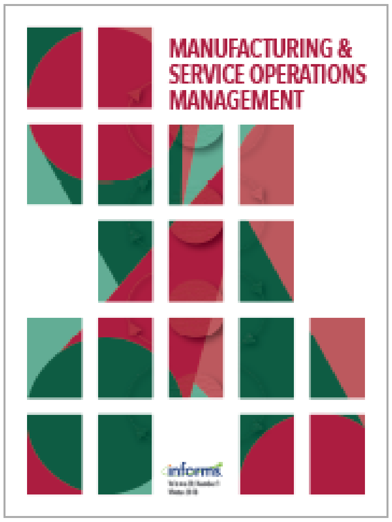


Manufacturing & Service Operations Management

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

Is Full Price the Full Story When Consumers Have Time and Budget Constraints?

Simin Li, Martin A. Lariviere, Achal Bassamboo

informs
https://pubsonline.informs.org/journal/msom

MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 26, No. 1, January–February 2024, pp. 370–388
ISSN 1523-4614 (print), ISSN 1526-5498 (online)

Is Full Price the Full Story When Consumers Have Time and Budget Constraints?

Simin Li,[a,*] Martin A. Lariviere,[b] Achal Bassamboo[b]

[a] Freeman School of Business, Tulane University, New Orleans, Louisiana 70118; [b] Kellogg School of Management, Northwestern University, Evanston, Illinois 60208

*Corresponding author

Contact: sli41@tulane.edu, https://orcid.org/0000-0001-8979-502X (SL); m-lariviere@kellogg.northwestern.edu, https://orcid.org/0000-0002-5005-9444 (MAL); a-bassamboo@kellogg.northwestern.edu, https://orcid.org/0000-0001-7758-4751 (AB)

Received: July 23, 2022
Revised: March 26, 2023; July 16, 2023; September 5, 2023
Accepted: September 8, 2023
Published Online in Articles in Advance: October 11, 2023

https://doi.org/10.1287/msom.2022.0357

Abstract. ***Problem definition*:** A canonical model in service management assumes that consumers base the purchase of a service on its full price, that is, a linear combination of the monetary price and the expected time commitment. Although analytically convenient, when this assumption holds is an unexplored question. ***Methodology/results*:** We present a model of consumers allocating their time and money between working, overhead activities that do not provide utility, one continuous leisure activity, and one discrete service. Both continuous leisure activity and discrete service increase utility. Consumers can allocate any nonnegative amount of time or money to the leisure activity. Consumption of the discrete service requires a specific amount of time and money. We examine when the decision to purchase the discrete service depends only on its full price. We show that the full-price assumption does hold in specific cases. To be precise, it depends on how consumers are paid. If consumers completely control the amount of time that they work and earn a constant wage, they base their purchase decision on the full price. If, however, they must work a fixed shift length, then the assumption fails, and the full price is not sufficient to determine the consumer's action. This leads to systematic differences in sellers' strategies when they serve consumers with different compensation structures. If the consumers must work longer than would be optimal if they controlled their schedule and earned the same hourly wage, that is, the consumers are overemployed shift workers, then a seller restricts sales (relative to selling to consumers who control their work hours), and the system is less congested. The reverse holds if the consumers would prefer to work longer at the offered wage; that is, the consumers are underemployed shift workers. ***Managerial implications*:** We show that sellers who fail to take prevailing compensation structures of the community they serve into consideration experience significant revenue loss. In some cases, we see losses in consumer surplus and social welfare as well.

Supplemental Material: The e-companion is available at https://doi.org/10.1287/msom.2022.0357.

Keywords: service operations • pricing and revenue management • economic models

1. Introduction

Services require commitments of both time and money. Getting one's hair styled requires not only a monetary price but also the time it takes to wait for the stylist to be free and the actual service time. Going to a boat tour requires not only purchasing a ticket but also two hours of one's time. How do consumers decide whether to purchase such services? Since Naor (1969), a standard assumption has been that consumers evaluate their purchase of a time-consuming service based on its full price: the sum of the explicit monetary cost and a weighting of the expected sojourn time for the service. This full-price regime is widely used in the literature because of its tractability; however, two underlying questions are worth reconsidering. First, is the full price always enough to describe the total costs consumers incur? Second, what determines the weighting of the expected time commitment? Although there is no ambiguity in the monetary cost, how consumers value their time is far less certain (Chen et al. 2021).

Many models in service management have built on the full-price regime. Apart from Naor (1969), many papers, including Edelson and Hildebrand (1975), Mendelson (1985), Mendelson and Whang (1990), Ha (1998), and Afèche (2013), have modeled the consumers' total expected cost to be the full price. Consumers' utility is then additive of service valuation and full price. Afèche and Mendelson (2004) assumed that service valuations are time dependent. But the consumers' utility remains a function of the full price. In service competition models, it is also common to assume that servers' demand rates depend on their respective full prices (Cachon and

Harker 2002). Under the full-price assumption, consumers evaluate a menu of high price and fast speed to be the same as a menu of low price and slow speed.

Additionally, there is a stream of literature that estimates the weightage on per-unit-time cost. In estimating per-unit-time cost, the rate of wage loss is a common approximation in both daily practice and academic research. In the popular press, we observe that people make time and money trade-offs based on their wage rates (Pesca 2018, Roberge 2021). In empirical research, patients' earnings are imputed to estimate their time cost while waiting for physicians (Keller and Laughhunn 1973, Mueller 1985). Shmanske (1993) and Gavirneni and Kulkarni (2016) used consumers' income as a proxy for heterogeneous waiting costs when pricing services. However, when Deacon and Sonstelie (1985) empirically estimated the value of consumers' time at a gas station and Allon et al. (2011) estimated the worth of consumers' waits at fast-food drive-throughs, both found discrepancies between consumers' value of time and their wage rates. In fact, a recent study conducted experiments that demonstrate that how consumers value their time depends on their compensation structures (Smitizsky et al. 2021), not merely their wage rates.

The full-price model has been questioned in the behavioral operations literature that studies how people might account for time and money differently (Allon and Kremer 2018, p. 329). Compared with money, time is less fungible, and the value of time is context dependent, that is, not constant, and more ambiguous (Leclerc et al. 1995, Okada and Hoch 2004). In the theory of constraints, the value of bottleneck time in a production environment depends on the mix of the production (Goldratt and Cox 2016). Liu et al. (2018) operationalized the cost of time as a patient's sensitivity to delay; they found that the data from experiments supported an S-shaped (i.e., a nonlinear) cost function over time. Ülkü et al. (2020) found that the amount consumers buy increases when they have experienced longer waits, a phenomenon they attribute to mental accounting, under which consumers amortize the sunk costs of time in the queue over larger quantities. Our approach to exploring the limitations of the full-price model differs from behavioral studies. We assume that consumers are wholly rational utility maximizers and examine whether the full price is a sufficient statistic for their choices. That is, we try to meet the literature spawned by Naor (1969) on its own terms.

Following the spirit of Becker (1965), we present a model of time and money allocation for a utility-maximizing consumer. We consider a world in which a consumer's utility depends on the time and money he or she has available for leisure. The consumer seeks to maximize his or her overall utility subject to constraints on both his or her time and monetary budget. Money is earned only by working. In addition to working and leisure, time must be spent on overhead tasks (e.g., cleaning the house or doing laundry). Overhead tasks are in some sense neutral. They neither increase nor decrease utility, and although we allow the consumer to spend money to reduce the time spent on overhead tasks (e.g., the consumer can hire someone to clean his or her house), overhead tasks do not necessarily require spending money.

We suppose the consumer can enjoy both continuous leisure activities and discrete services. Continuous leisure activities are the consumer's basic forms of leisure. How much to consume is a marginal decision that the consumer can fine-tune to maximize his or her utility. Discrete services are the ones whose time commitment is not fungible, for instance, having one's hair done and going on a boat tour. If the consumer decides to purchase the service, he or she will have to allocate a certain amount of time and money from his or her time and budget constraints. In return, the consumer's utility will increase by a fixed amount. That is, we do not a priori consider comparing the service valuation with the aggregated time and monetary expenses but rather endogenously determine the payoff of buying the service based on how the consumer compares the optimal utility of buying and not buying the discrete service.

This simplified setting allows us to study the impact of different compensation structures. Our premise is that not all jobs paying a given average hourly wage are the same. For example, one job may offer the consumer complete flexibility in terms of how long to work (akin to what the gig economy promises). Another employment opportunity may pay the same hourly wage but require the consumer to commit to a shift of S hours—no more, no less. According to a survey from the Bureau of Labor Statistics (2019), 43% of workers have an inflexible shift work schedule; that is, these workers were unable to vary or change the times that they began and stopped working.[1] Suppose that the consumer would work fewer than S hours if he or she had complete flexibility. Then, the consumer will be time-constrained in a setting where he or she must commit to S hours for the same hourly wage. Conversely, if the consumer would work more than S hours under conditions of flexibility, he or she will be earnings constrained when locked into a shift. Following Altonji and Paxson (1988), we will use term *overemployment* to describe the former setting and *underemployment* for the latter.

Even though the model we construct allows significant flexibility in the types of activities and services the consumers can choose from, we find that the full-price assumption holds only under limited circumstances. Specifically, our paper makes the following contributions.

- We prove that the full-price regime holds only for gig workers, that is, consumers who have complete flexibility in work time. See Proposition 1.

- We prove that the full-price regime does not capture consumers' decision making when they are either time constrained (i.e., overemployed) or earnings constrained (i.e., underemployed). See Lemma 2 and Proposition 2.
- We consider two models of discrete service providers: managing a single-server queue or a service with nonbinding capacity. We prove that, depending on whether its consumers are overemployed or underemployed shift workers, a revenue-maximizing discrete service provider would have to price its service so that it is systematically less congested (allowing fewer consumers to join the service) or more congested (allowing more consumers to join the service) than a service serving gig workers, respectively. See Section 4.
- We show numerically that pricing all consumers under a full-price regime and erroneously pricing shift workers as gig workers can cause substantial revenue losses. Likewise, incorrectly pricing gig workers as shift workers reduces revenues. In some cases, such mispricing can also diminish social welfare and consumer surplus. See Sections 5.2 and 5.3.

We organize the paper as follows. In Section 2, we model how consumers allocate their time and money across activities, including the discrete service being evaluated. We formally define the full-price regime based on the model. We subsequently investigate whether the full-price regime is enough to make purchase decisions under different compensation structures in Section 3. In Section 4, we move from the consumer's purchase decision to the seller's pricing problem. In Section 5, we show an extension of our model and also the managerial implications of assuming all consumers to be full-price decision makers. All proofs are relegated to the Appendix.

2. A Model of Allocating Time and Money

We examine how a utility-maximizing consumer allocates his or her time across three kinds of activities: (a) earning money, (b) necessary activities, and (c) utility-increasing activities. Necessary activities are the "overhead" burden of daily life (e.g., doing laundry or mowing the lawn). We assume that these are neither enjoyable nor remunerative. We assume that the consumer can spend money to reduce his or her overhead burden. Thus, in addition to allocating his or her time, the consumer can also allocate the money he or she earns between overhead reduction and leisure.

Utility-increasing activities are either continuous or discrete. The time and money allocated to continuous leisure activities (e.g., watching TV or hiking) can take essentially any nonnegative value. A discrete service, by contrast, is something like getting one's hair done or getting a coffee, which can be consumed only in its entirety or not at all, because the utility boost is rewarded only when the service is completed, that is, one's hair is done or the coffee is served. Whereas the consumer makes a marginal decision on how much money and time to spend on a continuous activity, a discrete service requires spending a specific amount of time and money.

In what follows, we consider a setting in which there is one continuous leisure activity and one discrete service and present a general framework for the consumer's problem. Based on the consumer's problem, we further formally define the full-price regime. We include one continuous leisure activity for convenience, but the model can be easily extended to include multiple continuous leisure activities. We discuss the extension of having multiple discrete services in Section 5.

2.1. The Utility-Maximizing Consumer's Problem

Consider a consumer who gains utility of $U(t_l, c_l)$ from leisure time t_l and leisure consumption c_l. We assume $U(t_l, c_l)$ to be continuous, nondecreasing, and concave in both arguments. In addition, $U(0, c_l)$ and $U(t_l, 0)$ are assumed to be nonnegative for all $c_l \geq 0$ and $t_l \geq 0$. A type θ consumer's problem is

$$
(P_C) \qquad \max_{(t_l, c_l, t_w, c_O, y)} \quad U(t_l, c_l) + Vy
$$
$$
\text{s.t.} \quad t_w + O(c_O) + t_l + \delta y \leq T,
$$
$$
c_O + c_l + ry \leq \Pi_\theta(t_w),
$$
$$
y \in \{0,1\}, t_l, c_l, t_w, c_O \geq 0.
$$

Following the classical theory, the utility-maximizing consumer is subject to two resource constraints, time and budget (Becker 1965). Suppose the consumer has T units of time per planning horizon. He or she spends time in three types of activities: work t_w, overhead activities O, and leisure, that is, the time constraint in (P_C). Instead of doing all overhead activities by him- or herself, the consumer might hire a professional to do some of the work or lease equipment to speed up the work. As a result, the time spent in overhead activities is a function of the consumer's spending c_O. We assume $O(c_O)$ to be decreasing and strictly convex in c_O; that is, there is a diminishing return in buying the overhead time back.

As described above, there are two types of utility-increasing activities that consumers might spend time doing. With a continuous leisure activity, the consumer can decide freely how much time he or she wants to put into, for instance, watching TV. The amount of time allocated to the continuous activity, t_l, can be made as granular as one wants. In contrast, a discrete service requires a block of time, δ, for instance, getting one's hair styled. Here, the consumer's decision is binary: either purchase and engage in the discrete service or not, that is, $y \in \{0, 1\}$. If the consumer decides to purchase, that is, $y = 1$, the activity consumes δ time units, rewards V units in utility, and costs r units of budget; otherwise, $y = 0$. That is, to the consumer, the question to consider is, given a known and fixed utility boost V, should he or she purchase the service at (r, δ)?

Note that we are agnostic between how the time δ is split between value-added time (having a stylist cut one's hair) and "dead" time (waiting for service to begin). This is consistent with Naor (1969) and essentially most of the literature that follows. Time spent chatting with the stylist while they work may be psychologically more enjoyable than sitting in the waiting room; this can be included in the utility gain V. However, we are interested in the economic cost of time. The consumer has a limited number of hours in the day, and any time needed for the discrete service—whether waiting or in service—counts against the available time budget and of necessity must displace some other activity.

The out-of-pocket price of the discrete service, r, comes from the consumer's total budget, which the consumer earns from working. Working also funds spending on continuous leisure consumption c_l and overhead activities c_O, that is, the budget constraint in (P_C). We denote the consumer's earnings from working as $\Pi_\theta(t_w)$, that is, the compensation function of a type θ consumer. The agent's earnings depend on both time spent working, t_w, and his or her type θ. We assume $\Pi_\theta(0) = 0$ and $\Pi'_\theta(t_w) \geq 0$. We further assume that the consumer has no savings, so his or her income comes entirely from working. Below, we will interpret θ as the consumer's wage, so it is natural to assume $\Pi_\theta(t_w) \geq \Pi_{\tilde{\theta}}(t_w)$ for all $\theta > \tilde{\theta}$.

The consumer plans for a typical day. We refer to the decision time horizon as a day for convenience. Depending on his or her wage type, and monetary cost and time required for the service, the consumer evaluates whether to purchase the discrete service $y^*(\theta, r, \delta)$ and how to allocate time (t_l^*, t_w^*) and budget (c_l^*, c_O^*) to achieve optimal utility u^* in advance. That is, we are thinking of a high-level planning problem involving the time and budget of the consumer. The consumer's plan is based on how he or she thinks his or her day is going to go and the expected time commitment of the discrete service. Furthermore, the consumer commits to his or her purchase decision. We effectively assume that the consumer pays for the service in advance. Formally, in summary, utility-maximizing consumers individually solve problem (P_C).

2.2. A Full-Price Decision Maker

We say that a consumer bases her decision on the full price if the out-of-pocket price r combined with a multiple of the time commitment for the discrete service δ is sufficient to characterize the decision whether to purchase the discrete service, that is, $y^*(\theta, r, \delta)$. As noted in Allon and Federgruen (2007), "[t]his is tantamount to assuming that all consumers assign a specific cost value to their waiting time and that the cost of waiting is simply proportional to the total waiting time."

More formally, we say that a discrete service price r and time commitment δ are feasible if there exist nonnegative t_l, t_w, c_l, and c_O that satisfy the constraints of (P_C); that is, (P_C) has a nonempty feasible set. Let $u^B(\theta, r, \delta)$ be the optimal utility from the *continuous* leisure activity for a type θ who *purchases* the discrete service at price r and time commitment δ. That is, $u^B(\theta, r, \delta) + V$ is the optimal total utility of a consumer that buys, that is, when $y = 1$.

Definition 1 (Full-Price Decision Maker). A consumer is a full-price decision maker if there exists an $\eta > 0$ such that $u^B(\theta, r, \delta) = u^B(\theta, \hat{r}, \hat{\delta})$ for every feasible $(\hat{r}, \hat{\delta})$ with $r + \eta\delta = \hat{r} + \eta\hat{\delta}$.

Thus, as Allon and Federgruen (2007) observed, full-price decision makers convert delay into dollars at a constant rate. Full price is a sufficient statistic for the consumer's purchase of the discrete service, because if the consumer buys under (r, δ), he or she must also buy under $(\hat{r}, \hat{\delta})$. However, this does not mean that other than purchasing, the consumer does the same things. Rather, the consumer adjusts his or her decision about how much to work and how much to spend on overhead reduction to achieve the same utility outside the discrete service. Define an iso-utility curve as the locus of the combinations of r and δ that will yield a constant level of utility from the continuous activity given that one buys the discrete service. (r, δ) and $(\hat{r}, \hat{\delta})$ are thus on the same iso-utility curve for a type θ consumer if and only if $u^B(\theta, r, \delta) = u^B(\theta, \hat{r}, \hat{\delta})$. This further means that on an iso-utility curve we must have $du^B = \frac{\partial u^B}{\partial r} dr + \frac{\partial u^B}{\partial \delta} d\delta = 0$ (see Chiang 1984).

Lemma 1. *A consumer is a full-price decision maker if and only if the iso-utility curves for the continuous activity of the buying consumers are linear in* (r, δ) *space.*

3. Consumers' Purchase Decisions of the Discrete Service

We begin by investigating whether consumers are ever full-price decision makers. We consider two common compensation structures observed in practice: first, a linear compensation structure under which the consumer earns a constant wage rate per time unit and can freely choose how long to work, and second, a shift pay compensation structure under which the consumer gets a lump sum payment by committing to a fixed number of working hours. One may think of the former as describing the gig economy and the latter as fitting office work or an assembly line.

3.1. When the Full Price is the Full Story: Linear Compensation Structure

We first consider the linear compensation structure, which models *gig workers* who are paid at a constant wage rate and can freely choose how long to work. It is given by

$$\Pi_\theta(t_w) = \theta t_w, \tag{1}$$

where θ is the wage per time unit and t_w is the number of time units the consumer chooses to work. We solve for consumers' optimal actions (t_w^j, c_O^j) through first-order conditions, where $j \in \{NB, B\}$ denotes a nonbuyer's and a buyer's decision, respectively. Define $\chi(t_l, c_l) = \frac{\partial U(t_l, c_l)}{\partial t_l} / \frac{\partial U(t_l, c_l)}{\partial c_l}$. By Lemma 1, the following proposition holds (See A.2 for proof).

Proposition 1. *Consumers, who are gig workers with compensation structure definition (1), are full-price decision makers. Furthermore, gig workers' marginal rate of substitution of time for money is their wages, that is,* $\chi(t_l^j, c_l^j) = \theta$, *for* $j \in \{NB, B\}$. *That is, gig workers have iso-utility curves with constant slopes, and that slope is equal to their wages.*

Remark 1. A common model in the pricing-of-queues literature assumes that consumers are full-price decision makers and that both their value of the service and their waiting cost depend on their type θ. Their utility when buying the service is then $v(\theta) - r - \kappa(\theta)\delta$, where $v(\theta)$ and $\kappa(\theta)$ are strictly increasing, positive functions, r is the monetary price of service, and δ the time commitment. Because both θ and $\kappa(\theta)$ are strictly increasing, we can without loss of generality assume that $\kappa(\theta) = \theta$. This structure is widely used. See, Afèche and Mendelson (2004), Nazerzadeh and Randhawa (2018), and Gurvich et al. (2019).

Although researchers have assumed this structure, the second part of Proposition 1—that the marginal rate of substitution of time for money is constant and equals the consumer's wage—provides a foundation for when it will hold. Let $f_\theta = r + \theta\delta$. We can then write the utility of a buying type θ consumer from the continuous activity as $u_\theta^B(f_\theta)$. Note that $u_\theta^B(f_\theta)$ does not depend on r or δ independently of f_θ. Let u_θ^{NB} denote the utility of a nonbuying type θ consumer from the continuous activity. The consumer is indifferent to buying the discrete service if $V + u_\theta^B(f_\theta) = u_\theta^{NB}$.

Assume that $u_\theta^B(f_\theta)$ is invertible, and let $\phi_\theta(\xi)$ denote its inverse. Indifference requires $f_\theta = \phi_\theta(u_\theta^{NB} - V)$, and the consumer strictly prefers buying when $\phi_\theta(u_\theta^{NB} - V) - r - \theta\delta > 0$. If we let $v(\theta) = \phi_\theta(u_\theta^{NB} - V)$, we recover the standard model. Note that in our setting, consumers with higher types do not per se value the discrete service more. All types would get the same bump in utility from the discrete service if it could be had for free and with no time commitment. In practice, however, different types have different utilities when not buying and make different adjustments to their leisure time and spending when they do buy. Hence, the utility from buying the discrete service varies with the consumer's type.

3.2. When the Full Price is Not Enough: Shift Pay Compensation Structure

In contrast to the gig economy, in which freelancers can make granular, minute-by-minute decisions about whether they want to work or play, office or factory workers generally have to commit to a threshold number of hours to get paid. See Fisman and Luca (2017). Additionally, they cannot unilaterally choose to work additional hours for additional pay, nor can they work fewer hours at lower compensation. Their compensation structure is thus a step function. We call them the *shift workers* hereafter. Denote S as the discontinuity point under this "shift pay" structure. The compensation function can be formulated as

$$\Pi_\theta(t_w) = \theta S \mathbb{I}_{\{t_w \geq S\}}, \tag{2}$$

where θ is the wage rate per time unit and S is the required shift length. Note that, under (P_C), it is never optimal for shift workers to work for $t_w \in (0, S) \cup (S, \infty)$. It only makes sense for the consumer working under a shift structure to choose either working for S hours to earn θS units of money or not working at all and earning nothing. Intuitively, owing to the discontinuity, consumers participating in the workplace might not always want to work exactly S hours. Either they might be coerced into working more than they would ideally want (i.e., they are overemployed), or they are willing to work longer hours but have no additional opportunities (i.e., they are underemployed).

Lemma 2 follows from the first-order conditions of (P_C) to solve for shift workers' optimal actions.

Lemma 2. *Under the shift pay compensation structure defined in (2),* $t_w = S$, *the consumer optimally chooses the overhead expenditure. In particular,*

1. *the optimal overhead expenditure for a consumer who does purchase the discrete service is found from* $O'(c_O^B) = -1/\chi(T - S - O(c_O^B) - \delta, \theta S - c_O^B - r)$, *and*
2. *the optimal overhead expenditure for a consumer who does not purchase the discrete service is found from* $O'(c_O^{NB}) = -1/\chi(T - S - O(c_O^{NB}), \theta S - c_O^{NB})$.

Unlike when he or she fully controls his or her working hours, the consumer's marginal rate of substitution of time for money in this case is not constant; that is, the iso-utility curves are not linear. Therefore, he or she is not a full-price decision maker. Instead, the slope of the iso-utility curve is a function of the optimal overhead expenditure, c_O, which in turn depends on whether the consumer is a buyer. Furthermore, for buyers, overhead expenditure also depends on the out-of-pocket price r and the time commitment δ. Consequently, full price is no longer a sufficient statistic to drive the consumer's purchase decision. Note that Proposition 1 and Lemma 2 still hold if the consumer has an initial endowment of wealth. In particular, a linear compensation structure leads to a consumer being a full-price decision maker, whereas a shift pay compensation structure does not. Hereafter, we denote τ_c as the consumer's compensation structure, where $\tau_c \in \{L, S\}$ denotes the linear compensation structure and the shift pay compensation structure with a shift length S, respectively.

We illustrate the differences between the linearly paid consumers and the shift-pay consumers in the marginal rate of substitution of time for money, which indicates how a consumer values his or her time, in Figure 1 (left), "Consumer's value of time." We see that for linearly paid consumers, $\chi_L(t_l^j, c_l^j)$ is always equal to θ, regardless of whether they buy a discrete service, as shown in Proposition 1. Note that, the x-coordinates of blue "+" and "•" mark the consumer types at which the linearly paid consumers start purchasing discrete services with price $r = 60$ or $r = 80$, respectively. By contrast, for shift-pay consumers, $\chi_S(t_l^j, c_l^j)$ are the same until they buy the service (left of black "+" or "•"). Once the consumer type θ becomes high enough for shift-pay consumers to start buying the discrete service (right of black "+" or "•"), their value of time depends on the discrete service price r as well as a time commitment δ. The constant slope of $\chi_L(t_l^j, c_l^j)$ confirms that gig workers are full-price decision makers, but shift workers are not. Given the same consumer type, the value of time for a shift worker can lie above or below the corresponding value for a gig worker.[2]

3.3. The Market for Services Depends on Consumers' Compensation Structures

We further elaborate on how consumers paid under different compensation structures evaluate discrete services differently in this section.

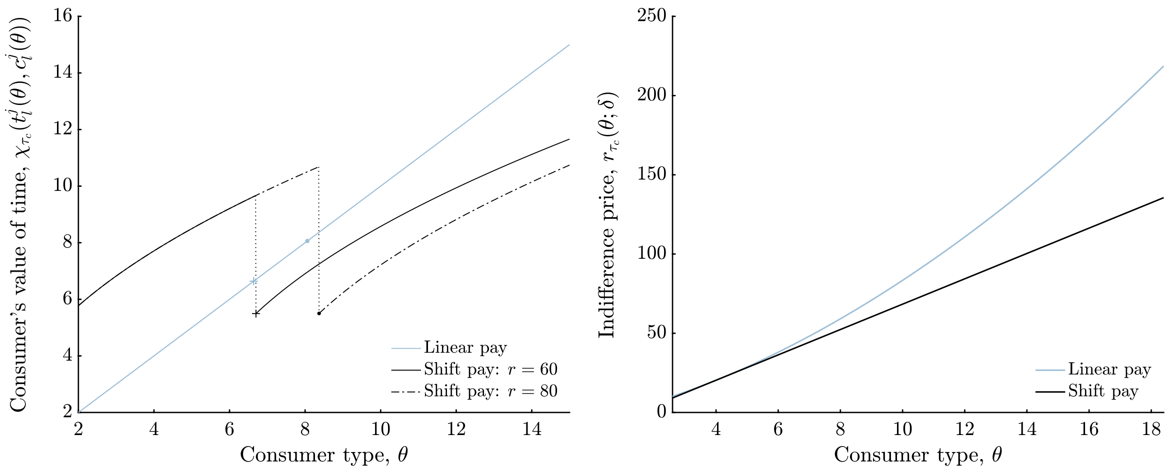
Consumer's Willingness to Pay for the Discrete Service. The different marginal rate of substitution of time for money suggests that given a time commitment, consumers' willingness to pay for the discrete service would also differ under the two compensation structures. Although shift workers have less flexibility in time compared with gig workers, they will also have either excess money or time, depending on their types. Hence, the comparison of who would be willing to pay more for the same service is not immediately clear. We start by quantifying consumers' willingness to pay.

Definition 2 (Indifference Price). Given time commitment δ, reward V, and the consumer's type θ, we define the indifference price to be the discrete service price that makes the consumer's optimal utility from buying and not buying the discrete service equal, that is, the r_{τ_c} that solves

$$u^{NB}_{\theta,\tau_c} = u^{B}_{\theta,\tau_c}(r_{\tau_c}, \delta) + V. \tag{3}$$

It is straightforward to show that $u^{B}_{\theta,\tau_c}(r_{\tau_c}, \delta)$ is strictly decreasing in r_{τ_c}. Hence, the indifference price is unique.

Figure 1. (Color online) Numerical Demonstration of the Consumer's Value of Time (Left), $\chi_{\tau_c}(t_l^j, c_l^j)$ and the Indifference Price $r_{\tau_c}(\theta;\delta)$ of Consumers Paid Under Linear and Shift Pay Compensation Structures



Notes. To construct the panel on the left, we fix S and δ, vary consumer type θ, and vary the discrete service monetary price in $r \in \{60, 80\}$. The "+" ["•"] symbols mark the consumer types θs at which the linear-pay and shift-pay consumers start purchasing the discrete service at price r = 60 [r = 80], respectively. To construct the panel on the right, we fix S and δ and plot the indifference prices varying across consumer type θ under linear- and shift-pay compensation structures.

A type θ consumer paid under compensation structure τ_c is the *indifferent consumer* at price r_{τ_c}. He or she is indifferent between buying or not buying the discrete service.

A consumer's willingness to pay is revealed by his or her indifference price. Recall from Remark 1 that, under the standard model, full-price consumers are indifferent in purchasing the product at $r = \phi_\theta(u_\theta^{NB} - V) - \theta\delta$, a function that depends only on V, δ, and θ. It is then interesting to see how indifference prices for the same service—same time commitment and same reward—differ between two consumers of the same type θ who are paid under different compensation structures, that is, $r_L(\theta;\delta)$ and $r_S(\theta;\delta)$. We suppress $r(\cdot)$'s dependence on V, because we hold V fixed. The analysis requires some structure. Denote θ_0 as the type at which the linearly paid indifferent consumer's optimal work time when not buying the service equals the shift length S. Denote θ_0' as the type at which the linearly paid indifferent consumer's optimal work time when buying the service equals the shift length S. $t^j_{w,\tau_c}(\theta)$ denotes the work time of a type θ consumer being paid under τ_c and making decision j. $u^j_{\theta,\tau_c,c}$ denotes the partial derivative of utility w.r.t. the leisure consumption c_l, of a type θ consumer being paid under τ_c and making decision j, evaluated at optimality. In addition, this partial derivative for buyers is evaluated at the consumer's indifference price, $r_{\tau_c}(\theta;\delta)$. We state the following.

Proposition 2.

1. *The indifference price for gig workers is higher than for shift workers in the neighborhood of* θ_0, *that is,* $r_L(\theta;\delta) \geq r_S(\theta;\delta)$ *for all* $\theta \in (\theta_0 - \epsilon, \theta_0 + \epsilon)$, *and for some* $\epsilon \geq 0$.
2. *The indifference price for gig workers is higher than for shift workers, that is,* $r_L(\theta;\delta) \geq r_S(\theta;\delta)$, $\forall\theta$, *under certain conditions*.

We relegate the conditions for Proposition 2, part 2, to hold to Appendix A.3.

Corollary 1. *Suppose the conditions in Proposition 2 are satisfied; then,* $r_L(\theta;\delta)$ *and* $r_S(\theta;\delta)$ *increase in* θ, *for all* $\theta > \min(\theta_0', \theta_0)$.

We illustrate Proposition 2 using a numerical study; see Figure 1, "Indifference price." Note that as one moves from low types to high types, the indifference prices first converge as θ approaches θ_0' and then diverge. At lower types, shift workers are overemployed and therefore are short on time. The additional pay is insufficient to compensate for the loss of time. Lower-type shift workers are thus unwilling to spend as much on discrete service as gig workers. At higher types, shift workers are underemployed and therefore short on cash. Although they have more free time, it cannot make up for the loss of cash compared with what gig workers earn. Higher-type shift workers are thus also unwilling to spend as much on the discrete service as the gig workers.

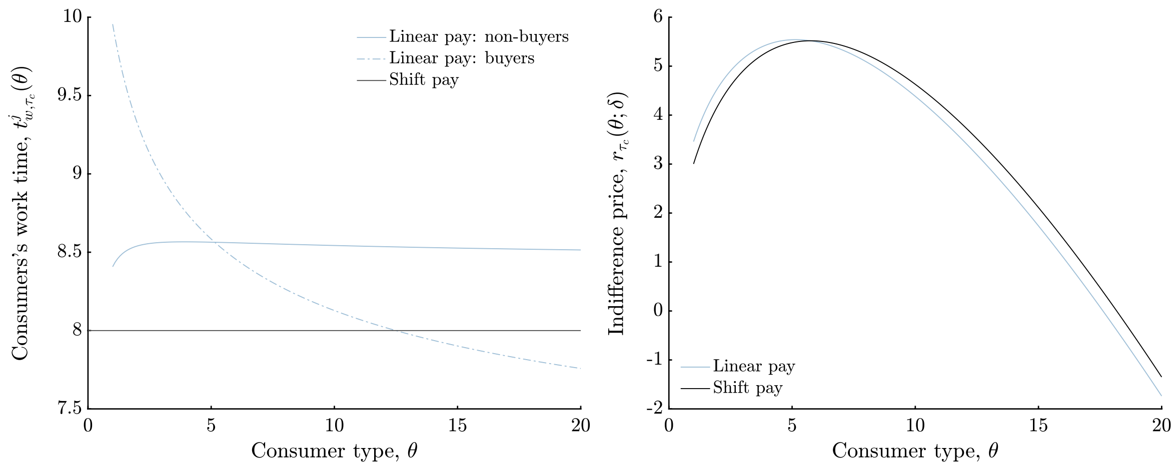
Remark 2. As noted in Remark 1, a standard model in the queue-management literature assumes that a type θ consumer's utility can be written as $v(\theta) - r - \theta\delta$ for some increasing function $v(\theta)$. More specifically, one needs the assumption that $v(\theta)$ increases "fast enough," that is, that $v(\theta)/\theta$ is increasing. When this holds, utility is increasing in the consumer's type if the utility is positive. That in turn implies that if a type θ consumer buys, so does any consumer of a higher type. Consequently, the indifference price $r(\theta;\delta)$ must be increasing in θ.

As Proposition 2 and Corollary 1 demonstrate, it takes some substantial conditions to assure this result. When the conditions of the proposition fail, one may have the indifference prices for both types of workers decrease or the indifference price of the shift worker exceeding that of the gig worker. This is illustrated in Figure 2, which assumes a Cobb-Douglas utility function, that is, $U(t_l, c_l) = Mt_l^{1-e_c}c_l^{e_c}$, for $1 > e_c > 0$.[3] As e_c increases, leisure consumption offers a higher return in boosting the consumer's utility; at lower values of e_c, leisure time more effectively lifts utility. Hence, $t^j_{w,L}(\theta)$ can be strictly increasing, strictly decreasing, or nonmonotone in θ. Indeed, as shown in Figure 2 (left), its behavior can depend on whether the consumer buys or does not buy the discrete service. Conditions of Proposition 2, part (ii), in A.3, are therefore violated.

The right-hand panel of Figure 2 shows the corresponding indifference prices for both linearly paid and shift-pay workers. For low θs, both prices are increasing, and the linear price is higher. However, at higher values of θ, both prices are decreasing, and shift workers are willing to pay a higher price for the service. Note that in this setting one can no longer simply assume that if a type θ consumer buys, everyone with a higher type must also buy. The set of buyers will be an interval, one defined by both an upper and a lower bound.

Finally, a distinction between the examples shown in Figures 1 and 2 is worth highlighting. The utility function used in Figure 1 has a finite limit, whereas the Cobb-Douglas utility in Figure 2 is unlimited. The finite limit favors having $v(\theta)$ increase sufficiently quickly. As θ climbs, a consumer has ample time and money even if they buy the discrete service, and those resources still provide a high utility because the marginal return on either tends to zero. In the Cobb-Douglas example, the marginal return on leisure time and spending remains high as θ increases. This diminution of marginal returns explains why Figure 1 shows indifference prices in the tens or hundreds of dollars when the value of the service is only 0.04 or

Figure 2. (Color online) An Example Where the Conditions of Proposition 2, Part 2, are Not Met



Notes. The panel on the left depicts the optimal work time for a linearly paid consumer when they buy (at the indifference price) or do not buy, as well as a shift length $S = 8$. The panel on the right displays the indifference prices for both linearly paid and shift-pay consumers varying consumer type θ.

0.001. Consumers are not paying more than the service is worth. Rather, they gain so little additional utility from other continuous services that they can afford to spend heavily on the discrete service.

4. Pricing the Discrete Service

We now turn to the pricing problem faced by a revenue-maximizing seller of the discrete service. We consider two types of common discrete services. First, we consider a classic queue setting for services. Specifically, we suppose homogeneous consumers who are served by an M/M/1 queue. Within this framework, we consider both a hidden and a visible queue. Second, we consider services with nonbinding capacity and a fixed time commitment being offered to a group of heterogeneous consumers, for instance, a boat tour. In both settings, we examine how consumer compensation structures affect the seller's pricing decision.

4.1. Managing a Single-Server Markovian Queue Serving Homogeneous Consumers

We first consider a classical operations model of a service provider pricing entry to a single-server Markovian queue. Edelson and Hildebrand (1975) assumed that the consumers could not see the current state of the queue but correctly anticipated the expected delay. Naor (1969) assumed that consumers saw the current state and thus based their decision on their conditional expected delay. The hidden queue clearly fits within our framework because buyers commit to purchasing based on a common expected delay. A visible queue is a shift from our planning model of decision making because actual purchasing is state dependent.

Depending on when the consumers make a purchase decision, we consider two cases under this setting. First, consumers plan their day and make their purchase decision before they arrive to the queue. Assume they always commit to their purchase decisions. Second, consumers evaluate the discrete service after they see the state of the queue. In this case, we assume that the consumers can make small adjustments in time allocation during the middle of the day but still achieve the same optimal utility as if they have planned the purchase ahead.

4.1.1. Pricing a Hidden Queue. Suppose the seller manages a single-server Markovian queue. Given the service price, consumers with homogeneous type θ have to make an irrevocable decision of whether to join the queue before they see the state of the queue. An example would be ordering and prepaying for a coffee through a mobile app for pickup when the coffee shop does not provide real-time delay information. Suppose the consumer always commits to his or her decision.

Consumers make a decision based on their expectations of the queue length. Denote the average service time for each consumer in the discrete service to be s and the consumer's arrival rate to be λ. The expected sojourn time in the system is then $\overline{\delta}(\lambda) = s/(1-\lambda s)$. Consumers make their purchase decisions without seeing the current state of the queue. Consequently, all type θ consumers will purchase the service at price r if $r \le r(\overline{\delta}(\lambda);\theta)$. As r increases, consumers employ a mixed strategy; they will purchase the service with probability β, not purchase with probability $1-\beta$. Consumers' arrival rate to the system is then $\beta\lambda$. The consumers purchase the service at a price $r = r_{\tau_c}(\overline{\delta}(\beta\lambda);\theta)$, where $r_{\tau_c}(\delta;\theta)$ denotes the indifference price function of the wait time δ. Inverting the function, we can find

$\beta(r)$ for $r \in \mathbf{R}^+$. The seller's problem is

$$\max_r R_{\tau_c} = \max_r r\lambda\beta_{\tau_c}(r;\theta). \tag{4}$$

The probability that a given consumer buys at optimality is then $\beta_{\tau_c}(r^*_{\tau_c})$. We show that the revenue function has a unique maximizer for both gig and shift workers under some conditions (See Appendix A.5).

We compare the *optimal admission threshold* $\beta^*_{\tau_c} := \beta_{\tau_c}(r^*_{\tau_c})$ and the maximum revenue, $R^*_{\tau_c}$, under the two compensation structures. That is, we contrast a market in which everyone has type θ and faces a linear compensation structure with one in which everyone has type θ but must work a shift of S time units. Define θ_e as the type at which an indifferent gig worker would choose to work S time units when he or she buys. We have the following propositions.

Proposition 3. *Suppose $c^B_{O,S}(\delta) > c_{O,L}(\delta)$ for all δ when $\theta \le \theta_e$, and $c^B_{O,S}(\delta) < c_{O,L}(\delta)$ for all δ when $\theta > \theta_e$. In contrasting a market of gig workers who buy and shift workers, we have*

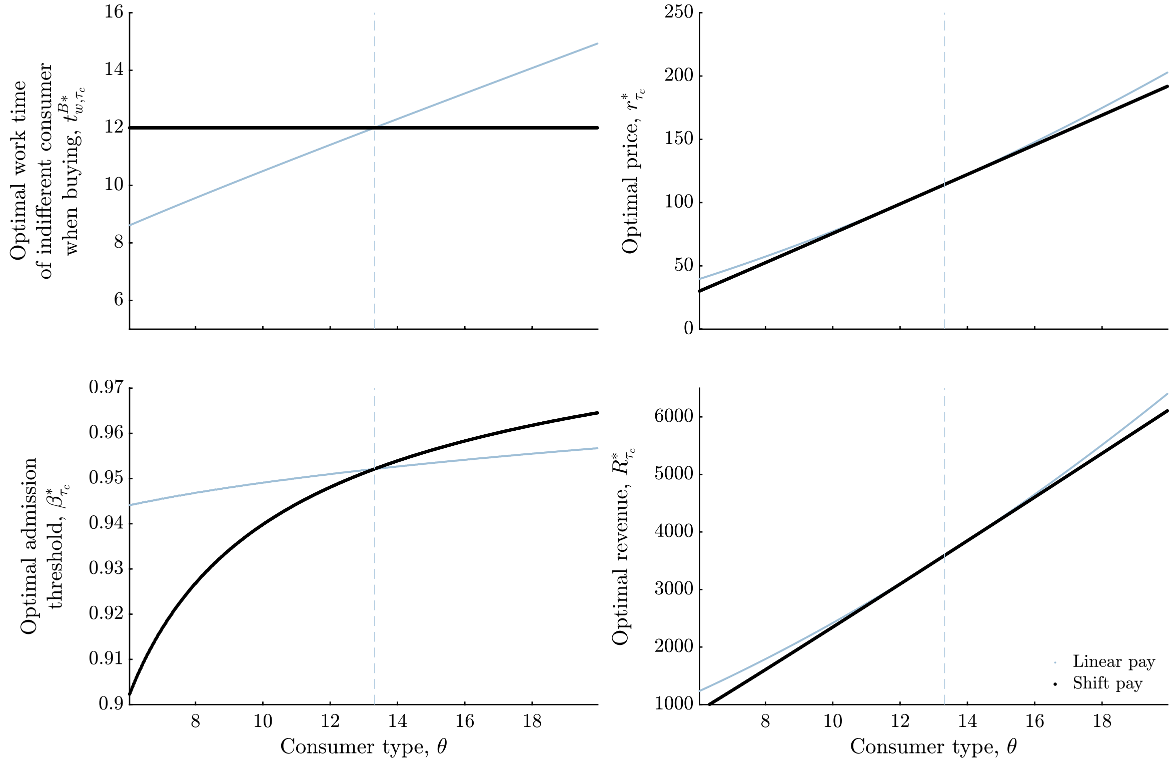
(i) *When $\theta \le \theta_e$, that is, when shift workers are overemployed (compared with gig workers who buy), the seller induces a smaller proportion of shift workers to join the system, that is, $\beta^*_L \ge \beta^*_S$.*

(ii) *When $\theta > \theta_e$, that is, when shift workers are underemployed (compared with gig workers who buy), the seller induces a larger proportion of shift workers to join the system, that is, $\beta^*_L < \beta^*_S$.*

Proposition 4. *Suppose all conditions in Proposition 2 are satisfied. The optimal revenue that a seller managing a single-server Markovian hidden queue can collect from serving gig workers is higher than what they can collect from serving shift workers. That is, for all θ, $R^*_L \ge R^*_S$.*

We illustrate Propositions 3 and 4 in Figure 3. When $t^{B*}_{w,L} > S$ (see Figure 3, top left), that is, on the right-hand side of the figure, the shift workers are underemployed, and the seller's optimal admission threshold (see Figure 3, bottom left) when managing shift workers is larger than when managing gig workers (i.e., $\beta^*_S > \beta^*_L$); the seller serves a more congested system, and vice versa. However, regardless of whether shift workers form a more congested or less congested system, the optimal revenue that the seller can collect from shift workers is always smaller than what the seller can collect from gig workers (see Figure 3, bottom right). We

Figure 3. (Color online) Pricing Problem of a Seller Managing a Single-Server Markovian Hidden Queue and Selling to Consumers with Linear or Shift-Pay Compensation Structures



Notes. We plot (1) the optimal work time of the indifferent consumer when he or she buys the service (top left), (2) the optimal admission threshold of the seller (bottom left), (3) the optimal discrete service monetary price to charge (top right), and (4) the optimal revenue from the service (bottom right). The dashed vertical lines indicate the consumer type θ_e.

defer the proofs of the two propositions to the next subsection.

4.1.2. Pricing a Visible Queue. We have thus far assumed no dynamics in consumers' decision-making process. That is, the consumers always plan ahead and stick to their decisions about whether to join a discrete service. However, our framework is also applicable to the scenario when consumers decide whether to join a service, which requires a relatively short amount of time compared with other planned activities, after they see the state of the system (for example, whether to join the queue at a food truck). That is, we allow dynamics in deciding whether to join short services in this section. Note that the service should be short enough that the consumer can make small adjustments in his or her allocation of time and money in the middle of the day but still achieve the same optimal utility as if he or she had planned ahead.

Consider the classic service system modeled as a single-server Markovian queue in Naor (1969). The queue is visible, and arriving consumers consequently trade off the value of service, its posted price, and their expected waits based on the number of consumers in line.

We again assume that the consumers arriving to the system are homogeneous with type θ. For a given out-of-pocket price, there is, then, a cutoff value n, and a consumer joins the queue only if the number of consumers in the system is below n upon his or her arrival. Denote the average service time for each consumer in the discrete service to be s. The nth consumer, that is, the indifferent consumer, would expect a total time commitment of $\delta = sn$. We show the monotonicity of the indifference price in time commitment in Appendix A.6 Lemma 5, which in turn guarantees a unique indifference price, $r_L(sn_L;\theta)$, and $r_S(sn_S;\theta,S)$ corresponding to the admission threshold n for gig workers and shift workers, respectively. Any arriving consumer seeing n consumers would find the price higher than their indifference price and balk.

Suppose also that all consumers have type θ and the shift workers all commit to a shift length of S hours. It is thus easy to see that a revenue-maximizing seller managing a single-server Markovian visible queue solves the following problem,

$$\max_n R_{\tau_c} = \max_n r_{\tau_c}(sn)\cdot\lambda\frac{1-\rho^n}{1-\rho^{n+1}},\ \forall n\in\mathbb{Z}^+, \tag{5}$$

for the optimal number of consumers to admit. The uniqueness of the solution follows from the strict concavity of the maximization problem (see Appendix A.7). Although there can be multiple discrete prices inducing the same admission threshold, there is a unique optimal threshold. We call $n^*_{\tau_c}$ the *optimal admission threshold*.

As in Propositions 3 and 4, the following propositions contrast markets with homogeneous type θ consumers that differ only in their compensation structures. Proofs are relegated to Appendices A.8 and A.9, which also include proofs of extensions for the two propositions for the previous hidden queue case.

Proposition 5. *Suppose $c^B_{O,S}(sn) > c_{O,L}(sn)$ for all n when $\theta \le \theta_e$, and $c^B_{O,S}(sn) < c_{O,L}(sn)$ for all n when $\theta > \theta_e$. In contrasting a market of gig workers who buy and shift workers, we have the following:*

i. *When $\theta \le \theta_e$, that is, when shift workers are overemployed (compared with gig workers who buy), the seller chooses a lower threshold, that is, $n^*_L \ge n^*_S$.*

ii. *When $\theta > \theta_e$, that is, when shift workers are underemployed (compared with gig workers who buy), the seller chooses a higher threshold, that is, $n^*_L < n^*_S$.*

Proposition 6. *Suppose all conditions in Proposition 2 are satisfied. The optimal revenue that a seller managing a single-server Markovian visible queue can collect from serving gig workers is higher than what they can collect from serving shift workers. That is, for all θ, $R^*_L \ge R^*_S$.*

We illustrate Propositions 5 and 6 in Figure 4 (left).[4] Suppose the prevailing wage among consumers, both gig and shift workers, is \$18 per hour. We see that shift workers can form a queue that is 11.1% more congested than gig workers can, which translates to a system with a 0.5% higher utilization rate when serving shift workers. Revenue, however, is 2.7% lower.

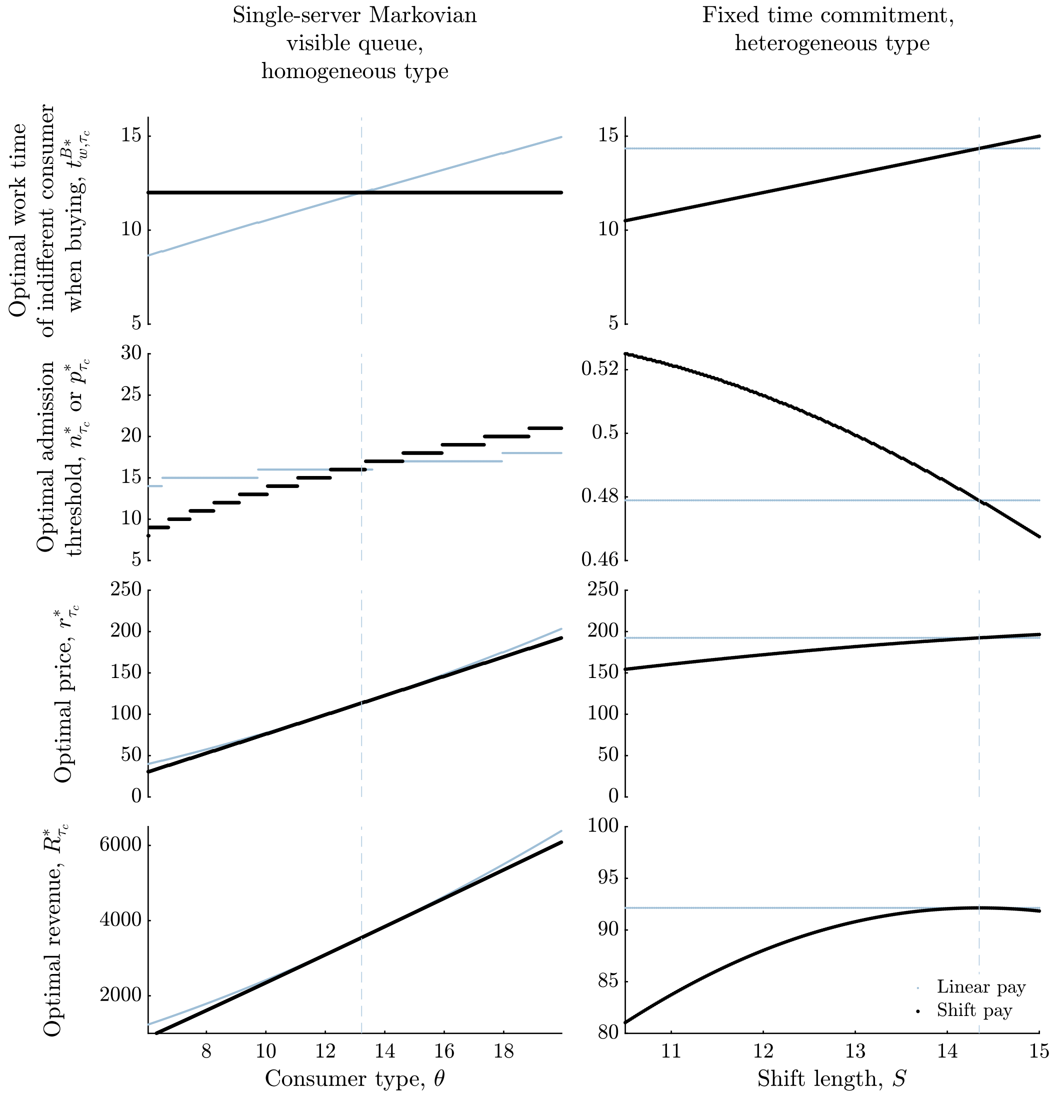
Propositions 5 and 6 provide an intuition similar to Propositions 3 and 4. Overemployed shift workers have ample money but are time constrained outside of working (compared with the gig workers with the same type). It is then optimal to restrict sales, which results in a less congested system (i.e., small β^*_S, small n^*_S). Underemployed shift workers have tight finances but have excess time outside of working (compared with the gig workers with the same type), which results in a more congested system (i.e., large β^*_S, large n^*_S).

4.2. Selling to Heterogeneous Consumers with a Fixed Time Commitment

Another type of service that allows the consumers to plan ahead is a service that requires a fixed time commitment independent of the other consumers' purchase decisions. Suppose the seller offers such a discrete service to a group of consumers who are heterogeneous in type, say $\theta \sim F$. One example can be a boat tour. The number of consumers upon arrival has a negligible effect on an individual consumer's time commitment in such services.

The seller trades off the price per consumer against the proportion of the total population to serve. The monotonicity of the indifference price in type from Corollary 1 indicates that for a fixed time commitment δ, there is a unique indifference price $r_{\tau_c}(\theta;\delta)$ for a type θ

Figure 4. (Color online) Pricing Problems of Two Types of Sellers Selling to Consumers with Linear or Shift-Pay Compensation Structures



Notes. We plot four metrics in four subfigures for each seller type as in Figure 3. Subfigures on the second row demonstrate Propositions 5 and 7. Subfigures in the fourth row demonstrate Propositions 6 and 8.

consumer. Furthermore, consumers buy the service at $r_{\tau_c}(\tilde{\theta};\delta)$ if and only if they have a higher type, that is, $\theta \geq \tilde{\theta}$ (see A.10). This suggests that a threshold value $\tilde{\theta}$ is sufficient to describe the consumer population that buys the service at a price $r_{\tau_c}(\tilde{\theta};\delta)$. We call the consumer of type $\tilde{\theta}$ the threshold, indifferent consumer. Denote $\overline{F}(\theta) = 1 - F(\theta)$ as the tail distribution of consumer types. Fix the market size at one and assume that capacity does not bind. We can formally state the seller's problem as follows:

$$\max_{\theta} R_{\tau_c} = \max_{\theta} r_{\tau_c}(\theta;\delta) \cdot \overline{F}(\theta). \tag{6}$$

The above revenue maximization problem's solution set has characteristics as shown in Lemma 3, and the proof can be found in Appendix A.11.

Lemma 3. *Suppose the consumer's indifference price $r_{\tau_c}(\theta;\delta)$ is log-concave in θ, and the consumer type distribution $F(\theta)$ has an increasing failure rate; that is, $f(\theta)/\overline{F}(\theta)$ increases in θ. We conclude the following*:

i. *The revenue maximization problem (6) has a unique optimum;*

ii. *The unique optimum when serving the shift-pay consumers, θ^*_S, increases in shift length S if for all $\theta < \theta'$: 1) $\frac{\partial r(\theta,S)}{\partial S} < \frac{\partial r(\theta',S)}{\partial S}$ and 2) $\frac{\partial r(\theta,S)/\partial S}{r(\theta,S)} < \frac{\partial r(\theta',S)/\partial S}{r(\theta',S)}$ whenever $0 < \frac{\partial r(\theta,S)}{\partial S}$.*

Next, consider the seller's problem (6) when its consumers are paid under linear compensation structures. First, denote $t^B_{w,L}(\theta^*_L, r_L(\theta^*_L))$ as the work time of the indifferent linearly paid consumer at optimality (i.e., type θ_L^* consumer) if they buy. Denote $S_{e_1} = t_{w,L}^B(\theta_L^*, r_L(\theta_L^*))$. Furthermore, there exists S_{e_2} such that $\theta_{S_{e_2}}^* = \theta_L^*$. Note that S_{e_1} and S_{e_2} are not necessarily equal. Denote the *optimal admission threshold* as $p_{\tau_c}^* := \overline{F}(\theta_{\tau_c}^*)$.

We can now compare optimal thresholds when the seller serves gig workers and shift workers with identical shift length S. Proposition 7 below immediately follows from Lemma 3.

Proposition 7. *Suppose the shift-pay consumers have heterogeneous types and an identical shift length S.*

i. *When $S \geq \max\{S_{e_1}, S_{e_2}\}$, that is, the threshold shift-pay consumer is overemployed (compared with gig workers who buy), $\theta_S^* \geq \theta_L^*$, and therefore, $p_L^* \geq p_S^*$.*

ii. *When $S \leq \min\{S_{e_1}, S_{e_2}\}$, that is, the threshold shift-pay consumer is underemployed (compared with gig workers who buy), $\theta_S^* \leq \theta_L^*$, and therefore, $p_L^* \leq p_S^*$.*

We further contrast the seller's revenue from serving a market of gig workers with a market of shift workers with the same type of distribution.

Proposition 8. *Suppose all conditions in Proposition 2 are satisfied. The optimal revenue that a seller managing a fixed time commitment service can collect from serving gig workers is higher than what the seller can collect from serving shift workers. That is, for all S, $R_L^* \geq R_S^*$.*

We illustrate Propositions 7 and 8 in Figure 4 (right).[5] We observe that a seller charges underemployed consumers less because they are earnings constrained. This is partially offset by selling to a larger proportion of consumers than in the gig economy case. Conversely, a seller charges overemployed consumers more; they have abundant cash but less time. A smaller proportion of consumers join the service than in the gig economy.

Propositions 7 and 8 provide an intuition similar to the results of Section 4.1. Overemployed shift workers have ample money but are time constrained. It is then optimal to restrict sales, which results in admitting a smaller proportion of the consumer population into the service, that is, resulting in a less congested system. Underemployed shift workers have tight finances but excess time, which results in admitting a larger proportion of the consumer population into the service, that is, resulting in a more congested system. Furthermore, in the fixed time commitment example, the seller cuts the price for the underemployed shift workers and recoups some of the price cut through higher volume in the fixed time commitment. However, neither case can make the seller whole relative to selling to gig workers.

Note that analysis of the case when the fixed-time-commitment discrete service is offered to a group of *homogeneous* consumers is trivial. That is, given r, either all workers join the service or none of them do.

Remark 3. To highlight the impact of consumers' compensation structures on sellers' decisions, we focus on selling to consumers with a *single* type of compensation structure thus far. In fact, our main insights on the suboptimality of pricing when all consumers are assumed to be full-price decision makers and the contrast in congestion levels can be extended to the case when consumers have *mixed* compensation structures, that is, when some consumers are gig workers whereas others are shift workers. Assuming the proportion of gig workers among consumers is $\alpha = 0.2$ (therefore, $1 - \alpha = 0.8$ shift workers), we conduct numerical analyses to find the revenue-maximizing pricing decisions for all three types of sellers that we studied above. We denote $\tau_c = mix$ as the mix-compensation-structure case.

First of all, not surprisingly, in general, $r_{mix}^* \neq r_L^*$. Therefore, pricing a system as if all consumers are gig workers leads to suboptimality. Second, comparing congestion levels formed by gig workers to congestion levels formed by a mixture of workers (see Figure 5), we see the same change in rank. When shift workers in the system are overemployed (i.e., small θ or large S), a seller managing mixed workers admits fewer consumers than when managing only gig workers. When shift workers are underemployed, (i.e., large θ or small S), a seller serving mixed workers admits more consumers than when serving only gig workers. We relegate the detailed analyses to Appendix B.

5. Extensions and Managerial Implications

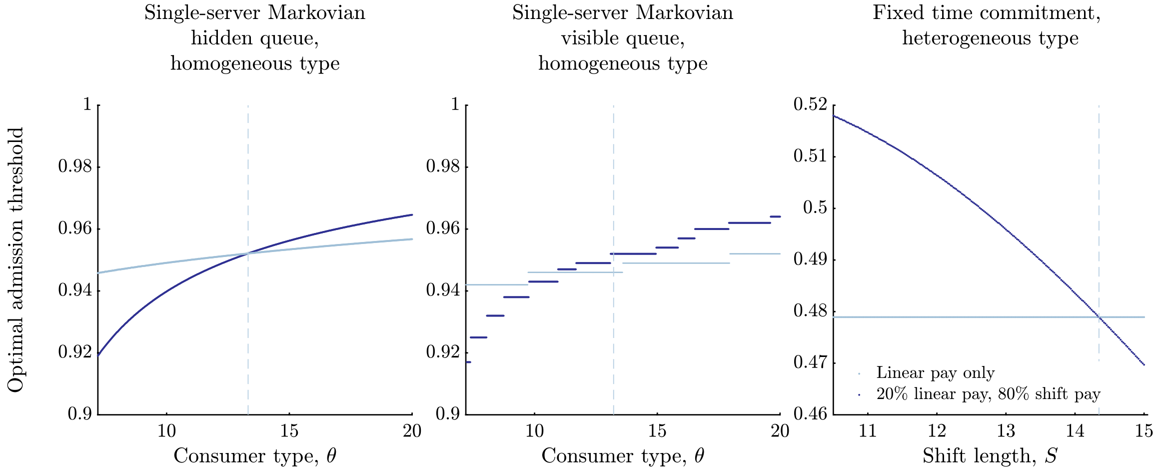
We now explore some variations on our basic model to derive additional insights. First, we consider having more than one discrete service. Next, we examine the implications of a seller mispricing their service because they do not fully understand how much control their consumers have over their work schedules.

5.1. Multiple Discrete Services

In the base model, we consider one discrete service but have now extended to suppose that there are multiple service providers to choose from. Therefore, we extend the model to contain two discrete services in this section. Although we focus on two services, the arguments can easily be extended to more than two discrete services. We include the model formulation and detailed discussion in Appendix C.

Suppose there are two discrete services, which require the same time commitment, $\delta_1 = \delta_2 = \delta$, reward the same utility boost, $V_1 = V_2 = V$, and have the same monetary price, $r_1 = r_2 = r$. Obviously, if a consumer were indifferent to buying Service 1, he or she should also be indifferent to buying Service 2. But does that imply that the consumer should be indifferent to buying both services? The conventional mode of modeling

Figure 5. (Color online) Plot of the Congestion Levels Observed By the Three Seller Types When They Serve Only Gig Workers and a Mixture of Gig and Shift Workers, Respectively



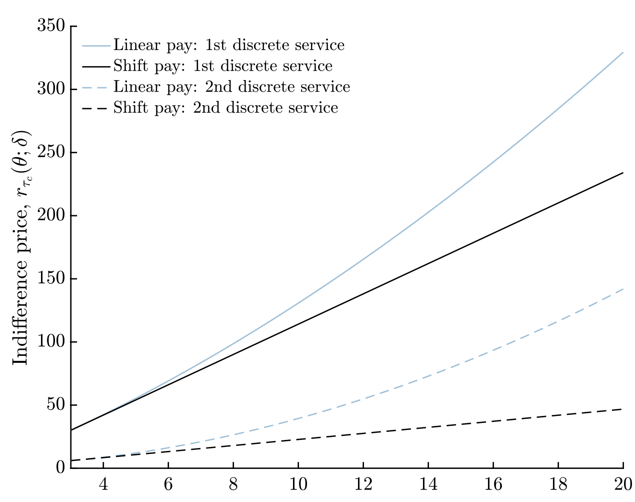
Notes. From left to right, the y-axis plots β^*, $1 - b^*$, and p^*. b^* is consumer's balking rate at optimality. More discussions on the mixed compensation structures setting are included in Appendix B.

full-price decision makers (see Remarks 1 and 2) assumes that the consumer's utility function is fixed independent of his or her purchase history. That would suggest the consumer should purchase both services. However, our results indicate otherwise. For both gig and shift workers, purchase history matters. Once one has committed time and money to a discrete service, the attractiveness of giving up more resources to a second service falls. Assuming that the time commitment of the second service cannot be reduced, the price of the second service must be reduced to induce the consumer to purchase.

We illustrate this using the following numerical study. Suppose price r is set to be 80% of the indifference price $r_S(\delta)$. Offered one of the services at $(0.8r_S(\delta), \delta)$, both gig and shift workers would strictly prefer to purchase it. Without loss of generality, we assume that the consumer sees the advertising of Service 1 first and purchases the service. The consumer then encounters Service 2 and calculates his or her indifference price for the same (δ, V), having already committed δ time and r dollars to the first service. We compare the consumer's indifference price for the first (δ, V) service to the consumer's indifference price for the second (δ, V) service given that the first one is purchased at the price $0.8r_S(\delta)$ (see Figure 6). We see that, in this study, at a consumer type $\theta = 12$, the seller has to offer a 67% price cut (on $r_L(\delta)$) to gig workers and an 80% price cut (on $r_S(\delta)$) to shift workers to motivate the purchase of Service 2.

The conventional models discussed in Remarks 1 and 2 ignore that the attractiveness of a discrete service depends on the consumer's available time and money. Once a consumer commits to one service, their available resources change, and so does their evaluation of a seemingly identical service. Consequently, services that do not seem to directly compete with one another based on their price or type do end up in competition for a consumer's limited time and cash. Given that the consumers bought Service 1, a substantial price cut is needed for the consumer to consider also buying Service 2. We thereby provide another benefit for sellers to offer advance sales. Consumers have limited time and money budgets, so it is good to lock in their time and money in advance. The later the service is presented to the consumers, the lower the price they would

Figure 6. (Color online) Plot of the Gig Workers' and Shift Workers' Indifference Prices Under Two Scenarios



Notes. The solid curves plot the indifference price of the discrete service with $\delta = 0.01$, 1/1,200 of shift length S. The dashed curves plot the indifference price of the discrete service with $\delta = 0.01$ when they have purchased the service at 80% of the shift workers' indifference price.

be willing to spend. The relevancy of a consumer's purchase history extends to the case when there are multiple discrete services that are not symmetric. We would expect the purchase decision problem of a consumer to be intractable when the number of discrete services is large. It involves a combinatorics question. The value of the next service to purchase depends on the sequence of purchases to date.

5.2. Revenue Loss When Service is Mispriced

We have seen that how consumers are paid should matter to a seller. But what if a seller fails to recognize the constraints on consumers' time allocations and earnings? In particular, what if the seller assumes that all consumers are full-price decision makers, that is, gig workers, and prices its service accordingly, but in fact the consumers are all shift workers? What if the seller assumes that all consumers are shift workers but in fact the consumers are gig workers? In this section, we use the seller managing a single-server, Markovian, hidden queue (see Section 4.1.1) as an example and investigate how large a revenue gap the seller might incur.

Given a consumer type θ, there is an optimal price if the seller thinks that all consumers are gig workers, $r_L^*(\theta)$. Suppose this price is imposed on a group of shift workers with type θ. It will induce an expected time commitment $\overline{\delta}_S(r_L^*(\theta))$, which leads to an arrival rate $\lambda\beta_S(r_L^*(\theta))$. Specifically, given $r_L^*(\theta)$, we solve Equation (3) for δ, which is $\overline{\delta}_S(r_L^*(\theta))$. We compare the revenue under mispricing, $r_L^*(\theta)\lambda\beta_S(r_L^*(\theta))$, to the optimal revenue the seller could get from serving type θ shift workers, that is, R_S^*. Recall that, at type θ_e, the consumers are indifferent in working under linear or shift compensation structure when they buy. We investigate the revenue gap relative to consumer type's deviations from the compensation-structure-indifferent type θ_e. We denote the deviation $\Delta\theta \triangleq (\theta - \theta_e)/\theta_e$.

We conduct a numerical analysis, and Figure 7 (left) plots the percentage revenue loss because of mispricing against percentage deviations from type θ_e. As the consumer type deviates further from θ_e, the revenue gap because of mispricing expands. Note that, in practice, the consumer's type can be anywhere on the horizontal axis. In this numerical study, we suppose that the shift workers must work for 12 hours, and $\theta_e \approx \$13$. The seller's revenue loss can be up to 17% if the consumer's wage rate is \$9, that is, a minimum wage worker, whereas the revenue loss can be up to 9% if the consumer's wage rate is \$18, a retail salesperson.

Furthermore, the numerical study suggests that the revenue loss incurred when consumers are overemployed or underemployed is not symmetric. That is, given the same percentage deviation in type, the seller serving underemployed ($\theta > \theta_e$) shift workers starts with a larger percentage revenue gap than the seller serving overemployed shift workers. This inequality flips as the percentage deviation expands further. Recall that $\overline{\delta} = s/(1-\lambda s)$ is the expected sojourn time in an M/M/1 queue. Rearranging the terms, we have $\lambda = (1/s - 1/\overline{\delta})$. Revenue is then $f(r,\overline{\delta}) := r\lambda = r(1/s - 1/\overline{\delta})$. When the service is mispriced, the mispriced revenue is essentially $f(r(1+\epsilon), \overline{\delta}(1+\xi))$ for some ϵ and ξ. A crude approximation of the percentage revenue loss can then be written as $\frac{(1/s-1/\overline{\delta})*r\epsilon+(r/\overline{\delta}^2)*\overline{\delta}\xi}{r(1/s-1/\overline{\delta})} = \epsilon + \frac{\xi}{\overline{\delta}/s-1}$ by a

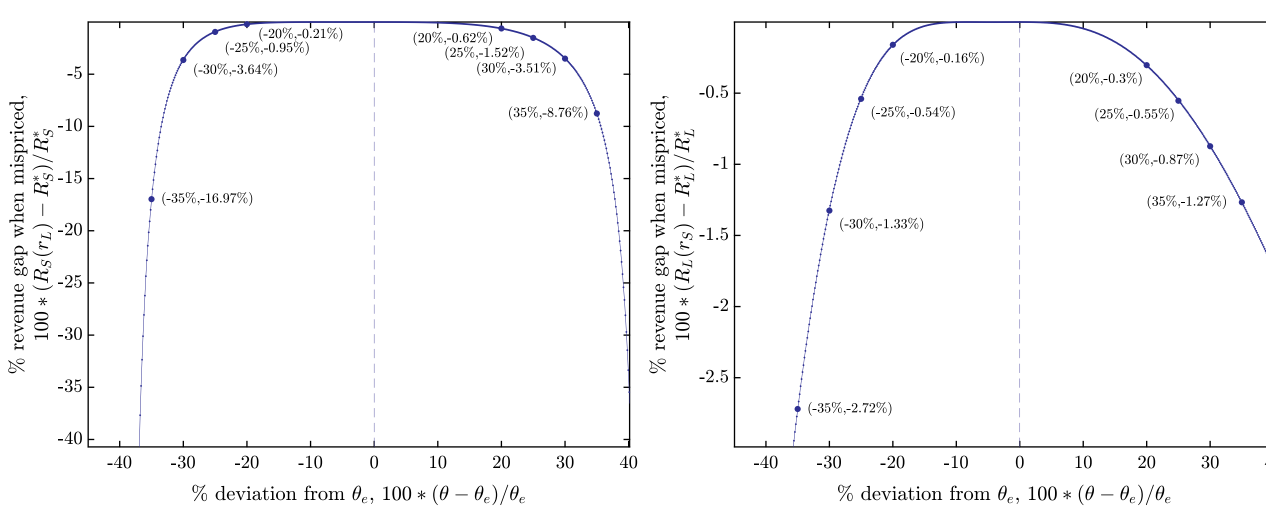


Figure 7. (Color online) Plot of the Percentage Revenue Gap Against the Percentage Deviations from the Compensation Structure Indifferent Type θ_e

Notes. From left to right, the percentage revenue gap resulting from mispricing the shift workers as gig workers and mispricing the gig workers as shift workers, respectively, are illustrated. To construct the two figures, we use the same parameters as in the numerical study in Section 4.1.1, Figure 3. In particular, shift workers are assumed to work a 12-hour shift in both numerical studies.

first-order Taylor expansion around $(r, \overline{\delta})$. When the shift workers encounter a suboptimal price ($\epsilon \triangleq r_L^*/r_S^* - 1 \neq 0$), they modify their overhead expenditure c_O to accommodate the suboptimal price, which further affects the overhead time $O(c_O)$. The impact on worker's utility that cannot be resolved by overhead expenditure is made up by updating $\overline{\delta}$ by ξ to keep up with the utility $u_{\theta,S}^{NB}(t_l, c_l)$. Therefore, we can intuitively see the asymmetry in revenue loss around θ_e as follows by considering how far θ deviates from θ_e.

First, in the close neighborhood of θ_e, that is, small $|\Delta\theta|$, there is a range that overemployed shift workers (i.e., $\theta < \theta_e$, left of the dashed line in Figure 7, left) are mispriced with a lower price (i.e., $r_L^* < r_S^*$), whereas the underemployed shift workers (i.e., $\theta > \theta_e$) are mispriced higher (i.e., $r_L^* > r_S^*$). For a similar $|\epsilon|$, the overemployed shift workers charged with a suboptimal lower price can increase, decrease, or keep their overhead expenditure constant to adjust to the lower price. By contrast, the underemployed shift workers cannot afford to further increase their overhead expenditure. Therefore, the percentage change in $\overline{\delta}$ is expected to be smaller for overemployed workers. Therefore, we see that the revenue loss is smaller for mispriced overemployed shift workers around θ_e.

Second, as $|\Delta\theta|$ expands, both overemployed and underemployed shift workers are overpriced (i.e., $r_L^* > r_S^*$), that is, $\epsilon > 0$ and $\xi < 0$. The difference in percentage revenue loss on both sides of θ_e depends on the comparison of ϵ and its induced ξ as well as $\overline{\delta}$. Note that ϵ starts with a smaller value for the overemployed shift workers. Additionally, they are rich in cash, so they can afford some reduction in leisure consumption c_l. Therefore, they make a smaller percentage of reduction in overhead expenditure c_O, that is, less impact on $O(c_O)$, and subsequently less impact on $\overline{\delta}$ compared with the underemployed shift workers who are already tight in finances. In this range, the percentage of revenue loss is still smaller for overemployed shift workers. However, as $|\Delta\theta|$ further expands, ϵ, the percentage price gap becomes larger for the overemployed shift workers, and the benefit of being rich in cash cannot make up for the large price gap, resulting in a larger ξ. Furthermore, the increasing difference in $\overline{\delta}$ at $\theta_e(1 \pm |\Delta\theta|)$ as $|\Delta\theta|$ increases amplifies the difference in $\xi(\overline{\delta}/s - 1)^{-1}$. Therefore, the inequality flips at some $|\Delta\theta|$, and we observe that the overemployed shift workers experience a larger revenue loss than the underemployed shift workers for large $|\Delta\theta|$.

The other direction of incorrect assumption—mistakenly assuming consumers to be shift workers—is not seen in the literature, but it is also interesting to study for practical purposes. What if the sellers assume that all consumers are shift workers with shift length S, but in fact the consumers are all gig workers? Following a similar argument as above, given θ and S, we compare the revenue under mispricing, $r_S^*(\theta)\lambda\beta_L(r_S^*(\theta))$, to the revenue under correct pricing, R_L^*, from solving (4). We use the same sets of parameters that construct Figure 7, left, to construct Figure 7, right.

We again see the expected revenue loss under model misspecification that in turn leads to mispricing. Moreover, the two numerical studies show that the revenue loss from mispricing gig workers as shift workers, that is, $(R_L(r_S) - R_L^*)/R_L^*$, is less pronounced than the revenue loss from mispricing shift workers as gig workers, that is, $(R_S(r_L) - R_S^*)/R_S^*$. Note that $R_L^* > R_S^*$, which could have resulted in $(R_S(r_L) - R_S^*)/R_S^*$ having a larger magnitude when the gaps in the numerators, that is, $(R_S(r_L) - R_S^*)$ and $(R_L(r_S) - R_L^*)$, are identical. But even if we consider only the gaps, $(R_S(r_L) - R_S^*)$ and $(R_L(r_S) - R_L^*)$, our numerical studies show that the former is larger in magnitude. This is expected because the mispriced gig workers have more flexibility in reallocating their time and budget to recover their theoretical optimal utility as much as possible when facing a suboptimal price than the shift workers. To be more specific, mispriced shift workers can only adjust their overhead expenditure, c_O, to recover the theoretical optimal utility. But mispriced gig workers can adjust both their work time, t_w, and overhead expenditure.

There is also asymmetry in revenue loss when consumers are overemployed or underemployed shift workers. In the close neighborhood of θ_e, $|\Delta\theta|$ is small and the denominator R_L^* is similar on both sides of θ_e. Therefore, the percentage revenue loss is dominated by the change in the numerator. The gig worker whose type is set to $\theta(1 + |\Delta\theta|)$, that is, on the right-hand side of θ_e, spends more in overhead expenditure than its counterpart on the left-hand side of θ_e, whose type is set to $\theta(1 - |\Delta\theta|)$. When mispriced with a suboptimal price, the gig worker adjusts his or her overhead expenditure c_O and work time t_w to recover the optimal utility. When c_O is large, that is, for the right-hand side gig worker, given the convexity we assumed for $O(c_O)$, the same changes in c_O would induce fewer changes in $O(c_O)$. Therefore, the right-hand side gig worker would have to induce a larger change in work time t_w and $\overline{\delta}$ so as to recover the optimal utility as much as possible. This leads to a larger magnitude of revenue loss, that is, the numerator of percent revenue loss, on the right-hand side of θ_e, which translates to a larger percent revenue loss than it is in the left-hand side of θ_e when $|\Delta\theta|$ is small. As $|\Delta\theta|$ gets larger, the magnitude of the denominator R_L^* dominates percent revenue loss. R_L^* increases in θ. Therefore, the inequality in percent revenue loss flips, and we see a larger revenue loss on the left-hand side of θ_e.

5.3. Implications on Consumer Surplus and Social Welfare

We further consider consumer surplus and social welfare when the system is mispriced. A consumer's surplus in our setting is defined as the difference between the consumer's indifference price to buy the discrete service and the (market) price of the discrete service, that is, $r_{\tau_c}(\theta;\delta) - r$. Consumer surplus is therefore the summation of all buying consumers' surplus. We consider the social welfare at price r, set by the seller for a service (δ, V). Social welfare is then the summation of all buying consumers' optimal utility at price r plus the seller's revenue at price r, that is, $\sum_{i\in I(B)} u^{B}_{i,\theta,\tau_c}(r,\delta) + R_{\tau_c}(r)$, where $I(B)$ denotes the set of buying consumers.

We start with a mispriced hidden queue. We see that when the shift workers are mispriced at $r^*_L(\theta)$, a corresponding expected time commitment $\overline{\delta}_S(r^*_L(\theta))$ is induced, such that at the menu $(r^*_L(\theta), \overline{\delta}_S(r^*_L(\theta)))$ individual consumers are indifferent in joining the system. That is, the consumers adjust their expected time commitment so as to make sure that the mispriced $r^*_L(\theta)$ is their indifference price, that is, the maximum price they are willing to pay, at $\overline{\delta}_S(r^*_L(\theta))$. Therefore, the consumer surplus is zero.

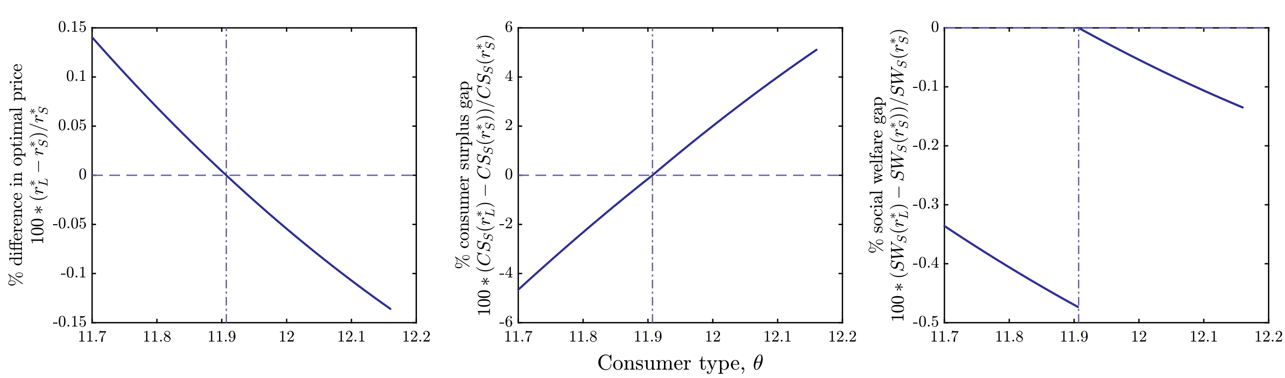
Furthermore, the definition of indifference price (refer to Definition 2) indicates that consumers' utility stays at $u^{NB}_{\theta,S}$. On the other hand, the seller does not achieve the highest possible revenue, as we showed in Section 5.2; therefore, the social welfare is lower when the seller fails to account for the existence of shift workers, that is, non-full-price decision makers.

We now turn to a mispriced visible queue. When the seller misprices the underemployed shift workers, they charge the underemployed shift workers a higher price. First, the system becomes less congested. Some of the consumers who would have joined in a correctly priced system now would not be willing to join. Consider further the consumers standing in the same position, the one standing in a mispriced system is paying a higher price for the same wait than his or her counterpart in the correctly priced system. Note that shift workers of the same type experiencing the same wait would have the same indifference price. Therefore, the consumer surplus of a mispriced system is lower than that of a correctly priced system. Additionally, shift workers in the mispriced system all paying a higher price for the same wait also leads to a lower consumer utility. Furthermore, the seller would not be able to achieve the maximum revenue in a mispriced system. Therefore, the total social welfare is lower than that of a correctly priced visible queue of shift workers.

For the overemployed case, the optimal price charged to gig workers can be either higher or lower. We focus on a neighborhood where we see both cases (see Figure 8). The subfigure on the left zooms in on the "optimal price r" of the "single-server Markovian visible queue" in Figure 4 and recalculates the y-axis to percentage differences. We consider a small range of consumer types θ. Note that when the gig workers are charged a higher price (left of the vertical dashed line), it follows the same argument as above, and both consumer surplus (center) and social welfare (right) are lower than the correctly priced system. If the gig workers are charged at a lower price (right of the vertical dashed line), the mispriced system is more congested than optimal. Furthermore, for the consumer standing in the same position the price is lower, whereas the wait stays the same. Therefore, both consumer surplus and total consumer utility are higher. Together with the lower—because of the suboptimal pricing—seller's revenue, the social welfare of the mispriced visible queue might increase or decrease. In this numerical study, we see lower social welfare (see Figure 8, right); the total increased consumer utility is not enough to cover the loss in seller revenue.

In short, we see that under the hidden queue, failing to account for the existence of shift workers leads to lower social welfare while keeping the consumer surplus at zero. In the mispriced visible queue case, the social welfare may increase or decrease. Furthermore,

Figure 8. (Color online) Plot of the Percentage Consumer Surplus Gap and Percentage Social Welfare Gap Resulting from Mispricing Shift Workers as Gig Workers Against Consumer Type θ



consumer surplus drops when $r_L^*(\theta) > r_S^*(\theta)$ but increases when $r_L^*(\theta) < r_S^*(\theta)$.

Lastly, consider a seller who sells a fixed-time-commitment service to a group of heterogeneous consumers. The consumer surplus is given by $CS(\theta^*_{\tau_c}) = \int_{\theta^*_{\tau_c}}^{\overline{\theta}} (r_{\tau_c}(\theta;\delta) - r_{\tau_c}(\theta^*_{\tau_c};\delta))dF$, where the support of distribution F is $\theta \in (\underline{\theta}, \overline{\theta})$, $\theta^*_{\tau_c}$ is the optimal solution for revenue maximization problem (6), and $r_{\tau_c}(\theta;\delta)$ is the indifference price of consumer type θ given the fixed service time δ solving from Equation (3). When shift workers, that is, consumers, are mispriced at $r_L(\theta_L^*;\delta)$, it induces a suboptimal threshold type $\tilde{\theta}_S$ among shift workers. Consumers who work under shift S with type $\tilde{\theta}_S$ are indifferent in joining the service at the menu $(r_L(\theta_L^*;\delta), \delta)$. We suppress δ in $r(\theta;\delta)$ hereafter. The gap in consumer surplus between the mispriced and correctly priced services is given by $CS_{gap}(S) = \int_{\tilde{\theta}_S}^{\overline{\theta}} (r_S(\theta) - r_L(\theta_L^*))dF - CS(\theta_S^*) = \int_{\tilde{\theta}_S}^{\overline{\theta}} (r_S(\theta_S^*) - r_L(\theta_L^*))dF - \int_{\theta_S^*}^{\tilde{\theta}_S} (r_S(\theta) - r_S(\theta_S^*))dF$.

On the left of the dotted line in Figure 4, "Fixed time commitment, heterogeneous type," are indifferent consumers that are underemployed shift workers, that is, $S < t_{w,L}^{B*}$. The system is mispriced higher, $r_L(\theta_L^*) > r_S(\theta_S^*)$, and a smaller proportion of consumers will join the service, $\tilde{\theta}_S > \theta_S^*$. This can be seen from the following. Shift worker type θ joins the service at price $r_L(\theta_L^*)$ when $r_S(\theta) \geq r_L(\theta_L^*)$. Following from Corollary 1, that $r_S(\theta)$ monotonically increases in θ, we get $\theta_S^* < \tilde{\theta}_S$. That is, when the indifferent consumer is underemployed, a smaller proportion of shift workers are admitted into the system. Also, each consumer that joins the service pays a higher price than he or she does in a correctly priced system. Therefore, the consumer surplus, decreases, that is, $CS_{gap}(S) < 0$, and each joining consumer's utility decreases. The same argument as above with reversed inequalities applies to the right of the dotted line, that is, an overemployed indifferent shift worker whose $S > t_{w,L}^{B*}$.

The seller does not achieve the highest possible revenue because the system is mispriced. Therefore, both consumer surplus and social welfare decrease when the seller misprices underemployed shift workers. The consumer surplus increases, and social welfare may increase or decrease when the seller misprices overemployed shift workers.

6. Conclusion

The full-price regime has been a common assumption in modeling how consumers evaluate purchasing time-consuming services for more than half a century. However, an explanation of when this standard assumption can be justified and when it fails has been lacking. Furthermore, there has been no study of the implications for sellers when consumers are not merely full-price decision makers. In this paper, we provide a framework to answer these questions.

The framework we propose follows the spirit of the classic time and money allocation model introduced by Becker (1965). We enrich the model in three directions: (1) the overhead function emphasizing the time and money trade-off, (2) a potentially non-smooth compensation structure determining the budgetary limit, and (3) the presence of a discrete service that requires a specified time commitment and money to purchase.

Our model shows that a consumer's evaluation of a particular service that requires both time and money cannot be isolated from the consumer's time and budget allocations regarding other daily activities, including leisure, working, and overhead activities. To buy the service or not is part of the consumers' utility-maximizing allocation of total time available and total budget available, whereas the full-price regime assumes that the purchase decision of the focal service is independent of consumers' other activities. In short, we show that the full-price regime does not hold in general.

In light of the two types of economies in which consumers can participate in today's world, gig and traditional economies, we compare the gig workers' and the shift workers' decisions. We show that the full-price regime is enough for decision making only for consumers who have full flexibility in choosing their work time, that is, a linear compensation structure for gig workers. In comparison with the linear compensation structure, a shift pay compensation structure, which allows less time flexibility, would affect consumers' evaluation of a service. A shift worker's value of time depends on whether he or she purchases the service, the discrete service price and time commitment, and how effectively shift worker can buy his or her time back from overhead activities. Gig workers and shift workers are therefore unlikely to make the same decision given the same discrete service.

This implies systematic differences in sellers' optimal strategies when they serve the two types of consumers. Compared with fully flexible gig workers, overemployed shift workers usually face low wages. Although relatively flush with cash compared with the gig workers of the same type, they are time constrained. Therefore, service providers should restrict sales and see a less congested system as optimal when they serve a group of overemployed shift workers as opposed to gig workers. On the other hand, consider service providers serving a community of underemployed shift workers, who could be a group of highly paid office workers. These workers are relatively short on cash compared with gig workers of the same type but will have plenty of time to enjoy the service. In this case, the service provider sees a more congested system formed by underemployed shift workers compared with the system formed by gig workers. We show that the contrast in congestion level is robust whether we compare a system serving gig workers to a system serving solely shift workers or to a system serving a mix of gig workers and shift workers.

Through numerical studies, we show that a revenue-maximizing seller's revenue can be substantially reduced if the seller fails to recognize the prevailing compensation structures of its consumers. This includes failing to properly understand the constraints on its consumers' earnings and time or overlooking the flexibility of its consumers' work schedules. In some cases, sellers mistakenly assuming all consumers as full-price decision-makers, that is, gig workers, also reduces consumer surplus and social welfare.

In general, the framework we provide offers a way to estimate consumers' value of time and their purchase decision with respect to time-consuming discrete services without having to make assumptions about the functional form of the total cost incurred by the consumer. Furthermore, we focus on exploring one determinant of the consumers' evaluation of service, the compensation structure. However, the flexibility of this framework opens up opportunities for investigating more factors that might influence consumers' purchase decisions. The multiple discrete services extension we discussed in the paper is one example, which implies that a consumer's purchase history affects his or her future purchase decisions. The relevance of purchase history is not captured by the full-price regime.

Endnotes

[1] Calculated from Table 7, "Workers by shift usually worked and selected characteristics, averages for the period 2017–2018." "Did not have flexible schedule (in thousands)"/"Total, 15 years and over (in thousands)" = 62,762/144,295 = 43%.

[2] To construct Figure 1, we suppose that consumers, regardless of the compensation structure, share an identical utility function, $U(t_l, c_l) = 1 - e^{-t_l\sqrt{c_l}} + Vy$, and face an identical overhead expenditure, $O(c_O) = \overline{O}\left(1 - \frac{\sqrt{c_O}}{\sqrt{\Omega}}\right)$. Also, consumers plan over $T = 24$ hours. The above-mentioned utility function, overhead expenditure function, and value of T are used in all numerical studies throughout the paper unless otherwise stated. Other parameters used to construct the figure on the left include the shift length of the shift workers $S = 12$ hours, discrete service's time commitment $\delta = 0.3$ hours, $V = 0.04$ if the consumer buys the service, the original overhead $\overline{O} = 16$, and the parameter to control for the efficiency of overhead reduction $\Omega = 120$. Other parameters used in the figure on the right include $S = 8$, $\delta = 3$, $V = 0.001$, $\overline{O} = 13$, and $\Omega = 120$.

[3] To construct Figure 2, we assume the overhead expenditure function to be $O(c_O) = \overline{O}/(1 + c_O)$. We let $\overline{O} = 4.8$, $e_c = 0.35$, $M = 0.7$, $V = 2$, $\delta = 2$, and the shift length of a shift worker $S = 8$.

[4] To construct Figure 3 and Figure 4 (left), the parameters we used included $S = 12$, $\overline{O} = 16$, $\Omega = 300$, $V = 0.04$, $\delta = 0.03$, and $\lambda = 33$.

[5] To construct Figure 4 (right), the parameters we used included $S = 15$, $\overline{O} = 8$, $\Omega = 120$, $V = 0.04$, $\delta = 5$, and consumer types $\theta \sim Unif\,(2, 30)$.

References

Afèche P (2013) Incentive-compatible revenue management in queueing systems: Optimal strategic delay. *Manufacturing Service Oper. Management* 15(3):423–443.

Afèche P, Mendelson H (2004) Pricing and priority auctions in queueing systems with a generalized delay cost structure. *Management Sci.* 50(7):869–882.

Allon G, Federgruen A (2007) Competition in service industries. *Oper. Res.* 55(1):37–55.

Allon G, Federgruen A, Pierson M (2011) How much is a reduction of your customers' wait worth? An empirical study of the fast-food drive-thru industry based on structural estimation methods. *Manufacturing Service Oper. Management* 13(4):489–507.

Allon G, Kremer M (2018) Behavioral foundations of queueing systems. *The Handbook of Behavioral Operations* (John Wiley & Sons, Inc., Hoboken, NJ), 325–366.

Altonji JG, Paxson CH (1988) Labor supply preferences, hours constraints, and hours-wage trade-offs. *J. Labor Econom.* 6(2):254–276.

Becker GS (1965) A theory of the allocation of time. *Econom. J.* 75(299):493–517.

Bureau of Labor Statistics USDoL (2019) Job flexibilities and work schedules summary. Accessed January 30, 2023, https://www.bls.gov/news.release/flex2.nr0.htm.

Cachon GP, Harker PT (2002) Competition and outsourcing with scale economies. *Management Sci.* 48(10):1314–1333.

Chen RR, Kumar S, Singhal J, Singhal K (2021) Note: The value and cost of the customer's waiting time. *Manufacturing Service Oper. Management.* 23(6):1539–1542.

Chiang AC (1984) *Fundamental Methods of Mathematical Economics* (McGraw-Hill Higher Education, New York).

Deacon RT, Sonstelie J (1985) Rationing by waiting and the value of time: Results from a natural experiment. *J. Political Econom.* 93(4):627–647.

Edelson NM, Hildebrand DK (1975) Congestion tolls for Poisson queuing processes. *Econometrica* 43(1):81–92.

Fisman R, Luca M (2017) Why we don't value flextime enough. *Wall Street Journal* (March 3), https://www.wsj.com/articles/why-we-dont-value-flextime-enough-1488547616.

Gavirneni S, Kulkarni VG (2016) Self-selecting priority queues with burr distributed waiting costs. *Production Oper. Management* 25(6):979–992.

Goldratt EM, Cox J (2016) *The Goal: A Process of Ongoing Improvement.* (Routledge, Great Barrington, MA).

Gurvich I, Lariviere MA, Ozkan C (2019) Coverage, coarseness, and classification: Determinants of social efficiency in priority queues. *Management Sci.* 65(3):1061–1075.

Ha AY (1998) Incentive-compatible pricing for a service facility with joint production and congestion externalities. *Management Sci.* 44(12-part-1):1623–1636.

Keller TF, Laughhunn DJ (1973) An application of queuing theory to a congestion problem in an outpatient clinic. *Decision Sci.* 4(3):379–394.

Leclerc F, Schmitt BH, Dube L (1995) Waiting time and decision making: Is time like money? *J. Consumer Res.* 22(1):110–119.

Liu N, Finkelstein SR, Kruk ME, Rosenthal D (2018) When waiting to see a doctor is less irritating: Understanding patient preferences and choice behavior in appointment scheduling. *Management Sci.* 64(5):1975–1996.

Mendelson H (1985) Pricing computer services: Queueing effects. *Commun. ACM.* 28(3):312–321.

Mendelson H, Whang S (1990) Optimal incentive-compatible priority pricing for the m/m/1 queue. *Oper. Res.* 38(5):870–883.

Mueller CD (1985) Waiting for physicians' services: Model and evidence. *J. Bus.* 58(2):173–190.

Naor P (1969) The regulation of queue size by levying tolls. *Econometrica* 37(1):15–24.

Nazerzadeh H, Randhawa RS (2018) Near-optimality of coarse service grades for customer differentiation in queueing systems. *Production Oper. Management* 27(3):578–595.

Okada EM, Hoch SJ (2004) Spending time vs. spending money. *J. Consumer Res.* 31(2):313–323.

Pesca M (2018) Odds, ends, and senators. Episode 915 of Slate's The Gist.

Roberge K (2021) I spend $725 a month on chores I could do myself, and it's one of the best money decisions I've ever made. Accessed January 30, 2023, https://www.businessinsider.com/personal-finance/spend-money-on-chores-no-regrets-2021-9.

Shmanske S (1993) A simulation of price discriminating tolls. *J. Transport Econom. Policy* 27(3):225–235.

Smitizsky G, Liu W, Gneezy U (2021) On the value(s) of time: Workers' value of their time depends on mode of valuation. *Proc. Natl. Acad. Sci. USA* 118(34):e2105710118.

Ülkü S, Hydock C, Cui S (2020) Making the wait worthwhile: Experiments on the effect of queueing on consumption. *Management Sci.* 66(3):1149–1171.