

Dynamic Incentives in Service Contracting

Xu Sun

University of Miami Herbert Business School, Coral Gables, Florida 33146, USA, xxs767@miami.edu

Xiaohan Zhu

School of Management and Engineering, Nanjing University, Nanjing, China, zhuxiaohan@nju.edu.cn

This paper concerns the design of a performance-based compensation scheme (or a “contract”) for a company that outsources customer service to a third-party service provider. The service process is represented as a queue, where the service provider acts as the single server, having the ability to adjust the service rate over time. The provider operates at a base service rate when only receiving a fixed base wage but can speed up the rate of service relative to the base value while incurring a quadratic effort cost for doing so. The company incurs disutilities in the form of congestion costs for delayed customer service. Despite not being able to observe the provider’s true service rate, the company can induce desired service rate adjustments by linking incentive payments to workload, which is assumed to be contractible. The company’s objective is to find a contract to minimize the total cost, which includes the congestion cost and the incentive payment. To gain clear-cut insights, we model the workload process as a reflected Brownian motion with an adjustable drift rate. By utilizing the calculus of variation, we derive the necessary conditions for a contract to be optimal. When congestion costs are quadratic, the derived necessary conditions generate a contract that is explicit and intuitive. We further analyze a scenario wherein the service provider is protected by limited liability. We provide an analytical characterization of the optimal contract in this scenario, which, in turn, allows us to expose how limited liability introduces contracting frictions.

Key words: service contracting; principal-agent framework; calculus of variation; Brownian queues

1. Introduction

Companies often opt to outsource their customer service operations to third-party service providers (Zhou and Ren 2010). The outsourced customer service market was valued at more than 70 billion USD in 2016 and is projected to experience a compounded annual growth rate of approximately 6% from 2017 to 2024. However, the practice of outsourcing customer service presents a significant challenge for business owners due to the inherent difficulty in overseeing the commitment level of these service providers, who might prioritize personal objectives. The lack of direct control over the service process makes it challenging for business owners to ensure that the customer services rendered are up to standard, in contrast to maintaining an in-house customer service team. This, in turn, creates a need for business owners to develop performance-based compensation to motivate third-party service providers to exert effort that benefits the business.

For example, continuous monitoring of call handling and managerial control, often implemented as an incentive-penalty system, is prevalent in today’s call center outsourcing (Poster 2007). To illustrate, consider a situation where a company enlists a freelance call center agent to handle customer requests. To align the agent’s incentives with the company’s objectives, a bonus-malus arrangement can be put in place, allowing for payments to be transferred based on agreed-upon performance metrics. If, say, both parties can monitor the workload (the time required to clear the current backlog of work) or reliably estimate it based on the number of jobs in the queue, incentive payments can be structured based on this contractible metric.¹ In particular, when the workload is minimal, implying that the agent is dedicating adequate effort to providing efficient customer service, the company could award a sizable bonus to the agent. Conversely, if the workload becomes excessive, the company might consider withholding the bonus or even imposing a malus payment on the agent to compensate for the service degradation. Another representative scenario involving a utility company outsourcing its IT services to a vendor was documented in Akkermans et al. (2019), where the authors consider buyer-supplier contracting using collaborative key performance indicators. In the described contractual relationship, the number of open tickets (i.e., requests yet to be resolved) serves as an observable and verifiable metric based on which a bonus-malus arrangement can be constructed. In essence, the provision of incentive payments holds the potential to incentivize the service provider to operate more efficiently, thereby granting the service “buyer” indirect influence over the customer service process.

Aiming to glean clear-cut insights into the design of dynamic incentives in the context outlined above, we introduce a stylized model that combines the principal-agent framework with a simple queueing system. In our model, the company assumes the role of the principal (“she”), whereas the service provider fulfills the role of the agent (“he”). The agent receives a base wage that is barely competitive enough to incentivize him to participate and exert the desired base service rate. Under this base rate of service, the workload evolves as a reflected Brownian motion (RBM), reflecting the virtual wait time for incoming jobs. The agent has the ability to adjust the service rate, impacting the workload’s drift rate, but at a cost. The principal incurs congestion costs tied to the workload. While direct observation of the agent’s effort is not possible, the principal can infer it from the workload. By linking payments to the workload, the principal can provide incentives for the agent to manage the speed of service in accordance with her operational objectives. Specifically, the principal seeks to motivate the agent to speed up as the workload increases. While doing so may result in increased effort costs for the agent in the short run, it has the potential to reduce

¹ In practice, real-time tracking of workload is unfeasible, yet accurate estimation is achievable using indicators like queue length, as demonstrated in works such as Ibrahim and Whitt (2009) and Ibrahim and Whitt (2011). However, our contract design problem does not necessitate real-time workload tracking; post-computation suffices.

congestion levels, leading to long-term monetary rewards for the agent under a performance-based compensation scheme.

The agent’s decision-making process begins with seeking a rate-control strategy to maximize his utility, which factors in both received incentive payments and effort costs. We refer to the agent’s desire to maximize his utility as the “incentive compatibility” (IC) constraint. If the agent determines that the highest attainable utility either meets or surpasses his reservation utility, he accepts the proposed arrangement. If not, he declines it. Meanwhile, there is a need to entice the agent into participating in this arrangement, leading to the emergence of the “participation compatibility” (PC) constraint.² Given the principal’s lack of direct visibility into the agent’s actual effort level, she designs a contract based on a contractible (i.e., observable and verifiable) quantity.³ In this study, the contractible quantity chosen is the workload. However, in principle, any other metric capable of reflecting the volume of pending tasks (e.g., the queue length) can serve this purpose, as long as it is observable and verifiable. A contract is said to be optimal if it induces a rate-control strategy that minimizes the principal’s overall cost while adhering to the IC and PC constraints.

For the principal, the search for an optimal contract proceeds in two steps. The first step is to consider each rate-control strategy that the agent could potentially employ and then deduce an incentive payment plan that induces such a strategy at the lowest possible cost. This is possible to achieve because the principal can deduce how the agent will react to a given payment plan given that she knows the agent’s cost structure and payoff function. The second step is to identify a rate-control strategy that, along with the cheapest way to induce such a strategy (as deduced in the first step), minimizes the principal’s long-run average cost. Importantly, the resulting optimal contract induces a “second-best” rate-control strategy, as opposed to the “first-best” policy the principal would choose if she directly controlled the service rate and incurred the effort costs herself.

Methodologically, we use an RBM to describe wait-time dynamics for both practical and tractability considerations. Brownian models are widely used in the operations management/research (OM/OR) literature due to their tractability and rich mathematical theories. An RBM proves to be a valuable choice as it effectively captures the stochastic variability in job arrivals and service completions through Brownian motion. It takes into account the agent’s strategic actions by incorporating an adjustable drift rate while also accounting for the server’s cumulative idleness through a one-sided reflecting process at the origin. The advantage of employing an RBM is particularly evident in the first step of the analysis. It allows for a clean characterization of the relationship

² In the related literature, the PC constraint is sometimes referred to as the “individual rationality” constraint.

³ In this research, “contract design” specifically refers to the design of performance-based incentives to be used to incentivize the agent to control the service rate in a manner that aligns with the principal’s best interests.

between a rate-control strategy and the most cost-effective incentive payment plan that induces that strategy. This simplification arises because the agent’s relative value function can be expressed as a solution to an ordinary differential equation, which *per se* brings to light the key factors influencing the agent’s decision-making, such as effort cost and incentive payment.

While the first step of the analysis can be eased to a great extent by the adoption of the Brownian model, the second step of the analysis presents immense technical difficulty. Specifically, we find that characterizing the optimal contract entails solving an instance of the calculus of variation, which is an infinite-dimensional optimization problem. In classical calculus of variation, the objective is to optimize a functional—a function of functions—expressed as a definite integral over a finite interval of integration, involving an unknown function (the quantity to be optimized) and its derivative. However, in our study, the variational problem takes the form of minimizing a functional expressed as the ratio of two integrals, each with an infinite interval of integration. Moreover, the functional involves the antiderivative of the unknown function in addition to the unknown function itself and its derivative. As a result, standard results pertaining to the calculus of variation do not directly apply to our unique problem. Nonetheless, by extending the classical theory of calculus of variation, we are able to demonstrate that the second-best rate-control strategy, provided it exists, must obey a Riccati equation (a first-order ordinary differential equation that is quadratic in the unknown function). For quadratic congestion costs, the solution to the Riccati equation is rather explicit, resulting in a contract that is neat and intuitive. The incentive payment plan has a carrot-and-stick structure. To be more precise, the “carrot” represents the maximum bonus that the agent can receive per unit of time for maintaining zero delays, and the “stick” represents a performance-based penalty interpreted as the amount of bonus that would have been paid out but was withheld by the principal due to poor service levels as measured by a large backlog of work. If the performance-based penalty exceeds the maximum bonus, the excess is interpreted as a malus payment (an amount that will be transferred from the agent to the principal).

Driven by practical considerations, we further analyze a scenario where the agent is protected by limited liability, that is, all incentive payments from the principal to the agent are non-negative. Characterizing the optimal contract in this scenario proves even more challenging as it requires solving a variational problem with point-wise constraints. To tackle this challenge, we employ the Lagrange multiplier approach. We first treat the value of the agent’s surplus (defined as the difference between the actual utility he derives from the contractual relationship and the agent’s reservation utility) as fixed. This allows us to integrate the Lagrange multiplier method and variational analysis to deduce a set of necessary conditions for a contract to be optimal (among those that provide the desired value of the agent’s surplus). Utilizing these conditions, we identify a contract expected to be optimal (again among those that deliver the desired surplus value). Then, by varying the value

of the agent’s surplus that the principal chooses to allocate to the agent, we deduce a parametric family of contracts. Finally, we show that this parametric family contains a cost-minimizer, which we expect to be an optimal solution to the contracting problem with the limited liability constraint.

1.1. Contributions

We consider the paper’s contributions to be trifold.

Modeling: We formulate a model that appears to be the first dynamic principal-agent problem that cannot be effectively solved using recursive methods. To tackle the challenge, we resort to the calculus of variation, a methodology that does not see a lot of use in the OM/OR literature. This solution approach, in our belief, supplements the recursive method that is typically applied to solving dynamic principal-agent problems, opening the door to its potential applicability in other dynamic contracting scenarios.

Methodology: To characterize the optimal contract, we extend the classical theory of the calculus of variations to deal with functionals expressed as a ratio of two definite integrals. These functionals involve a function of interest and its antiderivative and derivative functions. We prove that the associated Euler-Lagrange equation characterizing the “stationary points” of a given functional takes the form of a Riccati equation, subject to appropriate boundary conditions (Theorem 1). The arguments used are further generalized by incorporating a “Lagrange multiplier function” to characterize the optimal contract under limited liability. This approach, when fixing the value of the agent’s surplus, yields an equation jointly satisfied by the function of interest and the Lagrange multiplier function, subject to a set of constraints that mirror the KKT conditions in the theory of mathematical optimization. We prove that a solution can be found via that equation and the set of specified constraints (Theorem 2). There are several analytical results that can be of independent interest in the domain of optimal control for Brownian queues. For instance, Proposition 1 extends and generalizes a result from Kim and Randhawa (2018) by establishing the well-posedness of the Bellman equation associated with controlling a Brownian queue under a general cost function. Furthermore, Proposition 5 uncovers a phenomenon of diminishing returns on investments in flexibility if we interpret flexibility as the range of allowable rate perturbations one can exert on a Brownian queue. We believe this result and its proof are new. (More detailed discussions are provided after these results are formally stated.)

Managerial Implications: In addition to the aforementioned carrot-and-stick type of compensation, our study shows that both risk aversion and/or limited liability can create contracting frictions. These frictions imply that achieving the first-best outcome may not be feasible with the second-best solution. Specifically, we find that when the agent is risk-averse and not protected by limited liability, the original contracting problem involving a risk-averse agent can be transformed into a risk-neutral

equivalent, in which motivating the agent through monetary means becomes more costly. On the other hand, when the agent is risk-neutral but protected by limited liability, our analysis shows that a performance-based compensation structure can still be effective, but the agent is only willing to exert effort up to some maximum level. Although the principal has the ability to influence this maximum effort level, thereby gaining more indirect control over system congestion, it comes at a cost in the form of a greater surplus, which the principal must offer to the agent.

1.2. Organization

The remainder of this paper is organized as follows: Section 2 reviews the related literature. Section 3 formally introduces our base model, while Section 4 presents the analysis and the main results pertaining to the base model. Section 5 discusses some of the key model elements. Section 6 concerns the design of the optimal contract with limited liability. Section 7 concludes the paper. The mathematical proofs of all main results are relegated to §EC.1 of the e-companion.

2. Literature Review

Several recent papers have considered settings where servers are strategic, being able to choose their own service rates to optimize some utility function, leading to a game between servers and a firm that employs the servers. Gopalakrishnan et al. (2016) consider an $M/M/N$ queue where servers are strategic, seeking to balance the cost of effort with a preference for idleness. By characterizing the symmetric Nash equilibrium, the authors show that servers' strategic behavior can have a dramatic impact on the optimal staffing and routing policies. Zhan and Ward (2019) investigate how payment can be used to induce effort from servers, in addition to staffing and routing decisions to be made at the system manager's end, resulting in a joint staffing, routing, and payment problem in the context of an $M/M/N + M$ queue with selfish servers voluntarily choosing service rates to maximize their respected utility. Zhong et al. (2023) consider an $M/M/N/k$ queue with a finite buffer, where servers can choose the speed of service to optimize a utility function involving payment, effort cost, and idleness. To understand how job admission, staffing, and payment decisions impact servers' choices of action, the authors conduct an equilibrium analysis using both exact and asymptotic approaches. In particular, the paper reveals several interesting non-monotonic relationships between the service rates in equilibrium and key model inputs such as staffing level and buffer size. Our paper differs from these papers in two key aspects. First, while we consider a server's choice of effort level as "hidden actions," thereby leading to moral hazards, these papers consider the server's effort level to be observable by the system manager. Second, our model assumes the server has the ability to adjust the speed of service in real time, whereas the aforementioned studies assume a server's effort level is fixed once chosen. Put differently, the aforementioned studies and the present work focus on different time scales (one at a strategic or tactical level and one at an operational level).

This paper falls within the service contracting literature. Previous studies have explored various aspects of service contracts, such as capacity-setting and pricing decisions in call center outsourcing contracts (Akşin et al. 2008), contract design in service chains with observable or unobservable effort levels (Ren and Zhou 2008), and the influence of different contracts on call center capacity-setting decisions (Hasija et al. 2008). Additionally, there has been research on service contracting in situations with information asymmetry (Akan et al. 2011), performance-based contracting in medical services (Jiang et al. 2012), and contract performance in scenarios with private information and varying contractibility of effort (Zhang et al. 2018). In particular, the paper by Jiang et al. (2012) considers queueing. In their study, a service buyer (the principal) contracts with a service provider (the agent) and seeks to minimize costs while adhering to a service level target. The authors model the appointment dynamics as an $M/D/1$ queue and demonstrate that a linear performance-based contract, despite being able to achieve the first-best outcome without information asymmetry, cannot coordinate the service chain with asymmetric information. However, they show that a threshold-based contract can coordinate the chain even with information asymmetry. Our work, however, distinguishes itself from these papers in scope and methodology. It focuses on incentive provisioning at the operational level, operating on shorter time scales. This implies that both incentive payments and agent actions respond to the current system state. In contrast, the aforementioned studies primarily address long-term or steady-state system performance.

The present paper contributes to the growing literature on dynamic principal-agent problems. These problems are complex and require subtle economic reasoning. An approach that has proven successful is to use a recursive formulation with the agents' promised or continuation utility serving as the state variable. This approach was pioneered by Spear and Srivastava (1987) where the authors analyze a dynamic principal-agent model in discrete time and demonstrate that the solution can be found via a recursive method. Sannikov (2008) extends the approach to dynamic principal-agent problems in continuous time, where the agent's effort influences the drift of a Brownian motion, and the optimal contract is sought using stochastic optimal control techniques. Biais et al. (2010) employ similar techniques on a model where the agent's effort influences the rate of a Poisson process rather than a Brownian motion. The promised utility framework has also been successfully applied in several studies in the OM/OR literature. For instance, Li et al. (2013) study a dynamic principal-agent problem for managing critical suppliers via business volume incentives. Treating the suppliers' promised utility vector as the state-descriptor, the authors formulate the problem as a Markov decision process. Interestingly, their analysis identifies the "carrot" and "stick" as a set of "trapping" states where the suppliers choose low effort forever and a "recurrent" class of states where the suppliers engage in a competition exerting high effort forever, respectively. Wang et al. (2016) study voluntary disclosure of environmental hazards using a dynamic mechanism

design framework. In their study, the occurrence of the hazard is the firm's private information, and regulatory responses, which are costly to the firm, provide incentives for the firm to conceal the hazard. Their analysis results in explicit, easy-to-implement regulation policies. In particular, the optimal policy stipulates that inspections are performed periodically. Sun and Tian (2018) examine how to incentivize an agent to adopt the desired effort level over an infinite time horizon, while Chen et al. (2020) consider the optimal monitoring and payment mechanisms to incentivize an agent's effort in reducing the probability of adverse events occurring. Most recently, Tian et al. (2021) consider the design of dynamic incentives involving a principal that hires an agent to repair a machine when it is down and maintain it when it is up. The agent can make effort level adjustments over time, but the true effort level is not observable to the principal. Treating the agent's promised utility as the state variable, the authors derive a contract that is simple and intuitive. A very interesting finding of their research is that the IC constraint is not always binding. Although the promised utility framework provides an attractive analysis for a wide range of dynamic contracting problems, we find it difficult to apply this framework to our problem for two reasons. First, while defining the agent's promised utility is mostly straightforward in infinite-horizon discounted settings, since the agent in our model faces an ergodic risk-sensitive control problem, it is not immediately clear how to properly define the agent's promised utility. Second, the application of this framework typically requires the principal's objective to lack a state structure. For continuous-time models, this means that the time integral of the state process or some functional of it should not appear in the principal's objective, a feature our model lacks. These observations prompt us to consider another route, namely, the method of calculus of variation, which we believe complements the promised utility approach.

In a related line of work, Plambeck and Zenios (2000) study a dynamic principal-agent problem and show that the problem can be solved efficiently via dynamic programming without using the agent's promised utility as the state variable. Notably, the optimal contract the authors derived is history-independent and renegotiation-proof. However, their analysis and results rely on the assumption that the agent has exponential utility while having access to a perfect financial market for borrowing and lending to smooth cash flows. Building on this framework, Plambeck and Zenios (2003) analyze a make-to-stock model where a manufacturer bears the inventory-holding and back-order costs of the finished good while delegating production to a supplier who dynamically controls the production rate at some cost. Similar to their model, our work utilizes payments that are contingent on the observed system state. However, our paper differs from Plambeck and Zenios (2003) in a fundamental way: we do not assume that the agent has interactions with the financial market. This distinction is crucial in that we can no longer rely on a recursive method to characterize or devise the optimal contract.

Our literature search reveals limited use of the calculus of variations in the OM/OR literature. One example is Bensoussan et al. (2023), where the authors consider a decentralized supply chain involving a retailer facing random demand over time and having private inventory information. The supplier offers supply contracts to address this asymmetry. The authors derive the necessary conditions for optimizing long-term contracts under varied demand and belief distributions, applying them to a batch-order contract scenario. Like our work, the problem is formulated as an instance of the calculus of variation, utilizing the notion of Gâteaux derivative. However, there are two key distinctions between their work and ours. First, their problem deals with adverse selection, whereas we focus on addressing moral hazards. Second, our analysis leads to minimizing an unconventional functional, expressed as a ratio of two definite integrals, differing from the functional Bensoussan et al. (2023) try to optimize.

3. Base Model

In this section, we formulate our dynamic contracting model within a principal-agent framework and present the central question we will address. The model is structured along the lines outlined in the introduction and assumes that the agent controls the underlying state process by taking unobserved actions. Our model assumes that both parties have complete knowledge of each other’s objectives, which will be introduced shortly. Both the principal and the agent are aware of the specific mechanism by which the agent’s actions affect the evolution of the state process. Furthermore, the model assumes an infinite time horizon, which not only simplifies the decision space but also usefully captures the reality that contractual relationships typically do not have a fixed end date.

3.1. System Dynamics

Customer requests come from the outside world and are processed by a single agent who exerts a base service rate in the absence of any incentives (i.e., if receiving the base wage only). With this base service rate, the amount of work in the system, or the virtual wait time, is assumed to follow a driftless RBM. While we do not explicitly introduce the arrival and service completion processes, we provide a detailed justification for using an RBM to describe the wait-time dynamics of an M/G/· type queue in §EC.3.

The agent’s adjustments to the service rate result in a drift-rate control process, denoted by θ , which impacts the evolution of the workload process. Specifically, when considering a rate-control process θ , the evolution of the workload over time is described by the equation:

$$Z(t) = Z(0) - \int_0^t \theta(u) du + \sigma B(t) + L(t), \quad (1)$$

where $Z(0)$ represents the initial workload, B is a standard Brownian motion that introduces random shocks to the system, and L is a non-decreasing process ensuring that Z remains non-negative at

all times, with minimum effort. To establish a connection to a physical queue, one can envision L as a process that tracks the cumulative idle time of the server. Additionally, the Brownian term captures the inherent stochasticity in job arrivals and/or service completions in a physical queue, with the level of variability determined by the constant σ . It is assumed that both parties' decisions will be based on the state process Z .

3.2. Economic Factors

Transitioning to the economic aspect of the model, we posit that the principal incurs congestion costs at a rate of $k(z)$ when the workload Z is at level z . These costs encompass various forms of dissatisfaction, including the loss of goodwill arising from prolonged wait times. Importantly, these costs are absorbed by the principal rather than the agent. We assume that k is differentiable and grows to infinity at a polynomial rate, with $k(0) = 0$. We further impose a technical assumption on the cost function k .

ASSUMPTION 1. *For any $a_1 > 0$, there exists some $a_2 > 0$ such that $k(z) > a_1 k'(z) - a_2$ for $z \geq 0$.*

Intuitively, this assumption requires the cost function $k(z)$ to outgrow its derivative function as z grows large. It is satisfied by a broad spectrum of functions, including polynomial cost functions of the form $k(z) = az^b$ for $a > 0$ and $b \geq 1$. These cost functions have found extensive applications in performance analysis and optimization of queueing systems, particularly in those situations where a customer's cost of waiting exhibits non-linearity. For instance, a recent empirical investigation by Ding et al. (2019) discovered that the marginal waiting cost for critical patients in Canadian emergency departments could be approximated using a piece-wise linear function. In another study, Ouyang et al. (2022) showed that, under quadratic waiting cost functions, the choice between priority policies depends on the traffic intensity and the proportion of each type in the population. With this motivation, we choose to assume a general congestion cost function and believe that doing so will not only enrich the analysis and insights but also enhance the model's practical relevance.

Incentive rewards allocated to the agent correspond to costs for the principal. Both the principal and agent acknowledge a mutually agreed-upon incentive payment plan denoted as $S(\cdot)$. If $S(z) > 0$, it signifies payments being transferred from the principal to the agent at the rate $S(z)$ whenever the workload is at the level z . Conversely, if $S(z) < 0$, then the agent pays the principal at the rate $-S(z)$ whenever the workload is at the level z . Our notation assumes that the incentive payment is solely based on workload level, which has an intuitive interpretation. Since the principal only has access to the observed workload as a signal for the agent's efforts, the principal may seek to fully utilize this signal by specifying the incentive payment plan accordingly.

Apart from receiving payments from the principal, the agent incurs effort costs at a rate of $c(x)$ whenever the drift rate θ is at the level x . For the purposes of our study, we adopt the standard

specification $c(x) := \frac{\alpha}{2}x^2$. This specification is common in the literature. See, e.g., Zhang et al. (2018) where the service provider is also assumed to incur quadratic effort costs; see also (Huang and Gurvich 2018, section 5) for a demonstration of a rate control problem in the context of an $M/G/1$ queue where the rate control incurs a quadratic cost. Our rationale for this assumption is grounded in the fact that a quadratic function is arguably the most natural way to capture the notion of increasing marginal costs of effort as the agent raises his effort level. That said, we do not expect an extension of the model with a general convex effort cost function to either introduce fundamental difficulties to our analysis or alter our insights in a fundamental way. It should perhaps also be noted that this specification for effort cost implies that withholding some effort from the base value will also incur costs for the agent. We consider this assumption harmless for two reasons. From a practical perspective, it could mean that choosing an effort level below the base value can cause boredom and therefore disutility for the agent. From a methodological standpoint, inducing a slowdown is never optimal for the principal, indicating that only the non-negative portion of the function domain of c is relevant for our analysis. Therefore, going forward, we will refer to this cost as the effort cost.

3.3. The Agent's Problem

Taking into account these economic considerations, the agent evaluates his total payoff up to time t based on the following expression:

$$\mathcal{J}(t) := \int_0^t S(Z(u))du - \int_0^t c(\theta(u))du. \quad (2)$$

That is, the total payoff up to time t is the difference between the accumulated incentive payments and the effort costs incurred during that period. Note that the payoff defined above excludes the base wage the agent receives because it plays no essential role in the design of dynamic incentives other than serving to make the agent content with employing the base service rate.

We assume that the agent is risk-averse. Intuitively, being risk-averse means that one is more concerned with and hence more sensitive to losses than gains of the same magnitude. While there are a variety of ways to characterize risk aversion, we choose to capture it via an exponential utility function that links the total payoff over a time horizon to a utility value. Mathematically, this means that the agent calculates his total utility up to time t via $(1/\gamma) \log \mathbb{E} [e^{\gamma \mathcal{J}(t)}]$, where γ denotes the risk-sensitive parameter satisfying $\gamma \in (-\infty, 0)$. Then, given an incentive payment plan S , the agent seeks a control strategy θ to maximize the long-run average utility, defined as

$$\liminf_{t \rightarrow \infty} \frac{1}{\gamma t} \log \mathbb{E} [e^{\gamma \mathcal{J}(t)}], \quad (3)$$

where \mathcal{J} is defined by equation (2).

Problem (3) is essentially an ergodic risk-sensitive control problem. The majority of risk-sensitive control problems studied in the extant literature assume the risk-sensitive parameter to be positive. This is because these problems deal with cost minimization. In contrast, the agent’s problem in our model is to maximize his utility, and therefore being risk-averse entails a negative risk-sensitive parameter. For a general background on ergodic risk-sensitive control problems and the state-of-the-art development in this domain, we refer the reader to the comprehensive survey by Biswas and Borkar (2023). It may also be worth mentioning that risk-sensitive control has an intimate connection to robust control, a problem category that arises routinely when it is impossible to identify a true system model and is often formulated in the form of a two-player zero-sum game; see, e.g., Hansen et al. (2006) and Lim and Shanthikumar (2007). As a result, one can alternatively interpret the agent’s risk aversion as being uncertain about the probabilistic law governing external events that impact the system’s evolution.

Following the literary convention, we will henceforth refer to the agent’s desire to maximize (3) as the IC constraint. Note that as γ approaches zero, the expression (3) reduces to the following:

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \mathbb{E}[\mathcal{J}(t)]. \quad (4)$$

This observation allows us to extend the range of γ to $(-\infty, 0]$ with the understanding that the agent’s objective is to maximize the long-run average payoff (4) when $\gamma = 0$.

Given an infinite time horizon and state-dependent incentive payments, the agent can restrict attention to the class of Markov strategies; that is, at any given time t , the value of $\theta(t)$ links to t only through $Z(t)$. For this reason, we will henceforth treat each rate-control strategy as some deterministic function $\theta(z)$. Furthermore, we call θ an admissible control if it is non-negative and piecewise smooth, exhibits polynomial growth as z approaches infinity, and satisfies the requirement that $\theta(z) \geq \chi$ for some positive constant χ and all sufficiently large z . We note that the non-negativity requirement is imposed merely to streamline the analysis and is not crucial; it can be relaxed to require the function θ to have a uniform lower bound. Piecewise smoothness is imposed because when we apply the calculus of variations to characterize the optimal contract, θ serves as the unknown function of interest, and the minimal requirement for a variational problem to be well-posed is for the allowable functions to be piecewise smooth (Olver 2012, chapter 9). We require the function value of $\theta(z)$ to be strictly bounded away from zero for all large enough z in order to ensure that the system is stable when the agent employs θ as his control strategy. This requirement can be considered a restatement of the celebrated Foster-Lyapunov criterion for stability; see, e.g., (Xu 2023, chapter 5). Lastly, polynomial growth of θ is required to ensure that the relative value function associated with principal’s problem (to be introduced shortly) displays polynomial growth so that a desired verification theorem (part (ii) of Theorem 1) can be established.

In addition to having the ability to implement service rate control, the agent also decides whether to accept the principal's arrangement regarding the performance-based compensation scheme. If the agent finds the plan unappealing, he can refuse to enter into the contractual relationship and choose to pursue other options instead, giving rise to the PC constraint. Formally, this constraint means that the agent finds the arrangement acceptable if

$$\max_{\theta} \liminf_{t \rightarrow \infty} \frac{1}{\gamma t} \log \mathbb{E} [e^{\gamma \mathcal{J}(t)}] \geq 0. \quad (5)$$

The requirement that the left-hand side needs to be greater than or equal to zero stems from the premise that the value of the agent's outside option can be offset by a fixed wage.

3.4. The Principal's Problem

The principal is risk-neutral and calculates the total cost up to time t by adding the incentive payments to the congestion costs:

$$\mathcal{K}(t) := \int_0^t k(Z(u))du + \int_0^t S(Z(u))du. \quad (6)$$

As with the agent's decision problem, the objective defined above excludes the base wage provided to the agent, as the focus has been on the design of the performance-based compensation scheme.

We say that an incentive payment plan S induces a rate control strategy θ if the agent finds θ to be incentive- and participation-compatible under S . An incentive payment plan S is considered admissible if the induced rate-control strategy θ is admissible. The rate-control process induced by a given incentive payment plan, together with the random shocks generated by the Brownian motion, drives the evolution of the workload process, which in turn determines the congestion costs incurred over time. The principal's problem is to devise an admissible incentive payment plan that, when combined with the agent's rate-control strategy, minimizes her long-run average cost. Formally, the principal seeks to find some S that minimizes

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} [\mathcal{K}(t)]. \quad (7)$$

subject to the IC constraint (3) and PC constraint (5). At this stage, it may be worth reiterating that the rate-control strategy cannot be enforced, as the agent will only adopt it voluntarily if it satisfies both the IC and PC constraints.

3.5. Benchmark: A Single-Party Control Problem

As a benchmark, we present a simplified version of the problem in which the principal directly controls the service rate and bears the effort costs. In this case, the objective is to find an admissible θ that minimizes the long-run average cost for the principal, given by:

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t c(\theta(u))du + k(Z(u))du \right] \quad \text{subject to (1)}. \quad (8)$$

Alternatively, this objective arises when the principal can directly observe the agent's actions, removing the need to incentivize desired effort levels. Indeed, if the principal can monitor the agent's actions, she can threaten to impose severe penalties if the agent deviates from the desired actions. Consequently, compensating the agent's effort according to the function c becomes sufficient to fulfill the PC constraint.

4. Analysis

In this section, we will implement the two-step approach outlined in the introduction to characterize the optimal contract. The first step is to devise an incentive payment plan for each rate-control strategy that induces the strategy at the least expense to the principal. This can be accomplished since the principal is aware of the agent's payoff structure, which enables her to infer how the agent will respond to a given payment plan. The second step aims to identify an incentive payment plan that minimizes the overall cost to the principal, comprising both congestion costs and monetary incentives provided to the agent. To assist with the mathematical analysis and facilitate the presentation of the main results, we will adopt the notion of relative value functions.

It is instructive to begin our analysis with Problem (8), as its results are likely to provide inspiration for the analysis of the optimal contracting problem. Note that the analysis of Problem (8) will lead to the first-best rate-control policy, which can be derived using dynamic programming. The following result provides a formal statement.

PROPOSITION 1. *There exists a twice differentiable function v and a constant ρ^* that solve*

$$\min_{\theta} \left\{ \frac{\sigma^2}{2} v''(z) - \theta v'(z) + k(z) + c(\theta) \right\} = \rho^* \quad (9)$$

with $v'(0) = 0$, $v''(z) > 0$ for all $z \geq 0$, and that $v'(z)$ grows polynomially as $z \rightarrow \infty$. Furthermore, the minimum objective value of Problem (8) is ρ^ .*

The proof of this proposition can be found in the e-companion. Its proof reveals that equation (9) can be turned into the form of a Riccati equation. This type of equation has appeared and been analyzed in several studies. In studying joint dynamic pricing and leadtime quotation controls in the context of make-to-order systems, the paper by Çelik and Maglaras (2008) derives a Bellman equation in the form of a Riccati equation. By choosing to capture congestion-related disutilities through order rejection rather than explicitly modeling waiting costs, the authors obtain a closed-form solution to the Riccati equation. In considering joint pricing, outsourcing, and scheduling controls for a make-to-stock system in heavy traffic, Ata and Barjesteh (2023) characterize the near-optimal control policy via the solution to a Bellman equation also in the form of a Riccati equation. Considering the linear cost function with respect to on-hand inventory or back-order levels,

they provide a judicious construction of the solution on a finite interval (whose end points, however, are subject to optimization). Notably, the solution derived by the authors can be considered explicit in that it is expressed using Bessel and Airy functions. In studying pricing control in a single-server queue, the paper by Kim and Randhawa (2018) characterizes the near-optimal policy via the solution to a Riccati equation. Their Riccati equation, like ours, is defined on the entire non-negative real line. Interestingly, their equation can be seen as a special case of our Riccati equation. The distinction lies in the fact that their control problem involves a linear cost rate function, whereas we consider general cost rate functions, allowing for non-linear cases. The authors' proof of well-posedness relies on the assumption of a linear cost rate function, and as a result, we cannot directly apply their proof arguments to our case. In contrast, our proof of well-posedness builds on a condition (Assumption 1) that not only encompasses linear functions but also accommodates a wide range of polynomial cost functions.

Equation (9) conforms to the standard form of Bellman equations for average-cost stochastic control problems. The solution to this equation can be determined either explicitly or numerically, depending on the nature of the congestion cost function k . For example, when k is a quadratic function, part of the solution to equation (9), v , takes the form of a quadratic function, which in turn implies that the optimal control strategy for Problem (8) is linear in z . As a result, the service rate will increase linearly as the workload grows. Problem (8) provides a benchmark as it allows the principal to directly control the service rate.

Returning to the analysis of the optimal contracting problem, we begin by solving the agent's utility maximization problem while temporarily disregarding the PC constraint. Leveraging the standard results in ergodic risk-sensitive control theory (c.f. Eq. (3.9) in Biswas and Borkar (2023)), we can immediately write out the corresponding Bellman equation that characterizes the agent's strategy for maximizing his utility:

$$\max_{\theta} \left\{ \frac{\sigma^2}{2} U''(z) - \theta U'(z) + S(z) - c(\theta) + \frac{\gamma \sigma^2}{2} (U'(z))^2 \right\} = \eta \quad (10)$$

subject to the boundary condition $U'(0) = 0$.

In the language of dynamic programming, $U(z)$ is the relative value function associated with the agent's problem, and η corresponds to the agent's long-run average utility (provided the agent chooses to participate). Under this interpretation, the PC constraint can be formulated as $\eta \geq 0$. This leads us to the following result, which specifies the most cost-effective incentive payment plan that incentivizes a given rate-control strategy.

PROPOSITION 2. *If a rate-control strategy $\theta(\cdot)$, induced by a specific incentive payment plan, satisfies the PC constraint, then it must hold that $\theta(0) = 0$. Furthermore, there exists an incentive*

payment plan S that induces θ at the lowest cost to the principal. This payment plan stipulates monetary transfers according to the formula:

$$S(z) = \frac{\sigma^2}{2}\alpha\theta'(z) - \frac{1}{2}(\alpha - \gamma\alpha^2\sigma^2)\theta^2(z). \quad (11)$$

As per our model description, a higher value of α implies a higher cost of effort. It thus follows from Proposition 2 that the principal needs to provide higher-powered incentives to motivate an agent who values leisure more to adopt the desired service rate. Aside from exposing these operational implications, Proposition 2 serves as a critical building block for the subsequent analysis. It establishes the existence of an “induced cost function,” which quantifies the principal’s instantaneous cost of inducing a particular rate-control process θ . This means that the mission of identifying an optimal contract can be turned into an equivalent single-party decision problem. The following result contributes to this mission by establishing a connection between a specific pair of strategies employed by both parties and the long-run average cost incurred by the principal for that strategy pair.

PROPOSITION 3. *Fixing a pair of strategies (S, θ) , if some twice differentiable function V that and a constant κ collectively solve the following equation:*

$$\frac{\sigma^2}{2}V''(z) - \theta(z)V'(z) + S(z) + k(z) = \kappa, \quad (12)$$

subject to $V'(0) = 0$ and $\limsup_{t \rightarrow \infty} t^{-1}\mathbb{E}[V(Z(t))] = 0$, then κ is the long-run average cost of the principal under the pair of strategies (S, θ) .

To illustrate the value of Proposition 3, we substitute the expression for S from the right-hand side of (11) into (12), yielding

$$\frac{\sigma^2}{2}V''_{\theta}(z) - \theta(z)V'_{\theta}(z) + \frac{\sigma^2}{2}\alpha\theta'(z) - \frac{1}{2}(\alpha - \gamma\alpha^2\sigma^2)\theta^2(z) + k(z) = \kappa(\theta), \quad (13)$$

subject to the same conditions as specified in the proposition. We have appended the subscript θ to V' and V'' and written κ as $\kappa(\theta)$ in (13) to emphasize the dependence of these quantities on the choice of θ .

Suppose for now that equation (13) admits a solution pair $(V_{\theta}, \kappa(\theta))$ that satisfies all the stated conditions. Then Propositions 2 and 3 together imply that $\kappa(\theta)$ represents the principal’s long-run average cost when the agent chooses θ induced by the cheapest incentive payment plan $S(z)$. Thus, it is instructive to view (13) and its associated conditions as a functional (i.e., a function of functions) that maps each rate-control strategy θ to the principal’s long-run average cost. Consequently, the principal’s problem can be reframed as selecting an appropriate θ to

$$\text{minimize } \kappa(\theta). \quad (14)$$

Equation (13) departs from the standard form of Bellman equations, as it cannot be optimized point-wise due to the appearance of the first-order derivative of θ . This stands in contrast to the left-hand side of equation (9), which conforms to the standard form and allows for simpler point-wise optimization. Instead, the minimization problem defined by (14) needs to be viewed as a variational problem that optimizes an objective over a space of functions. Solving this problem yields the second-best solution, hence the optimal contract.

In preparation for presenting the main results, let us define $\phi(z)$ as the integral of $\theta(x)$ with respect to x from 0 to z , i.e., $\phi(z) := \int_0^z \theta(x) dx$. For notational convenience, also define $\tilde{\alpha}_\gamma := \alpha - \gamma\alpha^2\sigma^2$. The following theorem provides an explicit expression for $\kappa(\theta)$ and shows that the second-best rate-control strategy, if it exists, must obey a Riccati equation.

THEOREM 1. *(i) Provided $S(z)$ given by (11) is admissible, there exists a solution pair $(V_\theta, \kappa(\theta))$ to (13) subject to the specified conditions. Moreover, $\kappa(\theta)$ is the long-run average cost of the principal, admitting the following expression:*

$$\kappa(\theta) = \frac{\int_0^\infty e^{-\frac{2}{\sigma^2}\phi(x)} \left[\frac{\sigma^2}{2} \tilde{\alpha}_\gamma \theta'(x) - \frac{\tilde{\alpha}_\gamma}{2} \theta^2(x) + k(x) \right] dx}{\int_0^\infty e^{-\frac{2}{\sigma^2}\phi(x)} dx}. \quad (15)$$

(ii) If $\kappa(\cdot)$ attains its minimum at $\theta = \theta^$, then the minimizer, along with some constant $\tilde{\kappa}$, must obey*

$$\frac{\sigma^2}{2} \tilde{\alpha}_\gamma \theta'(z) - \frac{\tilde{\alpha}_\gamma}{2} \theta^2(z) + k(z) = \tilde{\kappa} \quad (16)$$

subject to the condition $\theta(0) = 0$ and the requirement that $\theta(z)$ is non-negative and exhibits polynomial growth as $z \rightarrow \infty$. The value $\tilde{\kappa}$ represents the principal's long-run average cost under the pair of strategies (S^, θ^*) , where S^* is determined by (11) with $\theta = \theta^*$.*

Based on Theorem 1, if the optimal contract exists, the induced rate-control strategy is expected to satisfy, along with some constant $\tilde{\kappa}$, equation (16). Note, however, that this equation provides a necessary rather than provably sufficient condition for optimality. A closely related question arises as to whether equation (16) admits a solution pair that fulfills all the specified conditions. The answer to this question is affirmative. To see this, it suffices to observe that equations (9) and (16) are structurally the same. That is, one can replicate the entire analysis that leads to the well-posedness of equation (9) to establish the well-posedness of equation (16), without making any essential changes to the arguments used. Further, from the proof of Proposition 1, we can assert that the solution θ to equation (16) is a non-negative, differentiable, and polynomially growing function in z , satisfying $\theta(z) \geq \chi$ for some positive constant χ , and for all sufficiently large z .

To illustrate the implications of Theorem 1, let us consider an example where the congestion cost function is quadratic, i.e., $k(z) = k_2 z^2$ for some positive constant k_2 . In this case, one can readily verify that equation (16), subject to the specified conditions, admits the following pair of solutions:

$$\theta(z) = \sqrt{\frac{2k_2}{\tilde{\alpha}_\gamma}} z \quad \text{and} \quad \tilde{\kappa} = (\sigma^2/2) \sqrt{2\tilde{\alpha}_\gamma k_2}.$$

Moreover, the incentive payment plan that induces this rate-control strategy θ is given by:

$$S(z) = \frac{\sigma^2}{2} \sqrt{2\tilde{\alpha}_\gamma k_2} - k_2 z^2. \quad (17)$$

The incentive payment plan given by (17) exhibits a simple, intuitive carrot-and-stick structure. The “carrot” component represents the highest bonus the agent can earn per unit of time for maintaining zero delays. The “stick” component is performance-based, reflecting the amount of bonus that would have been paid out but was withheld by the principal due to inadequate service levels, resulting in a substantial backlog of work. If the performance-based penalty exceeds the maximum bonus, the excess is regarded as a malus payment, deducted from the agent’s earnings, and transferred to the principal. The combined effect of the two components serves to motivate the agent to increase his speed in response to an escalation in the workload backlog.

Aside from the above-mentioned carrot-and-stick structure, one can also observe that the principal’s total cost increases with the value of α . This makes intuitive sense because to induce the same level of responsiveness, a lazier worker requires higher-power incentives. The degree of variability in the system, as measured by σ^2 , impacts the principal’s long-run average cost as well. This is also intuitive because as stochastic volatility increases, it makes it more challenging to identify whether stochastic fluctuation or shirking behavior is causing the observed service degradation, thereby making it more difficult for the principal to infer the agent’s hidden actions. To offset the increased “moral hazard” effect, the principal has to raise the power of incentives (by increasing the maximum bonus in this case) to induce the agent to take the desired actions.

5. Discussion

Having presented a model to understand the economic drivers influencing the design of an optimal contract, we will now briefly discuss the implications of the main results obtained and some of the model’s fundamentals.

5.1. Necessity vs. Sufficiency

The characterization of the optimal contract, as presented in Theorem 1, uses the calculus of variation, a mathematical technique concerned with identifying the stationary points of a given functional. It achieves this by setting the functional derivative (“first variation”) to zero, leading to

the associated Euler-Lagrange equation. Our proof of Theorem 1 relies on identifying a specific version of the Euler-Lagrange equation, which is derived through the first variation. Much like the first-order condition serves as a necessary condition for optimality in standard calculus theory, the Euler-Lagrange equation presents a necessary but not sufficient condition for a function to be an extremum of the relevant functional. However, the Euler-Lagrange equation oftentimes provides valuable insights into the solution of the variational problem. This has been demonstrated in Chen et al. (2020), where the solution to the Euler-Lagrange equation, despite not being verified to be a maximum of the corresponding functional, yields substantial insights into the optimal pricing scheme.

To confirm that the solution to equation (16) indeed yields the optimal contract, we must perform a second derivative test on the functional $\kappa(\theta)$ and demonstrate the positivity of its second derivative. In the calculus of variations, the second derivative of a functional is known as its *second variation*. Unfortunately, this verification procedure tends to be exceedingly technical, encompassing the notion of the central field of extremals as elucidated in (Olver 2012, chapter 8). For this reason, we have chosen not to pursue this endeavor.

Nevertheless, there are valid reasons to expect that solving equation (16) indeed gives us the second-best solution. We document two reasons below. First, there is only one value of $\tilde{\kappa}$ for which equation (16), viewed as a differential equation with respect to the unknown function θ , has a solution that satisfies all the stated requirements. To elaborate, we note that $\tilde{\kappa}$ is the smallest value of κ for which the equation

$$\frac{\sigma^2}{2}\tilde{\alpha}\gamma\theta'(z) - \frac{\tilde{\alpha}\gamma}{2}\theta^2(z) + k(z) = \kappa \quad (18)$$

with the boundary condition $\theta(0) = 0$ produces a non-negative solution function θ . More precisely, for all $\kappa < \tilde{\kappa}$, the solution function $\theta(z)$ to (18) will approach negative infinity as z grows large; and for all $\kappa > \tilde{\kappa}$, the solution to (18), albeit non-negative, will grow to positive infinity at least exponentially fast. Only when $\kappa = \tilde{\kappa}$ does the solution to (18) remain non-negative throughout the positive real line while growing to infinity at a polynomially fast rate. Second, in the special case where $\gamma = 0$, it can be readily verified that the rate-control strategy obtained by solving (16) aligns with the first-best policy derived from (9). Additionally, the constant $\tilde{\kappa}$ obtained from (16) coincides with the optimal objective value ρ^* as stated in Proposition 1. Put another way, the rate-control strategy obtained from (16) with $\gamma = 0$ attains the first-best outcome. Since this rate-control strategy is also admissible, it must represent the second-best solution.

5.2. Contracting Frictions

In studying principal-agent problems, it is common to assume the agent to be risk-averse (Dai and Chao 2013). It is also well-established in the principal-agent literature that when the agent is

risk-neutral, there is no “friction” between the principal’s desire to compensate the agent and the agent’s willingness to accept the arrangement, meaning that the cost to the principal of adopting the optimal contract is the same as the cost under the first-best solution. Therefore, it should come as no surprise that the solution derived from Theorem 1 can achieve the first-best outcome when we assume the agent to be risk-neutral (i.e., $\gamma = 0$). However, when the agent is risk-averse, the principal would need to pay a risk premium to counteract the agent’s aversion to risk, which would make the second-best solution differ from its first-best counterpart.

To further illustrate this point, note from (16) that the contracting problem with a risk-averse agent has a risk-neutral equivalent in which the agent is risk-neutral but has an effort cost function given by $\tilde{c}_\gamma(\theta) := (\tilde{\alpha}_\gamma/2)\theta^2$. When $\gamma < 0$, we have $\tilde{\alpha}_\gamma > \alpha$, which implies that additional compensation is needed for the agent to take desired actions.

5.3. More on the Brownian Motion Setup

It is known in the queueing literature that an RBM naturally arises as a process-level approximation of a *critically loaded* queueing system. In the present context, critical loading entails the assumption that the rate at which customer requests arrive matches the base service rate. While this assumption may seem a bit far-fetched, considering that real-world service systems can often be either underloaded or overloaded, we argue that this is the operational regime that is worth exploring. On the one hand, an underloaded system, where the base service capacity significantly exceeds the demand volume, is uninteresting, as the goal of incentive rewards is to alleviate congestion. However, an underloaded system seldom experiences congestion and instead results in excessive idleness. On the other hand, a severely overloaded system, where the base capacity falls substantially short of the demand volume, indicates that the incoming load is too much for one server to handle. We interpret this as a signal for the principal to contemplate recruiting multiple agents to share the load, which could be an interesting topic for future research. It is also plausible that the base capacity differs slightly from the demand volume. This scenario can be accurately captured by introducing a constant drift rate to the Brownian motion without introducing major hurdles to the analysis or fundamentally altering our main results. Lastly, we note that the assumed operational regime may result from capacity rationing. This situation arises when the principal has knowledge of the demand volume and can screen job candidates to find one whose base service capacity matches or approximates the demand volume.

6. Dynamic Incentives Under Limited Liability

The application of malus payments may encounter difficulties in certain practical scenarios where legal restrictions might hinder their implementation. This is especially evident when the contracted agent is categorized as a *de facto* employee. In this case, the agent could be eligible for a minimum

wage as dictated by labor laws. To illustrate, suppose that the base wage (used to induce the base service rate) coincides with the minimum wage. Since an employee’s pay cannot be reduced to below the minimum wage, it results in a “limited liability” constraint, which can be defined as follows:

$$S(z) \geq 0 \quad \text{for all } z \geq 0. \quad (19)$$

That is, the agent would, at worst, receive zero incentive payment.

This new constraint has significant implications for the design of incentive contracts. To get a sense of it, notice that if the principal again chooses to set η to zero in equation (10), then equation (2) would imply that θ must be a zero process. This observation compels the principal to only consider positive values of η in the presence of the newly added constraint (19). With this observation in mind, we devote the rest of the section to seeking some answers to the issue of contract design under this newly added constraint. To distill the unique effects of limited liability on the contract design, we will assume that both the agent and the principal are risk-neutral, so that $\gamma = 0$.

Formally, the principal’s problem becomes one that aims to minimize Problem (7), subject to the IC constraint, the PC constraint, and the limited liability constraint. As we have derived, without the limited liability constraint, the search for the optimal contract leads to solving an optimization problem in an infinite-dimensional space. With the newly added limited liability constraint, it becomes a constrained infinite-dimensional optimization problem, which is even more challenging relative to its unconstrained counterpart. To tackle the challenge, we adapt the Lagrange approach, a common technique for addressing finite-dimensional constrained optimization problems, to this infinite-dimensional problem.

Specifically, our solution approach draws inspiration from the celebrated Karush-Kuhn-Tucker (KKT) conditions in the theory of mathematical optimization. To apply this idea, we associate the non-negativity constraint on $S(z)$ with a Lagrange multiplier $\lambda(z)$. Because the limited liability constraint is imposed for each individual workload z , the collection of all such Lagrange multipliers forms a function of z . The introduction of the Lagrange multiplier function $\lambda(z)$ allows us to modify the objective functional by adding to it the pointwise constraints multiplied by their respective Lagrange multipliers, yielding the modified objective functional:

$$\tilde{\mathcal{K}}(t) := \int_0^t k(Z(u))du + \int_0^t S(Z(u))du - \int_0^t \lambda(Z(u))S(Z(u))du. \quad (20)$$

In addition, the KKT conditions motivate the “dual feasibility” constraint on the Lagrange multiplier function:

$$\lambda(z) \geq 0 \quad \text{for all } z \geq 0, \quad (21)$$

as well as the “complementary slackness” condition:

$$\lambda(z)S(z) = 0 \quad \text{for all } z \geq 0.$$

With these preparations, we can reformulate the principal’s problem as one that seeks some S and a Lagrange multiplier function λ to minimize

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\tilde{\mathcal{K}}(t) \right] \quad (22)$$

subject to the IC constraint (3), the PC constraint (5), the limited liability constraint (19) and the dual feasibility constraint (21), as well as the complementary slackness condition.⁴ We will refer to Problem (22) as the “Lagrange problem.” The following result provides theoretical justification for considering this relaxed problem.

PROPOSITION 4. *If some S^* solves the Lagrange problem (22) along with a Lagrange multiplier function λ , then S^* also solves the principal’s original problem (7) with the limited liability constraint (19).*

To determine the long-run average cost of the principal κ under the new objective (22) and a pair of fixed strategies (S, θ) , similar to equation (12), we consider the following equation:

$$\frac{\sigma^2}{2} V''(z) - \theta(z)V'(z) + S(z) + k(z) - \lambda(z)S(z) = \kappa, \quad (23)$$

subject to the boundary condition $V'(0) = 0$ and the limiting condition $\limsup_{t \rightarrow \infty} t^{-1} \mathbb{E}[V(Z(t))] = 0$.

As with the base model, a key step in solving the principal’s problem is to establish a connection between a rate control strategy and an incentive payment plan that induces the strategy at minimal cost. In the base model, this connection was established by solving the agent’s utility-maximizing problem (10) while setting $\eta = 0$. However, as explained at the beginning of the section, the introduction of the limited liability constraint would require η to be positive. For the time being, let us consider the value of η to be given and fixed. Following the analysis of Proposition 2, we can derive an incentive payment plan $S(z)$ that incentivizes a given rate control strategy θ at the lowest expense to the principal, which is expressed as follows:

$$S(z) = \frac{\sigma^2}{2} \alpha \theta'(z) - \frac{\alpha}{2} \theta^2(z) + \eta. \quad (24)$$

Upon substituting (24) into (23), we obtain the following equation:

$$\frac{\sigma^2}{2} V''_{\theta}(z) - \theta(z)V'_{\theta}(z) + \frac{\sigma^2}{2} \alpha \theta'(z) - \frac{\alpha}{2} \theta^2(z) + \eta + k(z) - \lambda(z) \left(\frac{\sigma^2}{2} \alpha \theta'(z) - \frac{\alpha}{2} \theta^2(z) + \eta \right) = \kappa(\theta) \quad (25)$$

⁴ In the language of KKT conditions, the limited liability constraint corresponds to the “primal feasibility” condition.

subject to the same set of conditions, namely, $V'_\theta(0) = 0$ and $\limsup_{t \rightarrow \infty} t^{-1} \mathbb{E}[V_\theta(Z(t))] = 0$. As in the base model, we use the subscript θ to indicate that V' and V'' depend on the choice of θ , and we write κ as $\kappa(\theta)$ to emphasize this dependence in equation (25).

Similar to part (i) of Theorem 1, equation (25), along with the specified boundary conditions, enables us to derive an expression for $\kappa(\theta)$ as follows:

$$\kappa(\theta) = \frac{\int_0^\infty e^{-\frac{2}{\sigma^2}\phi(x)} \left[\frac{\sigma^2}{2} \alpha \theta'(x) - \frac{\alpha}{2} \theta^2(x) + k(x) - \lambda(x) \left(\frac{\sigma^2}{2} \alpha \theta'(x) - \frac{\alpha}{2} \theta^2(x) + \eta \right) \right] dx}{\int_0^\infty e^{-\frac{2}{\sigma^2}\phi(x)} dx}.$$

With this expression at our disposal, we can apply the calculus of variations technique akin to the proof for part (ii) of Theorem 1. In particular, we can deduce that with a fixed Lagrange multiplier function λ , the function θ that minimizes $\kappa(\theta)$ must satisfy the ‘‘stationarity’’ condition as encapsulated in the following equation:

$$\frac{\sigma^2}{2} \alpha \theta'(z) - \frac{\alpha}{2} \theta^2(z) + k(z) - \lambda(z) \left(\frac{\sigma^2}{2} \alpha \theta'(z) - \frac{\alpha}{2} \theta^2(z) + \eta \right) + \frac{\sigma^2}{2} \alpha \left(\frac{\sigma^2}{2} \lambda''(z) - \theta(z) \lambda'(z) \right) = \tilde{\kappa}, \quad (26)$$

subject to the boundary condition $\theta(0) = 0$. A detailed derivation for equation (26) is provided in §EC.2.

Here, $\tilde{\kappa}$ represents the long-run average cost of the principal, subtracted by η , i.e., $\tilde{\kappa} = \kappa(\theta) - \eta$. In addition, the limited liability constraint and the complementary slackness condition can be translated into

$$\frac{\sigma^2}{2} \alpha \theta'(z) - \frac{\alpha}{2} \theta^2(z) + \eta \geq 0, \quad (27)$$

and

$$\lambda(z) \left(\frac{\sigma^2}{2} \alpha \theta'(z) - \frac{\alpha}{2} \theta^2(z) + \eta \right) = 0, \quad (28)$$

respectively. Note that we can substitute (28) into (26) to simplify the stationarity condition, yielding the following:

$$\frac{\sigma^2}{2} \alpha \theta'(z) - \frac{\alpha}{2} \theta^2(z) + \frac{\sigma^2}{2} \alpha \left(\frac{\sigma^2}{2} \lambda''(z) - \theta(z) \lambda'(z) \right) + k(z) = \tilde{\kappa}, \quad (29)$$

subject to the boundary condition $\theta(0) = 0$.

Now, our goal becomes to pinpoint a triple $(\theta, \lambda, \tilde{\kappa})$ that not only satisfies (29) but adheres to the three labeled constraints: (21), (27) and (28). This is a challenging endeavor due to the nature of (29), which does not conform to the structure of a conventional differential equation (it involves two unknown functions, θ and λ , plus one unknown constant). To be able to make further headway, our strategy is to formulate conjectures about the appropriate functional forms of θ and λ and proceed to verify that the conjectured solution meets all the stated requirements. This strategy

is rooted in the following insight: In the absence of the limited liability constraint, the principal retains the option to impose negative payments (penalties) as a means of exerting pressure on the agent when congestion levels are high, encouraging the agent to speed up in order to alleviate the congestion. However, in the presence of the limited liability constraint, the principal's ability to enforce penalties is compromised. This creates a situation where the principal's hands are tied, unable to levy a heavy penalty even during significant service level deterioration. What she can do best is stop providing incentive payments, and when that happens, the agent loses the impetus to increase speed further.

Based on the foregoing insight, we posit that a candidate minimizer θ has the following structural property: It is strictly increasing on the interval $[0, \bar{z})$ for some value \bar{z} and satisfies the equation

$$\frac{\sigma^2}{2}\alpha\theta'(z) - \frac{\alpha}{2}\theta^2(z) + k(z) = \tilde{\kappa} \quad \text{for all } z \in (0, \bar{z}), \quad (30)$$

subject to the boundary conditions $\theta(0) = 0$ and $\theta(\bar{z}) = b := \sqrt{2\eta/\alpha}$; and on the interval $[\bar{z}, \infty)$, θ is constant and equal to b . It is easy to verify that such a θ , if it exists, automatically fulfills (27). Moreover, the dual feasibility constraint (21) and the complementary slackness condition (28) imply that, for such a θ , the corresponding Lagrange multiplier function (if it exists) must satisfy the following requirement: $\lambda(z) = 0$ for $z \in [0, \bar{z})$, and

$$\frac{\sigma^2}{2}\alpha \left(\frac{\sigma^2}{2}\lambda''(z) - b\lambda'(z) \right) - \frac{\alpha}{2}b^2 + k(z) = \tilde{\kappa} \quad \text{for all } z \in [\bar{z}, \infty). \quad (31)$$

Returning to equation (30), we note that the two boundary conditions $\theta(0) = 0$ and $\theta(\bar{z}) = b$, combined with equation (30), are inadequate to determine the three unknowns θ , $\tilde{\kappa}$, and \bar{z} . Therefore, we need to identify additional conditions. To that end, let us first consider equation (25) on $[0, \bar{z})$. By substituting (28) and (30) into (25) and letting $\tilde{\kappa} = \kappa(\theta) - \eta$, we can simplify (25) to

$$\frac{\sigma^2}{2}V_\theta''(z) - \theta(z)V_\theta'(z) = 0. \quad (32)$$

Similarly, on $[\bar{z}, \infty)$, equation (25) can be simplified using (28) and the fact that θ is constant and equal to b on $[\bar{z}, \infty)$, yielding

$$V_\theta''(z) - bV_\theta'(z) - \frac{\alpha}{2}b^2 + k(z) = \tilde{\kappa}. \quad (33)$$

From (32) and the boundary condition $V_\theta'(0) = 0$, we can deduce that $V_\theta' \equiv 0$ on $[0, \bar{z})$. On the other hand, since V_θ must have a continuous first-order derivative, we can subject (33) to the boundary condition $V_\theta'(\bar{z}) = 0$, in addition to the limiting condition $\limsup_{t \rightarrow \infty} t^{-1}\mathbb{E}[V_\theta(Z(t))] = 0$.

To encapsulate, we have introduced one new unknown, V_θ satisfying (33), and added two new conditions, namely, $V_\theta'(\bar{z}) = 0$ and $\limsup_{t \rightarrow \infty} t^{-1}\mathbb{E}[V_\theta(Z(t))] = 0$. Together, these additions help increase both the total number of unknowns and the number of conditions to four.

THEOREM 2. (i) *There exist θ , $\tilde{\kappa}$, \bar{z} , and V_θ that, together, satisfy the derived conditions. (ii) For the θ , $\tilde{\kappa}$ and \bar{z} found in part (i), there exists a non-negative Lagrange multiplier function λ that satisfies $\lambda(z) = 0$ on the interval $z \in [0, \bar{z})$ and (31) on the interval (\bar{z}, ∞) , as well as the boundary conditions $\lambda(\bar{z}) = 0$ and $\lambda'(\bar{z}) = 0$.*

An inspection of the proof of Theorem 2 unveils an interesting connection between the solution encapsulated in part (i) of the theorem and the solution to the following *single-party* stochastic control problem: Choose an admissible θ to minimize

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t c(\theta(u)) du + k(Z(u)) du \right] \quad \text{subject to (1) and } \theta \in [0, b]. \quad (34)$$

Notably, this problem bears resemblance to Problem (8), differing solely in the confinement of θ to the interval $[0, b]$. It turns out that the θ and $\tilde{\kappa}$ identified in part (i) of Theorem 2, coincide, respectively, with the optimal policy and the optimal objective value of Problem (34). This connection is useful in that one can compute the desired θ and $\tilde{\kappa}$ by solving the single-party stochastic control problem (8) instead. As with Problem (8), the solution to Problem (34) can be conveniently characterized and computed through a Bellman equation. This equation seeks to identify a function v that is twice differentiable and a constant $\rho^*(b)$ that jointly satisfy the equation:

$$\min_{\theta \in [0, b]} \left\{ \frac{\sigma^2}{2} v''(z) - \theta v'(z) + k(z) + c(\theta) \right\} = \rho^*(b) \quad (35)$$

with $v'(0) = 0$, $v''(z) > 0$ for all $z \geq 0$, and that $v'(z)$ grows polynomially as $z \rightarrow \infty$.

Our discussions thus far have treated η as a given and fixed parameter. As a result, $\tilde{\kappa}$ given in Theorem 2 is a quantity dependent on η . To emphasize this dependence, we will henceforth write $\tilde{\kappa}$ as $\tilde{\kappa}(\eta)$. With the aim of minimizing costs in mind and given the family of contracts parameterized by η , the principal is naturally inclined to select a value for η to

$$\text{minimize } \eta + \tilde{\kappa}(\eta). \quad (36)$$

A related question is now in order: Does the optimization problem defined by (36) guarantee to admit a minimizer?

Before seeking an answer to this question, we numerically solved for the parameterized family of contracts, based on examples featuring linear and quadratic congestion cost functions. Let $\sigma^2 = 1$ and define the agent's effort cost as $c(\theta) = \theta^2$. Additionally, we set the linear congestion cost as $k(z) = z$ and the quadratic congestion cost as $k(z) = z^2$. The two graphs in Figure 1 illustrate the relationship between the values of $\tilde{\kappa}$ and η under linear and quadratic congestion cost functions, respectively. Both graphs demonstrate a decreasing trend in the principal's long-term average cost subtracted by η as η increases. In both cases, the function $\eta + \tilde{\kappa}(\eta)$ exhibits strict convexity in

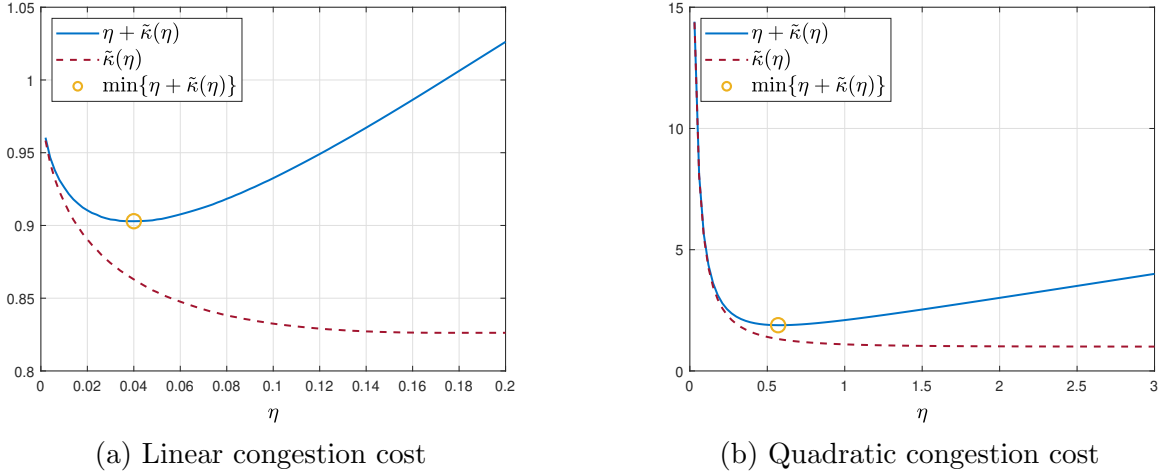


Figure 1 The minimum value of the long-run average cost for the principal

η , implying the existence of an optimal minimizing value, denoted as η^* . This, in turn, enables us to pinpoint a potential solution for the second-best rate control strategy, θ^* , along with the corresponding incentive payment plan, $S^*(z)$, which can be used to incentivize the adoption of θ^* .

Zooming in on the details, Figure 2 illustrates the rate-control strategy θ^* and the corresponding incentive payment plan, $S^*(z)$, for the specific examples. As depicted in the left panels, the incentive payment gradually decreases as congestion increases, and it suddenly drops to zero at a critical juncture, maintaining a value of zero thereafter. The right panels showcase how the agent's chosen service rate initially increases until congestion reaches the aforementioned critical point, beyond which it levels off, assuming its maximum value b .

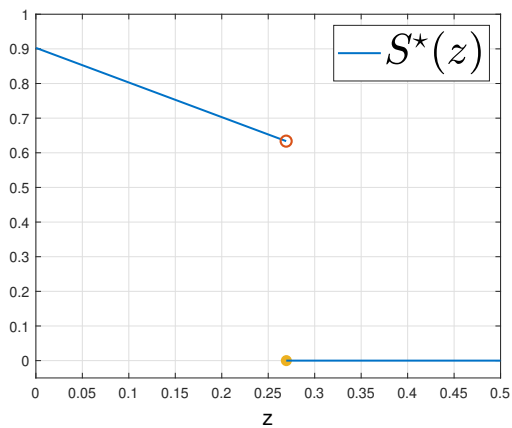
Returning to the question about the existence of a minimizer for Problem (36), given that $b = \sqrt{2\eta/\alpha}$ (by definition) and $\tilde{\kappa}(\eta) = \rho^*(b)$, we can see that the task of minimizing the expression in (36) is tantamount to selecting a value for b to

$$\text{minimize } \frac{\alpha b^2}{2} + \rho^*(b). \quad (37)$$

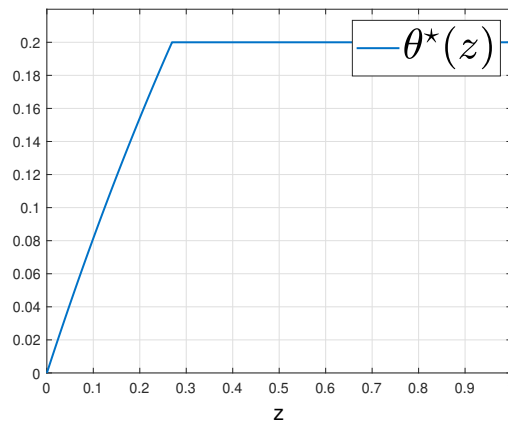
The first term is quadratic, thus exhibiting convexity with respect to b . As a result, if one can demonstrate that the second term is also convex in b , we can conclude that the expression in Problem (37) is strictly convex. This, in turn, guarantees the existence of a unique minimizer for Problem (37). The following proposition serves this purpose by establishing that $\rho^*(b)$ is indeed convex in b .

PROPOSITION 5. *The optimal objective value $\rho^*(b)$ of Problem (34) is convex in b .*

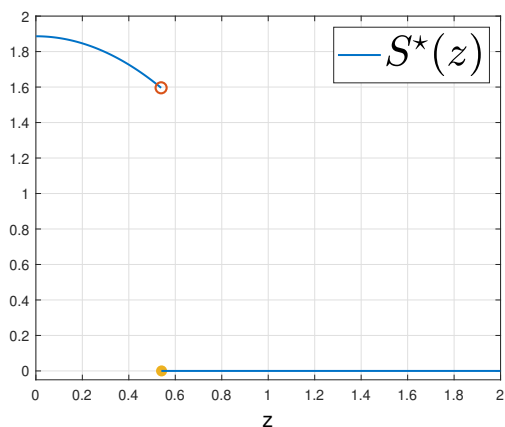
The implication of Proposition 5 extends beyond the specific contracting problem under consideration. For instance, imagine a service facility with a fixed base service capacity and the option



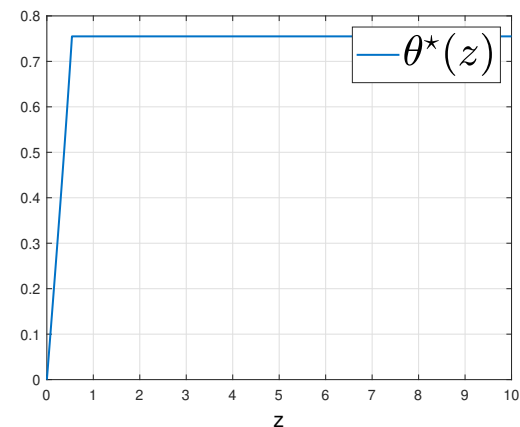
(a) Incentive payment (linear congestion cost)



(b) Rate control (linear congestion cost)



(c) Incentive payment (quadratic congestion cost)



(d) Rate control (quadratic congestion cost)

Figure 2 Illustration of the solution with linear and quadratic congestion costs when penalty payments are prohibited

to augment it with surge capacity. Utilizing surge capacity incurs operational costs and involves an upfront investment. If we assume that the facility’s workload follows a driftless RBM under the base capacity and that surge capacity can be utilized partially or fully, then Problem (34) becomes a quest for the optimal way to utilize the surge capacity, given that the maximum surge capacity is limited to b . In this context, Proposition 5 suggests that investing in surge capacity exhibits diminishing returns, and there exists an optimal value of b that minimizes the total cost, combining operational and investment costs. Proving Proposition 5 poses a challenge in that ρ^* seen as a function of b does not admit a closed-form expression (except for the case when $b = \infty$). To overcome the challenge, we devised a constructive proof, which we believe is novel in the context of extant literature.

At this point, it may be worthwhile to summarize what we have done so far. Fixing the value of η , we have derived a set of necessary conditions characterizing the optimal contract, where “optimal”

means that the contract minimizes the principal’s long-term average cost while ensuring the agent derives a utility of η from the contract. Utilizing the set of necessary conditions, we identify a contract that fulfills all the conditions for each fixed η . Then, varying the value of η gives rise to a family of contracts parameterized by η . Among this parametric family, we establish the existence of a contract that minimizes the principal’s objective, standing out as the “best-of-the-best.” Thus, a final question remains: Is this contract the optimal solution to our contracting problem under limited liability? The answer hinges on establishing the sufficiency of our “KKT conditions” for optimality. However, we point out that the task of establishing this desired sufficiency for optimality can be formidable in that such a challenge even pervades the domain of conventional constrained optimization problems confined to finite dimensions. Put another way, demonstrating that the KKT conditions suffice for optimality in most finite-dimensional cases can be intricate, let alone tackling the task of a constrained optimization problem with infinite dimensions.

Nonetheless, the contract identified as the best-of-the-best and showcased in Figure 2, aligns closely with one’s intuition regarding what an ideal contract should look like in the presence of the limited liability constraint. This alignment strongly suggests that it is likely to be the true optimal solution. Under this contract, the principal grants high rewards for favorable states while withholding rewards entirely for unfavorable ones. Unburdened by the prospect of severe negative consequences, as in the model, the agent displays a willingness to accelerate as long as the rewards continue, halting only when the principal ceases reward disbursements.

The minimization problem (36) reveals an inherent dilemma in designing dynamic incentives under limited liability. To raise the power of incentives (i.e., to enhance the agent’s responsiveness to the current system state), the principal needs to ensure a significant disparity between compensation levels for desirable and undesirable states. Without the limited liability restriction, achieving this would involve attaching high bonuses to favorable outcomes and levying hefty penalties for highly unfavorable ones. In the presence of the limited liability constraint, however, the principal loses the capacity to impose penalties; what she can do at most is offer zero rewards. As a result, to adequately differentiate the compensation associated with desirable and undesirable states, the principal is compelled to elevate the remuneration for favorable states. This adjustment, in turn, allows the agent to attain greater utility derived from the contract, as captured by an increase in the value of η . However, in exchange, the agent becomes more reactive to the change in the system state. Finally, while increasing the value of η holds the promise of providing high-powered incentives, the marginal benefit of doing so diminishes, which intuitively justifies the theoretical conclusion drawn in Proposition 5.

7. Concluding Remarks

This paper considers the optimization of contract design, focusing on a scenario where a principal hires an agent to manage customer requests in an uncertain environment. Although the agent can adjust his effort level over time, this action remains hidden and cannot be contractually enforced. To incentivize the agent to actively manage the service rate in a manner that the principal desires, the principal considers linking the observer system state to incentive rewards, resulting in a contracting problem. We illustrate that determining the optimal contract entails solving a variational problem, which leads to solving a Riccati equation with appropriate boundary conditions. In cases where congestion costs are quadratic, this characterization yields explicit formulas for both the incentive payment plan and the induced service rate control. In particular, the incentive payment plan holds a clear and intuitive interpretation, representing the maximum bonus amount for the agent minus a performance-based penalty. Additionally, we consider a scenario where negative or malus payments are restricted, leading to a contract design problem under the so-called limited liability condition. To tackle this problem, we devise an approach that blends variational analysis and the Lagrange multiplier method. This approach establishes a set of necessary conditions for an optimal contract. These conditions mirror the well-known KKT conditions commonly seen in solving conventional constrained optimization problems and allow us to deduce a parametric family of contracts that are expected to contain the optimal contract. We demonstrate that within this family of contracts, there is indeed the best one within that family, thereby serving as a strong candidate for the optimal contract.

There are several promising directions for future research. The first direction is to formally establish the sufficiency of the derived, necessary conditions for contract optimality. In the realm of variational calculus, well-established techniques like the Legendre and Jacobi tests are available for distinguishing between the minimum, maximum, and saddle points of a functional in standard problems (Liberzon 2011, section 2.6). Adapting and extending these methods to address our variational problem could yield valuable insights. Another avenue of exploration entails the development of dynamic incentives when dealing with multiple agents. While our current model primarily focuses on a single agent engaged with the principal, delving into scenarios where multiple agents participate in competitive or collaborative roles may unveil additional managerial insights. Finally, it may be worthwhile to consider impatient jobs that can abandon the queue while waiting. By incorporating this aspect, the resulting system model is likely to offer a more faithful representation of call center operations, potentially offering new insights into the contract design.

References

- Akan M, Ata B, Lariviere MA (2011) Asymmetric information and economies of scale in service contracting. *Manufacturing & Service Operations Management* 13(1):58–72.

- Akkermans H, Van Oppen W, Wynstra F, Voss C (2019) Contracting outsourced services with collaborative key performance indicators. *Journal of Operations Management* 65(1):22–47.
- Akşin OZ, De Véricourt F, Karaesmen F (2008) Call center outsourcing contract analysis and choice. *Management Science* 54(2):354–368.
- Ata B, Barjesteh N (2023) An approximate analysis of dynamic pricing, outsourcing, and scheduling policies for a multiclass make-to-stock queue in the heavy traffic regime. *Operations Research* 71(1):341–357.
- Bensoussan A, Sethi S, Wang S (2023) A stationary infinite-horizon supply contract under asymmetric inventory information. *Operations Research* .
- Biais B, Mariotti T, Rochet JC, Villeneuve S (2010) Large risks, limited liability, and dynamic moral hazard. *Econometrica* 78(1):73–118.
- Biswas A, Borkar VS (2023) Ergodic risk-sensitive control—a survey. *Annual Reviews in Control* .
- Browne S, Whitt W, Dshalalow J (1995) Piecewise-linear diffusion processes. *Advances in Queueing Theory, Methods, and Open Problems* 4:463–480.
- Çelik S, Maglaras C (2008) Dynamic pricing and lead-time quotation for a multiclass make-to-order queue. *Management Science* 54(6):1132–1146.
- Chen M, Sun P, Xiao Y (2020) Optimal monitoring schedule in dynamic contracts. *Operations Research* 68(5):1285–1314.
- Dai Y, Chao X (2013) Salesforce contract design and inventory planning with asymmetric risk-averse sales agents. *Operations Research Letters* 41(1):86–91.
- Ding Y, Park E, Nagarajan M, Grafstein E (2019) Patient prioritization in emergency department triage systems: An empirical study of the Canadian triage and acuity scale (ctas). *Manufacturing & Service Operations Management* 21(4):723–741.
- Gopalakrishnan R, Doroudi S, Ward AR, Wierman A (2016) Routing and staffing when servers are strategic. *Operations Research* 64(4):1033–1050.
- Hansen LP, Sargent TJ, Turmuhambetova G, Williams N (2006) Robust control and model misspecification. *Journal of Economic Theory* 128(1):45–90.
- Hasija S, Pinker EJ, Shumsky RA (2008) Call center outsourcing contracts under information asymmetry. *Management Science* 54(4):793–807.
- Huang J, Gurvich I (2018) Beyond heavy-traffic regimes: Universal bounds and controls for the single-server queue. *Operations Research* 66(4):1168–1188.
- Ibrahim R, Whitt W (2009) Real-time delay estimation in overloaded multiserver queues with abandonments. *Management Science* 55(10):1729–1742.
- Ibrahim R, Whitt W (2011) Wait-time predictors for customer service systems with time-varying demand and capacity. *Operations Research* 59(5):1106–1118.

-
- Jiang H, Pang Z, Savin S (2012) Performance-based contracts for outpatient medical services. *Manufacturing & Service Operations Management* 14(4):654–669.
- Kim J, Randhawa RS (2018) The value of dynamic pricing in large queueing systems. *Operations Research* 66(2):409–425.
- Li H, Zhang H, Fine CH (2013) Dynamic business share allocation in a supply chain with competing suppliers. *Operations Research* 61(2):280–297.
- Liberzon D (2011) *Calculus of variations and optimal control theory: a concise introduction* (Princeton university press).
- Lim AE, Shanthikumar JG (2007) Relative entropy, exponential utility, and robust dynamic pricing. *Operations Research* 55(2):198–214.
- Olver PJ (2012) The calculus of variations. *Applied Mathematics Lecture Notes. Sec 21*(4).
- Ouyang H, Tanik Argon N, Ziya S (2022) Assigning priorities (or not) in service systems with nonlinear waiting costs. *Management Science* 68(2):1233–1255.
- Plambeck EL, Zenios SA (2000) Performance-based incentives in a dynamic principal-agent model. *Manufacturing & Service Operations Management* 2(3):240–263.
- Plambeck EL, Zenios SA (2003) Incentive efficient control of a make-to-stock production system. *Operations Research* 51(3):371–386.
- Poster WR (2007) Who’s on the line? Indian call center agents pose as Americans for US-outsourced firms. *Industrial Relations: A Journal of Economy and Society* 46(2):271–304.
- Ren ZJ, Zhou YP (2008) Call center outsourcing: Coordinating staffing level and service quality. *Management Science* 54(2):369–383.
- Sannikov Y (2008) A continuous-time version of the principal-agent problem. *The Review of Economic Studies* 75(3):957–984.
- Spear SE, Srivastava S (1987) On repeated moral hazard with discounting. *The Review of Economic Studies* 54(4):599–617.
- Sun P, Tian F (2018) Optimal contract to induce continued effort. *Management Science* 64(9):4193–4217.
- Tian F, Sun P, Duenyas I (2021) Optimal contract for machine repair and maintenance. *Operations Research* 69(3):916–949.
- Wang S, Sun P, de Véricourt F (2016) Inducing environmental disclosures: A dynamic mechanism design approach. *Operations Research* 64(2):371–389.
- Xu K (2023) Drift method: from stochastic networks to machine learning. URL: <https://web.stanford.edu/~kuangxu/papers/driftmethod.pdf>. Last visited on 3(09).
- Zhan D, Ward AR (2019) Staffing, routing, and payment to trade off speed and quality in large service systems. *Operations Research* 67(6):1738–1751.

Zhang H, Kong G, Rajagopalan S (2018) Contract design by service providers with private effort. *Management Science* 64(6):2672–2689.

Zhong Y, Gopalakrishnan R, Ward AR (2023) Behavior-aware queueing: The finite-buffer setting with many strategic servers. *Operations Research* .

Zhou YP, Ren ZJ (2010) Service outsourcing. *Wiley Encyclopedia of Operations Research and Management Science* .

E-Companion

EC.1. Proofs of the Main Results

This part of the e-companion collects the proofs of the main results in the paper. We omit the proof of Proposition 3 because it is a fairly routine application of Itô's formula. The proofs of all the auxiliary results can be found in §EC.4.

Proof of Proposition 1. Towards establishing part (i), it is useful to define $w := v'$ and note that $\theta^* = v'/\alpha$. Thus, we can rewrite (9) as

$$\frac{\sigma^2}{2}w'(z) - \frac{(w(z))^2}{2\alpha} + k(z) = \rho^*.$$

For notational simplicity, in the proof of part (i), we further stipulate that $\sigma^2 = 2$ and $\alpha = 1/2$, so that we need to prove the existence of a pair (w, ρ^*) that solves

$$w'(z) - (w(z))^2 + k(z) = \rho^* \tag{EC.1}$$

with $w(0) = 0$, $w'(z) > 0$ for all $z > 0$, and $w(z) \leq C\sqrt{k(z)}$ for some $C > 0$ and all sufficiently large z . The preceding stipulations are made without loss of generality; that is, the arguments used below are still valid for general values of σ^2 and α . Now, to conduct the existence proof, we consider the following family of first-order differential equations parameterized by ρ :

$$w'_\rho(z) - (w_\rho(z))^2 + k(z) = \rho \quad \text{with} \quad w_\rho(0) = 0. \tag{EC.2}$$

Within this family, we intend to show that there exists a unique ρ^* such that the pair (w_{ρ^*}, ρ^*) solves (EC.1) subject to the stated requirements. To that end, define

$$z_\rho := \inf \left\{ z \geq 0 : \lim_{x \uparrow z} w_\rho(x) = \pm\infty \right\}.$$

Also, let $\mathcal{C}_\rho := \{z \in [0, z_\rho) : w'_\rho(z) \leq 0\}$ and define sets \mathcal{L} and \mathcal{U} that bisect the non-negative real line in the following way:

$$\mathcal{L} := \{\rho \geq 0 : \mathcal{C}_\rho \neq \emptyset\} \quad \text{and} \quad \mathcal{U} := \{\rho \geq 0 : \mathcal{C}_\rho = \emptyset\}.$$

To complete the proof of part (i), we need the following lemmas whose proofs are deferred to §EC.4.

LEMMA EC.1. *For any $\rho_1 < \rho_2$, we have that $w_{\rho_1}(z) < w_{\rho_2}(z)$ for $z \in [0, z_{\rho_1} \wedge z_{\rho_2})$.*

LEMMA EC.2. *For all $\rho \in \mathcal{L}$, w_ρ is quasi-concave and $\lim_{z \rightarrow \infty} w_\rho(z) = -\infty$.*

LEMMA EC.3. *Both \mathcal{L} and \mathcal{U} are nonempty.*

LEMMA EC.4. *Let $\rho^* := \sup \mathcal{L}$. Then $\rho^* \in \mathcal{U}$, and $w'_{\rho^*}(z) > 0$ for all $z \geq 0$.*

LEMMA EC.5. *For $\rho \in (0, \rho^*)$, we have that $w_\rho(z) \leq \sqrt{2k(z) + a}$ for some positive constant a .*

Continuing our proof of Proposition 1, Suppose that w_{ρ^*} does not satisfy the stated growth condition. Then, there exists some \tilde{z} such that

$$w_{\rho^*}(\tilde{z}) \geq \sqrt{2k(\tilde{z}) + a} + \epsilon.$$

On the other hand, by Lemma EC.5 we have that

$$w_\rho(\tilde{z}) \leq \sqrt{2k(\tilde{z}) + a} \quad \text{for all } \rho \in (0, \rho^*).$$

This, however, leads to a contradiction because $w_\rho(z)$ is jointly continuous in ρ and z . Hence, part (i) of the proposition follows.

To establish part (ii), let θ be an admissible rate-control strategy. A routine application of Itô's formula gives the following identity

$$\mathbb{E}[v(Z(t))] - v(Z(0)) = \mathbb{E} \left[\int_0^t \Gamma v(Z(u)) du + v'(0)L(t) \right],$$

where Γ is an operator and defined as $\Gamma f(z) = (\sigma^2/2)f''(z) - \theta(z)f'(z)$. Since $v'(0) = 0$ and v satisfies (9), we have that

$$\mathbb{E}[v(Z(t))] - v(Z(0)) \leq \mathbb{E} \left[\rho^* t - \int_0^t c(\theta(u)) du - \int_0^t k(Z(u)) du \right].$$

Dividing both sides by t and taking the limsup, we obtain

$$\rho^* \geq \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t c(\theta(u)) du + \int_0^t k(Z(u)) du \right] + \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E}[v(Z(t))].$$

Next, we intend to establish

$$\limsup_{t \rightarrow \infty} t^{-1} \mathbb{E}[v(Z(t))] = 0. \tag{EC.3}$$

To that end, note from part (i) that v has a polynomial growth rate. Thus, to establish (EC.3), it suffices to argue that $\limsup_{t \rightarrow \infty} t^{-1} \mathbb{E}[Z^q(t)] = 0$ for all $q > 0$. For this purpose, we need to construct a non-negative process \tilde{Z} such that

$$Z \leq_{s.t.} \tilde{Z} \tag{EC.4}$$

and \tilde{Z} has a steady-state distribution whose moments of all orders are finite. By the admissibility of θ , we know that $\theta(z)$ is strictly increasing in z , so there exists some $\bar{z} > 0$ and $\chi > 0$ such that $\theta(z) > \chi$ for all $z \geq \bar{z}$. This means that if we define \tilde{Z} as

$$\tilde{Z}(t) = Z(0) + \int_0^t b(\tilde{Z}(u)) du + \sigma B(t) + L(t),$$

where $b(z) := 0 \cdot 1_{\{0 \leq z < \bar{z}\}} - \chi \cdot 1_{\{z \geq \bar{z}\}}$ and L is a pushing process that uses a minimum effort to keep \tilde{Z} non-negative, then (EC.4) holds. Moreover, \tilde{Z} defined in this way is a piecewise-linear diffusion process with reflection at the origin. Using the techniques in Browne et al. (1995), we can write down its steady-state distribution and conclude that the distribution has finite moments of all orders. Taken together, we arrive at (EC.3). Hence,

$$\rho^* \geq \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t c(\theta(u)) du + \int_0^t k(Z(u)) du \right].$$

The desired optimality of θ^* follows immediately by noting that all the foregoing inequalities become equalities when $\theta = \theta^*$. This concludes the proof of part (ii). \square

Proof of Proposition 2. The maximizing θ on the left-hand of (10) can be expressed as $\theta = -\frac{U'}{\alpha}$. Substituting this expression back into (10) “removes” the max operator, yielding

$$\frac{\sigma^2}{2} U''(z) + \frac{1}{2\alpha} (U'(z))^2 + S(z) + \frac{\gamma\sigma^2}{2} (U'(z))^2 = \eta.$$

If the PC constraint is not binding, meaning that η is strictly positive, the principal can improve her objective without violating the agent’s PC constraint by choosing $\tilde{S}(z) := S(z) - \eta$. This implies that the payment plan, which induces θ at a minimum cost, must possess the following property:

$$\frac{\sigma^2}{2} U''(z) + \frac{1}{2\alpha} (U'(z))^2 + S(z) + \frac{\gamma\sigma^2}{2} (U'(z))^2 = 0. \quad (\text{EC.5})$$

Replacing U' in (EC.5) with $-\alpha\theta$ yields (11), as desired. \square

Proof of Theorem 1. To save us some ink, we will assume throughout the proof that $\gamma = 0$, resulting in $\tilde{\alpha}_\gamma = \alpha$. For the more general case where $\gamma \neq 0$, we can easily substitute all instances of α in the proof with $\tilde{\alpha}_\gamma$ without impacting the validity of any arguments presented.

Starting with part (i), we observe that (13) is, in essence, a first-order linear differential equation (by viewing V' as the unknown function). Therefore, we can apply the standard formula for first-order differential equations with the condition $V'(0) = 0$ to get

$$V'(z) = \frac{2}{\sigma^2} e^{\frac{2}{\sigma^2}\phi(z)} \int_0^z e^{-\frac{2}{\sigma^2}\phi(x)} \left(\kappa - \frac{\sigma^2}{2} \alpha \theta'(x) + \frac{\alpha}{2} \theta^2(x) - k(x) \right) dx, \quad (\text{EC.6})$$

where we have defined $\phi(z) := \int_0^z \theta(x) dx$. We intend to find some κ that, when plugged into the right-hand side, causes $V'(z)$ to grow polynomially as $z \rightarrow \infty$. This requirement entails that

$$\int_0^z e^{-\frac{2}{\sigma^2}\phi(x)} \left(\kappa - \frac{\sigma^2}{2} \alpha \theta'(x) + \frac{\alpha}{2} \theta^2(x) - k(x) \right) dx \rightarrow 0 \quad \text{as } z \rightarrow \infty.$$

The desired κ value, denoted as $\kappa(\theta)$, must be such that

$$\kappa(\theta) \cdot \int_0^\infty e^{-\frac{2}{\sigma^2}\phi(x)} dx = \int_0^\infty e^{-\frac{2}{\sigma^2}\phi(x)} \left(\frac{\sigma^2}{2} \alpha \theta'(x) - \frac{\alpha}{2} \theta^2(x) + k(x) \right) dx.$$

Rearranging yields (15). Now, letting $\kappa = \kappa(\theta)$ in (EC.6) yields

$$V'(z) = \frac{2}{\sigma^2} e^{\frac{2}{\sigma^2} \phi(z)} \int_z^\infty e^{-\frac{2}{\sigma^2} \phi(x)} \left(\frac{\sigma^2}{2} \alpha \theta'(x) - \frac{\alpha}{2} \theta^2(x) + k(x) \right) dx.$$

A direct application of L'Hospital's rule reveals that the right-hand side has a growth rate that is of the same order as

$$p(z) := \frac{\frac{\sigma^2}{2} \alpha \theta'(z) - \frac{\alpha}{2} \theta^2(z) + k(z)}{\theta(z)}.$$

By our assumption, $S(z)$ is admissible, so the numerator grows polynomially as $z \rightarrow \infty$, while the denominator is bounded below by some positive constant χ for all sufficiently large z . This implies that p has a polynomial growth rate, implying that V' and V exhibit polynomial growth as well. To show that $\kappa(\theta)$ is the long-run average cost of the principal, we must argue that $\limsup_{t \rightarrow \infty} t^{-1} \mathbb{E}[V(Z(t))] = 0$. The rest of the proof mimics those steps taken towards establishing part (ii) of Proposition 1. For completeness, we spell out the details below. Specifically, we need to find some non-negative process \tilde{Z} such that $Z \leq_{s.t.} \tilde{Z}$ and \tilde{Z} has a steady-state distribution whose moments of all orders are finite. By the hypothesis, θ is induced by an admissible payment plan, so there exist some $\bar{z} > 0$ and $\chi > 0$ such that $\theta(z) > \chi$ for all $z \geq \bar{z}$. This means we can choose \tilde{Z} to be one that satisfies the following dynamics:

$$\tilde{Z}(t) = Z(0) + \int_0^t b(\tilde{Z}(u)) du + \sigma B(t) + L(t),$$

where $b(z) := 0 \cdot 1_{\{0 \leq z < \bar{z}\}} - \chi \cdot 1_{\{z \geq \bar{z}\}}$ and the rest of the quantities are defined as before. Note that \tilde{Z} defined in this way is a piecewise-linear diffusion with reflection at the origin. Using the techniques in Browne et al. (1995), we can write down its steady-state distribution and conclude that the distribution has finite moments of all orders. This concludes the proof of part (i).

To establish part (ii), note that not only θ and ϕ determine each other, but $\phi' = \theta$ and $\phi'' = \theta'$. Thus, from (15) we can see that minimizing the right-hand side of (15) over all possible θ with the condition $\theta(0) = 0$ is equivalent to seeking some ϕ that minimizes the ratio

$$\frac{\int_0^\infty G(x, \phi(x), \phi'(x), \phi''(x)) dx}{\int_0^\infty H(\phi(x)) dx}, \quad (\text{EC.7})$$

subject to the conditions $\phi(0) = 0$ and $\phi'(0) = 0$, where we defined G and H as

$$G(x, p, q, r) := e^{-\frac{2}{\sigma^2} p} \left(\frac{\sigma^2}{2} \alpha r - \frac{\alpha}{2} q^2 + k(x) \right) \quad \text{and} \quad H(p) := e^{-\frac{2}{\sigma^2} p},$$

respectively. Let ϑ be an arbitrary, twice-differentiable test function that, along with its derivative function, vanishes at the boundaries. Now, let

$$F(\epsilon) := \frac{\int_0^\infty G(x, \phi(x) + \epsilon\vartheta(x), \phi'(x) + \epsilon\vartheta'(x), \phi''(x) + \epsilon\vartheta''(x))dx}{\int_0^\infty H(\phi(x) + \epsilon\vartheta(x))dx}.$$

Using the first-order condition, we deduce that if ϕ is a critical point of (EC.7), then $\frac{dF}{d\epsilon}|_{\epsilon=0} = 0$.

Define

$$g := \int_0^\infty G(x, \phi(x), \phi'(x), \phi''(x))dx \quad \text{and} \quad h := \int_0^\infty H(\phi(x))dx.$$

Using $\frac{dF}{d\epsilon}|_{\epsilon=0} = 0$ and with some rearrangement, we can deduce that

$$\frac{\int_0^\infty \left[h \left(\frac{\partial G}{\partial \phi} - \frac{d}{dx} \frac{\partial G}{\partial \phi'} + \frac{d^2}{dx^2} \frac{\partial G}{\partial \phi''} \right) - g \left(\frac{\partial H}{\partial \phi} \right) \right] \vartheta(x) dx}{h^2} = 0,$$

where we have omitted the arguments of the derivative functions. By the fundamental lemma of calculus of variations, the part of the integrand in the angle brackets is zero, yielding the corresponding Euler-Lagrange equation:

$$h \left(\frac{\partial G}{\partial \phi} - \frac{d}{dx} \frac{\partial G}{\partial \phi'} + \frac{d^2}{dx^2} \frac{\partial G}{\partial \phi''} \right) - g \left(\frac{\partial H}{\partial \phi} \right) = 0. \quad (\text{EC.8})$$

Using the definitions of G and H , we can deduce that

$$\begin{aligned} \frac{\partial G}{\partial \phi} &= -\frac{2}{\sigma^2}G, & \frac{d}{dx} \frac{\partial G}{\partial \phi'} &= \frac{2\alpha}{\sigma^2}e^{-\frac{2}{\sigma^2}\phi}(\phi')^2 - \alpha e^{-\frac{2}{\sigma^2}\phi}\phi'', \\ \frac{d^2}{dx^2} \frac{\partial G}{\partial \phi''} &= \frac{2\alpha}{\sigma^2}e^{-\frac{2}{\sigma^2}\phi}(\phi')^2 - \alpha e^{-\frac{2}{\sigma^2}\phi}\phi'' & \text{and} & \quad \frac{\partial H}{\partial \phi} = -\frac{2}{\sigma^2}e^{-\frac{2}{\sigma^2}\phi}. \end{aligned} \quad (\text{EC.9})$$

Plugging (EC.9) into (EC.8) and with some arrangement, we have that

$$h \left(\frac{\sigma^2}{2}\alpha\phi''(x) - \frac{\alpha}{2}(\phi'(x))^2 + k(x) \right) - g = 0.$$

Letting $\tilde{\kappa} := g/h$ in the preceding and noting that $\theta = \phi'$, we reach (16), as desired. \square

Proof of Proposition 4. Let Z^* denote the state process when the principal employs S^* , and the agent utilizes the rate control induced by S^* . For all admissible S ,

$$\begin{aligned} & \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t S(Z(u))du + \int_0^t k(Z(u))du - \int_0^t \lambda(Z(u))S(Z(u))du \right] \\ & \geq \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t S^*(Z^*(u))du + \int_0^t k(Z^*(u))du - \int_0^t \lambda(Z^*(u))S^*(Z^*(u))du \right]. \end{aligned}$$

Rearranging terms yields

$$\begin{aligned} \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t S^*(Z^*(u)) du + \int_0^t k(Z^*(u)) du \right] &\leq \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t S(Z(u)) du + \int_0^t k(Z(u)) du \right] \\ &+ \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t \lambda(Z^*(u)) S^*(Z^*(u)) du \right] - \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t \lambda(Z(u)) S(Z(u)) du \right]. \end{aligned} \quad (\text{EC.10})$$

By the complementary slackness condition, we see that

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t \lambda(Z^*(u)) S^*(Z^*(u)) du \right] = 0. \quad (\text{EC.11})$$

Using the primal and dual feasibility constraints gives

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t \lambda(Z(u)) S(Z(u)) du \right] \geq 0. \quad (\text{EC.12})$$

Combining (EC.10)–(EC.12) yields

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t S^*(Z^*(u)) du + \int_0^t k(Z^*(u)) du \right] \leq \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t S(Z(u)) du + \int_0^t k(Z(u)) du \right]$$

for all admissible S . The proof is therefore complete. \square

Proof of Theorem 2. To prove part (i), we introduce the following proposition:

PROPOSITION EC.1. *There exists a continuously differentiable function f and a constant ρ that satisfy*

$$\min_{\theta \in [0, b]} \left\{ \frac{\sigma^2}{2} f'(z) - \theta f(z) + k(z) + c(\theta) \right\} = \rho$$

with $f(0) = 0$, $f'(z) > 0$ for all $z \geq 0$, and $f(z)$ grows polynomially as $z \rightarrow \infty$.

This proposition is similar to Proposition 1, with the only difference being that in Proposition EC.1, θ is restricted to the range $[0, b]$, where we recall $b = \sqrt{2\eta/\alpha}$. Proposition EC.1 can be proven in almost the same fashion as in Proposition EC.1, so we omit the details of the proof.

Proposition EC.1 implies the existence of some ζ such that

$$\frac{\sigma^2}{2} f'(z) - \frac{1}{2\alpha} f^2(z) + k(z) = \rho \quad \text{for } z \in [0, \zeta]$$

and

$$\frac{\sigma^2}{2} f'(z) - b f(z) + \frac{\alpha}{2} b^2 + k(z) = \rho \quad \text{for } z \in [\zeta, \infty).$$

In particular, ζ is such that $f(\zeta) = \alpha b$. Now, set $\tilde{\kappa} = \rho$, $\bar{z} = \zeta$, and let $\theta = \alpha f$ on $[0, \bar{z})$ and $\theta \equiv b$ on $[\bar{z}, \infty)$, and $V'_\theta = f - \alpha b$ on $[\bar{z}, \infty)$. It is directly verifiable that the resulting quadruple $(\theta, \tilde{\kappa}, \bar{z}, V_\theta)$ meets all requirements arising from the stationarity condition, the primal and dual feasibility constraints, and the complementary slackness condition. The proof of part (i) is therefore complete.

To prove part (ii), it suffices to let $\lambda = 2V_\theta/(\sigma^2\alpha)$ and note that V_θ satisfies equation (33). \square

Proof of Proposition 5. Consider two positive values, b_1 and b_2 . Without loss of generality, assume that $b_1 < b_2$. In this context, $\rho^*(b_1)$ represents the optimal objective value of Problem (34) when b is set to b_1 , and for convenience, we refer to the resulting problem as \mathcal{P}_1 . Similarly, $\rho^*(b_2)$ corresponds to the optimal objective value of the same problem when b is assigned the value b_2 , and we refer to the problem as \mathcal{P}_2 . For an arbitrarily fixed $\beta \in (0, 1)$, define $b_3 := \beta b_1 + (1 - \beta)b_2$ and let the corresponding stochastic control problem be denoted as \mathcal{P}_3 . We aim to show that

$$\rho^*(b_3) \leq \beta \rho^*(b_1) + (1 - \beta) \rho^*(b_2). \quad (\text{EC.13})$$

For this purpose, let θ_1 denote the optimal policy for \mathcal{P}_1 and Z_1 the state process under the control policy θ_1 . Similarly, let θ_2 denote the optimal policy for \mathcal{P}_2 and Z_2 the state process under θ_2 . Let $\hat{\theta} := \beta \theta_1 + (1 - \beta) \theta_2$ and \hat{Z} the state process under $\hat{\theta}$. Clearly, $\hat{\theta}$ is a feasible control policy for \mathcal{P}_3 . By the definition of $\rho^*(b_3)$, we have

$$\rho^*(b_3) \leq \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t c(\hat{\theta}(u)) du + k(\hat{Z}(u)) du \right].$$

Thus, to prove (EC.13), it suffices to show that

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t c(\hat{\theta}(u)) du \right] \leq \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t (\beta c(\theta_1(u)) + (1 - \beta) c(\theta_2(u))) du \right] \quad (\text{EC.14})$$

and

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t k(\hat{Z}(u)) du \right] \leq \beta \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t k(Z_1(u)) du \right] + (1 - \beta) \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} [k(Z_2(u)) du]. \quad (\text{EC.15})$$

The inequality in (EC.14) immediately follows from the definition of $\hat{\theta}$ and the convexity of the effort cost function c . To establish (EC.15), we begin by defining

$$m_1 := \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t k(Z_1(u)) du \right] \quad \text{and} \quad m_2 := \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} [k(Z_2(u)) du].$$

Then, there exists some twice differentiable function ν_1 , in conjunction with m_1 , satisfies

$$\frac{\sigma^2}{2} \nu_1''(z) - \theta_1(z) \nu_1'(z) + k(z) = m_1 \quad (\text{EC.16})$$

subject to the conditions $\nu_1'(0) = 0$, $\nu_1''(z) > 0$ for all $z \geq 0$, and that $\nu_1'(z)$ grows polynomially as $z \rightarrow \infty$. Similarly, there exists some twice differentiable function ν_2 , along with m_2 , satisfies

$$\frac{\sigma^2}{2} \nu_2''(z) - \theta_2(z) \nu_2'(z) + k(z) = m_2 \quad (\text{EC.17})$$

subject to the conditions $\nu_2'(0) = 0$, $\nu_2''(z) > 0$ for all $z \geq 0$, and that $\nu_2'(z)$ grows polynomially as $z \rightarrow \infty$. We can show that

$$\nu_1'(z) > \nu_2'(z) \quad \text{for all } z > 0. \quad (\text{EC.18})$$

(The detailed proof of (EC.18) is given after the end of the current proof.)

Now, let $\nu_3 := \beta\nu_1 + (1 - \beta)\nu_2$. By utilizing (EC.16)–(EC.17), we can deduce that

$$\frac{\sigma^2}{2}\nu_3''(z) - \check{\theta}(z)\nu_3'(z) + k(z) = \beta m_1 + (1 - \beta)m_2,$$

where

$$\check{\theta}(z) := \frac{\beta\nu_1'(z)\theta_1(z) + (1 - \beta)\nu_2'(z)\theta_2(z)}{\beta\nu_1'(z) + (1 - \beta)\nu_2'(z)}.$$

This implies that the long-term average congestion cost under the control policy $\check{\theta}$ equals $\beta m_1 + (1 - \beta)m_2$. Thus, to establish

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\int_0^t k(\hat{Z}(u)) du \right] \leq \beta m_1 + (1 - \beta)m_2,$$

it suffices to demonstrate that $\check{\theta}(z) \leq \hat{\theta}(z)$ holds for all $z \geq 0$. This, however, follows directly from their respective definitions and (EC.18). The proof is therefore complete. \square

Proof of Inequality (EC.18): Let v_1 and v_2 denote the relative value functions associated with \mathcal{P}_1 and \mathcal{P}_2 , respectively. Consequently, we have

$$\min_{\theta \in [0, b_1]} \left\{ \frac{\sigma^2}{2} v_1''(z) - \theta v_1'(z) + k(z) + c(\theta) \right\} = \rho^*(b_1) \quad (\text{EC.19})$$

and

$$\min_{\theta \in [0, b_2]} \left\{ \frac{\sigma^2}{2} v_2''(z) - \theta v_2'(z) + k(z) + c(\theta) \right\} = \rho^*(b_2). \quad (\text{EC.20})$$

Additionally, we know that $\theta_1(z) = \max\{v_1'(z)/\alpha, b_1\}$ and $\theta_2(z) = \max\{v_2'(z)/\alpha, b_2\}$. Let ζ_1 (ζ_1') denote the point at which the function value of θ_1 (θ_2) first reaches the upper bound b_1 . Since $b_1 < b_2$, $\rho^*(b_1) > \rho^*(b_2)$. By applying the comparison principle for first-order differential equations to (EC.19) and (EC.20), we conclude that $\zeta_1 < \zeta_1'$, and $v_1'(z) > v_2'(z)$ for all $z \in (0, \zeta_1']$. Consequently, $\theta_1(z) > \theta_2(z)$ for all $z \in (0, \zeta_1')$. Applying the comparison principle for first-order differential equations again, but this time to (EC.16) and (EC.17), we deduce that $\nu_1'(z) > \nu_2'(z)$ for all $z \in (0, \zeta_1']$.

Hence, if (EC.18) is not true, then there must exist some $\tilde{z} > \zeta_1'$ such that

$$\nu_1'(\tilde{z}) = \nu_2'(\tilde{z}) \quad \text{and} \quad \nu_2''(\tilde{z}) > \nu_1''(\tilde{z}). \quad (\text{EC.21})$$

We define $\Delta(z) := \nu_2'(z) - \nu_1'(z)$. Then (EC.21) implies that $\Delta(\tilde{z}) = 0$ and $\Delta'(\tilde{z}) > 0$. We intend to demonstrate that, under the hypothesis (EC.21), Δ will exhibit exponential growth. This leads to a desired contradiction since both ν_1' and ν_2' have polynomial growth.

Since $\tilde{z} > \zeta_1'$, we must have $\theta_1(\tilde{z}) = b_1$ and $\theta_2(\tilde{z}) := \tilde{b} \in (b_1, b_2]$. Subtracting (EC.16) from (EC.17) and using the definition of $\Delta(z)$, we obtain

$$\begin{aligned} \frac{\sigma^2}{2} \Delta'(z) &= \theta_2(z)\nu_2'(z) - \theta_1(z)\nu_1'(z) + m_2 - m_1 \\ &= \theta_2(z)\nu_2'(z) - b_1\nu_1'(z) + m_2 - m_1 \\ &= b_1\Delta(z) + (\theta_2(z) - b_1)\nu_2'(z) + m_2 - m_1 \end{aligned} \quad (\text{EC.22})$$

for all $z \geq \tilde{z}$. By our hypothesis, we have

$$B := (\tilde{b} - b_1)\nu'_2(\tilde{z}) + m_2 - m_1 = \frac{\sigma^2}{2}\Delta'(\tilde{z}) - b_1\Delta(\tilde{z}) = \frac{\sigma^2}{2}\Delta'(\tilde{z}) > 0.$$

Noting that both θ_2 and ν'_2 are non-decreasing, we conclude that

$$(\theta_2(z) - b_1)\nu'_2(z) + m_2 - m_1 \geq B \quad \text{for all } z \geq \tilde{z}.$$

Combining this with (EC.22) yields

$$\frac{\sigma^2}{2}\Delta'(z) \geq b_1\Delta(z) + B \quad \text{for all } z \geq \tilde{z},$$

from which it follows that

$$\Delta(z) \geq \frac{B}{b_1} \left(e^{\frac{2b_1}{\sigma^2}(z-\tilde{z})} - 1 \right) \quad \text{for all } z \geq \tilde{z}.$$

This shows that Δ exhibits exponential growth, which is a contradiction. The proof is thus complete.

□

EC.2. Derivation of Equation (26)

Define

$$\tilde{G}(x, p, q, r) := e^{-\frac{2}{\sigma^2}p} \left(\frac{\sigma^2}{2}\alpha r - \frac{\alpha}{2}q^2 + k(x) - \lambda(x) \left(\frac{\sigma^2}{2}\alpha r - \frac{\alpha}{2}q^2 + \eta \right) \right).$$

Note that we still have the Euler-Lagrange equation:

$$h \left(\frac{\partial \tilde{G}}{\partial \phi} - \frac{d}{dx} \frac{\partial \tilde{G}}{\partial \phi'} + \frac{d^2}{dx^2} \frac{\partial \tilde{G}}{\partial \phi''} \right) - g \left(\frac{\partial H}{\partial \phi} \right) = 0.$$

Similar to equation (EC.9), using the definitions of \tilde{G} and H (where the definition of H can be found in §EC.1), we can deduce that

$$\begin{aligned} \frac{\partial H}{\partial \phi} &= -\frac{2}{\sigma^2}e^{-\frac{2}{\sigma^2}\phi}, & \frac{\partial \tilde{G}}{\partial \phi} &= -\frac{2}{\sigma^2}\tilde{G}, \\ \frac{d}{dx} \frac{\partial \tilde{G}}{\partial \phi'} &= \frac{2\alpha}{\sigma^2}e^{-\frac{2}{\sigma^2}\phi}(\phi')^2 - \alpha e^{-\frac{2}{\sigma^2}\phi}\phi'' + \alpha e^{-\frac{2}{\sigma^2}\phi}(\lambda'(x)\phi' + \lambda(x)\phi'') - \frac{2\alpha}{\sigma^2}e^{-\frac{2}{\sigma^2}\phi}(\phi')^2\lambda(x), & \text{(EC.23)} \\ \text{and } \frac{d^2}{dx^2} \frac{\partial \tilde{G}}{\partial \phi''} &= \frac{d}{dx} \frac{\partial \tilde{G}}{\partial \phi'} + \alpha\lambda'(x)\phi'e^{-\frac{2}{\sigma^2}\phi} - \frac{\alpha\sigma^2}{2}\lambda''(x)e^{-\frac{2}{\sigma^2}\phi}. \end{aligned}$$

Plugging equation (EC.23) back and with some arrangement, we have that

$$h \left(\frac{\sigma^2}{2}\alpha\theta'(z) - \frac{\alpha}{2}\theta^2(z) + k(z) - \lambda(z) \left(\frac{\sigma^2}{2}\alpha\theta'(z) - \frac{\alpha}{2}\theta^2(z) + \eta \right) + \frac{\sigma^2}{2}\alpha \left(\frac{\sigma^2}{2}\lambda''(z) - \theta(z)\lambda'(z) \right) \right) - g = 0.$$

Letting $\tilde{\kappa} := g/h$ in the preceding and noting that $\theta = \phi'$, we reach equation (26).

EC.3. Justifying the Use of an RBM to Model Wait-Time Dynamics

The demonstration below is largely adapted from the paper by Huang and Gurvich (2018), where the authors show that a universally near-optimal control policy for a single-server queue can be obtained via an intuitive Brownian control problem.

Consider an M/G/1 queue where a single server processes Poisson arrivals with service requirements distributed according to a general distribution. Let $A(t)$ denote the number of arrivals by time t , and let $\{\xi_i; t \geq 1\}$ denote the service requirements of consecutive jobs that enter the system. The process $A := \{A(t); t \geq 0\}$ is a Poisson process with some rate λ , and $\{\xi_i; t \geq 1\}$ are independent random variables that follow a common distribution G with a mean of $1/\mu$.

The workload, which corresponds to the virtual wait time in the M/G/1 queue, can be represented as a process that evolves according to the following equation:

$$\begin{aligned} W(t) &= W(0) + \sum_{i=1}^{A(t)} \xi_i - (t - I(t)) \\ &= W(0) + (\varrho - 1)t + I(t) + \left(\sum_{i=1}^{A(t)} \xi_i - \varrho t \right), \end{aligned}$$

where $\varrho := \lambda/\mu$, and $I(t)$ denotes the cumulative idle time of the server by time t . At each time t , the compound Poisson input of the M/G/1 queue can be approximated by a central limit theorem, given by:

$$\sum_{i=1}^{A(t)} \xi_i \approx \varrho t + \mathcal{N}(t),$$

where $\mathcal{N}(t)$ is a zero-mean normal random variable with variance $\lambda \mathbb{E}[\xi_1^2]t$. As a result, it is reasonable to replace the input process with $\varrho t + \sqrt{\lambda \mathbb{E}[\xi_1^2]}B(t)$, where $B := B(t), t \geq 0$ is a standard Brownian motion. This approximation leads to the Brownian queue:

$$Z(t) = Z(0) + (\varrho - 1)t + L(t) + \sqrt{\lambda \mathbb{E}[\xi_1^2]}B(t),$$

where W and I have been replaced by Z and L , respectively, to emphasize that they are approximations. In the case of exponentially distributed service requirements, we have that $\sqrt{\lambda \mathbb{E}[\xi_1^2]} = \sqrt{2\lambda/\mu^2}$. Assuming $\lambda = \mu$ and letting $\sigma := \sqrt{\lambda \mathbb{E}[\xi_1^2]}$, we obtain:

$$Z(t) = Z(0) + \sigma B(t) + L(t),$$

which is the uncontrolled version of equation (1).

EC.4. Proofs of Auxiliary Results

Proof of Lemma EC.1. The desired conclusion follows directly from the comparison principle for first-order differential equations. \square

Proof of Lemma EC.2. Since $w'_\rho(0) = \rho \geq 0$, it suffices to show that w_ρ does not have strict local minima. From (EC.2), we see that $w''_\rho(z) = -k'(z) < 0$ whenever $w'_\rho(z) = 0$. Hence, w_ρ does not have strict local minima. For the second part of the lemma, it suffices to demonstrate that w_ρ cannot be bounded from below by a finite constant. If so, then there exists a sequence $\{z_l\}$ such that $\lim_{l \rightarrow \infty} w'_\rho(z_l) = 0$ and $\lim_{l \rightarrow \infty} w_\rho(z_l) = b$ for some $b \in \mathbb{R}$. But this leads to a contradiction in view of (EC.2) and the fact that $\lim_{l \rightarrow \infty} k(z_l) = \infty$. \square

Proof of Lemma EC.3. By definition, we have that $0 \in \mathcal{L}$. To show that \mathcal{U} is not empty, we first claim that the set $\bar{\mathcal{U}}$ defined as $\bar{\mathcal{U}} := \{\rho \geq 0 : w_\rho(1) \geq 1\}$ is not empty. Since

$$w_\rho(1) = \rho + \int_0^1 (w_\rho(x))^2 dx - \int_0^1 k(x) dx \geq \rho - \int_0^1 k(x) dx,$$

we know that $\lim_{\rho \rightarrow \infty} w_\rho(1) = \infty$. It follows that $\bar{\mathcal{U}}$ is nonempty, establishing the claim. Also, if letting $\rho_1 := \inf \bar{\mathcal{U}}$, then $\rho_1 < \infty$. Suppose, for the sake of contradiction, that $\mathcal{U} = \emptyset$. Then for any $\rho > \rho_1$, there exists $z_{1,\rho} > 1$ such that $w_\rho(z_{1,\rho}) = 1$. For $z \in [1, z_{1,\rho}]$, we have that $(w_\rho(z))^2 \geq w_\rho(z)$, and therefore $w_\rho(z) \geq \underline{w}_\rho(z)$, where \underline{w}_ρ solves

$$\underline{w}'_\rho(z) - \underline{w}_\rho(z) + k(z) = \rho$$

with $\underline{w}_\rho(1) = w_\rho(1)$. A straightforward calculation yields

$$\underline{w}_\rho(z) = e^{z-1} w_\rho(1) + e^z \int_1^z e^{-x} (\rho - k(x)) dx.$$

The second term is strictly positive for sufficiently large ρ . Thus, $\underline{w}_\rho(z) > 1$ for all $z \in [1, z_{1,\rho}]$ when ρ is large enough. But this contradicts $w_\rho(z_{1,\rho}) = 1$. Therefore, \mathcal{U} is not empty. \square

Proof of Lemma EC.4. By Lemma EC.3, we know that $\rho^* < \infty$. We next show that $\rho^* \in \mathcal{U}$. To reach a contradiction, suppose that $\rho^* \in \mathcal{L}$. Then, by Lemma EC.2, we know that there exists a maximum point of w_{ρ^*} , denoted as z^* , and let $w^* := w_{\rho^*}(z^*) < \infty$. By the same lemma, for some given $\delta > 0$, there exists $\epsilon > 0$ such that $w_{\rho^*}(z^* + \epsilon) = w^* - \delta$. On the other hand, for all $\rho > \rho^*$, we have that $\rho \in \mathcal{U}$, and therefore,

$$w_\rho(z^* + \epsilon) > w_\rho(z^*) > w_{\rho^*}(z^*) = w^* \quad \text{if } \rho \in \mathcal{U}.$$

This is a contradiction since $w_\rho(z)$ is jointly continuous in ρ and z . Therefore, $\rho^* \in \mathcal{U}$ and $w'_{\rho^*}(z) > 0$ for all $z \geq 0$. \square

Proof of Lemma EC.5. First, let a be a positive constant such that

$$\sqrt{a}(k(z) + a) - k'(z) > 0 \quad \text{for all } z \geq 0.$$

Note that the existence of such a positive constant is ensured by Assumption 1. Next, for every $\rho \in (0, \rho^*)$, let \hat{z}_ρ be such that $w_\rho(\hat{z}_\rho) = \sup_{z \geq 0} w_\rho(z)$. Since $w'_\rho(\hat{z}_\rho) = 0$, we have that

$$w_\rho(\hat{z}_\rho) = \sqrt{k(\hat{z}_\rho) - \rho} < \sqrt{2k(\hat{z}_\rho) + a} \quad \text{for } \rho \in (0, \rho^*).$$

Now, let $f_\rho(z) := (w_\rho(z))^2 - 2k(z) - a$. We intend to argue that $f_\rho(z) \leq 0$ for all $z \in [0, \hat{z}_\rho]$. For that purpose, suppose by way of contradiction that this is not true. Then there must exist \underline{z}_ρ and \bar{z}_ρ with $0 \leq \underline{z}_\rho < \bar{z}_\rho < \hat{z}_\rho$ such that $f_\rho(\underline{z}_\rho) = f_\rho(\bar{z}_\rho) = 0$, and $f_\rho(z) > 0$ for all $z \in (\underline{z}_\rho, \bar{z}_\rho)$. In particular, the foregoing implies that $w_\rho(z) > \sqrt{2k(z) + a}$, and so

$$w'_\rho(z) = (w_\rho(z))^2 - k(z) + \rho > k(z) + a$$

for every $z \in (\underline{z}_\rho, \bar{z}_\rho)$ and $\rho \in (0, \rho^*)$. Thus,

$$f'_\rho(z) = 2w_\rho(z)w'_\rho(z) - 2k'(z) > 2\sqrt{a}(k(z) + a) - 2k'(z) > 0 \quad \text{for all } z \in (\underline{z}_\rho, \bar{z}_\rho).$$

This, however, contradicts our hypothesis which holds that $f_\rho(\underline{z}_\rho) = f_\rho(\bar{z}_\rho) = 0$. Therefore, $f_\rho(z) \leq 0$ for all $z \in [0, \hat{z}_\rho]$, from which the desired result follows. \square