informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–19
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Impact of Motivation and Workload on Service Time Components: An Empirical Analysis of Call Center Operations

Ahmad M. Ashkanani,[a,*] Benjamin B. Dunford,[b] Kevin J. Mumford[b]

[a] College of Business Administration, Kuwait University, Kuwait; [b] Krannert School of Management, Purdue University, West Lafayette, Indiana 47906
*Corresponding author

Contact: a.ashkanani@ku.edu.kw, https://orcid.org/0000-0003-4783-710X (AMA); bdunford@purdue.edu, https://orcid.org/0000-0002-3289-4729 (BBD); mumford@purdue.edu, https://orcid.org/0000-0003-0315-8471 (KJM)

Received: February 7, 2020
Revised: November 16, 2020; January 28, 2022; April 12, 2022
Accepted: April 14, 2022
Published Online in Articles in Advance: July 20, 2022

https://doi.org/10.1287/mnsc.2022.4491

Abstract. We study the joint effects of motivation and workload on human servers' service time. Using operational and survey data from a call center with a pooled queue structure and limited financial incentives, we examine how individual differences between servers' trait intrinsic motivation (IM) and extrinsic motivation (EM) impact their average offline, online, and total service times in response to changing workloads. We find significant differences in the patterns of workload and service time relationships across different stages of the service request between servers possessing different combinations of trait motivation. For example, servers with a combination of high IM and low EM were approximately 15% (161%) faster in processing the offline portion of service requests than their peers with the opposite combination (low and high) when workload levels were low (high), respectively. In contrast, servers with high IM-low EM were approximately 35% (5%) slower in processing the online portion of service requests than their low IM-high EM counterparts when workload levels were low (high), respectively. Our findings suggest important nuances in how servers with different trait motivation types respond to changing workload across different stages of the service request. The behavioral pattern shown by high IM-low EM servers is consistent with the preferences of productivity-seeking call center managers who favor *speedup* and *slowdown* at certain stages of the service request, conditional to workload. These findings underscore the importance of accounting for trait-based individual differences for a more complete understanding of the complex relationship between workload and service time.

History: Accepted by Charles Corbett, operations management.
Funding: This work was supported by the Regenstrief Center for Healthcare Engineering, Purdue University [Organizational Effectiveness Grant 300940].
Supplemental Material: The data files and e-companion are available at https://doi.org/10.1287/mnsc.2022.4491.

Keywords: behavioral operations • productivity • workload • motivation • call center operations

1. Introduction

Behavioral operations research has consistently shown that service times in queues are not exogenous to workload. Indeed, servers are considered to have some discretionary judgment in adjusting their work speed and productivity in response to fluctuating workloads in service queues (Allon and Kremer 2018, Delasay et al. 2019). Supporting this view, an impressive body of research has shown that instantaneous workload influences service time and other productivity variables in a variety of different patterns: linear (Kc and Terwiesch 2009), inverted U-shaped (Tan and Netessine 2014), and even N-shaped (Berry Jaeker and Tucker 2017) patterns. Behavioral operations research has illuminated a complex web of contingency factors that underlie a nuanced and multifaceted relationship between workload and service time. Delasay et al. (2019) observed: "There are mechanisms that are activated in different situations caused by different factors and have different effects … The answer to the question 'what is the effect of workload on service time?' is 'it depends.' The more difficult question is 'on what does it depend?'" (Delasay et al. 2019, p. 674).

Over the last several years, researchers have made great strides in answering that question. Multiple mechanisms and contingency factors have been identified such as queue design, queue length, queue visibility, incentive structure, and other environmental characteristics (Allon and Kremer 2018). Yet, despite this progress, no single perspective can explain all of the complexity, and additional research is still needed (Delasay et al. 2019). Behavioral operations research shows that human emotions, attitudes, and other "fixed effects" also play important roles in explaining performance in service queues (Rafaeli et al.

2020, Altman et al. 2021). For example, individual differences between servers have been identified as one potential factor that may help explain previously mixed findings about server responses to changing workload (Lau et al. 2014). However, to date, contingency factors influencing how servers respond to workload changes have been primarily environmental in nature. To our knowledge, most of the behavioral operations literature treats individual differences as fixed effects "to control for unobservable individual server effects that may significantly influence a [server's] productivity level" (Song et al. 2015, p. 3040), yet such approach does not examine the underlying psychological forces that drive between-server differences in productivity. Thus, "we need a better understanding of individual differences in behavioral operations, an area that has received only scant attention" (Croson et al. 2013, p. 4).

The interplay of individual and environmental determinants has defined social scientists' exploration of human behavior for most of the past century (Lewin 1935). Research supporting an "interactionist view" shows that individuals differ meaningfully in how they respond to the same environmental stimuli and that these individual differences are often based on enduring traits (Ekehammar 1974, Tett and Burnett 2003). Applying this logic, we reason that to more fully understand how servers respond to changes in workload, researchers should consider not only environmental factors but also, individual differences, in particular enduring trait differences between servers. To that end, we assert that adding trait-based individual differences to the study of how workload affects server behavior may enhance clarity in the operations literature that has shown mixed results.

For example, practitioner recognition (Carlaw et al. 2002), qualitative call center research (Mahesh and Kasturi 2006), and motivation theory suggest that trait-based differences in motivation may be one pivotal contingency factor that could explain why some studies have shown varying (e.g., linear, inverted U-shaped, and N-shaped) relationships between workload and productivity. In addition to differences in environmental settings across studies (such as queue structures and incentive systems), highly intrinsically motivated employees may respond much more productively to increased workflows than their less intrinsically motivated counterparts who are prone to conserve their energy (by slowing down) during critical times when service queues get overloaded.

Thus, the purpose of this study is to examine how trait-based individual differences in servers' intrinsic motivation (IM) and extrinsic motivation (EM) impact various phases of service time (offline, online, and total) in response to changing workloads in a U.S. call center with a pooled queuing structure and limited financial incentives.

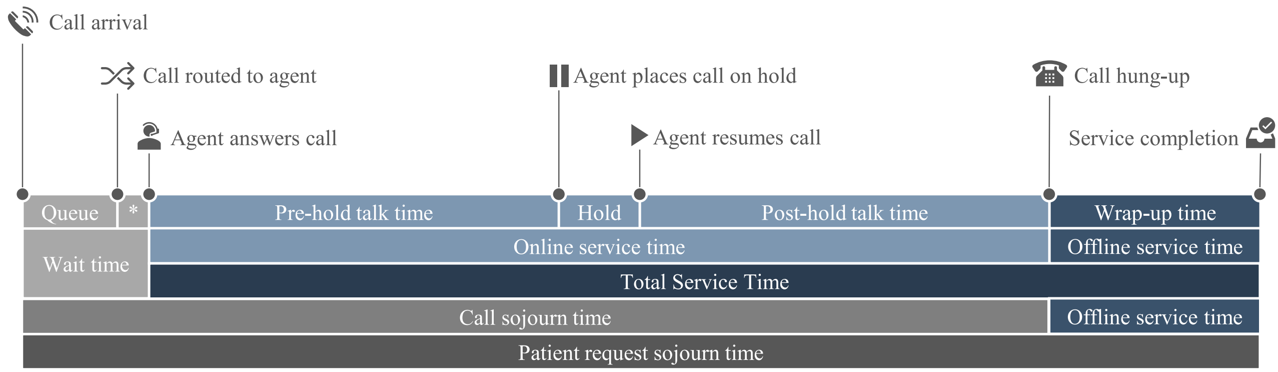
Our study makes several important contributions to the multidisciplinary literature on call centers. First, organizational behavior researchers have examined factors of various kinds associated with performance in call centers including hiring processes (Yakubovich and Lup 2006), high involvement work practices (Workman and Bommer 2004), identity attitudes (Raghuram 2013), and emotional labor perceptions (Diefendorff et al. 2019, Ashtar et al. 2021). We advance this research by emphasizing trait-based individual differences as pivotal factors explaining agent performance, extending previous research focused primarily on either practices or state-based individual differences (i.e., attitudes and emotions that fluctuate over time). Second, we propose an additional answer to the "on what does the workload-service time relationship depend" question in the behavioral operations literature. We suggest that it depends on both environment and trait-based individual differences. This is meaningful because previous findings may be missing important patterns in how different servers respond to changing workloads over time. Finally, we assert that trait-based differences in human behavior emerge under changing conditions of stress or high workload across various phases of the service call. Indeed, our results show that differences in server productivity are highly nuanced. For example, during the offline stage, differences between servers become more pronounced as workload increases. We find that these differences can be explained by both the type and magnitude of trait motivation. We also observe that servers with different types of trait motivation vary significantly in their response to workloads in the online stage and in total service time. In summary, we seek to advance our understanding of the relation between service time and workload beyond a "fixed effects" view (Ashkanani 2017), arguing that employees do not respond uniformly to changes in workload but vary based on enduring trait differences such as intrinsic and extrinsic motivation.

2. Empirical Setting

2.1. Operational Context

We use data from a U.S. call center that handles patient requests for healthcare services within a large healthcare system. The center processes an average of 57,000 calls per month, employs 82 service agents, and operates from Monday to Friday from 8 a.m. to 4:30 p.m. The center employs a pooled queuing structure, a single shift per day, and a fixed base pay structure (using hourly wage) with informal rewards loosely related to busy-period performance (e.g., promotions to leadership positions). We use the terms agent, worker, and server throughout the paper interchangeably.

Figure 1. (Color online) Example of a Service Request Flow



* Ring time

2.2. Process Flow

The center uses a standardized process flow (Figure 1). When a patient calls a service line, the system checks the availability of servers assigned to that line. If all servers are busy, then the call is placed in a "first come, first served" pooled queue. The call is assigned to the next available server who has access to a chat system (to consult with coworkers) and an electronic knowledge database (to access relevant medical information). Typical service tasks include scheduling appointments and reaching out to physicians and patients.

For each call, the system records *queue time* (patient's waiting time in the queue before assignment to a server), *ring time* (amount of time a server takes before answering an assigned call), *talk time* (time spent by a server talking to a patient), *hold time* (total time a call was placed on hold), and *wrap-up time* ("postcall" time spent by a server to finalize processing patient's request). Based on this process flow, we divide patient's request sojourn time into three main components: *waiting time* (queue + ring times), *online service time* (talk + hold times), and *offline service time* (wrap-up time). These measures are of interest to call center managers, and they are used to derive key performance indicators that help managers keep track of the performance of the call center and take necessary actions when needed.

Call centers generate large amounts of objective data, but managers and researchers alike recognize that not all of the data are meaningful and that what is most important is not easily observable (Taylor and Bain 1999). There is a trade-off between efficiency and quality that has important implications for measurement and research (Kinnie et al. 2000). Shortcuts in the interest of efficiency typically result in poorer customer service quality. Thus, it is important to explore the nuanced differences between online and offline service time to more fully understand the meaning of each metric.

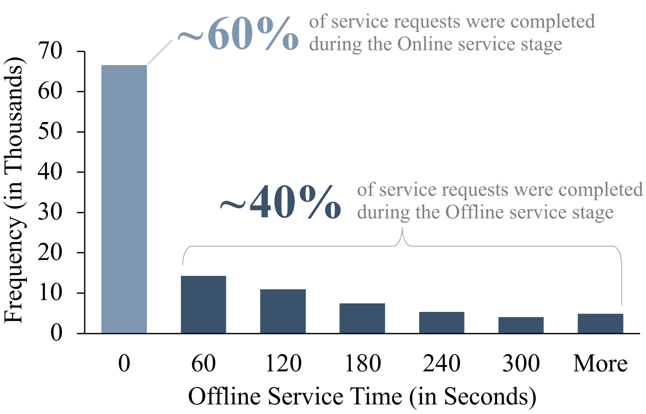
2.2.1. Online Service Time.

The online phase of the call is where customers have direct interaction with agents as the main point of contact with the organization. In this portion of the call, there is a dynamic trade-off between efficiency and quality that becomes more pronounced as workload increases. Under conditions of very low workload (i.e., where there are no other customers are waiting in the queue), managers prefer that agents spend time with customers than idly waiting for the next call. For example, all else equal, managers generally prefer that agents err on the side of giving more time to customers to be certain their needs are met rather than surfing the internet or reading a book. If there is no one else in the queue, extra time with customers typically enhances service quality without a loss in efficiency. However, as the workload rises, managers prefer for agents to be increasingly efficient in online time. To that end, most call centers provide agents with structured conversation scripts to follow while speaking with customers that provide a minimum standard for service quality, with allowance to speak longer if no one else is in the queue. Under conditions of very high workload, managers prefer online service time to be conducted as efficiently as possible to enable opportunities for more calls to be answered without sacrificing quality. Owing to the efficiency-quality trade-off, online talk time is considered to be a more noisy measure of agent productivity on its own (Knights and McCabe 1998, Taylor and Bain 1999, Gans et al. 2003). Call center managers interpret online service time conditionally within the context of workload.

2.2.2. Offline Service Time.

Offline service time is devoted to giving agents time to follow up on tasks that bring closure to the call but that do not require customer time. Agents are expected to extend a service request to the offline stage only when needed. This means the decision to extend the service request to the

Figure 2. (Color online) Distribution of Offline Service Times



offline stage is mainly driven by the nature of the call itself, and agents have little or no discretion in making that decision. For example, a service request that involves a caller seeking information about the availability of certain medical services is likely to be completed during the online stage (i.e., it does not require additional postcall processing). In contrast, more complex service requests may require additional postcall activities, such as "updating of the customer's history file or the processing of an order that the customer has requested" (Gans et al. 2003, p. 85). Approximately 60% of service requests in this call center were completed during the online stage of the call as illustrated in Figure 2, whereas the remaining service requests were extended to the offline stage.

Although agents have little or no influence over the decision to extend a request to the offline stage, there is more agent discretion in the *length* of offline time than there is in online time. Such discretion is granted to agents because the nature of offline tasks varies so much between calls that it is difficult to monitor. However, because offline time does not involve direct interactions with customers, managers generally prefer offline time to be as short as possible provided that customer needs are met (Knights and McCabe 1998). Unlike online time, there is no potential service benefit to the customer if the agent extends offline time beyond what is required to wrap up the call. Because idle time (time not spent on a service call) is actively tracked by management as a measure of productivity loss (Gans et al. 2003), agents will often discretely use offline time as a recovery or stress reduction mechanism (Knights and McCabe 1998, Kinnie et al. 2000). "One scenario that we see frequently is an advisor spending a significant amount of time in the post call wrap-up state, pretending to take notes or update customer relationship management (CRM) information for longer than necessary to delay the next call they receive" (*Contact Centre Magazine* 2020). Indeed, agents are noted to have more "opportunity" (Boudreau et al. 2003) to slow down (to conserve their effort) during the offline stage of the service request without exerting excessive additional effort or facing mistreatment from callers. This is problematic because offline service time impacts queue performance in call centers (e.g., queue wait times and queue abandonment rates), affecting an agent's availability to receive new calls (i.e., agents can receive new calls only after they complete processing a focal service request, including any offline activities).

Agents may also use their discretion to speed up during the offline portion of the call. As noted, speeding up during the online stage may compromise service quality (e.g., rushing the caller may lead to missing patient's medical history information or lowering customer satisfaction levels). Conversely, speeding up during the offline service stage is viewed by managers as helpful in reducing overall service time without affecting the focal caller's experience directly (because the caller is not directly involved in the offline activities). Managerial emphasis on reducing offline service time has become a point of contention with many workers who view it as their only chance to conserve their energy (Knights and McCabe 1998, Taylor and Bain 1999, Kinnie et al. 2000, Deery et al. 2004).

Consequently, the reduction of wrap-up time is a top priority for call center managers who consider it to be a more accurate measure of productivity than online service time, irrespective of workload (Knights and McCabe 1998; Deery et al. 2002, 2004). In other words, whereas the appropriateness of shorter online time is conditional upon workload, managers value shorter offline time in all workload conditions, provided that customer needs are addressed adequately in the offline portion. With less concern about quality-speed trade-offs (Kc and Terwiesch 2009) in the offline portion of calls, offline time is tracked in a large majority of call centers as a key performance indicator (Gans et al. 2003), and managers frequently issue company memos and conduct training programs to minimize or eliminate it (Brannan 2005).

2.2.3. Total Service Time. Total service time is defined as the sum of online and offline service times (see Figure 1). Although it is widely tracked in call centers as a measure of general productivity (Knights and McCabe 1998) and has been used as a measure of productivity in previous research, it aggregates information about complex processes, and call center managers know that it should be interpreted with caution. Because the ideal length of the online portion of calls varies so much according to workload, combining offline and online time into a single metric presents an oversimplified view of productivity (Gans et al. 2003).

Given the uniquely context-dependent nature of online time and limitations associated with the interpretation of total time, we focused our hypotheses on offline time. However, toward a more complete understanding of the joint effects of workload and trait motivation on service time, we measured and tested all phases of the service call (i.e., offline, online, and total time) and report the results in Section 6.

3. Literature Review and Hypotheses Development

3.1. Work Speed: A System-Level Perspective

We integrate insights from the behavioral queuing literature suggesting two contrasting work speed-related behaviors—speedup and slowdown—and draw on motivation theory to reason that these factors may work together to explain server behavior in the offline portion of the service call.

3.1.1. Slowdown. Empirical evidence suggests that servers sometimes slow down in response to higher levels of workload (Batt and Terwiesch 2017, Berry Jaeker and Tucker 2017), especially when overall workload levels are low, partly in an effort to increase service quality (Tan and Netessine 2014). Additional evidence suggests that servers slow down when working in pooled queues compared with parallel queues, which might be attributed to a sense of queue ownership in parallel queues (Song et al. 2015) or a social loafing (or free riding) effect in pooled queues (Wang and Zhou 2018), particularly when monitoring of effort is difficult (Delasay et al. 2019). Shunko et al. (2018) add that the slowdown effects associated with pooled queues were more evident under higher levels of workload and a fixed pay (rather than performance-based) incentives structure. In short, the slowdown effect is a nuanced behavior subject to contingencies that involve incentives, monitoring, queue structure, workload level, and social dynamics.

3.1.2. Speedup. On the other hand, servers sometimes speed up in response to higher levels of workload, presumably when congestion costs exceed speedup costs (Allon and Kremer 2018) and potentially at the expense of service quality (Kc and Terwiesch 2009). Delasay et al. (2019) argue that social pressure could influence the speed of slower servers where they "feel pressure to speed up in order to avoid delaying the service of others" and they "work faster when performance feedback is available" (Delasay et al. 2019, p. 679). Thus, when the system is highly congested, managers are more likely to increase monitoring behavior, especially when observable queue performance indicators become worse (e.g., longer queue waiting times, higher abandonment rates, lower service levels, etc.). However, social speedup dynamics may be nonlinear such that the speedup effect may be attenuated as server fatigue kicks in (Kc and Terwiesch 2009) or when speedup costs exceed waiting costs (Tan and Netessine 2014). Various studies show nonlinear relationships between workload and service time that involved a speedup effect for at least a portion of observed workload levels (Batt and Terwiesch 2017, Berry Jaeker and Tucker 2017). Thus, servers may engage in discretionary decreases and/or increases in service speed in response to numerous contextual factors.

3.1.3. Reinforcement Theory. This view is consistent with the basic prediction of reinforcement theory, which asserts that individuals increase the frequency of behavior that is rewarded and decrease the frequency of behavior that is punished or not rewarded (Skinner 2014). Reinforcement takes numerous forms in organizations including the granting of financial rewards (cash or short-term and long-term incentives), nonfinancial rewards (employee of the month, recognition, or praise), gifts in kind (food or prizes), and even the attention of or feedback from a supervisor. Even the measurement of a behavior can constitute reinforcement (Rynes and Gerhart 2000). Punishment in organizations typically takes the form of employees being held accountable for social loafing or not performing up to a certain standard on a task. Decades of research show that employees pay close attention to the behaviors that are reinforced in their workplace and those that are not; over time, employees avoid engaging in behavior that is not reinforced or punished (Rynes and Gerhart 2000).

Applying these foundational predictions of reinforcement theory, we reason that in our setting, when workload levels are low, servers might find more opportunities to slow down in response to an increase in workload to conserve their energy (by extending their offline service time and thus, diverting calls to other servers) because manager's monitoring behavior is lower and the fixed base pay structure does not award additional effort with direct performance-based incentives. In contrast, when workload levels are high, managers increase monitoring behavior, and servers might experience increased social pressure (Delasay et al. 2019) from their peers, prompting servers to speed up in response to an increase in workload to avoid delaying the service of others and/or avoid potential punishment. Thus, in a pooled queuing setting, with fixed base pay and increased monitoring at peak workload levels, we expect an inverted U-shaped relationship between workload and offline service time.

Hypothesis 1. *The relationship between workload and offline service time is curvilinear with an inverted U-shape.*

3.2. Work Speed: An Individual Differences Perspective

Motivation is defined as the direction, intensity, and persistence of effort in the performance of a task or behavior (Ryan and Deci 2000). Employee motivation represents an important individual difference that may help explain the complex dynamics between workload and service time. Theorists distinguish between two primary types of motivation in explaining why individuals put forth effort: trait extrinsic and intrinsic motivation.

Trait intrinsic motivation refers to a tendency for individuals to have internally regulated drivers of behavior, such as "doing an activity because [people] find it interesting and derive spontaneous satisfaction from the activity itself" (Gagné and Deci 2005, p. 331). Trait extrinsic motivation refers to a tendency for individuals to have externally induced motives for effort, including the receipt of tangible (e.g., money, promotions, etc.) and intangible (e.g., praise, recognition, etc.) rewards or the avoidance of punishment (Gagné and Deci 2005). We propose that individual differences in both trait intrinsic and extrinsic motivation impact offline service time and also impact how workers respond to workload demands in service queues.

3.2.1. Intrinsic Motivation. Motivating individuals to complete especially dull, routine, or repetitive tasks (like call center work) has been the subject of research for decades. Individuals differ in their trait-based ability to self-regulate their behavior on dull tasks by finding meaning, fun, and personal enjoyment (Amabile et al. 1994). Some individuals assigned to tedious work tend to seek out and find meaning and opportunities for creativity, challenge, and enjoyment (indicating a trait-based propensity for intrinsic motivation), leading them to engage in dull tasks with greater intensity of effort (Sansone et al. 1992).

Drawing on these insights, we reason that call center agents will differ in their trait propensity for being intrinsically motivated to perform their jobs; those who have higher trait intrinsic motivation are more likely to engage in tasks with greater intensity of effort, leading to higher productivity relative to those with less intrinsic motivation.

Hypothesis 2. *Intrinsic motivation is negatively related to offline service time.*

3.2.2. Extrinsic Motivation. Notwithstanding widespread debates in the literature, evidence consistently shows that financial and nonfinancial external rewards do indeed motivate people (Fang and Gerhart 2012). Employees pay close attention to what behaviors are monitored, measured, and rewarded in organizations (Rynes and Gerhart 2000); as noted, reinforcement theory asserts that individuals tend to engage in behaviors that are rewarded and extinguish behaviors that are not rewarded (Skinner 2014). Some individuals are more energized by and attend to pay, monitoring, and recognition than others (Amabile et al. 1994). Drawing on these perspectives, we reason that call center agents vary in trait propensity for being extrinsically motivated by rewards to perform their tasks.

In a pooled queue under a fixed base pay compensation system with limited monitoring, we reason that servers with greater trait propensity for extrinsic motivation would have little financial or external incentive to work more quickly and be more likely to engage in social loafing by increasing their offline service time in wrapping up calls. Thus, we propose that servers in our sample with higher trait extrinsic motivation are more likely to engage in social loafing by reducing their productivity in a fixed base pay scheme with limited monitoring in a pooled service queue.

Hypothesis 3. *Extrinsic motivation is positively related to offline service time.*

3.3. The Joint Effects of Motivation and Workload: A Cross-Level Perspective

As an outgrowth of reinforcement theory, self-determination theory (SDT) (Ryan and Deci 2000, Gagné and Deci 2005) attempts to explain why some individuals engage in tasks with greater intensity and persistence than others and how organizations can facilitate that effort. SDT suggests that employees can have both trait intrinsic and extrinsic motives but prefer to be self-directed rather than compelled by external rewards or pressures in their work, which is a common type of external regulation in call centers. Moreover, individuals respond to work pressure by either reducing the intensity of their effort to assert control over their work or increasing the intensity of their effort to meet increasing workload. However, increased effort in response to growing workload can only be sustained for a short period of time before role overload and fatigue undermine productivity and quality (Grant 2008, Kc and Terwiesch 2009).

Drawing on these insights, we predict that trait-based individual differences in both intrinsic and extrinsic motivation between agents will influence how they respond to changing workload over time. Servers with low intrinsic motivation may extend offline service time as the workload increases in an effort to conserve resources and fulfil needs for autonomy (Ryan and Deci 2000). In contrast, workload is likely to be exogenous to offline service time for highly intrinsically motivated agents because they enjoy the work itself.

Individuals high in extrinsic motivation pay close attention to performance-reward contingencies and avoid effort that is not likely to be rewarded. Thus, in pooled service queues with fixed base pay structures and limited monitoring, servers with high extrinsic motivation may increase service time as the workload increases up to a point at which the workload reaches a peak and supervisors initiate active monitoring of performance. Conversely, those with low extrinsic motivation are likely to be less influenced by workload because their motivation is less influenced by external factors.

Thus, we expect that offline service time will be highest and increase most prominently with workload when intrinsic motivation is low and extrinsic motivation is high. We predict that offline service time will be lowest and most exogenous to workload when intrinsic motivation is high and extrinsic motivation is low.

Hypothesis 4. *Intrinsic motivation and extrinsic motivation jointly influence the relationship between workload and offline service time. The higher (lower) the intrinsic motivation and the lower (higher) the extrinsic motivation, the shorter (longer) the duration of offline service time and the weaker (stronger) the relationship between workload and offline service time.*

4. Data

We test Hypotheses 1–4 by merging survey data (from February 2017) with archival call logs (from March and April 2017), as discussed.

4.1. Survey Procedure

Call agents were invited to complete an online survey during the last week of February 2017. Agents were presented with a $10 certificate for filling the survey (to encourage participation). Additionally, participants received an opportunity to enter a raffle for a chance to win one of five $100 gift certificates. Agents were assured of the confidentiality of their responses and that only aggregate-level reports were going to be shared with the leadership team. A total of 64 agents participated in the survey (78% participation rate). All survey items used seven-point Likert-type scales with anchors of one (strongly disagree) to seven (strongly agree). A description of the survey items used in this study is presented in Section EC.2 of the e-companion.

4.2. Archival Call Logs

We use archival data collected from the call center's information system, and the data include all customer calls processed over the period of March through April 2017, totaling 113,389 calls handled by 82 agents. The data include detailed information for each call such as date/time of call, identifier of the server who processed the call, service line type, and all wait time and service time components. To test Hypothesis 1, we dropped calls that were received outside regular working hours, had zero talk time, or had an extreme offline service time (i.e., the top percentile), leaving 109,796 calls and 82 agents. To test the remaining hypotheses, we dropped calls that satisfied the previous exclusion criteria in addition to calls that were handled by survey nonparticipants, leaving 88,024 calls and 64 agents. We use the following notations for defining our variables: call i, agent j, and service line s. Additionally, time period refers to the 30-minute time interval that call i was received in (e.g., a call arriving at 11:13 a.m. has a time period of 11:00–11:30 a.m.).

4.3. Measures

Table 1 includes summary statistics of our study variables. Operationalization of these measures is discussed. In addition, a data dictionary and additional distributional statistics of the main variables are available in the e-companion.

4.3.1. Dependent Variables.

4.3.1.1. Offline Service Time ($OFFLINE_{ij}$). Our main dependent variable is measured by calculating the number of "postcall" seconds spent by a server to complete processing of a patient's service request. Using service time as an inverse proxy for productivity is a common approach in the behavioral queuing literature (Delasay et al. 2019). We focus on offline service time as a proxy for server's productivity as discussed in Section 2.2.

4.3.1.2. Online Service Time ($ONLINE_{ij}$). We conduct parallel analysis on the online portion of a service request, which is measured by calculating the number

Table 1. Summary Statistics and Correlations

Variable	Unit	Mean	Standard deviation	1	2	3	4	5	6	7
1. Offline service time	Seconds	54.76	95.25	—						
2. Intrinsic motivation	Points	5.31	1.39	**−0.10**	(0.94)					
3. Extrinsic motivation	Points	5.65	1.43	**0.04**	−0.04	(0.91)				
4. Workload	Calls/agent.$\frac{1}{2}$-hour	3.13	1.24	**0.11**	**−0.07**	0.00	—			
5. Overwork$_{K=4}$	Calls/agent.$\frac{1}{2}$-hour	0.17	1.24	**0.12**	**−0.06**	**0.01**	**0.70**	—		
6. Number of agents	Agents	13.01	4.94	**−0.02**	**0.18**	**0.09**	**−0.21**	**−0.16**	—	
7. Online service time	Seconds	234.29	201.07	**0.05**	**0.04**	**−0.02**	**−0.03**	**−0.01**	**0.02**	—
8. Total service time	Seconds	289.05	226.50	**0.46**	0.00	0.00	**0.02**	**0.04**	0.01[a]	**0.91**

Notes. Bold denotes significance at the 1% level. Coefficient α estimates of reliability are in parentheses on the diagonal. Intrinsic motivation and extrinsic motivation are measured on a seven-point Likert scale ranging from one (strongly disagree) to seven (strongly agree).
[a]Significance at the 5% level.

of seconds spent by a server to process a patient's service request while the patient is on the line. This includes both the time spent by an agent talking to a patient and the duration of the time a patient is placed on hold (if any) (see Figure 1).

4.3.1.3. Total Service Time ($TOTAL_{ij}$). We also conduct parallel analysis on the total service time, which is operationalized as the summation of both online and offline service times (i.e., $TOTAL_{ij} = ONLINE_{ij} + OFFLINE_{ij}$). This measure captures the total time spent by a call center agent on completing a customer request, including any talk, hold, and wrap-up times (see Figure 1).

4.3.2. Independent Variables.

4.3.2.1. Workload (WL_{ij}). We define workload seen by server j during the time period call i was received as $WL_{ij} = \sum_{s \in S_{ij}} \frac{NC_{si}}{NA_{si}}$, where NC_{si} denotes the number of calls routed to service line s during time period i, NA_{si} denotes the number of servers assigned to service line s during time period i, and S_{ij} denotes a set of service lines assigned to server j during time period i. This workload measure is generated using the full sample and is adjusted for the number of coworkers servicing a given set of virtual queues. For example, if a server was assigned to serve two virtual queues that received 12 (25) calls and had six (five) servers assigned to it during a given time period, then the adjusted workload seen by the server is 7 calls/server. 1/2-hour $\left(= \frac{12\,calls}{6\,server.1/2\text{-}hour} + \frac{25\,calls}{5\,server.1/2\text{-}hour}\right)$.

4.3.2.2. Trait Intrinsic Motivation (IM_j). We measured trait-based intrinsic motivation using a four-item scale used in Grant (2008). We asked servers, "Why are you motivated to do your job?" Sample items include "[b]ecause I enjoy the work itself." This motivation scale assesses an agent's desire to exert effort because of enjoying the task itself. The mean intrinsic motivation score was 5.31 ($\alpha = 0.94$). A more detailed discussion of this scale is provided in Section EC.2 of the e-companion.

4.3.2.3. Trait Extrinsic Motivation (EM_j). We measured trait-based extrinsic motivation using a four-item scale used in Grant and Berry (2011). We asked servers, "Why are you motivated to do your job?" Sample items included "[b]ecause I need to earn money." This motivation scale assesses an agent's desire to exert effort because of external monetary factors. The mean extrinsic motivation score was 5.65 ($\alpha = 0.91$). A more detailed discussion of this scale is provided in Section EC.2 of the e-companion.

4.3.3. Controls. Our control variables include online service time (used in offline service time analysis), service line fixed effects, day of the week fixed effects, time of day, number of agents servicing a virtual queue, and $overwork_K$ (defined as average workload level seen by a server over the past K periods).

5. Econometric Specification

First, we estimate baseline regression models that examine the factors that impact offline service time. Then, we use instrumental variables to correct for potential endogeneity issues. We perform additional robustness checks of our main results. Finally, we conduct parallel analyses that examine the factors that impact online and total service time.

5.1. Offline Service Time Specification

Our aim is to estimate models that capture the effects of workload, intrinsic motivation, and extrinsic motivation on offline service time. Because our dependent variable takes on a value of zero for a significant fraction of the observations (see Figure 2) and is strictly positive and roughly continuous for the remaining observations, we need to use a model that accounts for censored data (Wooldridge 2020). Thus, we use a Tobit regression model (Tobin 1958) to account for qualitative differences between limit (zero) and nonlimit (strictly positive) observations (Greene 2018).[1]

First, we estimate a baseline Tobit model to provide preliminary estimates of the factors that impact offline service time. To test Hypothesis 1, we add linear and quadratic terms to capture the workload effect on offline service time. To test Hypotheses 2 and 3, we add intrinsic and extrinsic motivation terms to the model. Finally, to test Hypothesis 4, we add interaction terms for workload and intrinsic and extrinsic motivation to capture any cross-level interaction effects.

Let $OFFLINE_{ij}$ ($OFFLINE^*_{ij}$) denote the observed (latent) offline service time of call i handled by server j, WL_{ij} denote the mean-centered workload level seen by server j upon receiving call i, IM_j (EM_j) denote the mean-centered intrinsic (extrinsic) motivation level of server j, $\boldsymbol{X}_{ij}$ denote a vector of control variables including service line and day of the week fixed effects, and u_{ij} denote the error term. Then, our complete model (specification (3) of Table 2) is

$$OFFLINE^*_{ij} = \sum_{k=0}^{2}(\beta_{k0} + \beta_{k1}IM_j + \beta_{k2}EM_j + \beta_{k3}IM_j \times EM_j) WL^k_{ij} + \boldsymbol{X}_{ij}\Gamma_0 + u_{ij}, u_{ij} \sim N(0, \sigma^2) \quad (1)$$

$$OFFLINE_{ij} = max(0, OFFLINE^*_{ij}). \quad (2)$$

We compute Huber–White robust errors to account for potential heteroskedasticity issues (White 1980). We

Table 2. Joint Effects of Intrinsic Motivation, Extrinsic Motivation, and Workload on Offline Service Time

	Estimated by baseline Tobit models			Estimated by IV Tobit models		
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	−87.31	−81.38	−83.27	−143.37**	−119.72	−131.57
	(64.24)	(65.37)	(65.01)	(69.16)	(68.06)	(67.03)
WL	18.13***	12.28***	10.56***	59.69***	46.48***	50.17***
	(0.91)	(0.97)	(0.99)	(2.68)	(2.95)	(3.48)
WL^2	−7.37***	−10.28***	−9.61***	−11.23***	−18.59***	−16.76***
	(0.30)	(0.34)	(0.42)	(0.56)	(0.66)	(1.67)
IM	—	−4.02***	−7.29***	—	−6.54***	−12.09***
		(0.55)	(0.65)		(0.60)	(1.20)
EM	—	21.00***	19.23***	—	21.89***	15.49***
		(0.56)	(0.67)		(0.60)	(1.85)
$IM \times WL$	—	—	−1.82***	—	—	−4.99***
			(0.53)			(1.09)
$IM \times WL^2$	—	—	3.91***	—	—	7.13***
			(0.31)			(1.05)
$EM \times WL$	—	—	8.58***	—	—	13.74***
			(0.61)			(1.51)
$EM \times WL^2$	—	—	2.45***	—	—	6.60***
			(0.43)			(1.83)
$IM \times EM$	—	—	1.05***	—	—	−1.53
			(0.39)			(1.08)
$IM \times EM \times WL$	—	—	−2.91***	—	—	−16.85***
			(0.42)			(1.15)
$IM \times EM \times WL^2$	—	—	−1.91***	—	—	2.42
			(0.31)			(1.32)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Sigma	180.73	168.29	167.84	183.19	169.75	172.72
No. of calls	109,796	88,024	88,024	99,356	79,440	79,440
No. of agents	82	64	64	82	64	64
Model fit ($\Pr > \chi^2$)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Notes. Robust standard errors are in parentheses. The variables workload (*WL*), intrinsic motivation (*IM*), and extrinsic motivation (*EM*) are mean centered.

**Significance at the 5% level; ***significance at the 1% level.

ran these models using Stata 17 SE tobit procedure with the VCE(robust) option. As an additional robustness check, we compute clustered standard errors (see Section EC.3.2 of the e-companion) to account for potential within-cluster correlations among the error terms (Wooldridge 2020). Additional robustness checks to address potential model specification issues are included in Section EC.3.1 of the e-companion.

Although these models provide useful preliminary results, they do not address potential endogeneity issues. The included fixed effects and other control variables help to reduce endogeneity concerns, but they do not eliminate these concerns. There are many worker characteristics that are not captured by any of our measures and that could affect service time. These omitted variables are unlikely to be directly correlated with the workload experienced by a particular worker, conditional on the service line, day of the week, and time of day because worker behavior does not affect our measure of workload. However, one potential source of estimation bias is that managers may observe worker characteristics that are not captured by our measures, and they may use this information in assigning workers to specific service lines. For example, a savvy manager who knows to expect an unusually high workload for a particular service line on a particular day may assign more productive workers to that service line on that particular day. Some agents may have unobserved characteristics that produce lower average offline service time when facing high workload. If these agents are more likely to be assigned to work service lines that the manager believes will be likely to experience a spike in workload, this would induce negative correlation between the error term and workload and thus, would result in a negative bias in the baseline Tobit estimates, even after controlling for service line, day of the week, and time of day. Including worker fixed effects would eliminate this source of endogeneity, but it would also make it impossible to estimate the effects of intrinsic and extrinsic motivation, as these levels are measured only once for each worker. Thus, we adopt an instrumental variable Tobit (IV Tobit) approach (Newey 1987) as discussed in Section 5.2. The instrumental variables approach reduces this potential bias and should yield larger estimates of the effect of workload on offline service time. We also perform Wald tests of exogeneity and reject the null hypotheses that our workload, workload^2, and any interactions involving workload are exogenous.

5.2. Instrumental Variable Tobit Model

Following Tan and Netessine (2014), we propose using "the lagged values of the endogenous independent variables" (Tan and Netessine 2014, p. 1582) as our instrumental variables. First, we operationalize lagged workload (LWL_{ij}) as the average of workload levels observed over the past four weeks by server j during the same day of the week and same time period call i was received in. For example, if an agent received a call on Monday 3-APR-2017 9:15 a.m., then the lagged workload for that observation is calculated as the average of workload levels observed by the same agent on Mondays of the previous four weeks (i.e., 6-MAR-2017, 13-MAR-2017, 20-MAR-2017, and 27-MAR-2017) during the time period 9:00–9:30 a.m. We use the average over the past four weeks (rather than the value of the past week) to reduce missing values that would result if the agent did not work during the same time period of the previous week. Then, we compute lagged workload^2 (LWL_{ij}^2) using LWL_{ij}. In specifications where workload is interacted with the measures of intrinsic and/or extrinsic motivation, we need additional instruments and use the interactions of these measures with lagged workload. The first-stage estimates for the full specification are discussed in Section 6.4 and show that lagged workload is highly correlated with the current workload, even while controlling for the number of agents and the time of day as well as the day of the week and service line fixed effects. The advantage of the Tan and Netessine (2014) instrumental variables approach is that lagged workload is arguably exogenous because it removes contemporaneous shocks. Thus, these instruments should satisfy the exclusion assumption. Our IV Tobit estimation procedure, which runs both the first stage and second stage simultaneously, is as follows.

5.2.1. First Stage. Estimate WL and WL^2 (for models testing Hypotheses 1–3) in addition to any interaction terms involving workload (for the model that tests Hypothesis 4) using the instrumental variables (i.e., lagged versions of the aforementioned endogenous independent variables) with remaining exogenous controls (specified in Section 4.3.3) included.

5.2.2. Second Stage. Estimate the coefficients of models used to test Hypotheses 1–4 (i.e., models in specifications (4)–(6) of Table 2) using Tobit regression models with the predicted values of the endogenous independent variables generated from the first stage.

We ran these models using Stata 17 SE ivtobit procedure with the VCE(robust) option (clustered standard errors are included in Section EC.3.2 of the e-companion). We discuss the strength of our instrumental variables in Section 6.4.

5.3. Online Service Time Specification

Although not hypothesized, we conduct a parallel analysis on the online portion of a service request. We follow similar empirical strategies to the ones described in Sections 5.1 and 5.2 with the exceptions of using $ONLINE_{ij}$ (which denotes the observed online service time of call i handled by server j) as a dependent variable, using ordinary least squares (OLS) regression models (rather than Tobit) for the baseline models given the continuous nature of the online service time measure, using two-stage least squares (2SLS) models (rather than IV Tobit) for the endogeneity corrected models, using a linear term for the workload effect,[2] and excluding $ONLINE_{ij}$ from the control variables. We ran these models using Stata 17 SE regress (for OLS models) and ivregress 2sls (for 2SLS models) procedures.

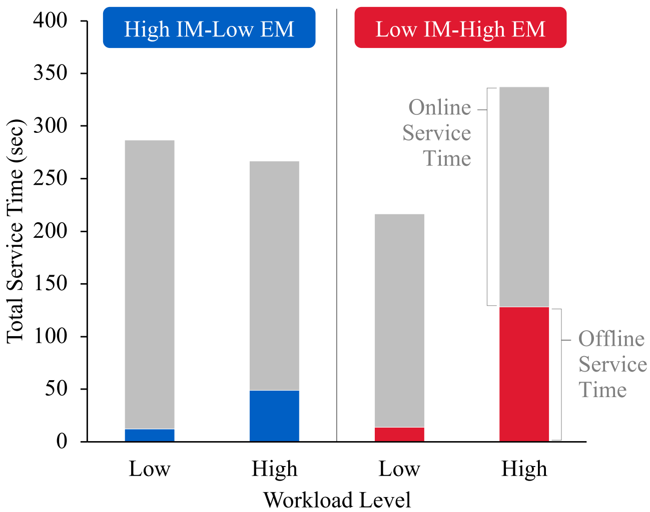
5.4. Total Service Time Specification

We also run parallel analysis on the total service time. We follow empirical strategies similar to the ones described in Section 5.3 with the exception of using $TOTAL_{ij}$ (denotes the observed total service time of call i handled by server j) as a dependent variable. We ran these models using Stata 17 SE regress (for OLS models) and ivregress 2sls (for 2SLS models) procedures.

6. Results

Prefacing our hypothesized and parallel study results, Figure 3 provides a summary of the estimated joint effects of workload and trait motivation on the components of service time.[3] Offline time is denoted by the shaded (blue and red) regions, and online time is denoted by the grey regions. The results illustrate different service time behaviors exhibited by agents with different trait motivation levels under different workload levels. In particular, agents with high intrinsic and low extrinsic motivation spent almost a constant amount of total service time irrespective of workload. Yet, the composition of total service time depended on workload levels where agents with high IM and low EM spent more (less) time interacting with customers online and less (more) time wrapping up service requests during the offline stage when workload levels were low (high). In contrast, agents with low IM and

Figure 3. (Color online) Expected Offline and Online Service Times as Functions of Intrinsic Motivation, Extrinsic Motivation, and Workload



high EM spent less time interacting with customers online irrespective of workload levels, whereas they increased their offline (and subsequently, total) service time as workload increased. Overall, agents with low IM and high EM spent less (more) total time servicing customers' requests than their high IM and low EM counterparts when workload levels were low (high). The detailed analyses of offline, online, and total service times are discussed next.

6.1. Offline Service Time Analysis

Table 2 provides the results of offline service time analysis. The linear and quadratic baseline Tobit estimates of workload on offline service time are shown in specification (1). The results suggest that the relation between workload and offline service time is curvilinear with an inverted U-shape as indicated by the positive linear (18.13, $p < 0.01$) and negative quadratic (–7.37, $p < 0.01$) coefficients of workload. This means that workload is initially associated with an increase in offline service time (slowdown effect) up to an inflection point of 1.2 units above the mean workload level, after which the offline service time decreases with the increase in workload level (speedup effect). This result supports Hypothesis 1. After correcting the endogeneity issue using the instruments, the IV Tobit estimates become 59.69 ($p < 0.01$) and –11.23 ($p < 0.01$) for the linear and quadratic coefficients of workload, respectively, as shown in specification (4) of Table 2. This means that compared with the baseline Tobit estimates, the workload effect in the IV Tobit model is larger as expected, and the slowdown effect lasts longer (with an inflection point of 2.7 units above the mean workload level) as illustrated in Figure 4. The results support Hypothesis 1.

Hypothesis 2 predicted a negative relationship between intrinsic motivation and offline service time. Supporting Hypothesis 2, the baseline Tobit coefficient of intrinsic motivation in specification (2) of Table 2 is negative (–4.02, $p < 0.01$). This result is consistent with the IV Tobit estimate of the intrinsic motivation effect on offline service time (–6.54, $p < 0.01$) as shown in specification (5) of Table 2. The results suggest that higher levels of intrinsic motivation are associated with shorter offline service times. However, the magnitude of the coefficient of intrinsic motivation indicates its partial effect on latent (rather than observed) values of the outcome (Wooldridge 2020). Therefore, we estimated the expected offline service time for different levels of intrinsic motivation while controlling for the remaining variables at their mean (or baseline) levels (see Figure 5).

Hypothesis 3 predicted a positive relationship between extrinsic motivation and offline service time. As shown in specification (2) of Table 2, the coefficient of extrinsic motivation is positive (21.00, $p < 0.01$). Similarly, the IV Tobit estimate of the extrinsic motivation effect is also positive (21.89, $p < 0.01$) as shown in specification (5) of Table 2. These results suggest that higher levels of extrinsic motivation are associated with longer offline service times, which supports Hypothesis 3. Figure 6 illustrates the relationship between extrinsic motivation and estimates of expected offline service time (controlling for the remaining variables at their mean or baseline levels). The increasing slope of the extrinsic motivation line indicates the increasing marginal effect of extrinsic motivation on expected offline

Figure 4. (Color online) Expected Offline Service Time as a Function of Mean-Centered Workload (Based on IV Tobit Estimates)

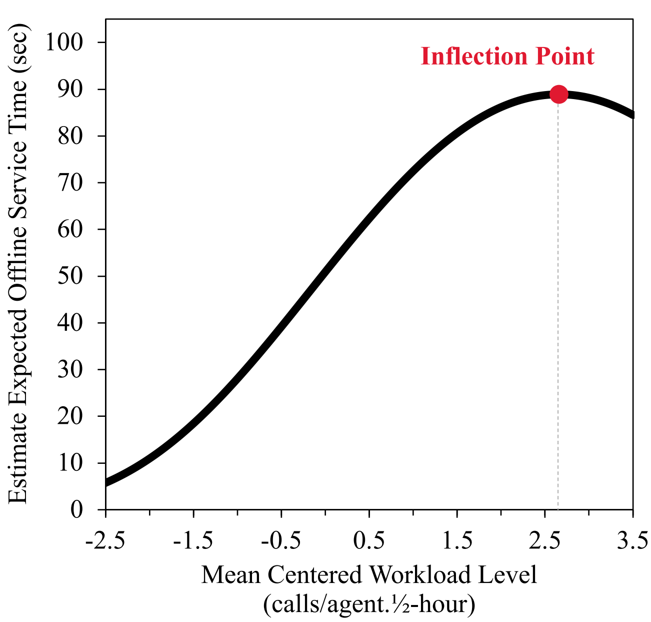
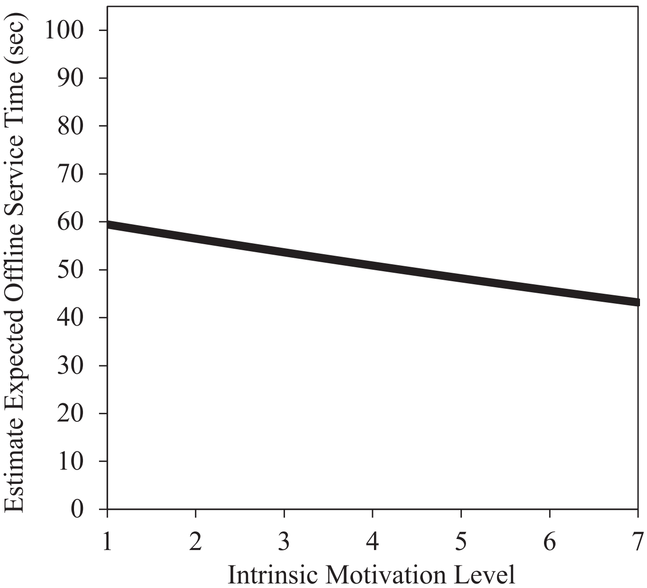


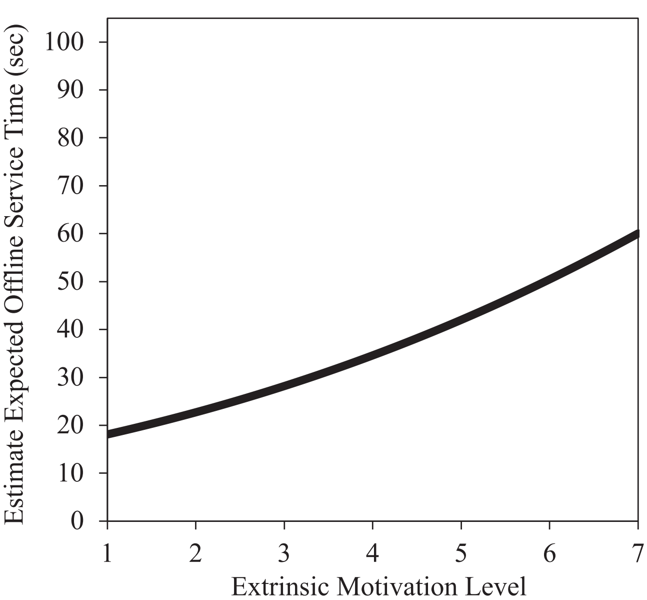
Figure 5. Expected Offline Service Time as a Function of Intrinsic Motivation (Based on IV Tobit Estimates)



service time, whereas the nonlinear shape of the line stems from the nonconstant nature of the marginal effect of an explanatory variable in Tobit models (Wooldridge 2020).

Finally, Hypothesis 4 predicted that the higher (lower) the intrinsic motivation and the lower (higher) the extrinsic motivation, the shorter (longer) the duration of offline service time and the weaker (stronger) the relationship between workload and offline service time. The baseline Tobit and the IV Tobit coefficients of the cross-level interaction effects in specifications (3) and (6) of Table 2, respectively, were statistically significant (with the exception of the coefficients of (a) the double interaction between intrinsic and extrinsic motivation and (b) the triple interaction between intrinsic motivation, extrinsic motivation, and the quadratic term of workload in specification (6)). To understand the result, we plotted estimates of expected offline service time as a function of mean-centered workload level for different intrinsic/extrinsic motivation configurations as illustrated in Figure 7(a). We defined low (high) levels of intrinsic/extrinsic motivation at one standard deviation below (above) the mean of intrinsic/extrinsic motivation. Slicing the data using the mean and standard deviation of survey variables is a common approach in the organizational behavior literature (e.g., Grant 2008). In line with our predictions in Hypothesis 4, the plot of line A in Figure 7(a) suggests that a configuration of low intrinsic motivation and high extrinsic motivation is associated with longer offline service times and an increasing marginal effect of workload on expected offline service time. In contrast, the plot of line B suggests shorter offline service times and a flatter marginal effect of workload on expected offline service time for a configuration of high intrinsic motivation and low extrinsic motivation. These results support Hypothesis 4.

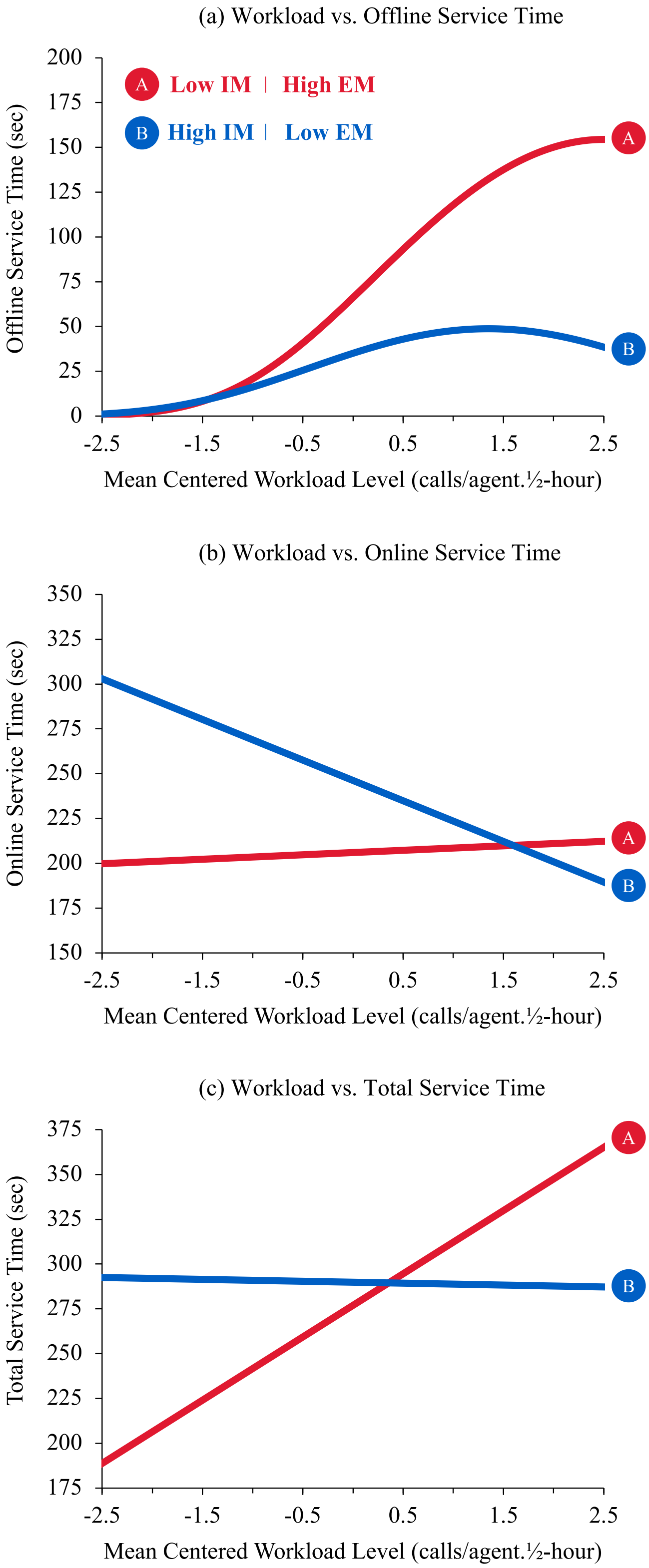
Figure 6. Expected Offline Service Time as a Function of Extrinsic Motivation (Based on IV Tobit Estimates)



6.2. Online Service Time Analysis

Table 3 provides the results of online service time analysis. The OLS results in specification (1) suggest that workload is associated with a decrease in online service time (speedup effect) where on average, a one-unit increase in workload is associated with an eight-second decrease ($p < 0.01$) in online service time. The 2SLS results (specification (4)) show a greater speedup effect as indicated by the larger negative coefficient of workload (–16.55, $p < 0.01$). Next, we examine the relationships between online service time and our trait-based motivation measures (specification (2) of Table 3). The OLS results suggest that intrinsic motivation is associated with higher levels of online service time (1.46, $p < 0.01$). In contrast, extrinsic motivation is associated with lower levels of online service time (–5.40, $p < 0.01$). The 2SLS results (specification (5)) show a larger intrinsic motivation effect (2.94, $p < 0.01$) and a similar extrinsic motivation effect (–5.94, $p < 0.01$) on online service time. Then, we examine the interplay between workload and the motivation measures in affecting online service time. Most of the cross-level interaction effects are statistically significant (except for the OLS coefficient of the interaction between extrinsic motivation and workload in specification (3) of Table 3). Furthermore, the OLS and 2SLS estimates have similar signs, whereas most of the 2SLS estimates have larger magnitudes. To understand the

Figure 7. (Color online) Expected Service Times as Functions of Mean-Centered Workload, Intrinsic Motivation, and Extrinsic Motivation (Based on IV Tobit and 2SLS Estimates)



interaction results, we plot estimates of online service time at high IM-low EM and low IM-high EM configurations using the 2SLS estimates from specification (6) of Table 3 as illustrated in Figure 7(b). The plot of line A suggests that a configuration of low intrinsic motivation and high extrinsic motivation is associated with shorter online service times and a flatter marginal effect of workload on expected online service time. In contrast, the plot of line B suggests longer online service times at lower workload levels and a decreasing marginal effect of workload on expected online service time for a configuration of high intrinsic motivation and low extrinsic motivation. Finally, we probe the model in specification (6) of Table 3 at different motivation levels (i.e., high IM-low EM and low IM-high EM) to examine the conditional effect of workload on online service time at each of the predefined motivation configurations (Hayes 2022). The results suggest a negative workload effect on online service time for the high IM-low EM (–22.69, $p < 0.01$) configuration and a statistically nonsignificant (ns) workload effect on online service time for the low IM-high EM (2.51, ns) configuration.

6.3. Total Service Time Analysis

Table 4 provides the results of total service time analysis. The OLS results in specification (1) suggest that workload is associated with a decrease in total service time (–3.44, $p < 0.01$). However, correcting for endogeneity using the instruments, the 2SLS estimate of workload (specification (4)) becomes positive (6.76, $p < 0.05$), suggesting an average slowdown (rather than speedup) effect. Next, we examine the relationships between total service time and our trait-based motivation measures. Both the OLS (specification (2)) and 2SLS (specification (5)) results suggest that intrinsic motivation is associated with lower levels of total service time (–2.29, $p < 0.01$ and –1.59, $p < 0.05$, respectively). In contrast, we fail to find any statistically meaningful relationship between extrinsic motivation and average total service time (in models that lack workload-motivation interactions) as indicated by the statistically nonsignificant coefficients of extrinsic motivation in specifications (2) and (5) of Table 4. Then, we examine the interplay between workload and the motivation measures in affecting total service time. The OLS estimates (specification (3)) show a negative workload effect for agents with average intrinsic and extrinsic motivation levels, whereas some of the cross-level interaction effects are statistically significant. In contrast, the 2SLS estimates (specification (6)) show a positive workload effect (9.44, $p < 0.01$) for agents with average intrinsic and extrinsic motivation levels. Furthermore, all of the 2SLS cross-

Table 3. Joint Effects of Intrinsic Motivation, Extrinsic Motivation, and Workload on Online Service Time

	Estimated by OLS models			Estimated by 2SLS models		
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	219.82*** (50.33)	252.32*** (53.45)	247.99*** (54.45)	246.28*** (55.20)	236.81*** (58.07)	245.33*** (58.78)
WL	−8.07*** (0.79)	−6.27*** (0.90)	−5.91*** (0.90)	−16.55*** (2.44)	0.61 (2.82)	−6.61** (3.07)
IM	—	1.46*** (0.55)	4.68*** (0.54)	—	2.94*** (0.59)	5.81*** (0.63)
EM	—	−5.40*** (0.56)	−7.38*** (0.57)	—	−5.94*** (0.60)	−8.44*** (0.62)
IM × *WL*	—	—	−2.48*** (0.42)	—	—	−3.89*** (0.68)
EM × *WL*	—	—	1.10 (0.56)	—	—	5.04*** (1.21)
IM × *EM*	—	—	10.33*** (0.37)	—	—	10.24*** (0.40)
IM × *EM* × *WL*	—	—	1.22*** (0.43)	—	—	1.75** (0.87)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. of calls	109,796	88,024	88,024	99,356	79,440	79,440
No. of agents	82	64	64	82	64	64
Model fit (Pr > χ^2)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Notes. Robust standard errors are in parentheses. The variables workload (*WL*), intrinsic motivation (*IM*), and extrinsic motivation (*EM*) are mean centered.
**Significance at the 5% level; ***significance at the 1% level.

level interaction effects are statistically significant at the 1% level. We plot 2SLS estimates of total service time at high IM-low EM and low IM-high EM configurations to better understand the interaction results (see Figure 7(c)). The results indicate a flat response to workload for the high IM-low EM group (line B of Figure 7(c)) and an increasing marginal effect of workload on total service time for the low IM-high EM group (line A of Figure 7(c)). Thus, at lower (higher) levels of workload, agents with a low IM-high EM motivation level spend less (more) time servicing patients' requests. Finally, we probe the model in specification (6) of Table 4 at different motivation levels (i.e., high IM-low EM and low IM-high EM) to examine the conditional effect of workload on total service time at each of the predefined motivation configurations. The results suggest a statistically nonsignificant workload effect on total service time for the high IM-low EM (–1.06, ns) configuration and a positive workload effect on total service time for the low IM-high EM (35.28, $p < 0.01$) configuration.

6.4. Strength of the Instrumental Variables

Table 5 shows the first-stage regressions of the instrumented variables for the full specification model. The first-stage regression results indicate that all lagged variables were positively correlated with their endogenous independent variable counterparts (i.e., *WL* and *LWL*, WL^2 and LWL^2, etc.), with the positive coefficients ranging between 0.317 and 0.748. The *F* statistics for the joint significance of the instrumental variables in the first stage are all well over 10, which indicates that this approach does not suffer from the weak instruments problem.

7. Discussion and Conclusion

Our findings echo evidence in the behavioral queuing literature that service time is not exogenous to workload and involves a complex set of contingency factors. We propose that clarification about mixed findings in prior research concerning the relationship between workload and service time (Delasay et al. 2019) may be enhanced by an interactionist view integrating (a) contextual factors such as incentive systems and queue structures that vary substantially across research settings (e.g., Song et al. 2015, Shunko et al. 2018), (b) enduring individual differences such as trait intrinsic and extrinsic motivation that vary across individuals (e.g., Grant 2008, Grant and Berry 2011), and (c) the nature of the service request itself (such as offline versus online services that lack/involve live interactions with customers, respectively).

In our sample, servers with high levels of intrinsic motivation and low levels of extrinsic motivation continued to display high levels of productivity in the *offline* portions of service requests and were less affected by workload levels. In contrast, the offline productivity of servers with high levels of extrinsic motivation

Table 4. Joint Effects of Intrinsic Motivation, Extrinsic Motivation, and Workload on Total Service Time

	Estimated by OLS models			Estimated by 2SLS models		
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	268.75***	305.89***	301.16***	266.37***	272.37***	276.01***
	(63.78)	(67.62)	(68.52)	(71.32)	(74.16)	(74.71)
WL	−3.44***	−4.60***	−4.44***	6.76**	14.72***	9.44***
	(0.89)	(1.03)	(1.03)	(2.76)	(3.22)	(3.51)
IM	—	−2.29***	1.08	—	−1.59**	2.85***
		(0.64)	(0.64)		(0.69)	(0.74)
EM	—	0.31	−1.09	—	-0.23	−1.77**
		(0.63)	(0.65)		(0.67)	(0.71)
IM × *WL*	—	—	−1.29***	—	—	−2.55***
			(0.47)			(0.78)
EM × *WL*	—	—	3.78***	—	—	10.21***
			(0.63)			(1.39)
IM × *EM*	—	—	9.35***	—	—	9.70***
			(0.43)			(0.46)
IM × *EM* × *WL*	—	—	0.19	—	—	−3.86***
			(0.48)			(1.00)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. of calls	109,796	88,024	88,024	99,356	79,440	79,440
No. of agents	82	64	64	82	64	64
Model fit (Pr > χ^2)	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Notes. Robust standard errors are in parentheses. The variables workload (*WL*), intrinsic motivation (*IM*), and extrinsic motivation (*EM*) are mean centered.
**Significance at the 5% level; ***significance at the 1% level.

and low levels of intrinsic motivation fluctuated more drastically in relation to workload. Specifically, servers with high extrinsic and low intrinsic motivation slowed down in offline time substantially as workload increased.

In nonhypothesized parallel analyses of *online* service times, we found that servers with high intrinsic and low extrinsic motivation spent much more online service time with customers (than their high EM and low IM counterparts) when workloads were low, as

Table 5. First-Stage Regressions of Instrumented Variables (Full Specification)

	WL	*WL*^{2}	*WL* × *IM*	*WL*^{2} × *IM*	*WL* × *EM*	*WL*^{2} × *EM*	*WL* × *IM* × *EM*	*WL*^{2} × *IM* × *EM*
LWL	0.317***	0.072***	−0.064***	−0.028	0.180***	0.006	0.235***	0.015
	(0.004)	(0.012)	(0.006)	(0.016)	(0.006)	(0.010)	(0.008)	(0.014)
LWL^2	−0.007***	0.525***	−0.009***	−0.041***	−0.018***	0.112***	0.044***	0.236***
	(0.002)	(0.008)	(0.002)	(0.010)	(0.003)	(0.007)	(0.005)	(0.011)
LWL × *IM*	0.007***	−0.027***	0.748***	0.225***	0.095***	0.036***	0.090***	0.210***
	(0.002)	(0.007)	(0.005)	(0.012)	(0.003)	(0.007)	(0.006)	(0.012)
LWL^2 × *IM*	−0.002	−0.043***	−0.002	0.594***	0.005**	0.135***	0.013***	0.038***
	(0.001)	(0.007)	(0.002)	(0.009)	(0.002)	(0.005)	(0.003)	(0.008)
LWL × *EM*	0.019***	−0.019***	0.111***	0.023**	0.557***	0.085***	0.007	0.108***
	(0.002)	(0.007)	(0.005)	(0.010)	(0.006)	(0.011)	(0.015)	(0.021)
LWL^2 × *EM*	−0.011***	0.063***	0.006**	0.199***	−0.023***	0.437***	−0.018**	−0.089***
	(0.002)	(0.007)	(0.003)	(0.009)	(0.004)	(0.008)	(0.008)	(0.014)
LWL × *IM* × *EM*	0.020***	0.035***	0.024***	0.072***	0.003	0.047***	0.532***	0.067***
	(0.002)	(0.005)	(0.004)	(0.008)	(0.005)	(0.008)	(0.015)	(0.022)
LWL^2 × *IM* × *EM*	0.003	0.083***	0.006**	−0.002	−0.001	−0.040***	0.013	0.426***
	(0.001)	(0.005)	(0.003)	(0.008)	(0.004)	(0.007)	(0.009)	(0.016)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	79,440	79,440	79,440	79,440	79,440	79,440	79,440	79,440
F statistic	877.3	788.6	5,022.3	778.3	3,145.1	1,033.7	858.6	498.7

Notes. Robust standard errors are in parentheses. The *F* statistic tests the joint significance of the instrumental variables for each first-stage model.
**Significance at the 5% level; ***significance at the 1% level.

evidenced by the high y intercept for line B in Figure 7(b), but significantly reduced their online time as workload increased. These findings are consistent with the preferences we noted by call center managers who under conditions of low workload, prefer agents to be on the phone spending extra time with customers rather than idly spending time off the phone between calls surfing the internet, chatting with coworkers, or checking email. Conversely, agents with high extrinsic motivation and low intrinsic motivation had much lower overall online service times irrespective of workload (see the lower y intercept and nearly flat line A in Figure 7(b)). These findings are consistent with concerns noted by managers that in very low-workload conditions, agents may rush the online portion of calls to get more idle time between calls, even at the expense of customer service.

Finally, we found evidence that *total* service time was effectively exogenous to workload for servers with high intrinsic and low extrinsic motivation but highly sensitive to workload for their low IM and high EM counterparts (as noted by the slopes and intercepts of line B and line A in Figure 7(c), respectively). It is interesting to consider these somewhat contrasting findings on total service time with those of offline and online service times. Perhaps the speed-up effect for the high IM, low EM servers found in Figure 7(b), line B (i.e., starting with a high y intercept at very low workload levels in online time) balanced out the relatively flat slowdown pattern in the offline portion (i.e., in Figure 7(a), line B). Conversely, our results suggest that servers with high EM and low IM slowed down as workload increased in offline call portions, whereas they had a relatively flat response to workload in online call portions (i.e., comparing line A in Figure 7(a) with line A in Figure 7(b)). As a whole, these findings suggested nuances and dynamics that we did not fully anticipate but are interesting. Servers with high IM and low EM appear to have been more likely to slow down and speed up when it was in the best interest of the customer, when it was in the best interest of the organization, and when the queue was empty (or near empty). Conversely, servers with high EM and low IM may have pursued more idle time and missed opportunities to boost service quality by going more swiftly through online calls at very low workloads. It also appears that these high EM, low IM servers may have had avoided getting back into the queue by extending the offline portion of the service call (as workload increased) up to a certain point.

More research is needed to better understand the complex dynamics of the various portions of the service call as they relate to trait motivation in workload. Indeed, total service time metrics may oversimplify nuanced differences between online and offline service times (Gans et al. 2003). However, it is helpful to be able to model offline, online, and total service times concurrently to get a more complete view of these relationships.

Together, our results suggest that trait-based differences in intrinsic and extrinsic motivation between servers may offer additional clarity in explaining why and when servers might speed up or slow down in response to workload fluctuations in service queues. These productivity differences may be partially driven by how servers perceive higher levels of workload given their different motives. For example, servers with a trait propensity to be intrinsically motivated by call center work might perceive busy times more favorably than their low IM counterparts because they enjoy the work itself. In describing how excess workload is perceived, one agent said: "The day goes by faster [when it is busy]. [I] like to stay busy." These servers may prefer to avoid idle time and meet the needs of customers.

Our study suggests the need to consider the interplay between individual differences and context to understand workplace phenomena. To better quantify the impact of our findings, we calculated differences in offline and online productivity between agents with different trait motivation levels working under various workloads. Using the results from specification (6) of Table 2, we found that servers with a combination of high intrinsic and low extrinsic motivation were approximately 15% (161%) faster in processing the offline portion of service requests than their peers with the opposite combination (low and high) when workload levels were low (high).[4] In contrast, using the results from specification (6) of Table 3, we found that servers with high IM and low EM were approximately 26% (5%) slower in processing the online portion of service requests than their low IM and high EM counterparts when workload levels were low (high). Given call center managers' preference for efficiency and general reduction in offline time *irrespective of workload* and their tendency to prioritize talking to customers when workload levels are low, these are meaningful differences.

Our findings raise important questions about how organizations might make practical use of evidence that trait-based individual differences influence how servers respond to changing workload. For example, could employers screen for trait motivation differences among job applicants to optimize hiring or team design decisions? For decades, personality and trait assessments have been used commonly in organizations for selection and placement purposes (Hurtz and Donovan 2000). Recent popular press estimates suggest that as many as 75% of large companies in the United States use personality tests in selection (NBC News 2021), and instruments like the Meyers Briggs

Type Indicator generate an estimated $20 million in yearly revenue (European CEO 2019).

However, research also strongly suggests need for caution, as some trait measures have been shown to be susceptible to faking and social desirability biases (Nunnally and Bernstein 1994), particularly when used for making critical administrative decisions (i.e., hiring or job placement) instead of for anonymous research purposes (Mount et al. 1999, Hogan et al. 2007). Consequently, researchers have developed and tested various approaches designed to overcome these biases and potential legal complications associated with them (Youngman 2017). Such methods include the use of peer and supervisor ratings, nonexplicit (or indirectly worded) survey items, assessment centers (using in-basket exercises and simulations), implicit association approaches, and statistical correction techniques involving social desirability markers (Sjöberg 2015). Evidence suggests that many of these methods are valid, supporting the popularity of their use (Collins et al. 2003, Yovel and Friedman 2013, Sjöberg 2015). The strengths, limitations, timing,[5] and legal implications of the multitude of approaches to measuring trait-based individual differences in the workplace are well documented (Martin 2014, Youngman 2017, Lundgren et al. 2019). This literature provides call center managers with ample guidance as they navigate the nuances of trait measurement to consider how they might utilize employee trait information to enhance operational performance.

We encourage future research to build on this study to explore the role of additional individual differences such as Big 5 personality, hardiness, or other traits as explanatory mechanisms for understanding agents' reactions to changing workloads. We further acknowledge that our study results may not generalize to other contexts such as those with robust pay for performance systems or those with parallel queues. Future research needs to examine whether and how our findings would change in service queuing systems with different financial incentives (e.g., if servers were rewarded for their performance during busy periods) and/or different queue structures (e.g., if each server was responsible for managing his or her own dedicated queue). Another threat to the generalizability of our findings concerns whether the distribution of worker types in our setting is also found in other settings. We note that this is a lesser concern in our study because the distribution of trait IM and EM scores found in our sample covered the full range of possible values (from a minimum of one point to a maximum of seven points), included various configurations of IM-EM as illustrated by Figure EC.2 of the e-companion, and had mean and standard deviation scores comparable with other studies that used similar motivation measures solicited from workers in various service and manufacturing organizations across the globe. We also note that although we control for weekly and daily temporal effects, we cannot rule out the possibility of missing other seasonal factors (e.g., monthly effects) that were not present during the examined time period. However, based on feedback received from the call center's leadership team, we note that this is not a major concern in our setting because the observed workload levels per agent (i.e., workload adjusted for number of servers) were mostly consistent across months because of the hiring practices employed by the call center, which accounted for potential seasonal changes in call volume. Finally, we note that we consider a trait-based view of intrinsic and extrinsic motivation, which means that we expect individual-level personality traits of intrinsic and extrinsic motivation to be stable over time. Hence, trait levels of IM and EM are relatively enduring and might be hard to manipulate in an experimental setting or via managerial interventions. Thus, we echo our previous assertions that moving beyond agent-level fixed effects brings additional insights into the complex relationship between workload and productivity.

Acknowledgments
The authors thank Department Editor Charles Corbett, the associate editor, and the reviewers for valuable comments that significantly shaped this manuscript. They also thank Kenneth Schultz, David Schoorman, Dhari Alrasheed, Mohammad Alyakoob, Salman Aljazzaf, and Jelle de Vries for valuable feedback on earlier versions of the manuscript.

Endnotes
[1] We include the results of alternative models with OLS estimates in Table EC.5 of the e-companion. The results of these models were consistent with our hypotheses.

[2] Unlike the offline service time models, both the online and total service time models had better model fits when using linear workload effects.

[3] The estimates in Figure 3 were obtained using the results from specification (6) of Tables 2 and 3.

[4] We defined low (high) levels of workload at one standard deviation below (above) the mean of workload.

[5] For example, employers are less likely to use trait-based selection tests in tighter labor markets.

References
Allon G, Kremer M (2018) Behavioral foundations of queueing systems. Donohue K, Katok E, Leider S, eds. *The Handbook of Behavioral Operations*, Wiley Series in Operations Research and Management Science (Wiley, Hoboken, NJ), 325–366.

Altman D, Yom-Tov GB, Olivares M, Ashtar S, Rafaeli A (2021) Do customer emotions affect agent speed? An empirical study of emotional load in online customer contact centers. *Manufacturing Service Oper. Management* 23(4):854–875.

Amabile TM, Hill KG, Hennessey BA, Tighe EM (1994) The work preference inventory: Assessing intrinsic and extrinsic motivational orientations. *J. Personality Soc. Psych.* 66(5):950–967.

Ashkanani AM (2017) Servers' effectiveness in service queuing systems: A behavioral operations management approach. PhD thesis, Purdue University, West Lafayette, IN.

Ashtar S, Yom-Tov GB, Akiva N, Rafaeli A (2021) When do service employees smile? Response-dependent emotion regulation in emotional labor. *J. Organ. Behav.* 42(9):1202–1227.

Batt RJ, Terwiesch C (2017) Early task initiation and other load-adaptive mechanisms in the emergency department. *Management Sci.* 63(11):3531–3551.

Berry Jaeker JA, Tucker AL (2017) Past the point of speeding up: The negative effects of workload saturation on efficiency and patient severity. *Management Sci.* 63(4):1042–1062.

Boudreau J, Hopp W, McClain JO, Thomas LJ (2003) On the interface between operations and human resources management. *Manufacturing Service Oper. Management* 5(3):179–202.

Brannan MJ (2005) Once more with feeling: Ethnographic reflections on the mediation of tension in a small team of call centre workers. *Gender Work Organ.* 12(5):420–439.

Carlaw M, Carlaw P, Deming VK, Friedmann K (2002) *Managing and Motivating Contact Center Employees: Tools and Techniques for Inspiring Outstanding Performance from Your Frontline Staff* (McGraw-Hill Companies, New York).

Collins JM, Schmidt FL, Sanchez-Ku M, Thomas L, McDaniel MA, Le H (2003) Can basic individual differences shed light on the construct meaning of assessment center evaluations? *Internat. J. Selection Assessment* 11(1):17–29.

Contact Centre Magazine (2020) 7 tricks that call centre employees play. Accessed September 2, 2021, https://www.callcentrehelper.com/7-tricks-that-call-centre-employees-play-67004.htm.

Croson R, Schultz K, Siemsen E, Yeo M (2013) Behavioral operations: The state of the field. *J. Oper. Management* 31(1–2):1–5.

Deery S, Iverson R, Walsh J (2002) Work relationships in telephone call centres: Understanding emotional exhaustion and employee withdrawal. *J. Management Stud.* 39(4):471–496.

Deery S, Iverson R, Walsh J (2004) The effect of customer service encounters on job satisfaction and emotional exhaustion. Deery S, Kinnie N, eds. *Call Centres and Human Resource Management* (Palgrave Macmillan, London), 201–222.

Delasay M, Ingolfsson A, Kolfal B, Schultz K (2019) Load effect on service times. *Eur. J. Oper. Res.* 279(3):673–686.

Diefendorff JM, Gabriel AS, Nolan MT, Yang J (2019) Emotion regulation in the context of customer mistreatment and felt affect: An event-based profile approach. *J. Appl. Psych.* 104(7):965–983.

Ekehammar B (1974) Interactionism in personality from a historical perspective. *Psych. Bull.* 81(12):1026–1048.

European CEO (2019) Testing times: Why personality assessments are being questioned. Accessed September 29, 2021, https://www.europeanceo.com/business-and-management/testing-times-why-personality-assessments-are-being-questioned/.

Fang M, Gerhart B (2012) Does pay for performance diminish intrinsic interest? *Internat. J. Human Resourse Management* 23(6):1176–1196.

Gagné M, Deci EL (2005) Self-determination theory and work motivation. *J. Organ. Behav.* 26(4):331–362.

Gans N, Koole G, Mandelbaum A (2003) Telephone call centers: Tutorial, review, and research prospects. *Manufacturing Service Oper. Management* 5(2):79–141.

Grant AM (2008) Does intrinsic motivation fuel the prosocial fire? Motivational synergy in predicting persistence, performance, and productivity. *J. Appl. Psych.* 93(1):48–58.

Grant AM, Berry JW (2011) The necessity of others is the mother of invention: Intrinsic and prosocial motivations, perspective taking, and creativity. *Acad. Management J.* 54(1):73–96.

Greene WH (2018) *Econometric Analysis* (Pearson Education, New York).

Hayes AF (2022) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (Guilford Publications, New York).

Hogan J, Barrett P, Hogan R (2007) Personality measurement, faking, and employment selection. *J. Appl. Psych.* 92(5):1270–1285.

Hurtz GM, Donovan JJ (2000) Personality and job performance: The big five revisited. *J. Appl. Psych.* 85(6):869–879.

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.

Kinnie N, Hutchinson S, Purcell J (2000) 'Fun and surveillance': The paradox of high commitment management in call centres. *Internat. J. Human Resource Management* 11(5):967–985.

Knights D, McCabe D (1998) 'What happens when the phone goes wild?': Staff, stress and spaces for escape in a BPR telephone banking work regime. *J. Management Stud.* 35(2):163–194.

Lau N, Hasija S, Bearden JN (2014) Newsvendor pull-to-center reconsidered. *Decision Support Systems* 58(2014):68–73.

Lewin K (1935) *A Dynamica Theory of Personality: Selected Papers* (McGraw-Hill, New York).

Lundgren H, Poell RF, Kroon B (2019) "This is not a test": How do human resource development professionals use personality tests as tools of their professional practice? *Human Resource Development Quart.* 30(2):175–196.

Mahesh VS, Kasturi A (2006) Improving call centre agent performance: A UK–India study based on the agents' point of view. *Internat. J. Service Indust. Management* 17(2):136–157.

Martin W (2014) The problem with using personality tests for hiring. *Harvard Bus. Rev.* (August 27), https://hbr.org/2014/08/the-problem-with-using-personality-tests-for-hiring.

Mount MK, Barrick MR, Strauss JP (1999) The joint relationship of conscientiousness and ability with performance: Test of the interaction hypothesis. *J. Management* 25(5):707–721.

NBC News (2021) Job hiring increasingly relies on personality tests, but that can bar people with disabilities. (March 4), https://www.nbcnews.com/think/opinion/job-hiring-increasingly-relies-personality-tests-they-can-bar-people-ncna1259466.

Newey WK (1987) Efficient estimation of limited dependent variable models with endogenous explanatory variables. *J. Econometrics* 36(3):231–250.

Nunnally J, Bernstein I (1994) *Psychometric Theory* (McGraw-Hill, New York).

Rafaeli A, Tov GBY, Ashtar S, Altman D (2020) Opportunities, tools, and new insights: Evidence on emotions in service from analyses of digital traces data. Härtel CEJ, Zerbe WJ, Ashkanasy NM, eds. *Emotions and Service in the Digital Age*, Research on Emotion in Organizations, vol. 16 (Emerald Publishing Limited, Bingley, United Kingdom), 105–133.

Raghuram S (2013) Identities on call: Impact of impression management on Indian call center agents. *Human Relations* 66(11):1471–1496.

Ryan RM, Deci EL (2000) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Amer. Psych.* 55(1):68–78.

Rynes SL, Gerhart B, eds. (2000) *Compensation in Organizations. Current Research and Practice* (Wiley, New York).

Sansone C, Weir C, Harpster L, Morgan C (1992) Once a boring task always a boring task? Interest as a self-regulatory mechanism. *J. Personality Soc. Psych.* 63(3):379.

Shunko M, Niederhoff J, Rosokha Y (2018) Humans are not machines: The behavioral impact of queueing design on service time. *Management Sci.* 64(1):453–473.

Sjöberg L (2015) Correction for faking in self-report personality tests. *Scandinavian J. Psych.* 56(5):582–591.

Skinner BF (2014) *Contingencies of Reinforcement: A Theoretical Analysis*, vol. 3 (BF Skinner Foundation, Cambridge, MA).

Song H, Tucker AL, Murrell KL (2015) The diseconomies of queue pooling: An empirical investigation of emergency department length of stay. *Management Sci.* 61(12):3032–3053.

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(6):1574–1593.

Taylor P, Bain P (1999) 'An assembly line in the head': Work and employee relations in the call centre. *Indust. Relations J.* 30(2):101–117.

Tett RP, Burnett DD (2003) A personality trait-based interactionist model of job performance. *J. Appl. Psych.* 88(3):500.

Tobin J (1958) Estimation of relationships for limited dependent variables. *Econometrica* 26(1):24–36.

Wang J, Zhou YP (2018) Impact of queue configuration on service time: Evidence from a supermarket. *Management Sci.* 64(7):3055–3075.

White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48(4):817–838.

Wooldridge JM (2020) *Introductory Econometrics: A Modern Approach* (Cengage, Boston).

Workman M, Bommer W (2004) Redesigning computer call center work: A longitudinal field experiment. *J. Organ. Behav.* 25(3): 317–337.

Yakubovich V, Lup D (2006) Stages of the recruitment process and the referrer's performance effect. *Organ. Sci.* 17(6): 710–723.

Youngman JF (2017) The use and abuse of pre-employment personality tests. *Bus. Horizons* 60(3):261–269.

Yovel I, Friedman A (2013) Bridging the gap between explicit and implicit measurement of personality: The questionnaire-based implicit association test. *Personality Individual Differences* 54(1): 76–80.