Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Cleaning House: The Impact of Information Technology Monitoring on Employee Theft and Productivity

Lamar Pierce, Daniel C. Snow, Andrew McAfee

MANAGEMENT SCIENCE
Vol. 61, No. 10, October 2015, pp. 2299–2319
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)
http://dx.doi.org/10.1287/mnsc.2014.2103

Cleaning House: The Impact of Information Technology Monitoring on Employee Theft and Productivity

Lamar Pierce
Olin Business School, Washington University in St. Louis, St. Louis, Missouri 63130, pierce@wustl.edu

Daniel C. Snow
Marriott School of Management, Brigham Young University, Provo, Utah 84602, dsnow@byu.edu

Andrew McAfee
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, amcafee@mit.edu

This paper examines how firm investments in technology-based employee monitoring impact both misconduct and productivity. We use unique and detailed theft and sales data from 392 restaurant locations from five firms that adopt a theft monitoring information technology (IT) product. We use difference-in-differences models with staggered adoption dates to estimate the treatment effect of IT monitoring on theft and productivity. We find significant treatment effects in reduced theft and improved productivity that appear to be primarily driven by changed worker behavior rather than worker turnover. We examine four mechanisms that may drive this productivity result: economic and cognitive multitasking, fairness-based motivation, and perceived increases of general oversight. The observed productivity results represent substantial financial benefits to both firms and the legitimate tip-based earnings of workers. Our results suggest that employee misconduct is not solely a function of individual differences in ethics or morality, but can also be influenced by managerial policies that can benefit both firms and employees.

Keywords: organizational studies; personnel; productivity; information systems; IT policy and management; judicial/legal; crime prevention; marketing; sales force; service operations

History: Received November 11, 2013; accepted October 20, 2014, by Serguei Netessine, operations management. Published online in *Articles in Advance* May 13, 2015.

1. Introduction

Employee theft and fraud are widespread problems in firms, with workers stealing roughly $200 billion annually from U.S. firms to supplement their income (Murphy 1993). A growing empirical literature clarifies when and how theft and other misconduct occur (e.g., Fisman and Wei 2009, Zitzewitz 2012),[1] but says little about the overall impact of firms' use of forensics to monitor and reduce theft. This is a significant gap in the literature, given the substantial investments made by firms in monitoring employees (Dickens et al. 1989) as well as the growing forensic and monitoring capabilities enabled by information technology (IT) systems and their demonstrated potential to improve labor productivity (Aral et al. 2012, Tan and Netessine 2014).

In this paper, we address three important yet unresolved questions about the impact of using IT to monitor employee crime. First, is employee monitoring indeed effective in reducing theft, as economics suggests (Becker 1968, Dickens et al. 1989), or does monitoring demotivate and constrain employees and thus negate gains from theft reductions (Cialdini 1996, Latham 2001, Bernstein 2012)? Second, do possible gains from monitoring result from changing worker behavior or from replacing unethical workers with more honest ones? Previous studies provide little insight on this question because they lack the detailed worker-level panel data needed to separate treatment from selection effects among employees. Third, if increased monitoring indeed reduces theft by existing workers, then through which mechanisms does productivity in other tasks change and what is the overall impact to the firm? Existing work suggests that theft monitoring might improve productivity through several mechanisms. These include economic multitasking that replaces lost income through other activities (Olken 2007, Yang 2008b), reduced cognitive impediments from multiple tasks (Schultz et al. 2003, Clark and Huckman 2012, Staats and Gino 2012

[1] The economics literature on other types of misconduct is vast. There is a large related economics literature on government corruption (e.g., Fisman and Miguel 2007, Yang 2008a, Niehaus and Sukhtankar 2013). There is a large related research stream on other misconduct at the firm and market levels, such as tax evasion (Fisman and Wei 2004) and stock option backdating (Heron and Lie 2007). See Zitzewitz (2012) for a survey of forensic economics.

KC 2014), perceptions of increased fairness among honest employees (Greenberg 1990), and broader perceptions of increased managerial oversight.

In this paper, we address these questions by examining the impact of improved theft monitoring from IT in a service setting, the American casual dining sector, using a data set that details employee-level theft and sales transactions at 392 restaurants in 38 American states. We focus on this setting for several reasons. First, detailed theft and sales data allow us to identify specific worker-level productivity, theft, and sorting responses to changes in firm monitoring. Second, unlike previous research on monitoring (e.g., Nagin et al. 2002, Duflo et al. 2012, Zitzewitz 2014, Bernstein 2012), our setting includes multiple firms with hundreds of geographically dispersed locations. Although previous experimental work by Nagin et al. (2002) studied the impact of perceived monitoring at one firm's 16 call center locations, our study examines five large chains (firms), 392 locations, and 22,150 employees, and employs a staggered implementation that significantly reduces concerns about contemporaneous shocks or policy changes. Finally, the implementation of monitoring in this setting is plausibly exogenous to local conditions, reducing concerns that monitoring is added only for locations that might most benefit from it. We use two years of detailed theft, sales, and operational data from restaurant chains that adopt an IT-based theft monitoring system that identifies and reports theft by specific employees. Restaurant servers (also called "waiters") use multiple techniques to steal from employers and customers, including voiding, "comping," and transferring sales after taking payment from customers. The system sends weekly reports that alert managers to repeated suspicious actions highly unlikely to result from mistakes, such as when one drink is virtually transferred in the system across 10 customers. These alerts represent the "tip of the theft iceberg," since the product is designed to identify instances of theft that are so obvious as to be indefensible by servers. Consequently, although the weekly alerts in the data average only \$108 per location, interviews with managers indicate the losses to be considerably larger.

We use difference-in-differences (DD) models to estimate the treatment effect of the monitoring technology on theft, sales productivity, customer service (tips), and employee turnover at the individual and restaurant levels. The data also allow us to control for time trends and location- and worker fixed effects. Our empirical models identify a 22% (or \$24 per week) decrease in identifiable theft after IT monitoring implementation. This treatment effect is persistent, growing from \$7 in the first month to \$48 in the third month. The effect on total revenue, however, is much larger. Total revenue increases by \$2,975 per week (approximately 7% of average revenue) following implementation, suggesting either an increase in employee productivity or a higher level of latent theft eliminated by the monitoring system. Furthermore, the implementation of theft monitoring increases drink sales (the primary source of theft) by \$927 per week (approximately 10.5%). This result is of particular economic importance because margins on drinks in casual dining are 60%–90%, representing approximately half of all profits at a typical restaurant. Furthermore, we observe an increase in average tip levels of 0.3%, which represents one-sixth of a standard deviation improvement from a base rate of 14.8%. This result is consistent with improvement in quality of customer service.

Individual worker models show that our restaurant-level model results are primarily explained by changed employees behavior rather than change in the group of employees working at the restaurant. Although the monitoring system does appear to increase turnover among workers who had previously stolen, we see no evidence that these workers are involuntarily terminated following theft reporting to managers. Instead, we find evidence that managers use staffing hours to reward honest employees—a meaningful but more muted way to provide credibility to the monitoring. The apparent rarity of termination also echoes Dickens et al. (1989) observation that firms infrequently employ the low-monitoring, high-punishment crime deterrence strategy argued to be efficient by Becker (1968).

We also examine the data for evidence of four mechanisms that could explain the productivity results. We find little evidence that economic multitasking or cognitive multitasking mechanisms can explain the observed productivity increases, but we see some support for a fairness-based mechanism among honest workers. Although we cannot directly measure employee perceptions of increased managerial attention, the evidence leads us to speculate that this is a plausible explanation for this result.

This paper has implications for several research streams. First, we contribute to research on the impact of technology on labor and firm productivity (Brynjolfsson 1993, Brynjolfsson and Hitt 1996, Athey and Stern 2002, Bresnahan et al. 2002, Aral et al. 2012, Rawley and Simcoe 2013, Brynjolfsson and McAfee 2014). Whereas other studies show that IT can improve productivity by reducing mild forms of misconduct such as shirking and absenteeism (Hubbard 2000, Baker and Hubbard 2003, Duflo et al. 2012), this paper is the first to show both the direct reduction in explicitly illicit behavior (theft) as well as the secondary effect of increased productivity. Our work is consistent with the Bloom et al. (2012) finding that the productivity gains from IT have been largest in industries—such as restaurants—with human resource policies that more strongly link pay and employment to employee performance.

Second, this paper contributes to growing work on worker behavior and productivity in service operations (e.g., Gino and Pisano 2008, Huckman et al. 2009, KC and Terwiesch 2009). The results are consistent with previous studies showing quality of service to be intimately tied to other performance metrics such as capacity or speed (Tucker and Edmondson 2003, KC and Terwiesch 2012). Our work is unique in showing a link between service quality (measured by tips) and the monitoring of the specific metric of theft.

Third, this research contributes to the literatures on forensic economics and corruption. Despite many studies on monitoring and crime in the fields of criminal justice, law, and economics (Lochner 2007, Draca et al. 2011), few studies focus on explicitly illegal behavior by employees in firms. And with few exceptions (e.g., Nagin et al. 2002, Pierce and Snyder 2008, Cohen et al. 2010, Gino and Pierce 2010b), these rely almost exclusively on firm-level data (Fisman and Wei 2004, 2009; Zitzewitz 2006; Heron and Lie 2007; DellaVigna and La Ferrara 2010; Chen and Sandino 2012; Pierce and Snyder 2012). These worker-level data allow us to disentangle firm-level misconduct from individual-level decisions that hurt firm profitability. The multifirm, longitudinal nature of these data also allows us to examine how monitoring affects both worker selection and behavior on both theft and other tasks. This detailed examination of service employee theft also shows how individual worker-level data might disentangle the mechanisms behind prior results on inventory inaccuracy and retail shrinkage (e.g., DeHoratius and Raman 2007, 2008).

Finally, this paper contributes to an older research stream in operations management and industrial relations (Deming 1938, 1993; Adler et al. 1997; Spear and Bowen 1999) that blames production problems on the operating system (broadly construed) rather than on individual worker incompetence or perfidy. Although this research stream has seen little growth in recent years (a rare example is Holweg 2007), its legacy continues to inform operations management teaching, practice, and popular culture. Our results are consistent with the view in these earlier works that worker misbehavior depends, to a large extent, on the operating environment created by management, a view also shared by a broad literature in behavioral and organizational ethics (e.g., Bazerman et al. 2002, Mazar and Ariely 2006, Tenbrunsel and Smith-Crowe 2008, Mead et al. 2009, Moore et al. 2012).

2. Theoretical Background

2.1. Monitoring and Misconduct

The implications of monitoring crime in the economics literature are clear—by increasing the likelihood of detection and punishment, monitoring motivates crime reduction (Becker 1968). A broad empirical literature supports this model outside the workplace, primarily focusing on how increased policing reduces crime (Lochner 2007, Draca et al. 2011) or on how auditing and inspections reduce corruption (Olken 2007). Yet, evidence on the impact of monitoring on workplace crime is more limited, with monitoring studies primarily focusing on shirking and absenteeism (Hubbard 2000, Duflo et al. 2012). To the best of our knowledge, the only paper to causally show the effect of monitoring on individual employee criminal behavior is Nagin et al. (2002), who used experimental data on misreported donation pledges at 16 call centers. Establishment-level studies from related fields support this relationship as well. Detert et al. (2007), for example, show a correlation between higher manager-to-employee ratios and lower restaurant-level food losses. DeHoratius and Raman (2008) similarly link monitoring with inventory accuracy, which partially reflects reduced theft.

Although this research stream supports the efficacy of theft monitoring, other work from management and psychology argues that the effects of monitoring are less clear. Using a field experiment in a Chinese mobile phone factory, Bernstein (2012) found that reducing monitoring (and thus increasing privacy) improved worker performance by encouraging experimentation and innovative problem solving and by reducing distraction and worker gaming in response to managerial oversight. This work builds on a literature in sociology on workers misrepresenting production activities when monitored by management (Roy 1952). The management literature on deviance further argues that monitoring employee misconduct can backfire by decreasing trust between workers and managers, potentially *increasing* the behavior it seeks to constrain (Litzky et al. 2006). This argument is backed by theory in psychology and economics that argues that excessive monitoring can decrease worker commitment, intrinsic motivation (Hochschild 1983, Frey 1993, Cialdini 1996, Bénabou and Tirole 2006), or trust (Frey 1993). Such growth in misconduct might also occur if increased managerial monitoring reduces job satisfaction or perceptions of fairness, both of which have been linked to increased employee theft (Greenberg 1990, Latham 2001, Chen and Sandino 2012).

Despite these theoretical disagreements about the efficacy of monitoring, most field studies support that monitoring improves desired behaviors and reduces misconduct, at least in the short run. It is less clear from the literature whether firms achieve these reductions by using this information to screen or remove dishonest workers or to change individual behavior. This question is part of a larger debate on whether employee misconduct primarily stems from individual differences or work environment. If individual differences drive most theft, then the largest monitoring gains should result from replacing the worst employees. Evidence comparing worker selection versus treatment hypotheses has been limited because it requires an individual-level data panel with many workers and a long time period and resulting turnover. Although a few papers suggest that misconduct can be reduced without replacing workers (Greenberg 1990), there is evidence that turnover can also be effective. Nagin et al. (2002), for example, find that although some workers change their behavior under increased monitoring, others do not. Unfortunately, their data do not allow for the comparison of which organizational mechanism—worker selection or worker treatment—is most effective in reducing theft. Furthermore, a large literature in economics and management also suggests that employee selection can be the critical mechanism through which firm policy changes impact performance (e.g., Banker et al. 2001, Campbell 2012, Lazear 1986, Lazear and Oyer 2011).

2.2. Spillover Effects from Theft Monitoring

Although the implementation of IT to monitor theft is likely to reduce theft, it may also have spillover effects on more general productivity through four possible mechanisms. A first mechanism is economic multitasking (e.g., Hölmstrom and Milgrom 1991). If workers under a pay-for-performance system (such as tips) can derive earnings from two tasks, in this case sales productivity and theft, then any reduction in the returns from one task might increase attention toward the other. Under such a model, earnings from each task are increasing and concave in attention (or effort). The cost of attention from each task is convex and increasing, but theft bears two additional costs. First, the employee will be detected and punished by management (the principal) with some probability p that is increasing in theft. Second, the employee may suffer moral or ethical costs based on identity or preferences that make theft costly even when it is effortless and unmonitored (e.g., Mazar et al. 2008, Bénabou and Tirole 2011, Dal Bó and Terviö 2013).

Such an economic multitasking mechanism has three implications. First it implies that any employee with existing nonzero theft levels will reduce attention and effort allocated to theft in response to increased monitoring by management. This is because monitoring increases the marginal cost of attention allocated to theft by raising the likelihood of detection and punishment. Second, the resulting decrease in earnings from decreased theft effort will motivate the worker to increase effort allocated toward productivity. This is because at low income levels (as in this setting), the utility function with regard to income can be quite concave. The loss of even small amounts of income can substantially impact the worker's ability to meet basic living expenses. Third, employees with existing nonzero theft levels will be more likely to leave the firm, as outside employment options become relatively more attractive than before.

A second mechanism that might generate improved productivity through IT theft monitoring is cognitive multitasking. Growing research in cognitive psychology has found that individuals are less able to focus attention on a specific task when forced to multitask (Charron and Koechlin 2010). These cognitive limitations mean that multitasking can create attention externalities to both tasks as workers alternate their focus between them. Constraining employees from a stealing task may, therefore, improve their performance in other more productive tasks, such as sales and customer service. Recent microproductivity studies on multitasking in healthcare (KC 2014) and banking (Staats and Gino 2012) support the existence of such negative spillovers from multitasking. Therefore, income and cognitive constraints suggest increased IT theft monitoring will both decrease theft and increase productivity among workers.

Third, productivity might increase after theft monitoring implementation, as honest employees become motivated by increased perceptions of fairness. Theft and other dishonest behaviors have frequently been linked to perceptions of fairness and inequity (Greenberg 1990; Gino and Pierce 2009, 2010a). If honest employees perceive the earnings advantage for dishonest workers to be unfair, it may demotivate them in a way similar to other forms of inequity or unfairness (Adams 1965, Larkin et al. 2012). The implementation of monitoring, in restoring perceptions of fairness to honest employees, might improve their productivity. Indeed, prior work in the management literature suggests perceptions of fairness and justice may be critical to monitoring's efficacy in improving worker behavior (e.g., Niehoff and Moorman 1993).

Finally, theft monitoring might also improve productivity through employees' perceptions (accurate or not) that theft monitoring represents a general increase in managerial attention toward all employee activities. In studying criminal behavior, Lochner (2007) notes that beliefs about the likelihood of detection can influence criminal behavior independent of true probability. This mechanism would imply that employees do not increase productivity because theft is constrained, but rather because they believe monitoring of productivity has also increased, whether accurate or not. This mechanism does not rely solely on employee perceptions, however. Managers, possessing new theft monitoring technology, actually may be able to increase their attention toward monitoring other activities because of reduced costs of theft monitoring.

3. Field Setting

The context of our study is a service setting—the "casual dining" segment of the United States restaurant industry, which is characterized by table service and midrange prices. Examples—not necessarily in this sample—include chains like Applebee's, Chili's, and Olive Garden. This segment is economically significant, generating approximately $33 billion of the annual revenue total of $110 billion in the American restaurant industry. Profit margins in the segment are thin, averaging 3.5% in 2010 (Sweeney and Steinhauser 2010). Much of this profit comes from sales of both alcoholic and nonalcoholic beverages, which have margins of 60% to 90%. Management policies that improve worker productivity are critical to the labor-intensive restaurant industry. Tan and Netessine (2014), for example, found that optimal staffing decisions can improve productivity and reduce staffing costs by 3% and 17%, respectively. Bernstein and Sheen (2014) similarly find that operational improvements from private equity buyouts reduce labor costs and improve survival.

Most casual dining restaurants employ point-of-sale (POS) systems that track orders, sales, and assignments for each server or bartender. When a server receives a food and beverage order from a customer or table (a "ticket"), the server enters it into a touch-screen panel, which then transfers the order to the kitchen through the POS system's database. After the customer has paid and left the restaurant, the server closes out the ticket. Servers typically have multiple tickets open simultaneously. All the restaurants in this sample use the basic POS product in each week in which they appear in the sample.

Important for our research, compensation for service staff at nearly all U.S. casual dining restaurants combines hourly wages with a performance-based component. In the United States, most servers in the casual dining segment are paid a fixed wage at or below the legal minimum wage. Subminimum wage pay is legally permissible so long as the legal minimum wage is exceeded when adding the pay-for-performance component in the form of customer tips. Social norms in the United States strongly suggest a minimum tip of 15% of the total check, with a lower percentage usually reserved for poor service or an unpleasant dining experience. Customers often increase the tip percentage to reward particularly good service. Since tip percentage is at the discretion of the customer, servers can increase their income through both increased sales revenue and improved customer service (tip percentage). Servers can increase revenue through effort toward selling additional drinks, desserts, or add-ons like side salads or extra bacon. Even though tipping behavior is relatively standardized in American culture, even simple efforts toward customer service can significantly raise percentages. Past studies find substantial increases in tipping due to touching customers, writing "thank you" on the check, and increased general service quality (Bodvarsson et al. 2003).

Although theft by servers and other restaurant workers is a constant problem, studies of this setting are rare. Jin and Leslie (2003, 2009) study food safety and foodborne illness, but present no evidence of intentional wrong doing. Two studies more directly examine the topic. Victor et al. (1993) use survey-based data on peer reporting in fast-food restaurants but do not directly observe theft. Detert et al. (2007) find restaurant-level food theft for on-site or outside consumption or resale to be associated with store-level characteristics, such as number of managers and abusiveness of managerial supervision. Perhaps the largest problem, and the focus of this paper, stems from servers stealing sales revenue either by not reporting the sale or by using one of a number of techniques to remove it from restaurants' IT systems.

Although there are many ways in which restaurant employees steal from employers, we focus on the three types detected by the data provider. These "scams" are well known in the industry, even having nicknames and books written about them (Francis and DeGlinkta 2004). The most common type of server theft in these data is called the "wagon wheel scam" wherein, following customer payment for a food or drink item, the server transfers that item in the POS system to another newly seated guest who has ordered the same item. The original check is then reprinted after the customer leaves and the waiter pockets the difference by taking cash from the register. The wagon wheel can be applied to both 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 other scams involve one of two techniques. The first involves "comping," or refunding meals of customers in the system after they have already paid but before the ticket has been closed. The second involves "voiding" a transaction as erroneous after charging a customer. In both cases, the customer pays for the meal. With a cash payment, the server pockets all or part of the payment rather than depositing it in the register. With credit card transactions, the server takes an equal amount from the register in cash as a fraudulent tip. Although the flexibility and worker discretion built into the POS system may seem extraordinary compared with other operational settings, our interviews with managers reveal that these attributes are standard in the industry. They are seen as critical in a setting in which each customer interaction is different and managers can be overwhelmed quickly when their approval is required for all adjustments.

Given their pay-for-performance compensation and opportunities for theft-based income, servers face a special type of multitasking problem of allocating effort or attention toward two task types: productive (sales and customer service) and corrupt (theft) tasks. Servers can earn income through both, but the restaurant benefits only from productive tasks and loses direct income from theft. If managerial monitoring creates a credible threat of theft detection and punishment, then increased monitoring through IT should reduced theft, but it may also improve productivity and workers seek to replace lost theft income with tips.

These data are from a large POS IT system used by many restaurants and other hospitality providers. The system vendor sells a theft monitoring add-on product to its core POS offering that utilizes proprietary algorithms to detect and report theft and fraud by servers to managers each week. The exact price was not provided by the vendor, but was characterized as being nonzero and less than $100 monthly per location. The algorithms are constructed with a strong bias toward false negatives because of the high cost of falsely accusing an employee of theft. This means that the system must observe repeated occurrences of suspicious activities before designating theft. Most restaurant managers we interviewed had intimate knowledge of the techniques through which servers stole money, primarily through their earlier work as servers.

At some point during the sample period, each of the restaurant locations in our sample received access to the theft monitoring system (as an addition to their existing POS system), allowing us to estimate its impact on theft and productivity. Each restaurant chain in the sample rolled out the monitoring system to its store locations in a piecemeal way that, according to interviews with managers at the IT provider, was not based on individual store needs or theft levels. Rather, the implementation was driven by the system vendor's rollout team's schedule. This rollout strategy allows us to treat implementation dates as plausibly exogenous and not driven by revenue or theft levels. Examination of this assumption is presented in §4.2.

Interviews with data provider employees and restaurant managers detail the impact of the IT monitoring product on restaurants. Although all managers could observe clear instances of theft from the weekly theft reports, the use of this information was not uniform. Some managers indicated a reluctance to fire identified thieves, since training new servers is time intensive and expensive. Similarly, some managers indicated that directly confronting detected thieves was difficult because accusations of theft can generate resentment, lower productivity, and even lead to lawsuits. Since managers often have worked as servers before promotion into management, the revelation of specific theft instances is rarely a surprise. Consequently, most managers appear to have either leaked information on the new product to staff through nonconfrontational discussions with servers or made announcements of the new product preemptively to reduce theft.

4. Empirical Strategy

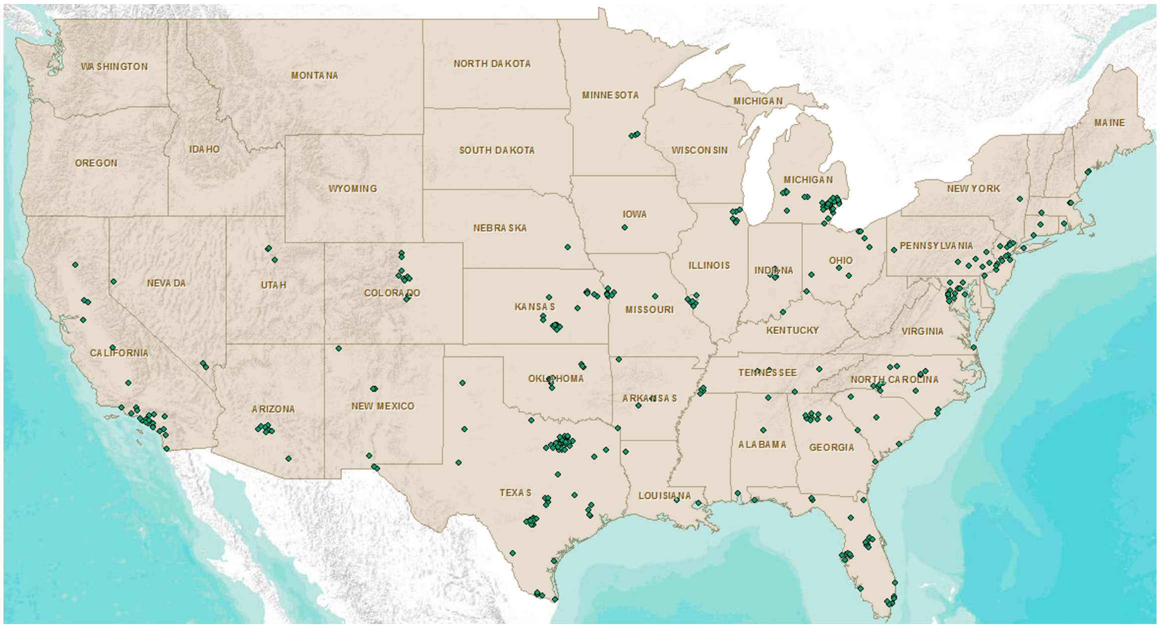
4.1. Data

The data for this study were provided directly by a restaurant IT system vendor. Since the company provides basic POS software services as well as the theft monitoring add-on, the data contain all transactions at each restaurant for up to two years. Our sample includes five restaurant chains (firms) with a total of 392 restaurant locations in 38 American states from March 2010 to February 2012. The data set includes the zip code of each restaurant and a coded identifier for each chain. Figure 1 presents a map of these locations.

The data include information on some food and beverage items sold, prices, tips, and customer count at each table. All transactions are time stamped, and they identify the specific server associated with the check. There are 22,150 employees who work exclusively as either servers or bartenders. Although the panel is unbalanced because of a combination of different POS system implementation dates (see the online appendix, available as supplemental material at http://dx.doi.org/10.1287/mnsc.2014.2103), our use of location fixed effects and alternative panel truncation dates (see the online appendix) reduces concerns that this might bias our results. Figure 2 presents the adoption dates of the theft monitoring system for each location in the data.

These transaction data are combined with weekly theft data. Since we are able to retroactively apply the theft algorithms to the entire sample regardless of when the theft monitoring system was implemented, we can observe preimplementation theft even though managers could not. After implementation, the alerts provided managers with weekly server-specific reports on the six categories of theft: transfers (i.e., the wagon wheel) and five types of

Figure 1 **(Color online) Geographic Distribution of All Restaurant Locations**

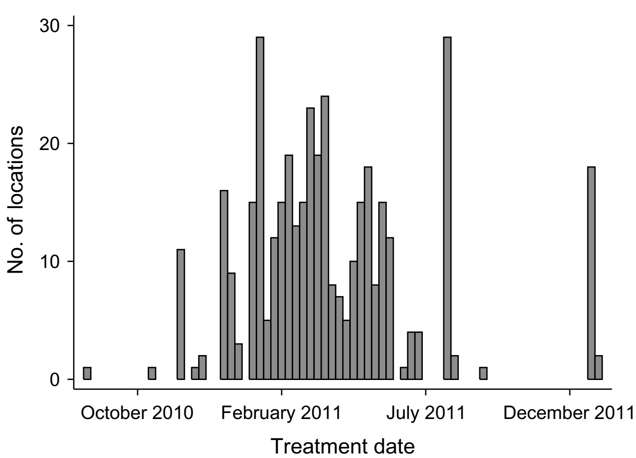


Notes. Data agreement does not allow for the representation of separate chains. Three Alaska locations and one Hawaii location are not shown.

sales voids and comps. An example of such a report is provided in Figure 3.

We sum these six categories to generate a measure of total weekly theft—*total losses*—that represent the total dollar amount of suspected theft reported to the manager. With a conservative estimate of theft, only instances that are highly suspicious are reported. This means that the instances of observed theft in the sample likely constitute only a portion of theft at each location. We aggregate the POS data to the weekly level in order to combine it with the weekly theft monitoring system data. This gives us a conservative estimate of weekly theft for each server and each location.

Figure 2 **Implementation Weeks by Location**



We use two such combined data sets in our analysis. The first is a weekly location-level data set that details total check revenue, drink revenue, credit card tips, and total losses for each restaurant. The second data set is a weekly server-level data set with the same measures plus an additional measure of total hours worked by the server. We restrict this data set to the 22,150 workers whose job descriptions in the data set are designated as server, waiter, or bartender. We exclude all takeout orders because they are not affected by server behavior. We calculate *tip percentage* as the weekly sum of credit card tips divided by the sum of credit card sales revenue. We are unable to observe tips for cash sales, so we must assume that any changes in credit card tips parallel changes in cash tips. Fortunately, credit card tips are difficult to misreport, whereas cash tips can be easily lied about. However, credit card tips may be manipulated when implementing the wagon wheel, void, or comp schemes on credit card transactions, as described earlier. Tip amounts are increased to account for the missing revenue in the total credit card charge. This likely biases downward the tip results, since decreased theft would mechanically decrease tip amounts, suggesting the impact of theft monitoring on tips might be even larger than we estimate. The data provider also gave us the precise date of implementation at each restaurant location. We designated the first full week after implementation (Sunday–Saturday) as the first treatment week. As mentioned previously, treatment dates are dispersed throughout the time period, even within chain,

Figure 3 (Color online) Example of Incident Report

RECENT INCIDENTS
The following table displays the most recent instances of this type of activity.

Date	Item	First Check	Last Check	# Transfers	Item Price	Potential Loss
02/07/2009	Soft Drink	1001	2032	2	1.79	3.58
02/07/2009	Iced Tea	1046	3012	5	1.79	8.95
02/07/2009	Iced Tea	1005	1014	3	1.79	5.37
02/05/2009	Coffee	2210	2215	2	1.59	3.18
02/05/2009	Iced Tea	1114	1154	3	1.79	5.37
02/05/2009	Soft Drink	3100	1146	6	1.79	10.74
02/05/2009	Soft Drink	1111	1121	2	1.79	3.58
02/04/2009	Soft Drink	1002	3004	4	1.79	7.16
02/04/2009	Chocolate Cake	1021	1033	2	4.99	9.98
02/03/2009	Coffee	2023	2040	3	1.59	4.77
02/03/2009	Soft Drink	2026	1141	3	1.79	5.37
02/03/2009	Iced Tea	2028	2033	4	1.79	7.16
02/03/2009	Iced Tea	2031	1100	2	1.79	3.58
02/03/2009	Side Salad	3114	2106	2	2.99	5.98
					Total Potential Loss:	**$84.77**

INCIDENT HISTORY ASSOCIATED WITH THIS EMPLOYEE ID
The following table displays a summary of all alerts generated for this employee ID.

Scam Type	Event Count	First Incident	Last Incident	Avg Loss / Incident	Total Potential Loss
Transfer	26	01/27/2009	02/07/2009	$6.06	$157.43
Comp after Print	3	01/29/2009	02/02/2009	$8.41	$25.23

which allows us to separate week-specific shocks and time trends from the treatment effect of the theft monitoring system. None of the restaurant locations drops the product in our sample period. The descriptive statistics for the weekly location and weekly server samples are presented in Table 1 (see the online appendix for statistics by chain).

There are 22,329 location-week observations and 439,838 server-week observations, with fewer tip observations because one chain does not report tips. Average weekly losses per location are $108.47, which emphasizes how little of the widespread theft detailed in interviews and anecdotes is reported by the system. Average weekly restaurant check revenue is $43,697, of which $8,879 come from drinks. Tip percentage averages 14.8%. Each restaurant has around 15.5 servers staffed per week. The worker data set shows comparable statistics, but has much wider variance across workers, particularly in theft and tip percentage. Although the data provider did not allow us to identify the chains, we were told that nearly all restaurant locations were corporate owned, with only a few franchises (which we cannot identify) in one of the chains.

Table 1 Descriptive Statistics for Week-Level Samples of Total Restaurant and Hourly Worker Variables

	Weekly location		Hourly worker	
Variable	Mean	SD	Mean	SD
Total losses ($)	108.47	207.24	0.15	0.92
Check revenue ($)	43,697	18,659	79	32
Drink revenue ($)	8,879	5,738	14	11
Tip percentage (%)	14.8	1.9	15.5	3.6
Treated	0.77	0.42	0.80	0.40
Workers/Day	15.5	4.3	—	—
Observations	22,329		439,838	

Notes. Total losses represent theft identified by the IT system. *Treated* is a dummy variable indicating the theft monitoring system has been implemented. *Tip percentage* has fewer observations because Chain 5 does not report tips.

4.2. Identification Strategy

We use a difference-in-differences design to estimate the impact of the implementation of the theft monitoring system on theft, total revenue, drink revenue, and customer service (measured by tip percentage) at the location and server levels. Our DD design treats the system implementation as the treatment on each location, using preimplementation locations as control groups. The DD design "differences out" variation not related to the passage of time in the locations and uses control groups to provide a counterfactual for what would have happened at the treatment location had it not adopted the system. DD models critically rely on the parallel trend assumption—that the control group represents the counterfactual change that would have occurred in the treatment group if it had not been treated. Several aspects of the data support this assumption. Although DD designs do not require control and treatment locations in any week to be identical, those in this sample are similar given the relative uniformity of locations within a given chain. Also, the locations in these data adopt the monitoring system at different times for what appear to be nonstrategic reasons. This allows us to account for time trends, since treatment is not highly correlated with a specific time treatment (as in many DD models). Furthermore, system implementation does not appear to be strongly correlated with preimplementation weekly theft or revenue levels (see the online appendix).

We use a standard DD specification to first estimate the impact of theft monitoring system implementation on the dependent variables (i.e., losses, sales revenue, tip percentage) for each location:

$$Y_{it} = \alpha_i + \beta_1 \times TREATED_{it} + \Upsilon_t + \varepsilon_{it},$$

where Y_{it} is the dependent variable for restaurant location i in week t, α_i is the set of restaurant fixed effects to account for unobserved location heterogeneity, Υ_t is the set of dummy variables for each week in the sample to control for time trends, and $TREATED_{it}$ is a dummy variable equal to 1 for all weeks t at restaurant i occurring after the week of implementation of the theft monitoring system. Our specification exploits the panel nature of the data by introducing a full set of fixed effects (described below). Each of our dependent variables is a continuous variable. Consequently, we estimate all regressions using ordinary least squares (OLS) regressions, clustering standard errors at the location level.

Our second specification uses a nearly identical DD approach at the individual worker level:

$$Y_{ijt} = \alpha_i + \beta_2 \times TREATED_{jt} + \Upsilon_{ijt} + \varepsilon_{jjt},$$

where Y_{jjt} is the dependent variable for worker j at restaurant i in week t. Our base specification includes restaurant fixed effects (α_i), since we cannot observe servers switching restaurants. We will also implement specifications with worker fixed effects (α_j) to estimate how much of the impact on restaurants is due to selection versus worker treatment. A time-varying dummy variable $TREATED_{it}$ is equal to 1 for all weeks t following the treatment date of restaurant i. The vector Υ_{ijt} includes weekly dummy variables and controls for servers' weekly shift assignments: the percentage of shifts from both dinners and weekends (Thursday–Saturday), which have higher customer traffic. We divide check revenue, losses, and drink revenue by the week's total hours to control for time worked, transforming the three measures into weekly averages of hourly productivity and theft. We cluster standard errors at the location level.

5. Results

5.1. Restaurant Treatment Effects

Table 2 presents the basic treatment effects of the theft monitoring system on individual restaurants.

Table 2 Impact of Information Technology on Weekly Restaurant Performance

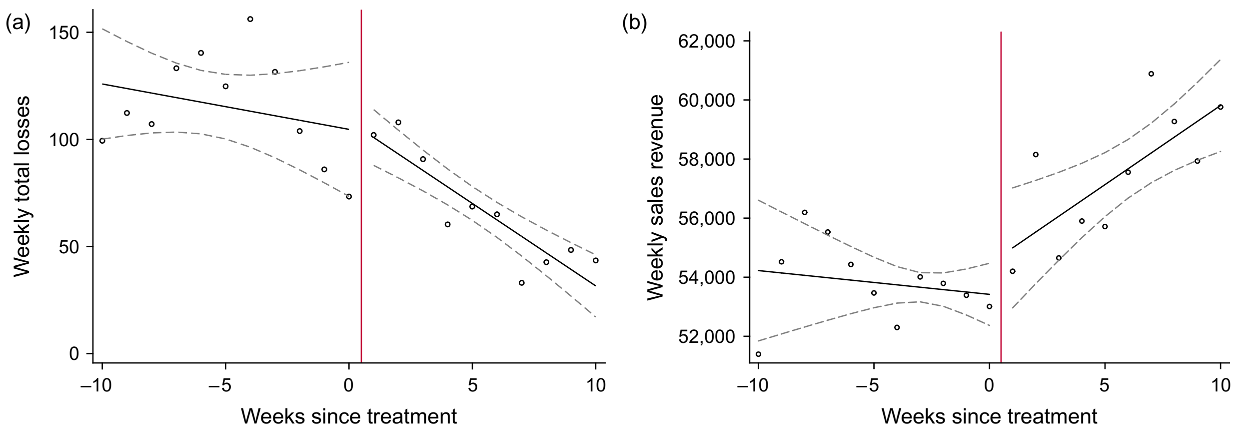
Dependent variable:	(1) *Losses*	(2) *Check revenue*	(3) *Drink revenue*	(4) *Tip percentage*
Treated with monitoring	−23.61***	2,975.21***	926.97***	0.0030***
	(7.03)	(508.91)	(148.14)	(0.0005)
Week fixed effects	Included	Included	Included	Included
Location fixed effects	Included	Included	Included	Included
Constant	122.31***	30,162.30***	7,182.84***	0.1475***
	(3.73)	(705.73)	(127.56)	(0.0011)
Adjusted R^2	0.180	0.817	0.891	0.726
Observations	22,329	22,329	22,329	20,912

Notes. Standard errors clustered at the location level in parentheses. Tip models have fewer observations due to some locations not tracking tips.
***Significant at the 1% confidence level.

The models include restaurant fixed effects that account for time-invariant restaurant heterogeneity, as well as week fixed effects to control for time trends. Column (1) shows that the system implementation reduced weekly theft losses by $23.61, or approximately 22% of the preperiod average. Again, we note that this represents only the portion of losses that can be proved definitively. The impact of IT monitoring on check revenue, presented in column (2), is much larger, accounting for an increase of $2,975.21, or approximately 7% of the average pretreatment revenue. This increase represents about one in four customers buying an additional dessert, appetizer, or salad and is in the range of the potential server productivity increases from workload management found by Tan and Netessine (2014). A similarly large impact on drink revenue is observed in column (3), with sales rising by $926.97, or about 10.5%. This result is important because drink sales are associated with high margins. If operating margins at these restaurants are near the 3.5% industry segment average, such an increase in drink revenues would represent at least a 36% increase in operating margins by itself.[2] We also observe a small, significant increase in tip percentage of 0.3%, which represents one-sixth of a standard deviation. Per Bertrand et al. (2004), we also block bootstrap the standard errors, which improves statistical significance (see the online appendix). These results are robust to both winsorized and logged dependent variables (see the online appendix). The improvements to sales and tips occur without an increase in the time customers spend at tables (see the online appendix)—another measure of server productivity (Tan and Netessine 2014).

[2] This increase is calculated based on conservative 60% margins on $927, or $556.20, in profit increases. Assuming the 3.5% segment average of operating margins (or $1,529.40 per week), this represents at least a 36% increase in profits solely from drink sales.

Figure 4 (Color online) Treatment Effect on Weekly Losses and Sales in Chain 1



Notes. Data are shown for purely illustrative reasons and do not adjust for location fixed effects or time trends. Each point in panel (a) represents the average weekly losses in USD for 28 restaurants in Chain 1. Dashed lines represent 95% confidence intervals. Each point in panel (b) represents the average weekly revenue in USD for 28 restaurants in Chain 1.

Figure 4 illustrates the effect of the monitoring system on losses and revenue for all 28 locations in one chain with a consistent sample of pretreatment and posttreatment locations (Chain 1 from the online appendix).

The figures represent the average weekly losses and revenue as linear spline functions of weeks after treatment, with 95% confidence intervals. The vertical line represents four separate treatment dates in the chain, which partially controls for time trends (see the online appendix for drink and tip figures). The impact of IT monitoring on all outcomes, even for the small sample of restaurants represented in the figures, suggests a persistent effect, which we examine next. We note that these figures are simply visual representations of the more robust regression analysis presented in Table 2.

5.2. Persistence of Restaurant Treatment Effects

One concern might be the existence of a Hawthorne effect, where a new policy has a fleeting impact due to (perceived or real) elevated managerial attention or organizational change (French 1950). To examine this, we adjust our DD specification to allow for separate treatment effect magnitudes in each of the first three months following system implementation. We present these results in Table 3.

Although we observe changes in the size of treatment effects across time, we see no evidence of attenuation during the window of the sample. Treatment effects are relatively consistent, although they increase for losses and tip percentage and decrease for check revenue. We note that one must be careful in interpreting these changes, since we cannot observe some restaurants in their third treatment month (see Figure 2).[3]

[3] See the online appendix for alternative models using only those locations with three months of posttreatment data, which produce similar but slightly larger coefficients.

5.3. Robustness Checks for Model Assumptions

Our DD models rely on several identifying assumptions that we address with additional tests. First, we address the potential concern that at the end of the sample, most restaurants have adopted the system, leaving few control observations. Although location and week fixed effects reduce concerns that this might bias our estimates, we alternatively truncate the sample at different end and start points to test for consistency in our estimated treatment effect. Given that the last week in the sample is 2,711, we rerun our models in Table 2 for samples truncated in weeks 2,701, 2,691, and 2,681. Each of these samples produces estimates that are statistically significant and directionally consistent with estimates in the main model, with most

Table 3 Persistence of Information Technology Impact

Dependent variable:	(1) *Losses*	(2) *Check revenue*	(3) *Drink revenue*	(4) *Tip* percentage
Treatment month 1	−6.75	2,189.034***	699.64***	0.0029***
	(8.92)	(629.96)	(178.11)	(0.0005)
Treatment month 2	−28.81***	2, 874.90***	935.51***	0.0032***
	(7.27)	(622.70)	(186.44)	(0.0006)
Treatment month 3	−47.78***	1,235.45**	546.12***	0.0043***
	(8.10)	(545.17)	(181.78)	(0.0006)
Week fixed effects	Included	Included	Included	Included
Location fixed effects	Included	Included	Included	Included
Constant	122.53***	30,238.04***	7,196.69***	0.1473***
	(4.01)	(781.70)	(135.46)	(0.0012)
Adjusted R^2	0.180	0.824	0.893	0.730
Observations	22,329	22,329	22,329	20,901

Notes. Standard errors clustered at the location level in parentheses. Tip models have fewer observations due to some locations not tracking tips.

*Significant at the 10% confidence level; **significant at the 5% confidence level; ***significant at the 1% confidence level.

estimated treatment effects actually being larger (see the online appendix for results).

Second, we further address the parallel trend assumption that trends among treated and untreated restaurants would be identical in the absence of treatment. To do so, we repeat our models in Table 2 with separate weekly time trends for treated and untreated restaurants. These results, presented in the online appendix, are also consistent with the main models, further supporting the assumption of parallel trends.

Third, we address concerns of endogenous treatment dates. Although the evidence suggests that treatment dates are not correlated with pretreatment theft or revenue, they cannot entirely dispel endogeneity concerns that increased local theft motivates IT adoption. We follow the approach of Granger (1969) and Autor (2003) by including both lags and leads that indicate the 20 weeks before and 20 weeks after the treatment date. The results (presented in the online appendix) are consistent with the monitoring system causing reduced theft. The one exception is an imprecise increase two weeks before implementation from one outlier restaurant that suffered losses of $3,273 and $5,470 in consecutive weeks. Given that IT implementation must be scheduled at least four weeks ahead of time, it is unlikely that these factors drove adoption patterns. Furthermore, our winsorized models (in the online appendix) reduce the impact of such an outlier.

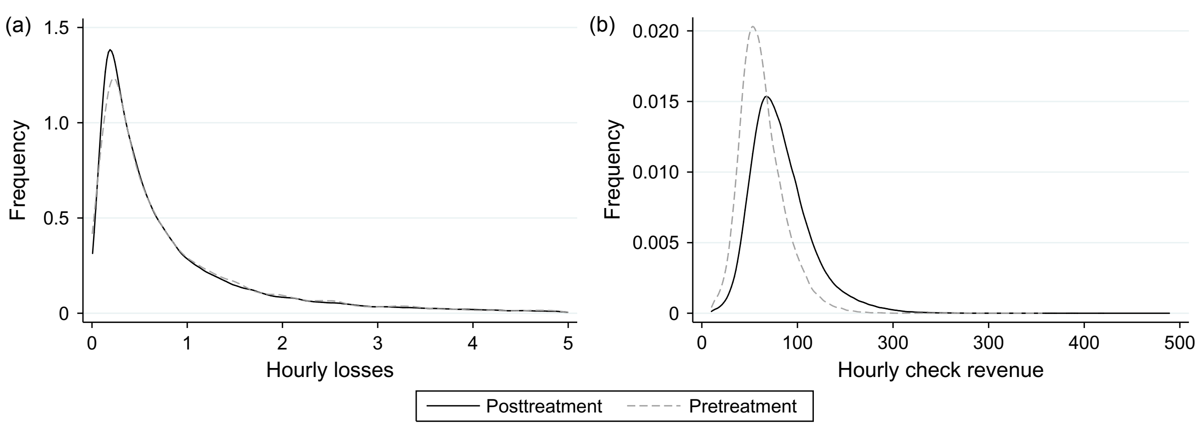
Finally, given the frequency of false positives in DD models (Bertrand et al. 2004), we implement placebo tests that demonstrate that our estimated treatment effects are not artifacts of the data structure. We randomly assign actual treatment dates from the data to each location, then repeat our primary models for each dependent variable from Table 2 with 60 placebo models. At most, one or two placebo models produce coefficients significant at the 5% level for any given dependent variable, and all are considerably smaller and with weaker statistical significance than our true data estimates (see the online appendix).

5.4. Worker Selection vs. Changes in Behavior

Although our initial DD models show clear effects on restaurant-level revenue and losses, they shed little light on the internal organizational mechanisms through which these improvements are achieved. The implementation of IT monitoring leads to reduced theft and larger revenue and profits, but through which mechanisms is it doing so? One possibility is that management is using the system to identify and replace dishonest servers, which would represent a selection mechanism in restaurant improvements. The alternative explanation is that the system is "treating" existing workers, changing their behavior through monitoring. To answer this question, we focus on the server-level data set of weekly losses and productivity. Figure 5 presents the pretreatment and posttreatment distributions of raw hourly theft and productivity for all workers using Epinechnikov kernel density functions; they show identifiable shifts posttreatment (see the online appendix for drink and tip figures).

We first focus on separating worker selection from worker treatment explanations by running pairs of regressions for each of our four dependent variables: *hourly revenue, hourly drink revenue, hourly losses*, and *tip percentage*. Each pair consists of one regression that includes restaurant fixed effects and one that includes worker fixed effects. This approach identifies selection effects as the coefficient in the location fixed effects model minus the coefficient from the individual fixed effects model (Lazear 2000, Hamilton et al. 2003). Since the individual fixed effects model estimates the treatment effect only on workers who span the treatment date, any productivity change caused by changes in the pool of workers should be absorbed by the individual fixed effects. In addition to week

Figure 5 **Distribution of Hourly Losses and Sales Pretreatment and Posttreatment**



Notes. Panel (a) presents Epinechnikov kernel density estimates of raw hourly loss averages at the worker week level for before and after treatment. Panel (b) presents hourly sales revenue averages.

Management Science 61(10), pp. 2299–2319, ©2015 INFORMS

Table 4 Impact of Information Technology on Hourly Measures of Worker Performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:	*Hourly losses*	*Hourly losses*	*Any losses*	*Any losses*	*Hourly revenue*	*Hourly revenue*	*Hourly drink revenue*	*Hourly drink revenue*	*Tip percentage*	*Tip percentage*
Treated with monitoring	−0.053*** (0.009)	−0.056*** (0.009)	−0.017*** (0.005)	−0.012*** (0.004)	1.054 (0.679)	2.690*** (0.557)	0.523*** (0.136)	0.883*** (0.125)	0.0027*** (0.0005)	0.0019*** (0.0004)
Week FE	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Shift controls	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Location FE	Included	—	Included	—	Included	—	Included	—	Included	—
Worker FE	—	Included		Included	—	Included	—	Included	—	Included
Constant	0.163*** (0.006)	−0.056*** (0.009)	0.163*** (0.009)	0.153*** (0.009)	65.028*** (0.830)	60.941*** (0.866)	13.868*** (0.256)	12.009*** (0.427)	0.1574*** (0.0015)	0.1574*** (0.0015)
Adjusted R^2	0.012	0.063	0.067	0.138	0.484	0.648	0.364	0.604	0.231	0.331
Observations	439,838	439,838	439,838	439,838	439,838	439,838	439,838	439,838	437,860	437,860

Notes. Standard errors in parentheses. Errors clustered at location level in all models except (4), where they are clustered at individual level by necessity. Tip models have fewer observations due to some locations not tracking tips. FE, fixed effects.

***Significant at the 1% confidence level.

fixed effects, we add controls for shift allocation. The variable *Highday* indicates the percentage of shifts from higher traffic weekend days, and *daypart* indicates the percentage of higher traffic dinner shifts.

We present these models in Table 4, with standard errors clustered at the restaurant level. Columns (1) and (2) show nearly identical treatment effects on losses of \$0.05 to \$0.06 per hour ($X^2 = 0.75, p = 0.38$), suggesting that theft reductions at the restaurant level are due to existing employees changing their behavior rather than better employees replacing thieves. Because 86.7% of all server-weeks involve *total losses* values of zero, we present linear probability models predicting any theft in columns (3) and (4) of Table 4, regressing a dummy representing any losses with worker or restaurant fixed effects. These models show nearly identical and substantial 1% to 2% decreases in the likelihood of having any total losses in both models against a base rate of 13.3%, although the restaurant fixed effect model coefficient is slightly larger ($X^2 = 4.61, p = 0.03$).

The effects on hourly revenue in columns (5) and (6) of Table 4 show large increases of \$1.05 to \$2.69 per server, although the restaurant fixed effect model is weakly identified ($p = 0.12$). Columns (7) and (8) of Table 4 show drink revenue increases of approximately \$0.52 to \$0.88. The individual fixed effect model coefficients are larger and more precise for both revenue ($X^2 = 44.28, p = 0.001$) and drinks ($X^2 = 18.70, p = 0.001$), suggesting a treatment effect on servers. Columns (9) and (10) of Table 4 present similar tip increases for both models, with the effect slightly larger in the restaurant fixed effect model ($X^2 = 17.19, p = 0.001$). Results are similar for individual fixed effect models using logged dependent variables (see the online appendix). Analysis using hourly revenue and drink sales per customer shows these gains to primarily result from additional sales per customer, not from added customers. The average customer spends an additional \$0.95 per hour posttreatment.

In sum, these results strongly suggest that the majority of productivity improvement and theft reduction is due to behavioral changes among existing workers rather than selection effects due to managers replacing problem workers revealed by the IT system. Selection does appear to play some role, however. Our models suggest that revenue improved more from existing worker changes than from turnover, whereas in contrast, tip percentages appear to have gained more from turnover.

Although our individual server models demonstrate distinct changes in the behavior of existing workers who choose to remain with their employers, they do not inform whether or not the monitoring system implementation induces turnover. Consequently, we further examine the selection mechanism by testing whether those workers with observable (to the researchers) pretreatment instances of theft are more likely than other workers to leave after monitoring system implementation. To test this, we split workers into two groups based on our observation of pretreatment theft. We designate the 4,034 servers with any preimplementation observable (only to the researchers) theft as "known thieves," and the remaining 7,700 servers without researcher-observable theft are designated "unknown," such that we must only include those workers with observable preimplementation hours. The implicit assumption in splitting workers based on preimplementation behavior is that those with observable (to the researchers) theft prior to implementation were also more likely to steal in the postimplementation period, which appears to be true (\$0.177 versus \$0.135 per hour). Furthermore, given the prevalence of unobservable thefts, many of the *unknown* servers are likely involved in theft, creating substantial measurement error in our sample split that biases against identifying differences in treatment effect and may generate false negatives.

We implement a Cox proportional hazards model, where workers are at risk for attrition immediately upon hire, including time-varying covariates that impact their likelihood to exit in any week. We include a quartic time trend to capture seasonal worker turnover and implement gamma-distributed restaurant location random effects as shared frailties to account for specific corporate policies. We interacted our treatment variable with a dummy variable representing a *known thief*. The hazard rate for *treated* represents the posttreatment change for *unknown* servers, and the interaction term represents any additional posttreatment hazard to *known thieves.* We present our results for these models in Table 5, column (1) as hazard ratios, with robust standard errors.

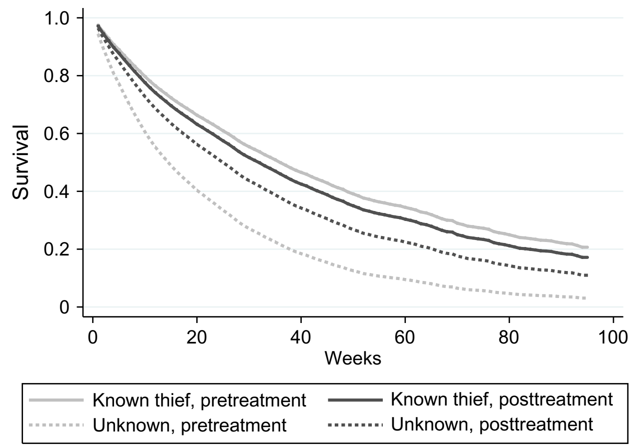
The odds ratio of 0.635 suggests decreased attrition following implementation for *unknown* workers, which is consistent with improved restaurant performance and hourly tip-based income. The interaction term of 1.767 indicates that *known thieves* suffer considerably higher odds of leaving than other workers after system implementation. The total treatment effect of *known thieves,* 1.12 (Wald: $p > X^2 = 0.045$), indicates that the likelihood of *known thieves* increases following implementation, which is consistent with either termination or voluntary attrition. We present pretreatment and posttreatment survival curves for both groups in Figure 6.

Table 5 Cox Models of Postimplementation Attrition

Sample:	(1) All	(2) All
Treated with monitoring	0.635*** (0.029)	0.819*** (0.036)
Known thief	0.450*** (0.020)	—
Treated × Known thief	1.767*** (0.092)	—
Theft occurred in previous two weeks	—	0.809*** (0.043)
Theft occurred and was reported to management in previous two weeks	—	1.070 (0.062)
Week fixed effects	Included	Included
Chain random effects	Included	Included
Subjects	18,885	18,885
Failures	11,255	11,255
Observations	335,681	335,681

Notes. Odds ratios presented, with standard errors in parentheses. All models account for right-hand censoring. The total hazard ratio of treatment on known thieves is $0.450 \times 1.767 = 1.122$ ($p = 0.045$).
***Significant at the 1% confidence level.

Figure 6 Cox Survival Functions for Known and Unknown Thieves



Notes. Survival functions based off Cox proportional hazard model with location-level shared frailties in Table 5, column (1), estimated at frailty = 1. Known thieves are designated by observable theft events prior to treatment date. Unknown includes everyone else, many of whom may be engaged in theft.

To investigate this slightly higher posttreatment attrition, we used another Cox model that interacts a dummy variable indicating any identified theft by the worker in the previous two weeks with the treatment dummy. If management is firing workers after observing theft, we should observe an increased hazard in the two weeks immediately following a theft report. We present this model in column (2) of Table 5. The odds ratio for the noninteracted dummy for recent theft indicates that those who recently stole are less likely to leave the firm in the absence of the monitoring system, possibly because of increased earnings from theft. The interaction, however, suggests that this hazard increases very little with the theft monitoring system. Results remain consistent using alternative postreport periods of between one and four weeks. This result casts doubt on the explanation that workers are being systematically fired following theft detection, and instead suggests that workers whose theft opportunities have decreased after implementation might be leaving for better theft opportunities elsewhere. We acknowledge, however, that since the choice to leave the firm is endogenously related to the choice to steal, we must be cautious in this conclusion.

If managers do not appear to be firing thieves, how do they make monitoring credible? We next examine whether managers punish or reward employees for dishonest or honest behavior by allocating hours differently. To test this, we repeat the worker fixed effects model, using weekly hours as our dependent variable. We test whether *known thieves* begin to receive fewer or worse hours than *unknown* workers once their behavior is revealed using split sample DD models. Columns (1) and (2) of Table 6, which present the total hours results, show that whereas the hours of

Table 6 Staffing Response of Managers to Information Technology Implementation

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Known thief	Unknown	Known thief	Unknown	Known thief	Unknown
Dependent variable:	*Weekly hours*	*Weekly hours*	*Weekend*	*Weekend*	*Dinner*	*Dinner*
Treated with monitoring	0.083	2.249***	0.005	0.002	−0.003	−0.004
	(0.265)	(0.242)	(0.004)	(0.004)	(0.008)	(0.006)
Week fixed effects	Included	Included	Included	Included	Included	Included
Shift controls	Included	Included	Included	Included	Included	Included
Worker fixed effects	Included	Included	Included	Included	Included	Included
Constant	13.778***	12.956***	0.016**	0.037***	0.323***	0.214***
	(0.651)	(0.483)	(0.008)	(0.013)	(0.019)	(0.019)
Adjusted R^2	0.418	0.493	0.302	0.364	0.635	0.646
Observations	123,091	177,541	123,091	177,541	123,091	177,541

Notes. Known thieves are designated by observable theft events prior to treatment date. Unknown includes everyone else, many of whom may be engaged in theft. Only those workers with at least 40 hours of preimplementation employment were included. Standard errors clustered at the location level in parentheses.

**Significant at the 5% confidence level; ***significant at the 1% confidence level.

known thieves do not change following monitoring system implementation, the hours of *unknown* workers increase by more than 10%.

Although we cannot observe why this occurs, the substantially lower growth in weekly hours for *known thieves* (Wald: $p > X^2 = 0.00$) suggests that managers may be allocating additional hours to servers who appear honest as their tenure at the restaurant grows, which could also explain why *known thieves* are more likely to leave posttreatment. However, columns (3)–(6), which predict percentage of time staffed for high-traffic weekend and dinner shifts, show no difference across the two worker types (Wald: $p > X^2 = 0.67$; $p > X^2 = 0.87$). We must, therefore, be careful in claiming that managers systematically punish thieves through shift assignment. But relative to other workers, *known thieves* lose hours following theft monitoring implementation.

Why does management not immediately fire employees who are detected stealing? One likely reason is the cost of hiring and training new servers. The average server in the sample sells only \$57.64 per hour in her first week, with productivity growing through the first four months to more than \$80 per hour. Management may tolerate the low levels of theft that occur after implementation of the monitoring technology, given the more substantial hourly sales that would occur with replacement workers. Still, managers appear to be allocating fewer hours to those who steal more. This may provide sufficient incentive to reduce theft behavior, and may be a less costly way than termination to rid a restaurant of the worst thieves.

5.5. Mechanisms Explaining Worker Productivity Gains

If productivity gains primarily result from the changed behavior of existing workers, and not selection, which mechanisms contribute to this behavioral modification? In this section we present evidence related to four mechanisms. We provide evidence that neither economic nor cognitive multitasking explains the majority of productivity increases, but that fairness may play some role. We are left to speculate that much of the observed productivity growth may be caused by increased managerial and worker attention.

Before exploring these mechanisms, we first show that the observed \$2,975 weekly revenue increases cannot be explained by previously unobserved theft appearing as recognized revenue after system implementation. Although some of the observed increase may reflect this effect, there are multiple reasons it represents at most only a fraction of the improvement. First, 88.3% of total losses involve drinks (e.g., soft drinks, alcohol, coffee; see Figure 3), meaning that although some of the \$927 weekly increase in drink sales could be mechanical, the much larger total revenue effect cannot be pure theft replacement. Second, the increase in tip percentage is inconsistent with this argument, since unobservable theft decreases the denominator in tip percentage (revenue) and, for credit card transactions, increases the numerator (see §2). Consequently, a decrease in unobservable theft should mechanically *decrease* tip percentage. Similarly, unobservable theft from servers trading free items (drinks, entrees, desserts) to customers for reciprocal tip increases is unlikely to explain our results. If the item never registers in the POS system, then the denominator on tip percentage is again artificially low. Deterring such theft would decrease tip percentage from its artificially high level before implementation. The finding of an increase in tip percentage, therefore, is inconsistent with the mechanical replacement explanation and instead likely reflects improved service quality.

Table 7 Testing the Relationship Between Changes in Posttreatment Loss and Productivity

Dependent variable:	(1) *Hourly revenue*	(2) *Hourly drink revenue*	(3) *Tip percentage*
Treated with monitoring	1.116*	0.547***	0.0271***
	(0.672)	(0.136)	(0.0005)
Total losses	1.019***	0.431***	−0.0006**
	(0.322)	(0.151)	(0.0003)
Treated × *Total losses*	−0.073	−0.012	0.0001
	(0.315)	(0.151)	(0.0003)
Week fixed effects	Included	Included	Included
Shift controls	Included	Included	Included
Location fixed effects	Included	Included	Included
Constant	64.869***	13.798***	0.1589***
	(0.830)	(0.258)	(0.0016)
Adjusted R^2	0.484	0.365	0.231
Observations	439,838	439,838	437,860

Notes. Regressions are the equivalent of the final step in mediation analysis, where the interaction is hypothesized to mediate the direct effect-treatment with monitoring. The coefficients for treatment can be compared with those in columns (1), (5), and (7) of Table 4, which represent the total effect. Insignificant coefficients for the interaction and similar coefficients for treated in Table 4 suggest limited or no mediation. Standard errors clustered at the location level in parentheses. Tip models have fewer observations due to some locations not tracking tips.

*Significant at the 10% confidence level; **significant at the 5% confidence level; ***significant at the 1% confidence level.

The first mechanism we examine is economic multitasking. We conduct regression analysis similar to the fourth step in the Baron and Kenny (1986) mediation analysis, separately regressing revenue, drink sales, or tip percentage on losses, our treatment dummy, and their interaction while controlling for week and restaurant fixed effects. In this model, the treatment dummy represents the equivalent of the direct effect of theft monitoring through other mechanisms. The interaction term of posttreatment losses acts as a mediator (change in losses) that might explain the relationship between treatment and productivity gains. If posttreatment loss decreases explain productivity gains, then the treatment coefficient should be much smaller than in the restaurant fixed effect models in Table 4 (which represent the total effect in mediation analysis), and the interaction effect should be negative and significant. The results, presented in Table 7, show no evidence of posttreatment changes in losses predicting changes to revenue, drink sales, or tip percentage.

Interaction coefficients are small in magnitude and are not statistically significant. The direct effect, represented by the coefficients on *treated*, continues to be strong, suggesting that other mechanisms explain the revenue gains. We caution, however, that mediation analysis is a rough, but not exact, fit with the equilibrium predictions from an economic multitasking model, since the mediator (losses) is an endogenous regressor, making the second step of mediation analysis (regressing productivity on losses) inappropriate and biased. With all those caveats, these results seem to indicate that economic multitasking is not a primary mechanism driving the observed productivity gains.

We test the second mechanism, cognitive multitasking, by identifying how theft monitoring affects workers during periods of different workloads. If the link between theft reduction and productivity gains is due to reduced cognitive costs, then we should see substantially larger increases during high workloads that challenge workers' cognitive limitations. Table 8 presents worker fixed effect models after splitting the sample by high- (Thursday–Saturday) and low-traffic (Sunday–Wednesday) days.

Table 8 Treatment Effects by Customer Traffic Level

Sample:	(1) High traffic	(2) Low traffic	(3) High traffic	(4) Low traffic	(5) High traffic	(6) Low traffic	(7) High traffic	(8) Low traffic
Dependent variable:	*Hourly losses*	*Hourly losses*	*Hourly revenue*	*Hourly revenue*	*Hourly drink revenue*	*Hourly drink revenue*	*Tip percentage*	*Tip percentage*
Treated with monitoring	−0.067***	−0.044***	2.312***	2.902***	0.788***	0.990***	0.0017***	0.0021***
	(0.013)	(0.010)	(0.573)	(0.550)	(0.146)	(0.136)	(0.0005)	(0.0005)
Week fixed effects	Included	Included	Included	Included	Included	Included	Included	Included
Shift controls	Included	Included	Included	Included	Included	Included	Included	Included
Worker fixed effects	Included	Included	Included	Included	Included	Included	Included	Included
Constant	0.172***	0.153***	74.363***	60.435***	14.648***	12.026***	0.1574***	0.1578***
	(0.008)	(0.012)	(0.883)	(0.912)	(0.520)	(0.398)	(0.0015)	(0.0016)
Adjusted R^2	0.026	0.076	0.532	0.584	0.529	0.531	0.216	0.260
Observations	396,962	427,023	396,962	427,023	396,962	427,023	394,745	424,818

Notes. Standard errors in parentheses. Errors clustered at location level in all models except (4), where they are clustered at individual level by necessity. Tip models have fewer observations due to some locations not tracking tips.

***Significant at the 1% confidence level.

Table 9 Split Sample Models by Observable Pretreatment

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Known thief	Unknown	Known thief	Unknown	Known thief	Unknown
Dependent variable:	*Hourly revenue*	*Hourly revenue*	*Hourly drink revenue*	*Hourly drink revenue*	*Tip percentage*	*Tip percentage*
Treated with monitoring	1.876***	4.513***	0.964***	1.028***	0.0028***	0.0011**
	(0.724)	(0.486)	(0.182)	(0.121)	(0.006)	(0.0005)
Shift controls	Included	Included	Included	Included	Included	Included
Week fixed effects	Included	Included	Included	Included	Included	Included
Worker fixed effects	Included	Included	Included	Included	Included	Included
Constant	62.690***	63.641***	12.924***	13.176***	0.154***	0.1587***
	(0.941)	(0.986)	(0.529)	(0.296)	(0.002)	(0.0020)
Adjusted R^2	0.657	0.654	0.626	0.618	0.377	0.290
Observations	123,091	177,541	123,091	177,541	122,800	177,093

Notes. Known thieves are designated by observable theft events prior to treatment date. Unknown includes everyone else, many of whom may be engaged in theft. Standard errors clustered at the location level in parentheses.
**Significant at the 5% confidence level; ***significant at the 1% confidence level.

Although high-traffic days enjoy slightly larger decreases in theft than do low-traffic days (Wald: $p > X^2 = 0.09$), there are small and statistically weak differences between the day-types in revenue (Wald: $p > X^2 = 0.07$), drink revenue (Wald: $p > X^2 = 0.13$), and tip percentage (Wald: $p > X^2 = 0.30$). In fact, these improvements are slightly larger in low-traffic days, which is inconsistent with cognitive multitasking.

The third mechanism we examine is whether productivity and tip gains might result from honest employees enjoying higher motivation after theft monitoring eliminated unfair income differences from theft. If improved motivation from fairness were driving productivity and service gains, we would expect much larger treatment effects for *unknown* workers than for *known thieves*. We repeated the core worker fixed effect models (losses, revenue, drinks, and tips) using split samples, with mixed results presented in Table 9.

Although the change in drink sales from the two groups is nearly identical (Wald: $p > X^2 = 0.72$), total revenue increases are larger for *unknown* workers (Wald: $p > X^2 = 0.00$), whereas tip percentage is smaller (Wald: $p > X^2 = 0.01$). Thus, although the revenue results are consistent with a fairness mechanism, tip results are not. These ambiguous results are also inconsistent with an economic multitasking mechanism, where *known thieves* might produce the largest revenue increases to compensate for lost theft. We note, however, that these models potentially suffer from mean reversion concerns common in split-sample DID models, where correlation between the sample-splitting criterion and dependent variable can generate spurious results.

Finally, we note that we do not directly observe how employees perceived managerial oversight and attention toward productivity may have changed following the system implementation. Interviews with managers revealed that news of the monitoring systems disseminated quickly among employees, although we were told that workers knew little about what types of theft it monitored and whether they had already been identified personally by the system. As we show in the next section, an employee rankings report bundled with the system appears to have had little impact on individual employees singled out for their good performance. Given the improved productivity for a group of employees not included in by the rankings report (i.e., bartenders), however, perceptions of increased managerial oversight tied to the theft monitoring system seem likely to play a key role in explaining the link between theft monitoring and productivity.

5.6. Robustness to Server Rankings Treatment

One potential confounding factor in the productivity results is an employee ranking report that was bundled with the theft monitoring system. This report gave a manager weekly productivity rankings for the top 10 servers (but not bartenders) in her location based on 10 criteria. Interviews revealed no evidence that reports were posted for viewing by employees, a policy sometimes employed in sales and manufacturing settings (Netessine and Yakubovich 2012). We perform a series of robustness checks to allay the concern that this potential second treatment might explain many of the productivity gains observed in our regression models based on two mechanisms summarized in Netessine and Yakubovich (2012). First, it might lead management to staff higher-ranked (and thus more productive) workers to higher-traffic shifts, thereby increasing overall worker productivity. If this shift allocation response indeed happened, we would expect higher-ranked workers to receive increased and better shift hours relative to others after system implementation. Second, rankings might motivate workers to achieve improved productivity either because of the shame of a low ranking or the status of a high ranking, or simply because they seek the financial rewards of winning more lucrative shift assignments. As Barankay (2014) notes, however, the empirical evidence on productivity effects from ranking systems is inconsistent because of the many theoretical mechanisms involved.

To examine this alternative explanation, we first exploit a subsample of workers who were affected by the theft monitoring but not by the productivity reports—bartenders. If theft monitoring is responsible for the productivity improvements in our earlier models, we should observe similar results for bartenders. If instead, productivity rankings are driving our results, we should see no changes in bartender productivity. We note that although bartender tickets may be transferred to a server's ticket (e.g., the bar patron takes his or her drink to a table for dinner without closing the ticket), the opposite rarely happens. Thus, although improved productivity or better service by a bartender might spill over to a server's data, a bartender's data almost certainly reflects only the bartender's actions. In other words, possible performance improvement from productivity rankings is highly unlikely to impact a bartender, which allows us to isolate the impact of the theft monitoring system. We repeat our DD models with individual fixed effects and present them in Table 10.

We see very similar results for the models in Table 10. Although the number of observations is greatly reduced, we see large and significant increases in check revenue, drinks, and tip percentage, as well

Table 10 **Models for Bartenders Who Were Not Part of Ranking System**

Sample	(1) Bartenders	(2) Bartenders	(3) Bartenders	(4) Bartenders
Dependent Variable:	*Hourly losses*	*Hourly revenue*	*Hourly drink revenue*	*Tip percentage*
Treated with monitoring	−0.046 (0.032)	3.314*** (1.004)	1.013*** (0.376)	0.0031* (0.018)
Week fixed effects	Included	Included	Included	Included
Shift controls	Included	Included	Included	Included
Worker fixed effects	Included	Included	Included	Included
Constant	0.443*** (0.168)	79.099*** (1.963)	8.307*** (1.063)	0.0979*** (0.0130)
Adjusted R^2	0.016	0.584	0.622	0.353
Observations	33,753	33,753	33,753	33,492

Notes. Bartenders' samples include only those employees who worked the entire week under a bartender job code. Most bartenders at casual dining restaurants also serve food at the bar. Standard errors clustered at the location level in parentheses. Tip models have fewer observations due to some locations not tracking tips.

*Significant at the 10% confidence level; ***significant at the 1% confidence level.

as a weakly identified decrease in theft. Since bartenders are unaffected by the server ranking system, these results cast considerable doubt that server productivity ranking reports drive our main results.

We next test whether management responded to the rankings by giving desirable shifts and hours to highly ranked employees. Using the ranking algorithm provided by the data provider, we calculated the ranking sent to management each week for each server. We split all employees who worked the first four weeks following implementation into four groups based on whether their average weekly ranking during this period was 1 to 5, 6 to 10, 11 to 15, or 16 and above. A server with an average rank of 3.5, for example, would be in the first group, and one with an average of 13.5 would be in the third group. Although only the top 10 servers showed up in the rankings each week, many of those in the third tier achieved this top 10 status in at least one week. We interact the treatment dummy variable with dummy variables representing each of the top three ranking groups and estimate the impact of ranking on employee hours and high-traffic dinner and weekend shifts. We present the results in columns (1)–(3) of Table 11.

Since the omitted interaction is for the lowest-ranked group, the coefficient on *treated* represents their treatment effect, and the interaction terms represent any *additional* impact of the monitoring system to higher-ranked groups. We observe no statistically significant differences in shift assignment between the top three groups following the ranking implementation. The bottom-ranked group is weakly different in hours from the second quartile group ($p < 0.1$). Managers are clearly not using the rankings to assign shifts or hours, which casts strong doubt that the ranking system had a major effect on the restaurants.

We also examine whether the rankings have any relationship with productivity changes after implementation of the monitoring system. If the ranking system affects worker motivation, we might observe differences in productivity increases across ranking levels, although as we note earlier, the theoretical predictions on this relationship are multitudinous and the existing empirical evidence inconsistent (Barankay 2014). We present results for these models in columns (3)–(7). For the top three groups (i.e., the top 15 employees), we find few statistically significant differences.[4] The only differences are for the lowest-ranked employees, who have lower total revenue increases and higher tip percentage increases after implementation.

[4] T-tests tested the equality of coefficients for each group's treatment effect. For the top three groups, only the change in drink sales for group 3 was significant at the 10% level.

Table 11 Tests for Alternative Productivity Rankings Treatment

Dependent variable:	(1) *Hours*	(2) *Dinner*	(3) *Weekend*	(4) *Losses*	(5) *Check revenue*	(6) *Drink revenue*	(7) *Tip percentage*
Treated	0.907**	−0.011	0.010**	−0.062***	0.918	0.825***	0.0031***
	(0.371)	(0.014)	(0.005)	(0.016)	(0.844)	(0.250)	(0.0007)
Treated × *Top five*	0.473	0.002	−0.009	0.029	2.395***	−0.292	−0.0028***
	(0.292)	(0.017)	(0.007)	(0.019)	(0.847)	(0.356)	(0.0008)
Treated × *Second five*	0.671*	0.020	−0.010	0.003	2.208***	−0.030	−0.0017***
	(0.363)	(0.015)	(0.006)	(0.019)	(0.670)	(0.271)	(0.0007)
Treated × *Third five*	0.351	0.009	−0.006	−0.004	2.817***	0.450	−0.0020***
	(0.403)	(0.018)	(0.007)	(0.018)	(0.698)	(0.318)	(0.0008)
Week fixed effects	Included	Included	Included	Included	Included	Included	Included
Shift controls	Included	Included	Included	Included	Included	Included	Included
Worker fixed effects	Included	Included	Included	Included	Included	Included	Included
Constant	13.56***	0.278***	0.362***	0.159***	61.565***	11.366***	0.1556***
	(0.427)	(0.019)	(0.014)	(0.011)	(0.870)	(0.434)	(0.0015)
Adjusted R^2	0.065	0.600	0.358	0.065	0.660	0.580	0.315
Observations	406,085	406,085	406,085	406,085	406,085	406,085	404,368

Notes. Standard errors clustered at the location level in parentheses. Tip models have fewer observations due to some locations not tracking tips.
*Significant at the 10% confidence level; **significant at the 5% confidence level; ***significant at the 1% confidence level.

The results reported in Tables 10 and 11 provide evidence that the productivity rankings did not play a direct role in the overall observable treatment effect.

6. Discussion and Conclusion

In this paper, we show evidence that the use of information technology can both substantially improve the productivity of employees and organizations as well as reduce the corrupt behavior of employees. Furthermore, our results suggest that the majority of improvement in organizational performance and productivity stems from the improved behavior of existing employees, not from the firing of those engaged in theft. Although worker selection may also play a role in our setting, systematic attrition (whether voluntary or termination) cannot explain our results.

We argued that a combination of four individual-level mechanisms might explain our results on productivity and service quality improvements: economic multitasking, cognitive multitasking, motivation from improved fairness, and perceptions of increased productivity monitoring. We presented analysis focused on isolating each mechanism. Our results cast significant doubt on both the cognitive and economic multitasking mechanisms, and provide mixed evidence on fairness concerns. Although we cannot directly test for perceptions of increased productivity monitoring, this explanation seems most consistent with our results. Improved monitoring technology, by reducing the managerial attention needed to stop theft, may free the manager to focus more on both directly monitoring and facilitating employee productivity. Management also faces a multitasking challenge, with fundamental trade-offs between misconduct and productivity. The theft monitoring system, by reducing necessary effort toward reducing theft, allows additional managerial effort to be focused toward improving productivity. We note that this characterization of the manager's multitasking challenge comports well with the one in DeHoratius and Raman (2007). In that case, however, the implementation of managerial incentives toward one task (productivity) produces negative spillovers to the other (theft). This highlights one of the key benefits of using technology-based solutions to address operations management problems: they can lower costs while avoiding many of the agency problems that result from solutions based in financial incentives.

We caution that other cost-based worker activities remain unobservable in these data. We cannot, for example, observe whether reducing one type of theft (stealing revenue) through monitoring increases other forms of theft or misconduct such as inventory shrinkage. Given the Olken (2007) results on substitution across types of corruption, such costs may very well exist and, thereby, reduce the profit gains from monitoring. This is particularly a concern in the restaurant industry, where inventory losses can be substantial. Future work could integrate POS and inventory data to test for this spillover effect. We also note that the transaction-based monitoring in this setting is relatively unobtrusive compared to alternatives, such as camera, managerial shadowing, or email and Internet usage tracking. We, therefore, caution that excessive employee monitoring or surveillance might indeed reduce employee productivity, as theory suggests (Cialdini 1996, Bénabou and Tirole 2006). Furthermore, some of the theft in these data may be tolerated or even encouraged by local managers for financial or social reasons, so one must be careful not to assign ethical responsibility to servers and bartenders for actions facilitated by local management.

Finally, our results represent average effects that ignore organizational and managerial characteristics that could change how monitoring affects a given restaurant location. Although the data provider told us that almost all of the locations are corporate owned, there may be differences in how IT-based theft monitoring is used in a franchise, where the higher-powered incentives for local managers to reduce theft might make monitoring more effective. The sample limits our ability to explore heterogeneity across locations, but this would be a fruitful path for future research.

The results in this paper are important for several reasons. First, they represent the measurement of an important economic activity—employee theft—that largely has been observed only indirectly or anecdotally. Although there is a large literature on corruption (e.g., Olken and Barron 2009), direct behavioral evidence on illicit behavior by individual employees is rare in the economics, management, and operations literatures. Nagin et al. (2002) is a rare exception. We are able to show not only the direct effect of monitoring on theft, as they do, but also the secondary employee adjustments to other productive tasks to account for lost income. Furthermore, our multifirm setting, long panel, and staggered treatment dates provide additional contributions in understanding employee turnover and controlling for omitted time- and firm-specific factors.

Second, our results suggest a counterintuitive and hopeful pattern in human behavior—employee theft is a remediable problem *at the individual level*. Although individual differences in moral preferences may indeed exist, subtle or inexpensive changes in managerial practices or technology can have powerful effects in reducing misconduct. This runs counter to a common view in the human resource management literature that productivity and integrity are about selection rather than managerial practice or technology (Ones et al. 1993). We show that firms can use information about employee theft not simply to fire the culprits, but rather to alter their behavior in ways that improve productivity and increase legitimate worker income. Our models suggest that the theft monitoring system increases average hourly tip income by \$0.58 per worker.[5] This is consistent with evidence and theory from lean manufacturing (Deming 1938, Adler et al. 1997) and behavioral ethics (Mazar et al. 2008) that theft and other misconduct is the work of many individuals stealing relatively small amounts rather than a few "bad apples" who can be eliminated to remove the problem.

[5] This income increase is calculated from the coefficients of our worker fixed effects models in Table 4, which include an additional 0.2% in tips on hourly sales of \$79 plus the average tip of 15.7% for an additional \$2.02 in hourly sales.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mnsc.2014.2103.

Acknowledgments

The data were provided by NCR Corporation. NCR provided no financial support or compensation to the authors, nor did the company retain the right to edit or block publication of the paper. The authors thank Arthur Brief, Ernesto Dal Bó, Rui de Figueiredo, Barton Hamilton, Andrew Knight, Ian Larkin, Lars Lefgren, and Jason Snyder, as well as participants in seminars at Harvard Business School, Massachusetts Institute of Technology, University of California at Berkeley, HEC Lausanne, the University of Minnesota, Oxford University, Boston University, Washington University in St. Louis, and Brigham Young University for their comments and suggestions on the paper. Jeff Hindman and Scott Walton were critical to acquiring and understanding the data, and Thomas Bennett, David Hall, and John Holbrook provided invaluable research support.

References

Adams JS (1965) Inequity in social exchange. Berkowitz L, ed. *Advances in Experimental Social Psychology* (Academic Press, New York), 267–299.

Adler PS, Goldoftas B, Levine D (1997) Ergonomics, employee involvement, and the Toyota Production System: A case study of NUMMI's 1993 model introduction. *Indust. Labor Relations Rev.* 50:416–437.

Aral S, Brynjolfsson E, Van Alstyne M (2012) Information, technology, and information worker productivity. *Inform. Systems Res.* 23:849–867.

Athey S, Stern S (2002) The impact of information technology on emergency health care outcomes. *RAND J. Econom.* 33:399–432.

Autor DH (2003) Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *J. Labor Econom.* 21:1–42.

Baker GP, Hubbard TN (2003) Make versus buy in trucking: Asset ownership, job design, and information. *Amer. Econom. Rev.* 93:551–572.

Banker R, Lee S, Potter G, Srinivasan D (2001) An empirical analysis of continuing improvements following the implementation of a performance-based compensation plan. *J. Accounting Econom.* 30:315–350.

Barankay I (2014) Rank incentives: Evidence from a randomized workplace experiment. Working paper, The Wharton School, University of Pennsylvania, Philadelphia.

Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J. Personality Soc. Psych.* 51:1173–1182.

Bazerman M, Loewenstein G, Moore D (2002) Why good accountants do bad audits. *Harvard Bus. Rev.* 80:96–103.

Becker G (1968) Crime and punishment: An economic approach. *J. Political Econom.* 76:169–217.

Bénabou R, Tirole J (2006) Incentives and prosocial behavior. *Amer. Econom. Rev.* 96:1652–1678.

Bénabou R, Tirole J (2011) Identity, morals, and choices. *Quart. J. Econom.* 126:805–855.

Bernstein E (2012) The transparency paradox: A role for privacy in organizational learning and operational control. *Admin. Sci. Quart.* 57:181–216.

Bernstein S, Sheen A (2014) The operational consequences of private equity buyouts: Evidence from the restaurant industry. Working paper, Stanford University, Stanford, CA.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust difference-in-differences? *Quart. J. Econom.* 119:249–275.

Bloom N, Sadun R, Van Reenen J (2012) Americans do IT better: U.S. multinationals and the productivity miracle. *Amer. Econom. Rev.* 102:167–201.

Bodvarsson O, Luksetich W, McDermott S (2003) Why do diners tip: Rule-of-thumb or valuation of service? *Appl. Econom.* 35:1659–1665.

Bresnahan T, Brynjolfsson E, Hitt L (2002) Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quart. J. Econom.* 117:339–376.

Brynjolfsson E (1993) The productivity paradox of information technology. *Comm. ACM* 36:66–77.

Brynjolfsson E, Hitt L (1996) Paradox lost? Firm-level evidence on the returns to information systems spending. *Management Sci.* 42:541–558.

Brynjolfsson E, McAfee A (2014) *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies* (WW Norton & Company, New York).

Campbell D (2012) Employee selection as a control system. *J. Accounting Res.* 50:931–966.

Charron S, Koechlin E (2010) Divided representation of concurrent goals in the human frontal lobes. *Science* 328:360–363.

Chen C, Sandino T (2012) Can wages buy honesty? The relationship between relative wages and employee theft. *J. Accounting Res.* 50:967–1000.

Cialdini R (1996) The triple tumor structure of organizational behavior. Messic DM, Tenbrunsel AE, eds. *Codes of Conduct* (Russell Sage Foundation, New York), 44–58.

Clark JR, Huckman RS (2012) Broadening focus: Spillovers, complementarities, and specialization in the hospital industry. *Management Sci.* 58:708–722.

Cohen L, Frazzini A, Malloy C (2010) Sell-side school ties. *J. Finance* 65:1409–1437.

Dal Bó E, Terviö M (2013) Self-esteem, moral capital, and wrongdoing. *J. Eur. Econom. Assoc.* 11:599–663.

DeHoratius N, Raman A (2007) Store manager incentive design and retail performance: An exploratory investigation. *Manufacturing Service Oper. Management* 9:518–534.

DeHoratius N, Raman A (2008) Inventory record inaccuracy: An empirical analysis. *Management Sci.* 54:627–641.

DellaVigna S, La Ferrara E (2010) Detecting illegal arms trade. *Amer. Econom. J.: Econom. Policy* 2:26–57.

Deming WE (1938) *Statistical Adjustment of Data* (John Wiley & Sons, New York).

Deming WE (1993) *The New Economics for Industry, Government, Education* (MIT Press, Cambridge, MA).

Detert JR, Trevino LK, Burris ER, Andiappan M (2007) Managerial modes of influence and counterproductivity in organizations: A longitudinal business-unit-level investigation. *J. Appl. Psych.* 92:993–1005.

Dickens W, Katz L, Lang K, Summers L (1989) Employee crime and the monitoring puzzle. *J. Labor Econom.* 7:331–347.

Draca M, Machin S, Witt R (2011) Panic on the streets of London: Police, crime, and the July 2005 terror attacks. *Amer. Econom. Rev.* 101:2157–2181.

Duflo E, Hanna R, Ryan S (2012) Incentives work: Getting teachers to come to school. *Amer. Econom. Rev.* 102:1241–1278.

Fisman R, Miguel E (2007) Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *J. Political Econom.* 115:1020–1048.

Fisman R, Wei SJ (2004) Tax rates and tax evasion: Evidence from "missing imports" in China. *J. Political Econom.* 112:471–496.

Fisman R, Wei SJ (2009) The smuggling of art, and the art of smuggling: Uncovering the illicit trade in cultural property and antiques. *Amer. Econom. J.: Appl. Econom.* 1:82–96.

Francis P, DeGlinkta RC (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).

French JRP (1950) Field experiments: Changing group productivity. Miller JG, ed. *Experiments in Social Process: A Symposium on Social Psychology* (McGraw-Hill, New York), 79–96.

Frey B (1993) Does monitoring increase work effort? The rivalry with trust and loyalty. *Econom. Inquiry* 31:663–670.

Gino F, Pierce L (2009) Dishonesty in the name of equity. *Psych. Sci.* 20:1153–1160.

Gino F, Pierce L (2010a) Lying to level the playing field: Why people may dishonestly help or hurt others to create equity. *J. Bus. Ethics* 95:89–103.

Gino F, Pierce L (2010b) Robin Hood under the hood: Wealth-based discrimination in illicit customer help. *Organ. Sci.* 21:1176–1194.

Gino F, Pisano G (2008) Toward a theory of behavioral operations. *Manufacturing Service Oper. Management* 10:676–691.

Granger C (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37:424–438.

Greenberg J (1990) Employee theft as a reaction to underpayment inequity: The hidden cost of pay cuts. *J. Appl. Psych.* 75: 561–568.

Hamilton BH, Nickerson JA, Owan H (2003) Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *J. Political Econom.* 111:465–497.

Heron RA, Lie E (2007) Does backdating explain the stock price pattern around executive stock option grants? *J. Financial Econom.* 83:271–295.

Hochschild AR (1983) *The Managed Heart: Commercialization of Human Feeling* (University of California Press, Berkeley).

Hölmstrom B, Milgrom P (1991) Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *J. Law, Econom., Organ.* 7:24–52.

Holweg M (2007) The genealogy of lean production. *Production Oper. Management* 25:420–437.

Hubbard TN (2000) The demand for monitoring technologies: The case of trucking. *Quart. J. Econom.* 115:533–560.

Huckman RS, Staats BR, Upton DM (2009) Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Sci.* 55(1):85–100.

Jin G, Leslie P (2003) The effect of information on product quality: Evidence from restaurant hygiene grade cards. *Quart. J. Econom.* 118:409–451.

Jin G, Leslie P (2009) Reputational incentives for restaurant hygiene. *Amer. Econom. J.: Microeconomics* 1:237–267.

KC DS (2014) Does multitasking improve performance? Evidence from the emergency department. *Manufacturing Service Oper. Management* 16:168–183.

KC DS, Terwiesch C (2009) Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Sci.* 55(9):1486–1498.

KC DS, Terwiesch C (2012) An econometric analysis of patient flows in the cardiac intensive care unit. *Manufacturing Service Oper. Management* 14:50–65.

Larkin I, Pierce L, Gino F (2012) The psychological costs of pay-for-performance: Implications for the strategic compensation of employees. *Strategic Management J.* 33:1194–1214.

Latham G (2001) The importance of understanding and changing employee outcome expectancies for gaining commitment to an organizational goal. *Personnel Psych.* 54:707–716.

Lazear EP (1986) Salaries and piece-rates. *J. Bus.* 59:405–431.

Lazear EP (2000) Performance pay and productivity. *Amer. Econom. Rev.* 90:1346–1361.

Lazear E, Oyer P (2011) Personnel economics: Hiring and incentives. Ashenfelter O, Card D, eds. *Handbook of Labor Economics*, Vol. 4B (North-Holland, Amsterdam), 1769–1823.

Litzky BE, Eddleston KA, Kidder DL (2006) The good, the bad, and the misguided: How managers inadvertently encourage deviant behaviors. *Acad. Management Perspect.* 20:91–103.

Lochner L (2007) Individual perceptions of the criminal justice system. *Amer. Econom. Rev.* 97:444–460.

Mazar N, Amir O, Ariely D (2008) The dishonesty of honest people: A theory of self-concept maintenance. *J. Marketing Res.* 45: 633–644.

Mazar N, Ariely D (2006) Dishonesty in everyday life and its policy implications. *J. Public Policy and Marketing* 4:117–126.

Mead N, Baumeister R, Gino F, Schweitzer M, Ariely D (2009) Too tired to tell the truth: Self-control resource depletion and dishonesty. *J. Experiment. Soc. Psych.* 45:594–597.

Moore C, Detert J, Trevino L, Baker V, Mayer D (2012) Why employees do bad things: Moral disengagement and unethical organizational behavior. *Personnel Psych.* 65:1–48.

Murphy K (1993) *Honesty in the Workplace* (Brooks/Cole, Belmont, CA).

Nagin D, Rebitzer J, Sanders S, Taylor L (2002) Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *Amer. Econom. Rev*. 94:950–873.

Netessine S, Yakubovich V (2012) The Darwinian workplace. *Harvard Bus. Rev.* 90:25–28.

Niehaus P, Sukhtankar S (2013) Corruption dynamics: The golden goose effect. *Amer. Econom. J.: Econom. Policy* 5:230–269.

Niehoff B, Moorman R (1993) Justice as a mediator of the relationship between methods of monitoring and organizational citizenship. *Acad. Management J.* 36:527–556.

Olken BA (2007) Monitoring corruption: Evidence from a field experiment in Indonesia. *J. Political Econom.* 115:200–249.

Olken BA, Barron P (2009) The simple economics of extortion: Evidence from trucking in Aceh. *J. Political Econom.* 117:417–452.

Ones D, Viswesvaran C, Schmidt F (1993) Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *J. Appl. Psych.* 78:679–703.

Pierce L, Snyder JA (2008) Ethical spillovers in firms: Evidence from vehicle emissions testing. *Management Sci.* 54:1891–1903.

Pierce L, Snyder JA (2012) Discretion and manipulation by experts: Evidence from a vehicle emissions policy change. *BE J. Econom. Anal. Policy* 12(3) Article 1. DOI: 10.1515/1935-1682.3246.

Rawley E, Simcoe TS (2013) Information technology, productivity and asset ownership: Evidence from taxicab fleets. *Organ. Sci.* 24:832–845.

Roy D (1952) Quota restriction and goldbricking in a machine shop. *Amer. J. Sociol.* 57:427–442.

Schultz KL, McClain JO, Thomas LJ (2003) Overcoming the dark side of worker flexibility. *J. Oper. Management* 21:81–92.

Spear S, Bowen H (1999) Decoding the DNA of the Toyota Production System. *Harvard Bus. Rev.* 77:96–108.

Staats BR, Gino F (2012) Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Management Sci.* 58(6):1141–1159.

Sweeney D, Steinhauser S (2010) *Restaurant Industry Operations Report* (National Restaurant Association, Washington, DC).

Tan TF, Netessine S (2014) When does the devil make work? An empirical study of the impact of workload on worker productivity. *Management Sci.* 60:1574–1593.

Tenbrunsel AE, Smith-Crowe K (2008) Ethical decision making: Where we've been and where we're going. *Management Ann.* 2:545–607.

Tucker A, Edmondson A (2003) Why hospitals don't learn from failures: Organizational and psychological dynamics that inhibit system change. *Calif. Management Rev.* 45:55–72.

Victor B, Trevino L, Shapiro D (1993) Peer reporting of unethical behavior: The influence of justice evaluations and social context factors. *J. Bus. Ethics* 12:253–263.

Yang D (2008a) Integrity for hire: An analysis of widespread customs reform. *J. Law Econom.* 51:25–57.

Yang D (2008b) Can enforcement backfire? Crime displacement in the context of customs reform in the Philippines. *Rev. Econom. Statist.* 90:1–14.

Zitzewitz E (2006) Nationalism in winter sports judging and its lessons for organizational decision making. *J. Econom. Management Strategy* 15:67–99.

Zitzewitz E (2012) Forensic economics. *J. Econom. Literature* 50: 731–769.

Zitzewitz E (2014) Does transparency reduce favoritism and corruption? Evidence from the reform of figure skating judging. *J. Sports Econom.* 15:3–30.