Yenkey, Christopher B. 2018. Fraud and market participation: social relations as a moderator of organizational misconduct. *Administrative Science Quarterly* **63**(1) 43–84.

Yin, Yue, Itai Gurvich, Stephanie McReynolds, Debora Seys, Jan A Van Mieghem. 2018. Learning by doing versus learning by viewing: An empirical study of data analyst productivity on a collaborative platform at ebay. *Proceedings of the ACM on Human-Computer Interaction* **2**(CSCW) 193.

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

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.

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

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

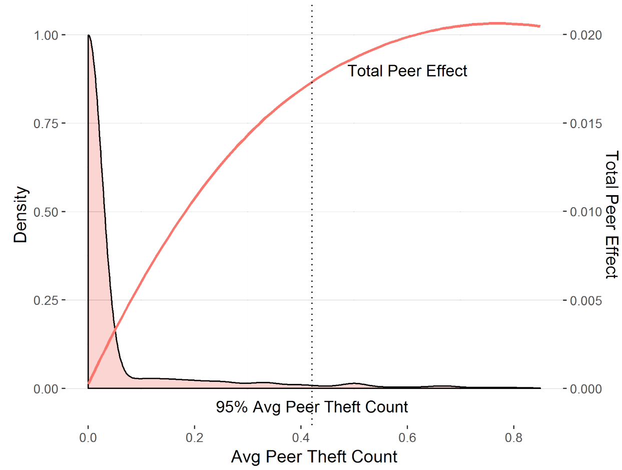
	DV:Server Add-on Sales		DV:Server Drink Sales	
	OLS	IV	OLS	IV
Avg Peer Amount	0.20^{***}	0.169^{***}	0.20^{***}	0.19^{***}
	(0.005)	(0.029)	(0.02)	(0.04)
Observations	3769578	3769578	3769578	3769578
R^2	0.652	0.647	0.558	0.556
1st Stage F Statistics		151.3^{***}		280.4^{***}
Difference in Estimate	-0.031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

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.

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 **??**. 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

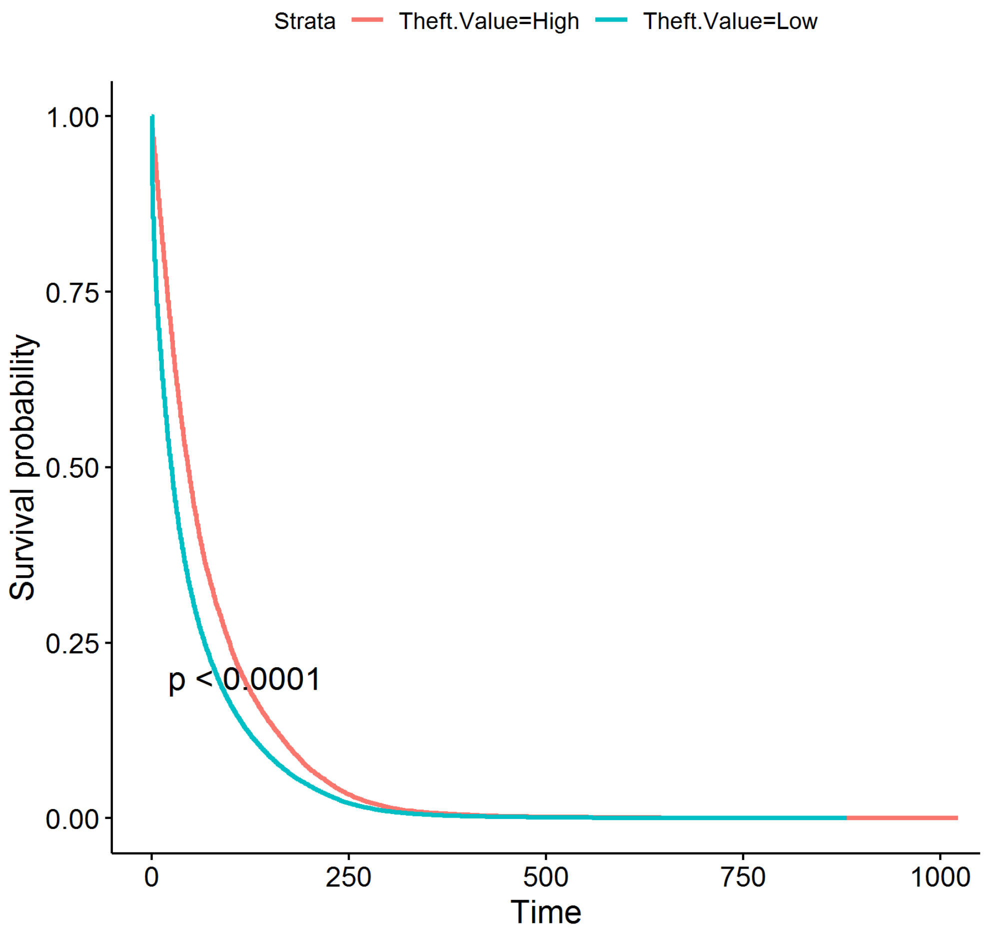


Figure shows raw survival data for all high- and low-theft workers, defined by whether their average daily theft value is below or above the median worker.

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.

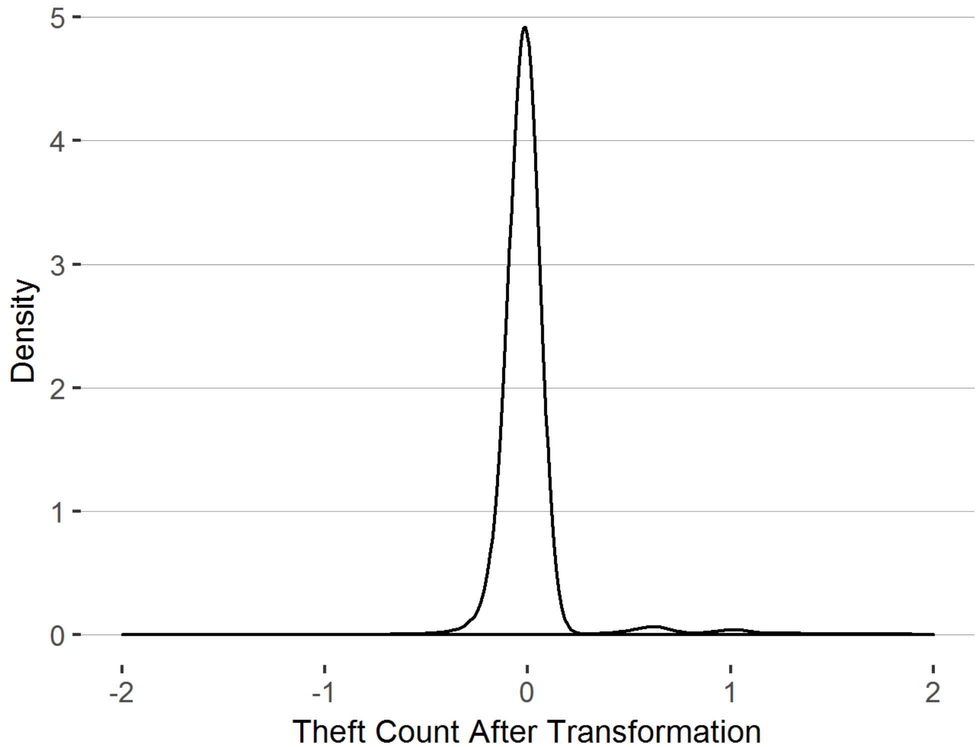
D. Calculating Total Impact of Peer Effects

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 (7).

$$\begin{cases} Y_i = \beta Y_j + \alpha_i \\ Y_j = \beta Y_i + \alpha_j \end{cases} \tag{7}$$

Solving the equation, we have

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



$$\begin{pmatrix} Y_i \\ Y_j \end{pmatrix} = \begin{pmatrix} \frac{\beta\alpha_j+\alpha_i}{1-\beta^2} \\ \frac{\beta\alpha_i+\alpha_j}{1-\beta^2} \end{pmatrix} \tag{8}$$

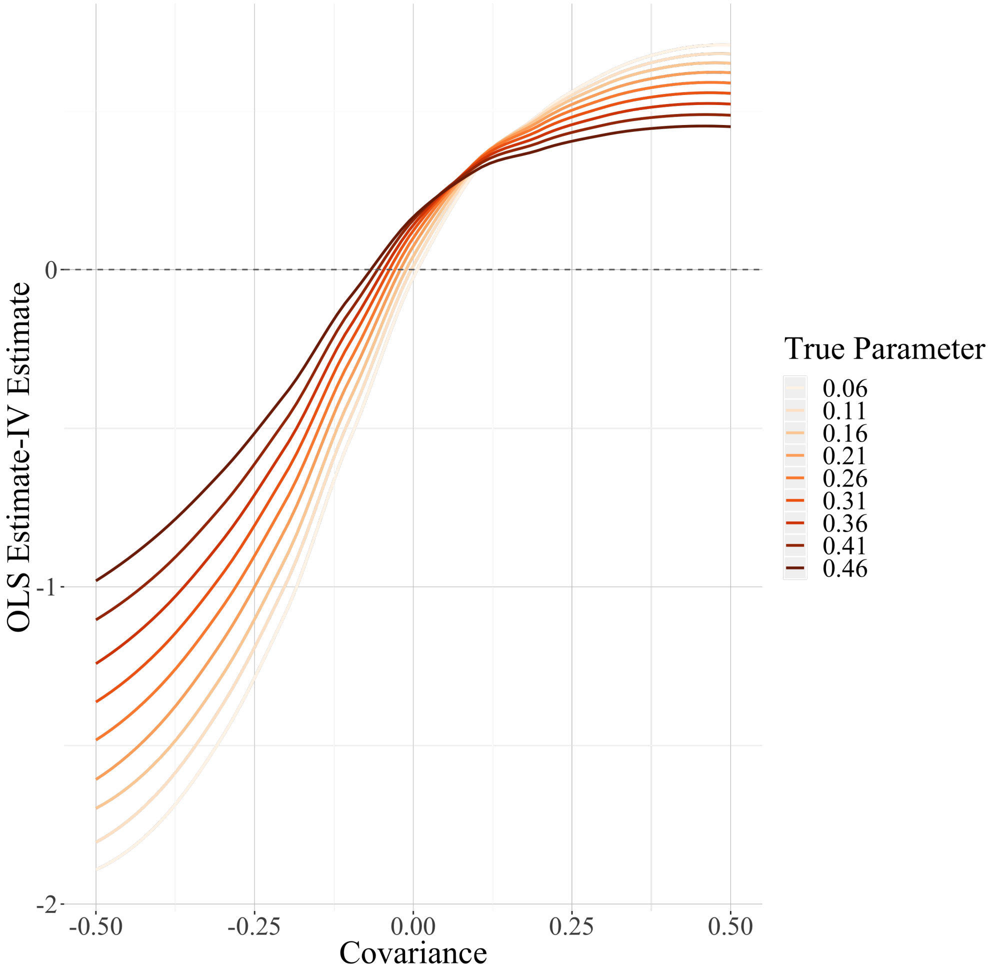
Now suppose the intrinsic theft level of worker i, α_i increases by an amount of Δ. Solving Equations (7) again, we have the new theft level $(Y_i^{'}, Y_j^{'})$ as

$$\begin{pmatrix} Y_i^{'} \\ Y_j^{'} \end{pmatrix} = \begin{pmatrix} \frac{\beta\alpha_j+\alpha_i+\Delta}{1-\beta^2} \\ \frac{\beta(\alpha_i+\Delta)+\alpha_j}{1-\beta^2} \end{pmatrix} \tag{9}$$

Comparing the theft levels as shown in Equations (8) and (9), 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 (7); 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.

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

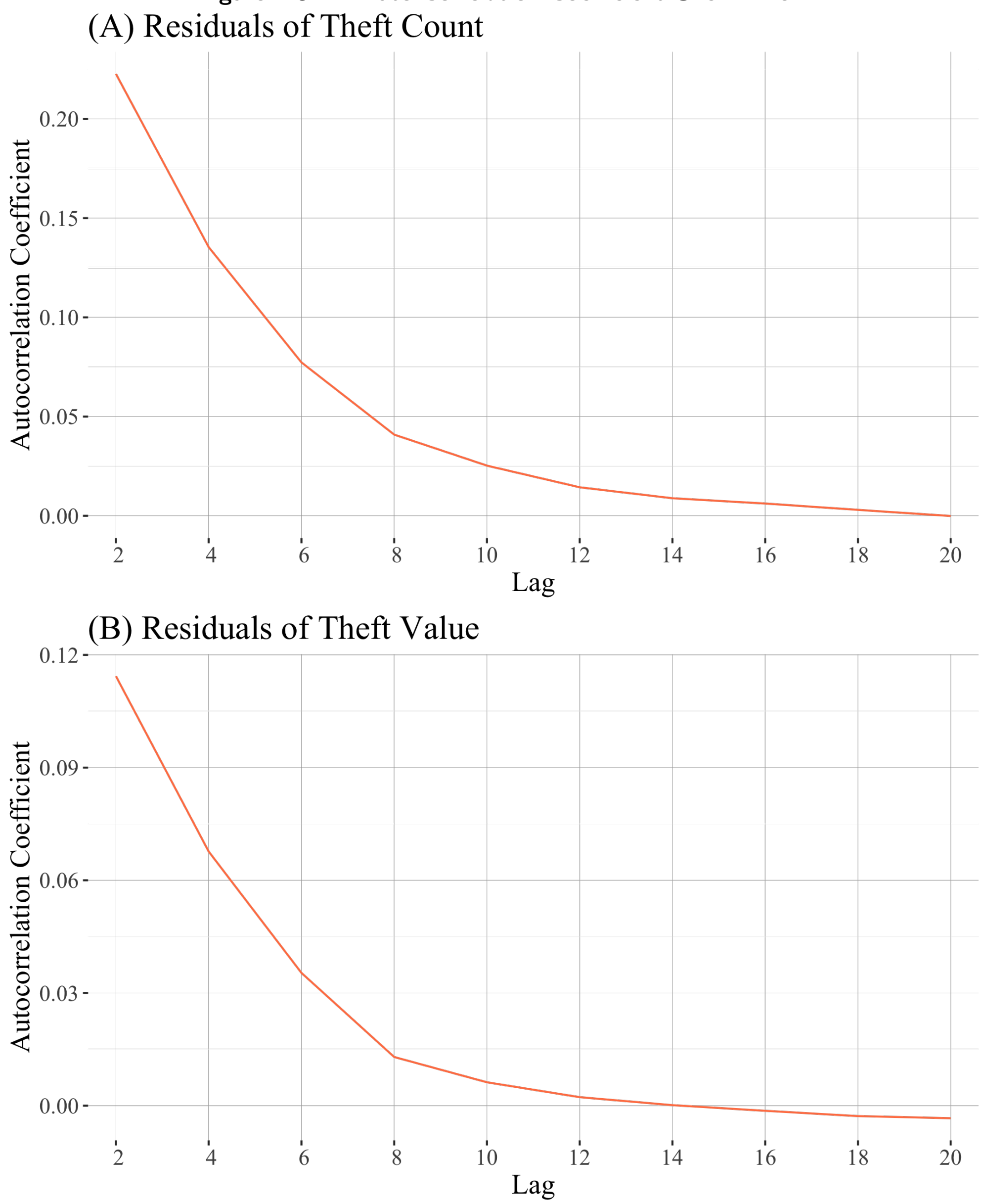


This figure shows how the OLS bias (OLS - IV estimates) varies with different covariance in residuals for true parameters ranging from 0.06 to 0.46. Each point in each of the nine curves includes 1,000 simulations

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.

Figure E.5 Auto Correlation Coefficient Over Time



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.

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.

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 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?