Submitted to
manuscript

Driver Surge Pricing

Nikhil Garg
Stanford University
nkgarg@stanford.edu

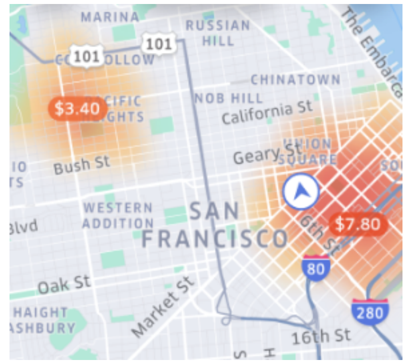
Hamid Nazerzadeh
USC Marshall and Uber Technologies
nazerzad@usc.edu

Uber and Lyft ride-hailing marketplaces use dynamic pricing, often called surge, to balance the supply of available drivers with the demand for rides. We study pricing mechanisms for such marketplaces from the perspective of drivers, presenting the theoretical foundation that has informed the design of Uber's new additive driver surge mechanism. We present a dynamic stochastic model to capture the impact of surge pricing on driver earnings and their strategies to maximize such earnings. In this setting, some time periods (surge) are more valuable than others (non-surge), and so trips of different time lengths vary in the opportunity cost they impose on drivers. First, we show that multiplicative surge, historically the standard on ride-hailing platforms, is not incentive compatible in a dynamic setting. We then propose a structured, incentive-compatible pricing mechanism. This closed-form mechanism has a simple form, and is well-approximated by Uber's new additive surge mechanism.

1. Introduction

Ride-hailing marketplaces like Uber, Lyft, and Didi match millions of riders and drivers every day. A key component of these marketplaces is a *surge* (dynamic) pricing mechanism. On the rider side of the market, surge pricing reduces the demand to match the level of available drivers and maintains the reliability of the marketplace, cf., (Hall et al. 2015), and so allocates the rides to the riders with the highest valuations. On the driver side, surge encourages drivers to drive during certain hours and locations, as drivers earn more during surge (Lu et al. 2018, Hall et al. 2017, Chen and Sheldon 2016). Castillo et al. (2017) show that surge balances both sides of this spatial market by moderating the demand and the density of available drivers, hence avoiding so called "Wild Goose Chase" equilibria in which drivers spend much of their time on long distance pick ups. Due to these effects, surge pricing – along with centralized matching technologies – is often considered the primary reason that ride-sharing marketplaces outperform traditional taxi services on many metrics, including driver utilization and overall welfare (Cramer and Krueger 2016, Buchholz 2017, Ata et al. 2019).

(a) *Multiplicative surge heatmap.* "1.6x" on the map means that the standard fares for trips from the corresponding area are increased by 60%.



(b) *Additive surge heatmap.* "\$7.8" on the map means that \$7.8 is added to the standard fares for trips from the corresponding area.

Figure 1 **Example driver surge heatmaps with multiplicative and additive surge, respectively. On Uber, drivers can see a heatmap of surge when they are logged in but not on a trip. The heatmap is meant to guide drivers to higher earning opportunities, signaling each location's value to drivers (Lu et al. 2018). Structural simplicity is essential to clearly communicate payments to drivers, and additive and multiplicative surge represent the two simplest options.**

However, variable pricing (across space and time) must be carefully designed, since it can create incentives for "cherry-picking" and rejecting certain trip requests. Such behavior increases earnings of strategic drivers at the expense of other drivers, who may then disproportionately receive such trip requests after they are rejected by others, cf., (Cook et al. 2018). It also reduces overall platform reliability, inconveniencing riders who may have to wait longer before receiving a ride.

Uber recently revamped its driver surge mechanism to simplify such decisions, in an attempt to improve the driver experience and making earnings more dependable (Uber 2019b). The main change is making surge "additive" instead of "multiplicative." Under **multiplicative surge**, the payout of the driver from a surged trip scales with the length of the trip. In contrast, under **additive surge**, the surge component of the payout is constant (independent of trip length), with some adjustment for very long trips (Uber 2019d). Figure 1 depicts the heat-map of surge on the driver app for each type of surge. We show that this change directly addresses the issue that drivers who strategically reject trip requests may earn more than drivers who do not, even as total payments remain the same.

Contributions

We consider the design of *incentive compatible* (IC) pricing mechanisms in the presence of surge. Trips differ by their length $\tau \in \mathbb{R}_+$, and the platform sets the payout $w(\tau)$ for each trip in each world state (i.e., surge vs non-surge). Drivers decide which trip requests $\sigma \subseteq \mathbb{R}_+$ to accept in each world state, in response to the payout function w.[1] The technical challenge is to design a IC pricing mechanism w, for which accepting all trips is an earning maximizing strategy for drivers over a long horizon, i.e. where $\sigma = (0, \infty)$ in each world state maximizes driver earnings.

We first study a continuous-time, infinite horizon *single-state model*, where there is only one world state and trip requests arrive over time according to a stationary Poisson process. We show that in this model, multiplicative pricing — where the payout of a trip is proportional to the length of that trip — is incentive compatible. To obtain this result, we show in Theorem 1 that the best response strategy of a driver to function w, to maximize her earning, is a threshold strategy where she accepts all trips with payout rate $\frac{w(\tau)}{\tau}$ above that threshold. Hence, an mechanism that equalizes the payout rate of all trips is incentive compatible.

We then present a model where the world state stochastically transitions between surge and non-surge states, with trip payments, distributions, and intensity varying between states. In such a temporally *dynamic* system, completing a given trip affects a driver's earnings beyond just the length of the trip, i.e. imposes a future-time externality on the driver that is a function of the trip length. The driver's trip opportunity cost thus includes both what occurs during a trip, and a continuation value. This externality implies that *multiplicative surge is not incentive compatible in the presence of surge* (Theorem 2), in contrast to the single-state model. Namely, drivers can benefit from rejecting long trips in a non-surge state, and short trips in the surge state.

In our main result, Theorem 3, we propose a class of incentive compatible pricing functions that are described as closed forms of the model primitives. The prices account for the driver's temporal externalities; e.g., during surge, short trips pay more per unit time than do long trips.

[1] Drivers' level of sophistication and experience varies, cf. Cook et al. (2018). An IC mechanism aligns the incentives of drivers to accept all trips, for any level of strategic response to pricing strategies.

Finally, we consider additive and multiplicative surge in our dynamic model. Through numerical simulations,[2] we show that the new additive surge mechanism has desirable incentive compatibility properties when compared to multiplicative surge. More specifically, we observe that under additive surge, drivers may benefit only from rejecting a small fraction of long trips, supporting the practical remedy that makes adjustments for certain long trips (Uber 2019d).

To our knowledge, ours is the first ride-hailing pricing work to incorporate dynamic (non-constant), stochastic demand and pricing. This component is essential to uncover the way through which a particular trip imposes substantial temporal externalities on a driver's future earnings.

Related Work

We discussed some of the related work on surge pricing above. Here, we briefly review the lines of research closest to ours. We refer the reader to a recent survey by Korolko et al. (2018) for a broader overview of the growing literature on ride-hailing markets.

Driver spatio-temporal strategic behavior. Several works model strategic driver behavior in a spatial network structure, and across time in a single-state. Ma et al. (2018) develop spatially and temporally smooth prices that are welfare-optimal and incentive compatible in a deterministic model. Their prices form a competitive equilibrium and are the output of a linear program with integer solutions. We similarly seek to develop incentive compatible pricing schemes, and both works broadly construct VCG-like prices that account for driver opportunity costs. Our focus is on structural aspects (e.g. multiplicative in trip length) in a non-deterministic model.

Bimpikis et al. (2016) show how the platform would price trips between locations, taking into account strategic driver re-location decisions, in a single-state model with discrete locations. They show that pricing trips based on the origin location substantially improves surplus, as well as the benefits of "balanced" demand patterns. Besbes et al. (2018b) consider a continuous state space setting and show how a platform may optimally set prices across the space in reaction to a localized demand shock to encourage drivers to relocate; their model has driver cost to re-locate, but no explicit time dimension. They find that localized prices have a global impact, and, e.g., the optimal pricing solution incentivizes

[2] Whenever possible, we fit the parameters of the Uber's data.

some drivers to move away from a demand shock. Afèche et al. (2018) consider a two state model with demand imbalances and compare platform levers such as limiting ride requests and directing drivers to relocate, in a two-state fluid model with strategic drivers. They upper-bound performance under these policies, and find that it may be optimal for the platform to reject rider demand even in over-supplied areas, to encourage driver movement. A similar insight is developed by Guda and Subramanian (2019). Finally, Yang et al. (2018) analyze a mean-field system in which agents compete for a location-dependent, time-varying resource, and decide when to leave a given location. They leverage structural results — agents' equilibrium strategies depend just on the current resource level and number of agents — to numerically study driver decisions to relocate between locations as a function of the platform commission structure.

Dynamic pricing in ride-sharing and service systems. There is a growing literature on queuing and service systems motivated in part by ride-sharing market. For example, Besbes et al. (2018a) revisit the classic square root safety staffing rule in spatial settings, cf., Bertsimas and van Ryzin (1991, 1993). Much of the focus of this line of work is how pricing affects the arrival rate of (potentially heterogeneous) customers, and thus the trade-off between the price and rate of customers served in maximizing revenue.

Banerjee et al. (2015) consider a network of queues in which long-lived drivers enter the system based on their expected earnings but cannot reject specific trip requests. Under their model, dynamic pricing *cannot* outperform the optimal static policy in terms of throughput and revenue, but is more robust. Cachon et al. (2017) argue on the other hand that surge pricing and payments are welfare increasing for all market participants when drivers decide when to work. Similar in spirit to our work, Chen and Hu (2018) consider a marketplace with forward-looking buyers and sellers who arrive sequentially and can wait for better prices in the future. They develop strategy-proof prices whose variation over time matches the participants' expected utility loss incurred by waiting.

One of the most related to our work in modeling approach, Kamble (2018) studies how a freelancer can maximize her long-term earnings with job-length-specific prices, balancing on-job payments and utilization time. In his model, a freelancer sets her own prices for a discrete number of jobs of different lengths and, with assumptions similar to our single-state model, it is optimal for the freelancer to set the same price per hour for all jobs. We further discuss the relationship of this work to our single-state model below.

Organization. The rest of the paper is organized as follows: In Section 2, we formally present our model. In Section 3, we formulate a driver's best response strategy to affine pricing functions in each model. In Section 4, we present incentive compatible pricing functions for our surge model. Finally, in Section 5, we numerically compare the IC properties of additive and multiplicative surge.

2. Model

We consider a large ride-hailing market with decoupled pricing, from the perspective a *single driver*. This driver receives trip requests of various lengths, whose rate, distribution, and payment are known to the driver but determined exogeneously to her decisions to accept or decline requests. We do not consider spatial heterogeneity in our setting, to abstract away the impact of location and focus on temporal opportunity cost and continuation value based on a length of the trip.[3]

In this section, we first describe the primitives of our two models, a single-state model (Section 2.1) and a dynamic model with surge pricing (Section 2.2). Then we describe the driver's strategy space and the platform's pricing design challenge (Section 2.3).

2.1. Single-state model

We start with a model where there is a single world state, i.e. all distributions are constant over time. Time is continuous and indexed by t. At each time t, the driver is either *open*, or *busy*. While the driver is open, she receives job (trip) requests from riders according to a Poisson process at rate λ, i.e., the time between requests is exponential with mean $\frac{1}{\lambda}$. Job lengths, denoted by τ, are drawn independently and identically from continuous distribution F.

If the driver accepts a job request of length τ at time t (as discussed below), she receives a payout of $w(\tau)$ at time $t+\tau$, at which time she becomes open again. If the request is not accepted, the driver remains open. Note that our model of a trip is a simplification from practice, where a given job has two components: the time it takes to pick-up the rider, and the time while the rider is in the driver's vehicle. To simplify the presentation, we combine these two components into trip length. When these components are separated, the space of driver strategies becomes richer, but results remain qualitatively the same.

[3] We believe our insight can be extended to a spatial setting where the price can be decomposed to a time-based component, based on the length of the trip, and a spatial component based on the destination of the trip. However, this would be beyond the scope of this work, cf., Bimpikis et al. (2016).

Pricing function w is assumed to be continuous and non-decreasing, and such that the hourly payment as a function of the trip length is asymptotically bounded, $\exists c : \liminf_{\tau\to\infty} \frac{w(\tau)}{\tau} \leq c$.

2.2. Dynamic model with surge pricing

A model with fixed pricing and arrival rates of jobs is not a realistic representation of ride-hailing platforms. In particular, rider demand (both in intensity and in distribution) may vary substantially over time. To study how this **dynamic** nature affects driver decisions, we consider a model with two states, $i \in \{1,2\}$, where $i = 2$ denotes the *surge* state. (At a high level, the surge state provides a higher earnings rate to the driver. The precise definition of what distinguishes the surge state is presented in Section 3.2, after we formulate the driver's earnings rate in each state).

The world evolves stochastically between the two states, as a Continuous Time Markov Chain (CTMC). When the world is in state i, the state changes to j according to a fixed exponential clock that ticks at rate $\lambda_{i\to j}$, independently of other randomness.

When open in state i, the driver receives job requests at rate λ_i with lengths $\tau \sim F_i$, and collects payout according to payment function w_i, which is presumed to have the same properties as w in the single-state model. Note that the state of the world may change while a driver is on trip. The driver receives payments according to the state of the world i when she *starts* a trip. We will use $w = \{w_1, w_2\}$ to denote the overall pricing mechanism.

2.3. Driver strategies and earnings

In our model, the driver can decide whether to accept the trip request, with no penalty.[4] In the single-state model, let $\sigma \subset \mathbb{R}_+ = (0,\infty)$ denote the driver's (deterministic) strategy, where $\tau \in \sigma$ implies that a driver accepts job requests of length τ. In the dynamic model, the driver follows deterministic policy $\sigma = \{\sigma_1, \sigma_2\}$, where $\sigma_i \subset \mathbb{R}_+$ indicates the jobs she accepts in state i. We assume that driver policies are measurable with respect to F (corresponding F_i in dynamic model); additionally, for technical reasons, in the dynamic model we also assume that σ_i consist of a union of open intervals, i.e. are open subsets of $\mathbb{R}_+$. When we write equalities with policies σ, we mean equality up to changes of measure 0.

The driver is long-lived and aims to maximize her lifetime average hourly earnings on the platform, including both open and busy times. Let $R(w,\sigma,t)$ denote the (random) total

[4] This assumption follows Uber's current practice. We further discuss the driver's information set below.

earnings from jobs accepted from time 0 up to time t if she follows policy σ and the payout function is w; see the appendix for a more formal definition. Then, the driver's lifetime *earnings rate* is

$$R(w,\sigma) \triangleq \liminf_{t\to\infty} \frac{R(w,\sigma,t)}{t}.$$

This earnings rate is a (deterministic) function of driver policy σ, pricing function w, and the primitives. We can now define notions of an optimal driver policy and incentive compatible pricing.

A driver policy σ^* is **optimal** (best-response) with respect to pricing function w if it maximizes the lifetime earnings rate of the driver among all policies: $R(w,\sigma^*) \geq R(w,\sigma)$, for all valid policies σ (i.e. measurable with respect to F, F_i, with σ_i open sets). Then, pricing function w is **incentive compatible (IC)** if accepting all job requests is optimal with respect to w, i.e. $\sigma = (0,\infty)$ in the single-state model or $\sigma = \{(0,\infty),(0,\infty)\}$ in the dynamic model is optimal with respect to w. In other words, pricing function w is incentive compatible if an earnings-maximizing driver (who knows all the primitives and w) accepts every trip request. In Section 5, we also consider an *approximate* notion of incentive compatible pricing, defined as the fraction of trips that are accepted under an optimal driver policy with respect to a given pricing function.

Information structure. We assume that the platform reveals the total trip length to the driver at the time of request, and that the driver can freely reject it without penalty. We note that in current practice in ride-hailing markets, drivers often cannot see the rider's destination or the trip length until they pick up the rider (but they can reject a request based on the pick-up time to the rider, without penalty). Some drivers call ahead to find out the rider's destination or even cancel the trip at the pick-up location, creating negative experiences for both the rider and the driver.[5] Our notion of incentive compatibility is *ex-post*, implying that drivers would accept all trips, even if the trip length is not revealed, and so this setting from practice is covered as well.

Decoupled pricing. Our setting is *decoupled*: rider and driver prices are determined separately. Namely, changes in the driver payout function and decisions do not change the

[5] We note that destination discrimination is against Uber's guidelines and could lead to deactivation (Uber 2019a).

trip request rate or the distribution of the trips. This modeling assumption follows the current practice (Uber 2019e) and furthermore allows us to focus on the drivers' perspective, without further complicating the analysis.[6]

3. Driver Reward and Affine Driver Pricing

In this section, we present the driver's lifetime earnings rate $R(w,\sigma)$, using the renewal reward theorem on an appropriately defined renewal process. For the dynamic model, this process is not immediate from the primitives defined; we break down the lifetime driver's earning rate into: (1) a function of the fraction of her time she spends in each such state and (2) the earnings rates while the driver is open in each state or on a job that started in that state. This decomposition allows us to analyze the driver's (best response) strategy to payout w in our dynamic model.

We are especially interested in *affine* pricing schemes, where $w_i(\tau) = m_i\tau + a_i$, with $m_i \geq 0$ (in the single-state model: $w(\tau) = m\tau + a$, with $m \geq 0$). Such pricing functions can be clearly communicated as time and distance rates (see, e.g., (Uber 2019c)), or otherwise be displayed on a surge heat-map. We refer to the case with $a_i > 0$ ($a_i < 0$) as positive (negative) affine pricing.

Below, we first characterize the driver's best-response strategy with respect to any pricing function w in the single-state model. We observe that multiplicative pricing — a special case of affine pricing where $w(\tau) = m\tau$ — is incentive compatible. In contrast, in Section 3.2, we show that in the dynamic model multiplicative pricing may no be longer incentive compatible. We further derive the structure of optimal driver policies in each state with respect to affine or multiplicative pricing, which will enable numerical study of the (approximate) incentive compatibility properties of additive and multiplicative surge in Section 5. Section 3.4 discusses the key differences in the two models, setting up Section 4 where we derive incentive compatible pricing functions.

3.1. Single-state model

In the single-state model, the primitives of our model directly induce a renewal reward process, where a given renewal cycle is defined as being from the time a driver is newly open to the time she is open again after completing a job. Let $W(\sigma)$ denote the mean

[6] Coupled pricing imposes more constraints on the pricing functions chosen by the platform. For example, Bai et al. (2018) find that the platform should adjust its payout ratio with demand – an example of decoupled pricing – to maximize profit or overall welfare.

earnings on trips $\tau \in \sigma$, i.e. the expected earning in a renewal cycle; let $T(\sigma)$ be the sum of the expected wait time to an accepted trip and the expected length of a trip, and thus the expected renewal cycle length. Then, the lifetime mean hourly earnings (earnings rate) for a driver is given by

$$R(w,\sigma) = \frac{W(\sigma)}{T(\sigma)} = \frac{\frac{1}{F(\sigma)}\int_{\tau\in\sigma} w(\tau)dF(\tau)}{\frac{1}{F(\sigma)\lambda} + \frac{1}{F(\sigma)}\int_{\tau\in\sigma} \tau dF(\tau)}$$

where $F(\sigma)$ denotes the probability of the driver receiving a job request $\tau \in \sigma$. The first equality follows from the Renewal Reward Theorem, and holds with probability 1 (see, e.g., Gallager (2013)). We view the earnings rate as a constraint for the platform in its pricing. Given some demand model, the platform receives revenue at a rate that depends on the prices it sets for riders. This revenue rate yields a earnings payout rate target R, at which the platform needs to pay drivers. Then, the platform task is to find an incentive compatible pricing function w such that the actual earnings rate for drivers who accept every trip meets the target, $R(w,(0,\infty)) = R$. This is close to how decoupled surge pricing is set in practice, where the revenue from the rider-surge is viewed as a target for average driver surge earnings.

Our first result is that, in the single-state model, the driver's optimal policy has a simple form.

Theorem 1. *In the single-state model, for each w there exists a constant $c_w \in \mathbb{R}_+$ such that the policy $\sigma^* = \left\{\tau : \frac{w(\tau)}{\tau} \geq c_w\right\}$ is optimal for the driver with respect to w.*

Theorem 1 establishes that, in a single-state model with Poisson job arrivals, the *length* of the job is not important, only the hourly rate while busy on the job. Note, however, that the optimal c_w in the policy is not necessarily $c_w = \sup \frac{w(\tau)}{\tau}$: drivers must trade off the earnings rate while on a trip with their utilization rate; the more trips that a driver rejects, the longer she must wait for an acceptable trip. In the appendix we prove the result by, starting at an arbitrary policy σ, making changes to the policy that increase the earnings rate while on a job *without decreasing the utilization rate*. Thus, each such change improves the reward $R(w,\sigma)$, and the sequence of changes results in a policy of the above form, for some c'. Then, this minimum on-job earnings rate c' can be optimized, leading to an optimal policy of this form.

An immediate corollary of Theorem 1 is that $w(\tau) = m\tau$, for $m > 0$, is IC. In other words, if the platform pays a constant hourly rate $\frac{w(\tau)}{\tau} = m$ to busy drivers then in the single-state model it is in the driver's best interest to accept every trip. This result is driven by the following insight: while *receiving* long trip requests is more beneficial to drivers in the single-state setting as they increase one's utilization rate (the driver is busy for a longer time until her next open period), rejecting short trips to cherry-pick long trips decreases utilization by the same amount.[7] Further note that, given a earnings rate target R, calculating the multiplier m and thus an IC pricing policy is trivial.

On the other hand, affine pricing may not be incentive compatible because short trips are worth more per hour than are long trips: $\frac{w(\tau)}{\tau} = m + \frac{a}{\tau}$. The optimal policy may be to accept trips in $\sigma^* = (0, T)$ for some T. However, our next proposition establishes that affine pricing is incentive compatible if the additive component stays small enough as a function of the request arrival rate:

PROPOSITION 1. *In the single-state model, $w(\tau) = m\tau + a$ is incentive compatible if $0 \le a \le \frac{m}{\lambda}$.*

The sufficient condition has a simple intuition: when open, the expected amount of time the driver must wait for her next request is $\frac{1}{\lambda}$; if on-trip time is valued at m per unit-time, then with $a = \frac{m}{\lambda}$ the additive component can be interpreted as paying for the driver's expected waiting time. Thus, while a driver may earn more per hour for a short trip than a long trip with affine pricing, such a short trip is not worth the time the driver must wait for her next trip request. We further note that the condition in the proposition is not a necessary one; however, deriving necessary and sufficient conditions in closed form requires specifying the trip distribution F.

As we'll see in the next sub-section, the structure of optimal driver policies in reaction to affine pricing differs sharply in the dynamic model.

3.2. Dynamic model

We start our analysis of the dynamic model by characterizing the lifetime driver earnings rate, $R(w, \sigma)$. Here, we can no longer directly use the renewal reward theorem as in the

[7] As discussed in the related work, this corollary and insight is similar to a result of Kamble (2018); however, the proof is more involved in our setting as a driver's strategy σ is a *subset* of $\mathbb{R}_+$ denoting the job requests she accepts, as opposed to a discrete set of prices she charges. Further, in our setting, the driver responds to the platform's prices instead of setting her own prices, enabling a wider range of IC pricing mechanisms.

single-state model, with a renewal cycle containing just a single trip. The driver's earning on a given trip is no longer independent of her earnings on other trips: given a job that starts in the surge state, the driver's next job is more likely to start in the surge state. Given whether each job started in the surge state, however, job earnings are independent. We can use this property to prove our next lemma, which gives the mean hourly reward overall in the dynamic model.

LEMMA 1. *The overall earning rate can be decomposed into the earnings rate $R_i(w_i, \sigma_i)$ and fraction of time $\mu_i(\sigma)$ spent in state i. The following equality holds with probability* 1*:*

$$R(w,\sigma) \triangleq \liminf_{t\to\infty} \frac{R(w,\sigma,t)}{t} = \mu_1(\sigma)R_1(w_1,\sigma_1) + \mu_2(\sigma)R_2(w_2,\sigma_2).$$

As in the single-state model, $R_i(w_i,\sigma_i) = \frac{W_i(\sigma_i)}{T_i(\sigma_i)}$*, where*

$$W_i(\sigma_i) = \frac{1}{F_i(\sigma_i)}\int_{\tau\in\sigma_i} w_i(\tau)dF_i(\tau), \qquad T_i(\sigma_i) = \frac{1}{\lambda_i F_i(\sigma_i)} + \frac{1}{F_i(\sigma_i)}\int_{\tau\in\sigma_i} \tau dF_i(\tau)$$

We prove the result as follows. We define a new renewal process, in which a single reward renewal cycle is: the time between the driver is open in state 1 to the next time the driver is open in state 1 *after* being open in state 2 at least once. In other words, *each* renewal cycle is composed of a number of sub-cycles in which the driver is open in state 1 and then is open in state 1 again after a completed trip; one sub-cycle which starts with the driver open in state 1 and ends with her open in state 2 (either after a completed trip or a state transition while open); a number of sub-cycles in which the driver is open in state 2 and then is open in state 2 again after a completed trip; and finally one sub-cycle starting in state 2 and ending with the driver open in state 1.

Now, given the number of such large renewal reward cycles completed up to each time t, the total earnings on trips starting in each state (earnings in each sub-cycle) are independent of each other, resulting in the separation. Then, we use Wald's identity (Wald 1973) to separate $\mu_i(\sigma)$ and $R_i(\sigma_i)$.

Note that $T_i(\sigma_i)$ is not exactly the expected length of time in a single sub-cycle in a state given σ_i, but rather is proportional to it; the multiplicative constant $\frac{1}{\lambda_i F_i(\sigma_i)+\lambda_{i\to j}}$ cancels out with the same constant in the expected earnings in a single sub-cycle in a state given σ_i. This constant emerges from our surge primitives: when the driver is open in state i, there are two competing exponential clocks (with rates $\lambda_i F_i(\sigma_i)$ and $\lambda_{i\to j}$, respectively) that determine whether the driver will accept a trip request in state i before the world state changes to state j. Furthermore, we can now precisely define what it means for $i=2$ to be the surge state: it has a higher potential earning rate than state 1; $\exists\sigma_2$ such that $R_2(w_2,\sigma_2) > R_1(w_1,\sigma_1), \forall\sigma_1 \subseteq \mathbb{R}_+$. This assumption is not the same as $w_2(\tau) \geq w_1(\tau), \forall\tau$, and neither implies the other; it is a condition jointly on λ_i, F_i, w_i, (but not $\lambda_{i\to j}$), indicating that a driver can, following some policy, earn more while in the surge state than she can following any policy in the non-surge state. Throughout, we set $i=2$ as the surge state according to this definition. When constructing pricing functions w_1, w_2, we will *also* require the following constraint to be met: $R_i(w_i,(0,\infty)) = R_i$, for some exogenous $R_2 > R_1$, analogously to the single-state model constraint.

What does $\mu_i(\sigma)$ look like? We defer showing the exact form to Section 4.1 in advance of developing incentive compatible pricing. Here, we provide some intuition: the trips that a driver accepts in each state determines the portion of her time she spends on trips started in each state. For example, if a driver never accepts trips in the non-surge state, she will be open and thus available for a trip as soon as surge begins. Inversely, if a driver accepts a long surge trip immediately before surge ends, she will be paid according to the surge payment function w_2 even though surge has ended. Surprisingly, given the complex formulation of the reward $R(w,\sigma)$ as it depends on $\sigma = \{\sigma_1,\sigma_2\}$, we can find the structure of optimal policies as they depend on the pricing structure w_i, as well as incentive compatible pricing functions. We begin this analysis in the next subsection, deriving optimal driver responses to multiplicative and affine pricing.

3.3. Driver's Best-Response Strategy to Affine Prices

In the single-state model, multiplicative pricing is incentive compatible; a driver cannot benefit in the future by rejecting certain trips if all trips have the same hourly earning rate. In contrast, we now show that the same insight does not hold for the dynamic model, as a driver can influence her future trips through her decision to accept or reject certain trips.

THEOREM 2. *Consider pricing $w=\{w_1,w_2\}$. Then, there exists an optimal policy $\sigma = \{\sigma_1,\sigma_2\}$ (i.e. that maximizes $R(w,\sigma)$), defined with parameters $t_1,t_2,t_3,t_4,t_5,t_6 \in [0,\infty)\cup\{\infty\}$, such that*

- *Non-surge state driver optimal policy σ_1:*

— *If w_1 is multiplicative or positive affine, σ_1 rejects long trips, i.e.* $\sigma_1 = (0, t_1)$.

— *If w_1 is negative affine, σ_1 rejects short and long trips, i.e.* $\sigma_1 = (t_2, t_3)$.

- *Surge state driver optimal policy σ_2:*

— *If w_2 is multiplicative or negative affine, σ_2 rejects short trips, i.e.* $\sigma_2 = (t_4, \infty)$.

— *If w_2 is positive affine, σ_2 rejects medium length trips, i.e.* $\sigma_2 = (0, t_5) \cup (t_6, \infty)$.

Furthermore, there exist settings where t_i's take positive finite values. Hence, multiplicative pricing is not incentive compatible in general.

We discuss the intuition in the next section. In the appendix, we prove the result for each case as follows: fixing σ_j for $j \neq i$, we start with an arbitrary open set $\sigma_i = \cup_k^\infty (\ell_k, u_k)$, recalling that open sets can be written as a countable union of such disjoint intervals. Then, we find $\frac{\partial}{\partial u_k} R(w, \sigma)$, the derivative of the set function $R(w, \sigma)$ with respect to one of the interval upper end-points of σ_i, i.e. u_k. This derivative is the infinitesimal change in the overall reward if σ_i is expanded by increasing u_k, and it has useful properties. In the surge state with multiplicative pricing, for example, $\frac{\partial}{\partial u} R(w, \sigma)$ has the same sign as a function that is *increasing* in u, for each fixed σ. With affine pricing, it has the same sign as a *quasi-convex* (positive affine in the surge state) or *quasi-concave* (negative affine in the non-surge state) function in u, for a fixed σ. Such properties enable constructing a sequence of changes to σ_i that each do not decrease the reward $R(w, \sigma)$, with the limit being a policy of the appropriate form. In particular, we can show that any policy that is not of the appropriate form above has $\frac{\partial}{\partial u_k} R(w, \sigma) \geq 0$ for some u_k, allowing local improvements until adjacent intervals $(\ell_k, u_k), (\ell_{k+1}, u_{k+1})$ can be combined or expanded to infinity.

Further note that the results of rejecting long trips in non-surge (and short trips in surge) extend to arbitrary pricing functions where $\frac{w_1(\tau)}{\tau}$ is non-increasing (respectively, $\frac{w_2(\tau)}{\tau}$ is non-decreasing), as the same properties hold. The other two results do not hold with such generality, as the behavior of the derivative may be arbitrarily complex.

3.4. Why is multiplicative surge pricing not incentive compatible?

"I thoroughly dislike short trips ESPECIALLY when I'm picking up in a waning surge zone"

Anonymous driver

What explains the difference between multiplicative pricing being incentive compatible in the single-state model but not in the dynamic model? In the latter, a driver's policy affects not just her earnings while she is busy, but also the fraction of her time at which she is busy during the lucrative surge state. In particular, it turns out, accepting short trips during surge may *reduce* the amount of time that a driver is on a surge trip! Figure 2 shows in an example how the fraction of time in the surge state $\mu_2(\sigma)$ changes as a function of how many short trips the driver rejects. The anonymous driver we quote above identifies the key effect: when surge is short-lived, a driver may only have the chance to complete one surge trip before it ends. Thus, the driver may be better off waiting to receive a longer trip request, as with multiplicative surge she is paid a higher rate for the full duration of the longer trip. (Of course, there is a trade-off as if she rejects too many trip requests, she may not receive any acceptable request before surge ends). In the surge state, then, multiplicative pricing does not compensate drivers enough to accept short trips that may reduce their future surge earnings. In the non-surge state, analogously, multiplicative pricing under-compensates long trips that may prevent taking advantage of a future surge.

Affine pricing is a first, reasonable attempt at fixing these issues. In the surge state, the additive value makes the previously under-compensated short trips comparatively more valuable, as the earnings per unit time $\frac{w_2(\tau)}{\tau} = m_2 + \frac{a_2}{\tau}$ (with $a_2 > 0$) are now higher for short trips. Unfortunately, with such pricing the structure for the surge optimal policy becomes $\sigma_2 = (0, t_5) \cup (t_6, \infty)$ – if the values m_2, a_2 are not balanced correctly, the additive value is enough to make accepting extremely short trips $(0, t_5)$ profitable; for medium-length trips $\tau \in (t_5, t_6)$, however, the additive value is not large enough to make up for the fact that accepting the trip prevents accepting another surged trip before surge ends. Similarly, negative affine pricing in the non-surge state, $w_1(\tau) = m_1\tau + a_1$, (with $a_1 < 0$) is now too harsh on very short trips but potentially not enticing enough for long trips.

We expand further on such effects in the next section, where we fix these issues by constructing true incentive compatible pricing schemes for both states. Then, in Section 5 we use the structural results derived in this section to perform numerical simulations comparing (approximate) incentive compatibility of additive and multiplicative surge.

4. Incentive Compatible Surge Pricing

In this section, we present our main result regarding the structure of incentive compatible pricing in the dynamic model. As discussed earlier, consistent with decoupled pricing, we distinguish between the two states by assuming that the surge state has a higher potential

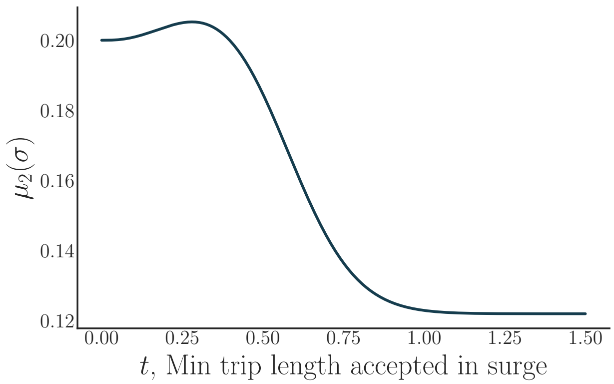


Figure 2 **Fraction of time spent in surge state, $\mu_2(\sigma)$, with driver policy $\sigma = \{\sigma_1 = (0,\infty), \sigma_2\}$, where $\sigma_2 = (t,\infty)$, i.e. t is the minimum trip length accepted in the surge state. The primitives are as follows: $\lambda_1 = \lambda_2 = 12, \lambda_{1\to 2} = 1, \lambda_{2\to 1} = 4$; in both states, trip lengths are distributed according to a Weibull distribution with shape 2 and mean $\frac{1}{3}$. These parameters reflect realistic average trip to wait time values, and that surge tends to be short-lived compared to non-surge times. Note that the driver can increase the time she spends in the surge state by rejecting short surge trips.**

earnings rate than the non-surge state. We further require earning rate constraints as before: the platform must set pricing function w_i to satisfy $R_i(w_i,(0,\infty)) = R_i$, i.e. the state i earnings rate for drivers who accept every trip in that state is set to R_i, for (exogenously) given $R_2 > R_1$. We focus on the driver's decision on accepting or rejecting a trip, observing that this decision depends on the "opportunity cost" of performing that trip. A driver's expected value for accepting a trip is the payout from that trip, plus the continuation value which depends on the world state when the trip is completed and she becomes open again. If the driver rejects a trip, she can accept other trips during the time it would have taken to complete that trip. If the state transitions in the meantime, some of those trips could be lucrative surge trips. Intuitively, we find incentive compatible prices that compensate drivers for such opportunity costs.

To this aim, in Section 4.1, we characterize, $\mu_i(\sigma)$, how much time the driver spends in each state. In Section 4.2, we present incentive compatible prices, under mild conditions on the ratio of earning rates between the two states. Section 4.3 contains a discussion on the intuition of the IC pricing structure in terms of the driver's opportunity cost, and Section 4.4 contains a proof sketch.

4.1. Transition probabilities and expected time spent in each state

The expected fraction of time spent in each state, $\mu_i(\sigma)$, depends both on the evolution of the world state and the trips a driver accepts in each state. In order to quantify the effects previewed in Section 3.4, we first need to analyze the evolution of the CTMC that determines the surge state.

LEMMA 2. *Suppose the world is in state i at time t. Let $q_{i\to j}(s)$ denote the probability that the world will be in state $j\neq i$ at time $t+s$. Then,*

$$q_{i\to j}(s)=\frac{\lambda_{i\to j}}{\lambda_{i\to j}+\lambda_{j\to i}}\left[1-e^{-(\lambda_{i\to j}+\lambda_{j\to i})s}\right]$$

Note that $q_{i\to j}(s)$ is not just the probability that the world state transitions once during time $(t,t+s)$, but the probability that it transitions an odd number of times. This formulation emerges through a standard analysis of two-state CTMCs, in which this probability can be found through the inverse of the Laplace transform of the inverse of the resolvent of the Q-matrix for the system. Incorporating this value in closed form is the main hurdle in extending our results to general systems with more than two states. Using this formulation, the following lemma shows $\mu_i(\sigma)$.

LEMMA 3. *Let $T_i(\sigma_i)$ be as defined in Lemma 1. The fraction of time a driver following strategy $\sigma=\{\sigma_1,\sigma_2\}$ spends either open in state i or on a trip started in state i is*

$$\mu_i(\sigma)=\frac{\lambda_i F_i(\sigma_i)T_i(\sigma_i)Q_j(\sigma_j)}{\lambda_j F_j(\sigma_j)T_j(\sigma_j)Q_i(\sigma_i)+\lambda_i F_i(\sigma_i)T_i(\sigma_i)Q_j(\sigma_j)}$$
$$\textit{where}\quad Q_i(\sigma_i)=\lambda_{i\to j}+\lambda_i\int_{\tau\in\sigma_i}q_{i\to j}(\tau)dF_i(\tau)$$

We prove this lemma by finding the expected number of sub-cycles in each state i, i.e., within a larger renewal reward cycle as defined, the expected number of sub-cycles that start with the driver being open in state i. This expectation is the mean of a geometric random variable parameterized by the probability that the driver will next be open in state j, given she is currently open in state i. $Q_i(\sigma_i)$ is proportional to this probability. (As in the case of $T_i(\sigma_i)$, there is a normalizing constant $\frac{1}{\lambda_i F_i(\sigma_i)+\lambda_{i\to j}}$); the larger it is, the fewer sub-cycles that are spent in state i. It has two components: the first is the probability that the state changes before the driver accepts a trip request; the second is the probability that the world state is j when the driver completes a trip. Thus, the numerator in $\mu_i(\sigma)$ is proportional to the length of a sub-cycle in state i, times the fraction of such cycles that are started in state i. The larger $Q_j(\sigma_j)$ or $T_i(\sigma_i)$ is, the more time the driver spends in state i.

4.2. Incentive Compatible pricing

How can the platform create incentive compatible pricing given the previously described effects? Our main result establishes when such IC prices exist, and reveals their form.

THEOREM 3. *Let R_1 and R_2 be target earning rates during non-surged and surge states, respectively. There exist prices $w = \{w_1, w_2\}$ of the form*

$$w_i(\tau) = m_i\tau + z_i q_{i\to j}(\tau),$$

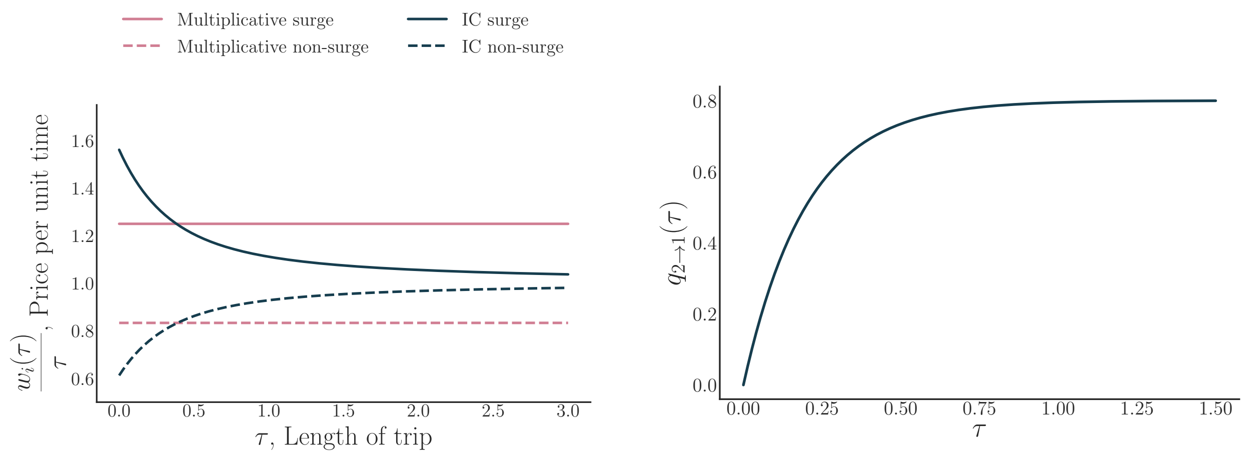
where $m_1, m_2, z_2 \geq 0$ (but z_1 may be either positive or negative), such that the optimal driver policy is to accept every trip in the surge state and all trips up to a certain length in the non-surge state. Furthermore, for $\frac{R_1}{R_2} \in [C, 1]$, there exist fully incentive compatible prices of this form, where

$$C = 1 - \frac{1}{T_1}\frac{Q_2(\lambda_{12}T_1 - Q_1) + Q_1(T_2\lambda_{1\to 2} + Q_2)}{Q_2(\lambda_{1\to 2}T_1 - Q_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2)} \in [0, 1),$$

and $T_i = \lambda_i F_i(\sigma_i) T_i((0,\infty))$, and $Q_i = Q_i((0,\infty))$.

We discuss a sketch of the proof in Section 4.4, with the complete proof in the appendix. To convey intuition, Figure 3a shows pricing functions in each state, plotting $\frac{w_i(\tau)}{\tau}$ against τ. Compared to multiplicative pricing with constant $\frac{w_i(\tau)}{\tau}$, IC surge pricing pays more for short trips, to compensate drivers for their opportunity cost, and less for short trips. Inversely, IC non-surge pricing pays more for long trips than it does for short trips. Further, as τ increases, $w_1(\tau)$ approaches $w_2(\tau)$, reflecting the fact that the opportunity cost for long trips does not depend as strongly on the state in which it started (as discussed in Section 4.3). Next, observe that IC surge pricing $w_2(\tau) = m_2\tau + z_2 q_{2\to 1}(\tau)$ is approximately affine: $q_{2\to 1}(\tau)$ (plotted in Figure 3b) is upper bounded by $\frac{\lambda_{2\to 1}}{\lambda_{1\to 2}+\lambda_{2\to 1}}$, and so is eventually approximately constant even as τ increases. The two components of pricing, m_i and z_i, thus balance the comparative benefit of long and short trips.

We note that, rather surprisingly and contrary to the focus of platform designers, it is the *non-surge* state that is difficult to make incentive compatible. Our result establishes that there always exist pricing schemes, for any target driver earning rates $R_1 < R_2$, such that accepting every trip in the surge state is driver optimal; we cannot say the same for the non-surge state. We give further intuition for the form of the pricing function w_i and the range $[C, 1]$ in the next subsection, showing how they emerge from the driver's opportunity cost.



(a) Price per unit time $\frac{w_i(\tau)}{\tau}$ for trips of different lengths τ in the each state for Incentive Compatible and multiplicative pricing when $R_2 = 1$ and $R_1 = \frac{2}{3}$.

(b) $q_{2\to1}(\tau)$ when $\lambda_{1\to2} = 1, \lambda_{2\to1} = 4$. The shape suggests that IC surge pricing is well-approximated by an affine function: $z_2 q_{2\to1}(\tau)$ remains essentially constant for longer trips.

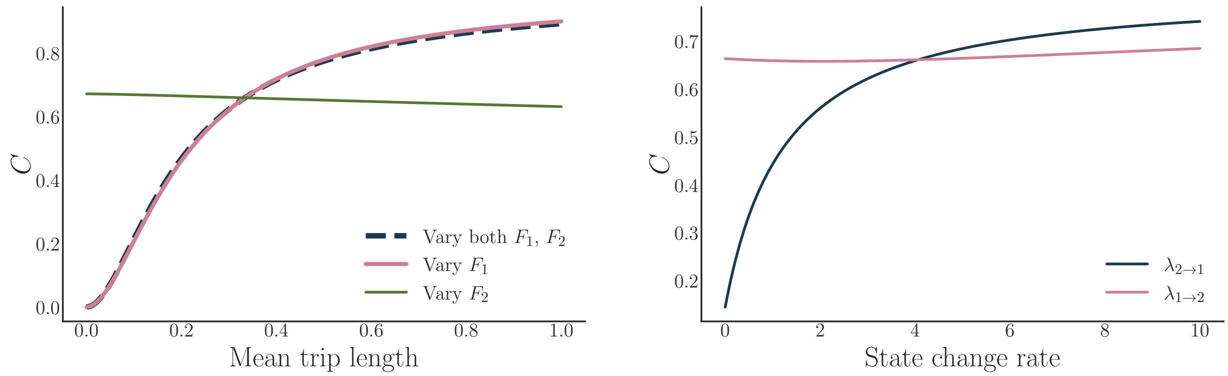
Figure 3 **Using the same model primitives as in Figure 2: (1) how the incentive compatible pricing compares to multiplicative pricing, and (2) $q_{2\to1}(\tau)$**

Finally, for a given feasible R_1, R_2, there is a range of m_i, z_i that form an incentive compatible pricing scheme. Why? If a driver rejects a trip request, she waits to receive another request, during which time she does not earn. This wait time tilts the driver toward accepting any trip request to maximize earnings. Thus, there is flexibility in the balance between short and long trip earnings. The same insight drives Proposition 1; even in the single-state model, trips do not have to have the same earnings per unit time, $\frac{w(\tau)}{\tau}$, as long as they meet some minimum threshold, $\frac{w(\tau)}{\tau} \geq c_w$.

4.3. Opportunity cost intuition for incentive compatible pricing

We now present some intuition to understand Theorem 3 and our incentive compatible pricing scheme. To do so, we introduce a weaker property than incentive compatibility, **trip indifference**: given a specific trip request length τ, in expectation the driver is at least as well off accepting the request as she is rejecting it, assuming she will accept any future request.[8]

[8] If w is incentive compatible, then it also satisfies trip indifference. The latter criterion is a *local* condition in which the driver cannot infinitesimally improve her earnings by instead following strategy $\sigma' = (0, \infty) \setminus \{\tau\}$ ("infinitesimally", as trip length τ has measure 0 by assumption); in contrast incentive compatibility is a *global* condition on the space of driver policies, requiring that the policy σ that accepts all trips globally maximize the earnings rate, i.e. $R(w, \sigma) \geq R(w, \sigma')$, for all σ'.



(a) C as the mean trip length changes. When not varied, the mean trip length in a given state is $\frac{1}{3}$.

(b) C as $\lambda_{i\to j}$ change. When $\lambda_{2\to1}$ is varied, $\lambda_{1\to2} = 1$. When $\lambda_{1\to2}$ is varied, $\lambda_{2\to1} = 4$.

Figure 4 How C, the ratio R_1/R_2 at which IC pricing is feasible from Theorem 3, changes (1) with respect to the mean trip length, and (2) with respect to $\lambda_{i\to j}$. Except for those that are varied in each plot, the primitives are fixed to those used in Figure 2.

Trip indifference allows us to illustrate the various features that must be incorporated into any IC pricing scheme: given an accepted trip request, it is simple in our model to formulate theoretically the driver's counter-factual expected earnings in a certain time window if she had instead rejected the request, i.e. her opportunity cost for accepting the trip.[9] Intuitively, the amount $w_i(\tau)$ that the driver is paid for the trip must account for this opportunity cost, i.e. in a VCG-like manner. Of course, this opportunity cost itself depends on the pricing scheme w. We now break down parts of this opportunity cost.

On-trip opportunity cost. While the driver is on-trip, the world state continues to evolve: surge might end or start, and such changes affect the opportunity cost. We call this component the "network minutes" cost.

Let $\phi_i^k(\tau)$ be the expected amount of time that the world is in state k during time $(t, t+\tau)$, given that it is in state i at time t. Then, by integrating $q_{i\to j}(s)$ from 0 to τ:

$$\phi_i^i(\tau) = \left[\frac{\lambda_{j\to i}}{\lambda_{i\to j} + \lambda_{j\to i}}\right]\tau + \left[\frac{1}{\lambda_{i\to j} + \lambda_{j\to i}}\right] q_{i\to j}(\tau)$$
$$\phi_i^j(\tau) = \left[\frac{\lambda_{i\to j}}{\lambda_{i\to j} + \lambda_{j\to i}}\right]\tau - \left[\frac{1}{\lambda_{i\to j} + \lambda_{j\to i}}\right] q_{i\to j}(\tau) = \tau - \phi_i^i(\tau)$$

What does this tell us about the opportunity cost? Let us define $\tilde{R}_i$ as the driver's earnings rate while the world state is i (whether the driver is open, or on a trip that started in

[9] Similarly, it is easier to measure empirically: counter-factual driver earnings for a accepted trip request can be approximated by measuring the future earnings of other nearby drivers who did not receive that trip request. Verifying incentive compatibility, on the other hand, requires full off-policy learning and estimation.

either state). $\tilde{R}_i$ is close to but not exactly R_i, which instead is the earnings rate counting open time and trips that *start* in state i. Then, the driver's opportunity cost during time $(t, t+\tau)$, starting in state i is

$$\tilde{R}_i \phi_i^i(\tau) + \tilde{R}_j \phi_i^j(\tau) = \left[\frac{\lambda_{j\to i}\tilde{R}_i + \lambda_{i\to j}\tilde{R}_j}{\lambda_{i\to j} + \lambda_{j\to i}}\right]\tau + \left[\frac{\tilde{R}_i - \tilde{R}_j}{\lambda_{i\to j} + \lambda_{j\to i}}\right] q_{i\to j}(\tau)$$

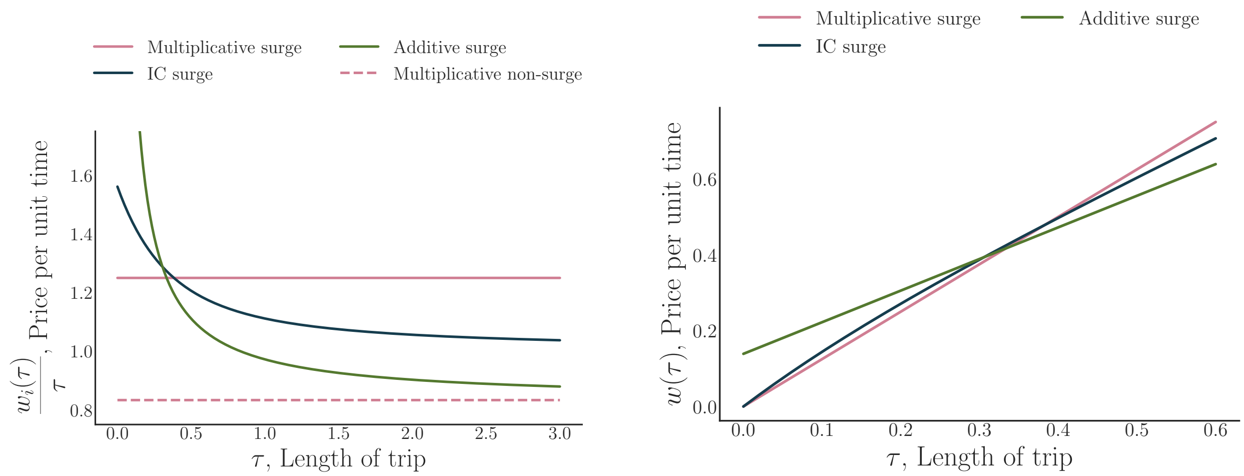
Though $\tilde{R}_i$ is not a simple expression in terms of R_i, several insights emerge:

One. The network minutes opportunity cost is of the form, $m_i'\tau + z_i' q_{i\to j}(\tau)$, for some m_i', z_i'. This matches the shape of our IC pricing scheme, which has different m_i, z_i that incorporate complications ignored here.

Two. As trip length $\tau \to \infty$, the first component $\left[\frac{\lambda_{j\to i}\tilde{R}_i + \lambda_{i\to j}\tilde{R}_j}{\lambda_{i\to j}+\lambda_{j\to i}}\right]\tau$ dominates the opportunity cost. This component is the same whether the given trip request starts in state $i = 1$ or $i = 2$, i.e. the stationary distribution of a positive recurrent CTMC does not depend on the starting state. This fact implies that we cannot always construct incentive compatible prices, for any R_1, R_2: as $\tau \to \infty$, the trip's opportunity cost does not depend on the starting state i, and so the trip's payments must be similar, $w_1(\tau) \approx w_2(\tau)$. When all trips in the non-surge state are long, i.e. F_1 is concentrated around large values, the earnings rate in each state must be similar, $R_1 \approx R_2$.

C exactly encodes such constraints, as shown in Figure 4. As the mean of $\tau \sim F_1$ goes to 0, then $\lambda_{12}T_1 - Q_1 \to 0$ and so $C \to 0$, and so the range of feasible $\frac{R_1}{R_2}$ expands. Similarly, $\lambda_{2\to 1}$ also plays a large role. When small, the surge state is long. Thus, regardless of how long a driver's last non-surge trip is, she will receive many trips during surge – and so long trips during non-surge are no longer constrained to be highly paid compared to short trips.

Continuation value opportunity cost The previous discussion misses a crucial detail: it is not sufficient to consider just the opportunity cost for the duration of the trip. A driver's counter-factual earnings by rejecting the trip depends on future trips that she accepts. Such counter-factual trips both (1) pay the driver according to their starting state even after a world state transition, i.e. the difference between R_i and $\tilde{R}_i$ above; and (2) potentially are still in progress past time $t+\tau$, when the current trip ends. This second complication is illustrated in Figure 2, where a driver can extend the time she spends on trips starting in the surge state by rejecting short surge trips. The effect depends on the lengths of future potential trips, i.e. $T_i(\sigma_i)$, and state transitions during those trips, $Q_i(\sigma_i)$, and is incorporated in both C and the pricing scheme.



(a) Price per unit time $\frac{w_i(\tau)}{\tau}$ (b) Price $w_2(\tau)$

Figure 5 **Using the same model primitives as in Figure 2: the payment function $w_i(\tau)$ for various surge mechanisms plotted two ways, when $R_2 = 1$ and $R_1 = \frac{2}{3}$ for drivers who accept every trip.**

4.4. Proof sketch of Theorem 3

The result is shown in the appendix by manipulating the derivative of the reward function with respect to the policy σ. In particular, when the pricing function is of the given form with the appropriate constants m_i, z_i, then *any* policy $\sigma = \{\sigma_1, \sigma_2\}$ can be locally improved by adding more trips to it, i.e. the overall reward is non-decreasing as the driver accepts more trips: $R(\sigma') \geq R(\sigma), \forall \sigma \subseteq \sigma'$. This result follows from $\frac{\partial}{\partial u} R(\sigma) \geq 0$, for all u, σ, given the constraints, where u is an upper endpoint of the policy in a state, $\sigma_i = \cup_k (\ell_k, u_k)$.

The key step is finding sufficient constraints for this derivative to be positive with a pricing function of the given form, given any σ_i, as opposed to just $\sigma_i = (0, \infty)$. This difficulty emerges because incentive compatibility is a global condition on the set function $R(w, \sigma)$. In particular, we need to express these constraints simply – e.g. as a function of just $T_i((0, \infty)), Q_i((0, \infty))$, instead of the values $T_i(\sigma_i), Q_i(\sigma_i), \forall \sigma_i \subseteq \mathbb{R}_+$. The C presented in the theorem statement results from such a set of constraints on m_i, z_i.

5. Approximate Incentive Compatibility with Additive Surge

We now analyze surge policies that reflect practice at ride-hailing platforms today, as they are simple to communicate through a heat-map. Non-surge pricing is typically purely *multiplicative*, i.e $w_1(\tau) = m_1 \tau$, where m_1 is the base time (and distance) rate for a ride.

We consider two types of affine *surge* pricing w_2, which differ in their relationship to w_1 through a single parameter:

$$\textbf{Multiplicative surge:} \quad w_2(\tau) = m_2\tau \qquad m_2 \geq m_1$$

$$\textbf{Additive surge:} \quad w_2(\tau) = m_1\tau + a_2 \qquad a_2 \geq 0$$

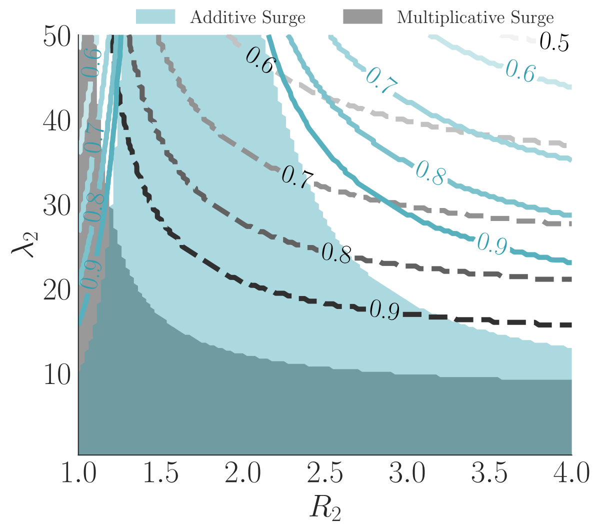
In multiplicative surge, a higher multiplier is used than the base fare m_1, and $\frac{m_2}{m_1}$ is reported on the heatmap as in Figure 1a; in additive surge, the *same* base fare multiplier m_1 on the trip length is used in both surge and non-surge times, with an additive factor a_2 during surge that is reported on the heatmap in Figure 1b. These surge functions are simple to calculate, given fixed primitives and target earnings rate R_2 in the surge state: m_2 or a_2 are determined given these values.

Figure 5 shows these types of pricing, compared to the incentive compatible pricing function. Multiplicative surge has constant $\frac{w_2(\tau)}{\tau}$ and so under-pays short trips and over-pays long-trips compared to IC pricing. Additive surge asymptotically (for large τ) pays the same as multiplicative non-surge pricing, i.e. $\lim_{\tau\to\infty} \frac{w_2(\tau)}{\tau} = \lim_{\tau\to\infty} \frac{w_1(\tau)}{\tau} = m_1$. As a result, it over-pays short trips and under-pays long trips compared to IC surge pricing.

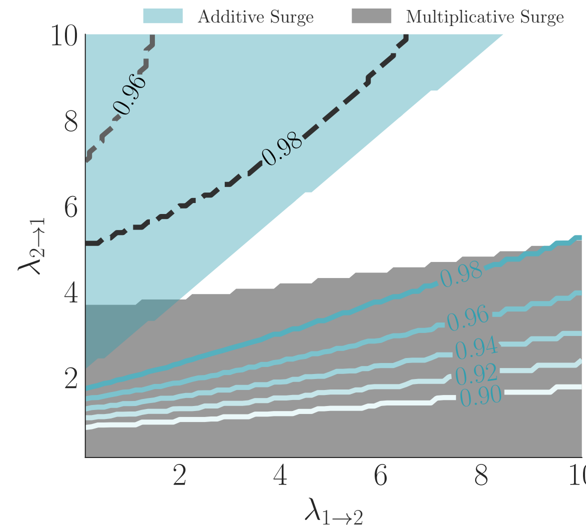
Uber has recently started a transition from multiplicative to additive surge. In this section, we argue that the additive component is more important than the multiplicative component for incentive compatibility, motivating Uber's transition.

Computing optimal driver policies Recall that Theorem 2 establishes that multiplicative pricing (and, more generally, affine pricing) may not be incentive compatible. However, we still wish to compare the various types of surge pricing. We thus compare how approximately compatible these pricing functions are, in the sense of what fraction of trips is accepted by an optimal driver policy with respect to the pricing function.

However, to do this comparison, one needs to calculate optimal driver policies with respect to a pricing function. Recall that the optimal driver policy in each state σ_i is some subset of $\mathbb{R}_+$. Finding such optimal subsets for general pricing functions w is computationally intractable. Thus, Theorem 2 is particularly important for *computational* reasons. It establishes that, for any affine pricing structure in the surge state, there exists a driver optimal policy of the form $(0, t_1) \cup (t_2, \infty)$, for some t_1, t_2. Thus, we only need to find the values for these parameters that maximize the driver reward among sets of this form, and the resulting policy is optimal. Deriving closed forms is still intractable, but we can



(a) With R_2, earnings rate in the surge state, and λ_2, arrival rate of jobs in the surge state. $R_2 \in [1.1, 3]$ is common in practice.



(b) With $\lambda_{2\to1}, \lambda_{1\to2}$, rates for world state changing. $\lambda_{2\to1} \gg \lambda_{1\to2}$ is common in practice.

Figure 6 **How the fraction of trips accepted in the surge state in the optimal policy varies with primitives. The shaded regions are areas where the respective surge type is fully incentive compatible in the surge state ($\sigma_2 = (0, \infty)$ is optimal). The plots also include contour lines, with values embedded. When not varied, the primitives are $\lambda_1 = \lambda_2 = 10, \lambda_{1\to2} = 1, \lambda_{2\to1} = 4, R_2 = 1, R_1 = 0.3$, and the mean trip length is 0.3. We assume every trip is accepted in the non-surge state.**

computationally find them through grid search and numeric integration. Note that the proposition does not establish uniqueness of the driver optimal policy; we thus choose the policy that maximizes the fraction of trips accepted in our computations.

Approximate incentive compatibility We now study how (approximately) incentive compatible the surge mechanisms are, i.e. the fraction of trips accepted in the surge state by the optimal policy. Figure 6 shows how this fraction changes with the primitives. The shaded regions correspond to areas where the surge pricing function is fully incentive compatible in the surge state ($\sigma_2 = (0, \infty)$ is optimal), and the lines are contour lines for approximate incentive compatibility, indicating the fraction of trips accepted. For example, when $R_2 = 3, \lambda_2 = 30$, then about 90% are accepted with additive surge, and 70% are accepted with multiplicative surge.

Overall, we note that additive surge is far more approximately incentive compatible in the most common parameter regimes for ride-hailing platforms such as Uber: (1) surge is between 1.1 and 3 times more valuable than non-surge; (2) surge is short-lived compared to non-surge periods ($\lambda_{2\to1} \gg \lambda_{1\to2}$); (3) and in a typical surge the driver will be able to receive several trip requests ($\frac{\lambda_2}{\lambda_{2\to1}} > 1$, but small) but may only be able to complete one or two such trips ($\frac{1}{\lambda_{2\to1}} \approx$ mean trip length). In each of these regimes, additive surge is either fully incentive compatible or more approximately IC than is multiplicative surge. For example, with $R_2 = 2$ (i.e. a surge multiplier of 2), every trip is accepted with additive surge for any λ_2 in our range, whereas up to 40% of trips are rejected with multiplicative surge. These simulations thus support Uber's recent shift from multiplicative to additive surge. We can also draw qualitative insights in terms of sensitivity to the primitives, similar in spirit to effects in the form of C in Theorem 3.

Figure 6a shows how the approximate IC properties of additive and multiplicative surge change with λ_2 and R_2. As the arrival rate of jobs in the surge state, λ_2, increases, both types of surge become less incentive compatible: "cherry-picking" becomes easier, as the driver is likely to receive many more trip requests before surge ends. Similarly, as surge becomes increasingly more valuable compared to non-surge (R_2 increases), the incentive to reject non-valuable trips in the surge state increases (short trips with multiplicative surge, long trips with additive surge).

For additive surge, an interesting non-monotonicity with R_2: when $R_2 \gg R_1$, the effect above dominates, and long trips are rejected. When the surge state is moderately more valuable than non-surge, additive surge effectively balances the payments for different trip lengths, and so is incentive compatible. When the two states become almost equally valuable, however, again the optimal driver policy with additive surge rejects long trips: the system approximates our single-state model, and so additive surge may not be incentive compatible, cf. Theorem 1.

Figure 6b shows the effects of the relative lengths of surge and non-surge. Note that, here, the two types of surge are incentive compatible in exactly opposing regimes. When $\frac{\lambda_{2\to1}}{\lambda_{1\to2}}$ is large, surge is comparatively rare and short, and so short trips are naturally undervalued — accepting them decreases the time spent in the surge state — and additive surge is more incentive compatible. With long-lasting surge (small $\frac{\lambda_{2\to1}}{\lambda_{1\to2}}$), on the other hand, the world almost seems unchanging in the surge state, and so multiplicative surge nears incentive compatibility. In modern ride-hailing platforms, the scenario with short, in-frequent surge is more common.

6. Conclusion

In this work, we studied the problem of designing incentive compatible mechanisms for ride-hailing marketplaces. We presented a dynamic model to capture essential features of these environments. Even-though our model is simple and stylized, it highlights how driver incentives and subsequently dynamic pricing strategies would change in the presence of stochasticity. We hope our work inspires other researchers in this area to incorporate such uncertainty in their models, as it is one of the biggest challenges faced in practice.

An important direction for extending our work is studying matching and pricing polices jointly, i.e. how to best match open drivers to riders in the presence of such effects, cf. (Özkan and Ward 2016, Banerjee et al. 2017, Feng et al. 2017, Zhang et al. 2017, Banerjee et al. 2018, Hu and Zhou 2018, Korolko et al. 2018, Ashlagi et al. 2018, Kanoria and Qian 2019). In this work, we look at incentive compatible pricing. The platform, in addition to pricing, can use matching policy to align incentives.

Acknowledgments

We would like to thank Uber's driver pricing data science team, in particular Carter Mundell, Jake Edison, Alice Lu, Michael Sheldon, Margaret Tian, Qitang Wang, Peter Cohen, and Kane Sweeney for their support and suggestions without which this work would have not been possible. We also thank Leighton Barnes, Ashish Goel, Ramesh Johari, Vijay Kamble, Hannah Li, and Virag Shah. This work was funded in part by the Stanford Cyber Initiative, the Office of Naval Research grant N00014-15-1-2786, and National Science Foundation grant 1544548.

References

Afèche P, Liu Z, Maglaras C (2018) Ride-Hailing Networks with Strategic Drivers: The Impact of Platform Control Capabilities on Performance. *SSRN Electronic Journal* ISSN 1556-5068, URL `http://dx.doi.org/10.2139/ssrn.3120544`.

Ashlagi I, Burq M, Jaillet P, Saberi A (2018) Maximizing efficiency in dynamic matching markets. *arXiv preprint arXiv:1803.01285* URL `https://arxiv.org/pdf/1803.01285.pdf`.

Ata B, Barjesteh N, Kumar S (2019) Spatial Pricing: An Empirical Analysis of Taxi Rides in New York City. *Working Paper* .

Bai J, So KC, Tang CS, Chen XM, Wang H (2018) Coordinating Supply and Demand on an On-Demand Service Platform with Impatient Customers. *Manufacturing & Service Operations Management* URL `http://dx.doi.org/10.1287/msom.2018.0707`.

Banerjee S, Gollapudi S, Kollias K, Munagala K (2017) Segmenting two-sided markets. *Proceedings of the 26th International Conference on World Wide Web*, 63–72.

Banerjee S, Kanoria Y, Qian P (2018) State Dependent Control of Closed Queueing Networks with Application to Ride-Hailing URL http://arxiv.org/abs/1803.04959.

Banerjee S, Riquelme C, Johari R (2015) Pricing in Ride-Share Platforms: A Queueing-Theoretic Approach. *SSRN Electronic Journal* ISSN 1556-5068, URL http://dx.doi.org/10.2139/ssrn.2568258.

Bertsimas DJ, van Ryzin G (1991) A Stochastic and Dynamic Vehicle Routing Problem in the Euclidean Plane. *Operations Research* 39(4):601–615, ISSN 0030-364X, 1526-5463, URL http://dx.doi.org/10.1287/opre.39.4.601.

Bertsimas DJ, van Ryzin G (1993) Stochastic and Dynamic Vehicle Routing in the Euclidean Plane with Multiple Capacitated Vehicles. *Operations Research* 41(1):60–76, ISSN 0030-364X, 1526-5463, URL http://dx.doi.org/10.1287/opre.41.1.60.

Besbes O, Castro F, Lobel I (2018a) Spatial Capacity Planning. *SSRN Electronic Journal* ISSN 1556-5068, URL http://dx.doi.org/10.2139/ssrn.3292651.

Besbes O, Castro F, Lobel I (2018b) Surge Pricing and Its Spatial Supply Response. *SSRN Electronic Journal* ISSN 1556-5068, URL http://dx.doi.org/10.2139/ssrn.3124571.

Bimpikis K, Candogan O, Saban D (2016) Spatial Pricing in Ride-Sharing Networks. SSRN Scholarly Paper ID 2868080, Social Science Research Network, Rochester, NY, URL https://papers.ssrn.com/abstract=2868080.

Buchholz N (2017) Spatial Equilibrium, Search Frictions and Efficient Regulation in the Taxi Industry URL https://scholar.princeton.edu/sites/default/files/nbuchholz/files/taxi_draft.pdf.

Cachon GP, Daniels KM, Lobel R (2017) The Role of Surge Pricing on a Service Platform with Self-Scheduling Capacity. *Manufacturing & Service Operations Management* 19(3):368–384, ISSN 1523-4614, URL http://dx.doi.org/10.1287/msom.2017.0618.

Castillo JC, Knoepfle D, Weyl G (2017) Surge Pricing Solves the Wild Goose Chase. 241–242 (ACM Press), ISBN 978-1-4503-4527-9, URL http://dx.doi.org/10.1145/3033274.3085098.

Chen MK, Sheldon M (2016) Dynamic Pricing in a Labor Market: Surge Pricing and Flexible Work on the Uber Platform URL http://dx.doi.org/10.1145/2940716.2940798.

Chen Y, Hu M (2018) Pricing and Matching with Forward-Looking Buyers and Sellers. SSRN Scholarly Paper ID 2859864, Social Science Research Network, Rochester, NY, URL https://papers.ssrn.com/abstract=2859864.

Cook C, Diamond R, Hall J, List J, Oyer P (2018) The Gender Earnings Gap in the Gig Economy: Evidence from over a Million Rideshare Drivers URL http://dx.doi.org/10.3386/w24732.

Cramer J, Krueger AB (2016) Disruptive Change in the Taxi Business: The Case of Uber. *American Economic Review* 106(5):177–182, ISSN 0002-8282, URL http://dx.doi.org/10.1257/aer.p20161002.

Feng G, Kong G, Wang Z (2017) We are on the way: Analysis of on-demand ride-hailing systems URL https://dx.doi.org/10.2139/ssrn.2960991.

Gallager RG (2013) *Stochastic Processes: Theory for Applications* (Cambridge University Press), ISBN 978-1-107-03975-9.

Guda H, Subramanian U (2019) Your uber is arriving: Managing on-demand workers through surge pricing, forecast communication, and worker incentives. *Management Science* 65(5):1995–2014, URL http://dx.doi.org/10.1287/mnsc.2018.3050.

Hall JV, Horton JJ, Knoepfle DT (2017) Labor Market Equilibration: Evidence from Uber URL https://eng.uber.com/research/labor-market-equilibration-evidence-from-uber/.

Hall JV, Kendrick C, Nosko C (2015) The effects of Uber's surge pricing: A case study URL https://eng.uber.com/research/the-effects-of-ubers-surge-pricing-a-case-study/.

Hu M, Zhou Y (2018) Dynamic type matching. *Rotman School of Management Working Paper* (2592622).

Kamble V (2018) Revenue Management on an On-Demand Service Platform URL http://arxiv.org/abs/1803.06797.

Kanoria Y, Qian P (2019) Near Optimal Control of a Ride-Hailing Platform via Mirror Backpressure URL http://arxiv.org/abs/1903.02764.

Korolko N, Woodard D, Yan C, Zhu H (2018) Dynamic Pricing and Matching in Ride-Hailing Platforms. *SSRN Electronic Journal* 40, ISSN 1556-5068, URL http://dx.doi.org/10.2139/ssrn.3258234.

Lu A, Frazier PI, Kislev O (2018) Surge Pricing Moves Uber's Driver-Partners. *Proceedings of the 2018 ACM Conference on Economics and Computation*, 3–3, EC '18 (New York, NY, USA: ACM), ISBN 978-1-4503-5829-3, URL http://dx.doi.org/10.1145/3219166.3219192.

Ma H, Fang F, Parkes DC (2018) Spatio-Temporal Pricing for Ridesharing Platforms URL http://arxiv.org/abs/1801.04015.

Özkan E, Ward A (2016) Dynamic Matching for Real-Time Ridesharing. *SSRN Electronic Journal* ISSN 1556-5068, URL http://dx.doi.org/10.2139/ssrn.2844451.

Uber (2019a) Community Guidelines. URL https://www.uber.com/legal/community-guidelines/us-en/.

Uber (2019b) Dependable Earnings. URL https://www.uber.com/drive/resources/dependable-earnings/.

Uber (2019c) How are fares calculated. URL https://help.uber.com/riders/article/how-are-fares-calculated?nodeId=d2d43bbc-f4bb-4882-b8bb-4bd8acf03a9d.

Uber (2019d) New Driver Surge. URL https://www.uber.com/blog/your-questions-about-the-new-surge-answered/.

Uber (2019e) Service Fee. URL https://marketplace.uber.com/pricing/service-fee.

Wald A (1973) *Sequential analysis* (Courier Corporation).

Yang P, Iyer K, Frazier P (2018) Mean field equilibria for resource competition in spatial settings. *Stochastic Systems* 8(4):307–334.

Zhang L, Hu T, Min Y, Wu G, Zhang J, Feng P, Gong P, Ye J (2017) A taxi order dispatch model based on combinatorial optimization. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2151–2159 (ACM).

Appendix A: Proofs of main text theorems and lemmas

In this Appendix, we provide proofs of the theorems and lemmas in the main text. The proofs rely on claims proved in Appendix Section B.

A.1. Notation and assumptions

Notation.

- In general for the dynamic model, we use σ in the function argument when the function depends on policies in both states, and σ_i when it only depends on the policy in state i.
- In the dynamic model, let $\Delta(\sigma_i, \sigma_j) = R_i(w_i, \sigma_i) - R_j(w_j, \sigma_j)$.
- All policy equalities are up to measure 0.
- We use $\propto$ to denote *has the same sign as*, rather than *proportional to*

Assumptions (repeat from main text).

- Distribution of jobs F, F_i is a continuous probability measure, i.e. f, f_i bounded.
- Payment functions w, w_i are continuous.
- We assume that there exists a policy in state 2 that dominates state 1: $\exists \sigma_2$ such that $\Delta(\sigma_2, \sigma_1) > 0, \forall \sigma_1 \subseteq (0, \infty)$.
- σ, σ_i constrained to be measurable with respect to F, F_i, and σ_i are open.

A.2. Single-state model theorem and propositions proofs

A.2.1. Proof of Theorem 1 We now prove Theorem 1, regarding the form of the optimal policy in the single-state model – where the *length* of a trip does not matter, only the earnings rate. The optimal policy trades off the earnings rate while on a trip with the driver's utilization rate. At a high level, the proof proceeds as follows: starting from any policy that is not of the appropriate form, we replace trips in the policy with those with a higher earnings rate, while keeping the utilization rate exactly the same. Such replacements result in a policy that is *almost* of the correct form, except there may be an earnings rate c such that only a *subset* of $\{\tau : \frac{w(\tau)}{\tau} = c\}$ is in the policy. The remainder of the proof is showing that such a policy can be improved to form a policy of the appropriate form.

THEOREM 1. *In the single-state model, for each w there exists a constant $c_w \in \mathbb{R}_+$ such that the policy $\sigma^* = \left\{\tau : \frac{w(\tau)}{\tau} \geq c_w\right\}$ is optimal for the driver with respect to w.*

Proof. Let $\gamma(\tau) \triangleq \frac{w(\tau)}{\tau}$. Assume that $F(\{\tau : w(\tau) > 0\}) > 0$. Otherwise any policy is optimal and so the result is trivial.

Start at $\sigma \subsetneq (0,\infty)$. We first show that there exists $c \in \mathbb{R}_+$ such that $\sigma_c = \{\tau : \gamma(\tau) \geq c\}$ or $\sigma_c = \{\tau : \gamma(\tau) > c\}$ such that $R(w,\sigma_c) \geq R(w,\sigma)$. Assume that $0 < F(\{\tau : w(\tau) > 0\} \cap \sigma) < 1$. (If $F(\{\tau : w(\tau) > 0\} \cap \sigma)$ is either 0 or 1, we are done, as $R((0,\infty)) \geq R(w,\sigma)$ and is of the desired form.)

1. First we construct $\tilde{\sigma}_c = \{\tau : \gamma(\tau) > c\} \cup C$, where $C \subseteq \{\tau : \gamma(\tau) = c\}$ and $R(w,\tilde{\sigma}_c) \geq R(w,\sigma)$.

   For the given σ, c, let

$$
\begin{aligned}
A_c &= \{\tau : \tau \notin \sigma, \gamma(\tau) \geq c\} \\
B_c &= \{\tau : \tau \in \sigma, \gamma(\tau) < c\} \\
L(X) &= \int_{x \in X} \tau dF(\tau) \qquad\qquad X \subseteq (0,\infty)
\end{aligned}
$$

   A_c is a set of trips that pay more than c per unit time but are not in σ, and B_c is the set of the trips that pay less than c per unit time but are not in σ. $L(X)$ is the mean extra utilization that trips in X contribute in a renewal cycle. The idea is that if we find sets A, B such that $L(A) = L(B) > 0$ and $\gamma(a) > \gamma(b), \forall a \in A, b \in B$, then $\sigma' = \sigma \cup A \setminus B \implies R(w,\sigma') > R(w,\sigma)$: the denominator of the reward stays the same, and the numerator increases. A few facts that follow from assumptions:

   - $L(A_0) > 0$
   - $\exists c : L(B_c) > 0$
   - $L(B_c)$ is non-decreasing as c increases, and $L(B_0) = 0$
   - $L(A_c)$ is non-increasing as c increases, and $\lim_{c\to\infty} L(A_c) = 0$
   - $L(A_c)$, $L(B_c)$ both are continuous from the left in c.
   - The above imply that $\exists c'$ such that $L(A_c) < L(B_c), \forall c > c'$.
   - Thus, there exists $c_0 = \max\{c' : L(A_{c'}) \geq L(B_{c'})\}$

   If $L(A_{c_0}) = L(B_{c_0})$, then we are done with this part: let $\tilde{\sigma}_{c_0} = \sigma \cup A_{c_0} \setminus B_{c_0} = \{\tau : \gamma(\tau) \geq c\}$.

   Otherwise if $L(A_{c_0}) > L(B_{c_0})$ (which can happen if there is a point mass at $\gamma(\tau) = c$):

   - By max, for all $c > c_0$: $L(A_c) < L(B_c)$. Then

$$
L(B_{c_0}) < L(A_{c_0}) < L(B_{c_0}) + L(\{\tau : \tau \in \sigma, \gamma(\tau) = c_0\})
$$

   - let $C \subseteq \{\tau : \tau \in \sigma, \gamma(\tau) = c_0\}$ such that $L(C) + L(B_{c_0}) = L(A_{c_0})$. Such C exists by F continuous.
   - Let $\tilde{\sigma}_{c_0} = \sigma \cup A_{c_0} \setminus (C \cup B_{c_0})$

   We now have $\tilde{\sigma}_{c_0} = \{\tau : \gamma(\tau) \geq c_0\} \setminus C$, where $C \subseteq \{\tau : \tau \in \sigma, \gamma(\tau) = c_0\}$, and $R(w,\tilde{\sigma}_{c_0}) > R(w,\sigma)$, unless σ already was of the form $\tilde{\sigma}_{c_0}$ for some c_0.

2. Next, we construct $\sigma_{c_0} = \{\tau : \gamma(\tau) \geq c_0\}$ or $\sigma_{c_0} = \{\tau : \gamma(\tau) > c_0\}$ such that $R(w,\sigma_{c_0}) \geq R(w,\sigma)$.

- Suppose $c_0 \geq R(w, \tilde{\sigma}_{c_0})$. Then

$$R(w, \{\tau : \gamma(\tau) \geq c_0\}) = \frac{\lambda \int_{\tau \in \tilde{\sigma}_{c_0}} w(\tau) dF(\tau) + \lambda \int_{\tau \in C} w(\tau) dF(\tau)}{1 + \lambda \int_{\tau \in \tilde{\sigma}_{c_0}} \tau dF(\tau) + \lambda \int_{\tau \in C} \tau dF(\tau)}$$
$$\geq R(w, \tilde{\sigma}_{c_0}) \tag{1}$$

  Where the inequality follows from $R(w, \tilde{\sigma}_{c_0}) = \frac{\lambda \int_{\tau \in \tilde{\sigma}_{c_0}} w(\tau) dF(\tau)}{1 + \lambda \int_{\tau \in \tilde{\sigma}_{c_0}} \tau dF(\tau)}$,
  $\frac{\lambda \int_{\tau \in C} w(\tau) dF(\tau)}{\lambda \int_{\tau \in C} \tau dF(\tau)} = \frac{\lambda \int_{\tau \in C} \frac{w(\tau)}{\tau} \tau dF(\tau)}{\lambda \int_{\tau \in C} \tau dF(\tau)} = c_0$, and $\frac{w}{y} \leq \frac{x}{z} \implies \frac{w+x}{y+z} \geq \frac{w}{y}$.
  Then let $\sigma_{c_0} = \{\tau : \gamma(\tau) \geq c_0\}$

- Similarly, suppose $c_0 < R(w, \tilde{\sigma}_{c_0})$. Then

$$R(w, \{\tau : \gamma(\tau) > c_0\}) = \frac{\lambda \int_{\tau \in \tilde{\sigma}_{c_0}} w(\tau) dF(\tau) - \lambda \int_{\tau \in \{\tau : \tau \in \sigma, \gamma(\tau) = c_0\} \setminus C} w(\tau) dF(\tau)}{1 + \lambda \int_{\tau \in \tilde{\sigma}_{c_0}} \tau dF(\tau) - \lambda \int_{\tau \in \{\tau : \tau \in \sigma, \gamma(\tau) = c_0\} \setminus C} \tau dF(\tau)}$$
$$> R(w, \tilde{\sigma}_{c_0}) \tag{2}$$

  Where the inequality follows from $\frac{w}{y} > \frac{x}{z} \implies \frac{w-x}{y-z} > \frac{w}{y}$.
  Then let $\sigma_{c_0} = \{\tau : \gamma(\tau) > c_0\}$ (choosing arbitrarily if $c_0 = R(w, \tilde{\sigma}_{c_0})$)

Thus, we have shown that for all σ, there exists $\sigma_{c_0} = \{\tau : \gamma(\tau) > c_0\}$ or $\sigma_{c_0} = \{\tau : \gamma(\tau) \geq c_0\}$ such that $R(w, \sigma_{c_0}) \geq R(w, \sigma)$.

Let $\sigma_c^> = \{\tau : \gamma(\tau) > c\}$, and $\sigma_c^\geq = \{\tau : \gamma(\tau) \geq c\}$. We finish the proof by showing that $\exists c^*$ such that $\forall c, R(w, \sigma_{c^*}^\geq) \geq \max(R(w, \sigma_c^\geq), R(w, \sigma_c^>))$.

By assumption of $w(\tau)/\tau$, the reward is bounded: $0 \leq R(w, \sigma_c^>)$, $0 \leq R(w, \sigma_c^\geq)$; further, there exists C such that $\forall c > C$: $R(w, \sigma_c^>) < R((0, \infty)), R(w, \sigma_c^\geq) < R((0, \infty))$ (which follows from F a continuous distribution, and so as $c \to \infty$, $F(\{\tau : \gamma(\tau) \geq c\}) \to 0$.

Further, $R(w, \sigma_c^>)$ is continuous from the right in c, and $R(w, \sigma_c^\geq)$ is continuous from the left in c, and the two functions have the same points of discontinuities: c such that $F(\{\tau : \gamma(\tau) = c\}) > 0$ (and these are their only points of disagreement). Thus, the function $\max(R(w, \sigma_c^\geq), R(w, \sigma_c^>))$ of c attains its maximum at some $c^* \in [0, C]$.

In other words, there exists c^* such that $\forall c, \max(R(w, \sigma_{c^*}^\geq), R(w, \sigma_{c^*}^>)) \geq \max(R(w, \sigma_c^\geq), R(w, \sigma_c^>))$.

We finish by proving that $R(w, \sigma_{c^*}^\geq) \geq R(w, \sigma_{c^*}^>)$.

- Suppose $c^* \geq R(w, \sigma_{c^*}^>)$. Then, by the same argument as line (1), $R(w, \sigma_{c^*}^>) \leq R(w, \sigma_{c^*}^\geq)$.
- Suppose $c^* < R(w, \sigma_{c^*}^>)$.
  - If $\exists B : c^* < B$ such that the mass $F(\{\tau : \gamma(\tau) \in (c^*, B]\}) = 0$, then note that $\sigma_{c^*}^>$ is equal to $\sigma_B^\geq$ up to a set of measure 0, and so $R(w, \sigma_{c^*}^>) = R(w, \sigma_B^\geq)$.
  - Otherwise, let $B : c^* < B < R(w, \sigma_{c^*}^>)$, and note that $F(\{\tau : \gamma(\tau) \in (c^*, B]\}) > 0$. Then, by the same argument as in line (2), $R(w, \sigma_{c^*}^>) < R(w, \sigma_B^\geq) \leq \max(R(w, \sigma_{c^*}^\geq), R(w, \sigma_{c^*}^>)) = R(w, \sigma_{c^*}^\geq)$.

Thus, $R(w, \sigma_c^\geq) \geq R(w, \sigma_{c^*}^>)$, for some c.

Thus, for some c^*, the policy $\sigma_{c^*}^\geq = \{\tau : \gamma(\tau) \geq c^*\}$ is optimal. □

A.2.2. Proof of Proposition 1

PROPOSITION 1. *In the single-state model, $w(\tau) = m\tau + a$ is incentive compatible if $0 \leq a \leq \frac{m}{\lambda}$.*

Proof. Let $T = \int_{\tau \in (0,\infty)} \tau dF(\tau)$. Let $\sigma' = (0,\infty) \setminus \sigma$, for some σ.

$$R((0,\infty)) = \frac{\lambda \int_{\tau \in (0,\infty)} w(\tau) dF(\tau)}{1 + \lambda T}$$

$$R(\sigma') = \frac{\lambda \int_{\tau \in (0,\infty)} w(\tau) dF(\tau) - \lambda \int_{\tau \in \sigma} w(\tau) dF(\tau)}{1 + \lambda T - \lambda \int_{\tau \in \sigma} \tau dF(\tau)}$$

$$\implies R(\sigma') \leq R((0,\infty)) \iff \frac{\lambda \int_{\tau \in (0,\infty)} w(\tau) dF(\tau)}{1 + \lambda T} < \frac{\int_{\tau \in \sigma} w(\tau) dF(\tau)}{\int_{\tau \in \sigma} \tau dF(\tau)}$$

Where the last line follows from $\frac{w}{y} \leq \frac{x}{z} \iff \frac{w}{y} \geq \frac{w-x}{y-z}$.

Thus, a necessary and sufficient condition for incentive compatibility is that

$$\frac{\lambda \int_{\tau \in (0,\infty)} w(\tau) dF(\tau)}{1 + \lambda T} \leq \frac{\int_{\tau \in \sigma} w(\tau) dF(\tau)}{\int_{\tau \in \sigma} \tau dF(\tau)} \qquad \forall \sigma$$

Thus $w(\tau) = m\tau$ is immediately incentive compatible, for all F continuous.

Further, suppose $w(\tau) = m\tau + a$. Then, for $a \leq \frac{m}{\lambda}$:

$$\frac{\lambda \int_{\tau \in (0,\infty)} w(\tau) dF(\tau)}{1 + \lambda T} = \frac{\lambda(mT + a)}{1 + \lambda T} \leq \frac{m(1 + \lambda T)}{1 + \lambda T} = m < \frac{mT_2 + F(\sigma)a}{T_2}, \forall T_2 = \frac{\int_{\tau \in \sigma} w(\tau) dF(\tau)}{\int_{\tau \in \sigma} \tau dF(\tau)}$$

□

A.3. Dynamic model theorem proofs

A.3.1. Proof of Theorem 2 We now prove a theorem regarding the structure of optimal policies in reaction to various pricing functions. This theorem, alongside the non-incentive-compatibility examples presented in our numeric simulations, directly implies Theorem 2. The proof of Theorem 3 uses this theorem as well.

The general structure of the theorem is as follows: In Section B.1, we prove properties of the derivative of the reward function for multiplicative and affine pricing; and in Appendix Section B.2, we prove such properties for the Incentive Compatible pricing form. Further, in Section B.3, we show how such properties imply existence of optimal policies in each state of the appropriate form.

In this theorem:

- Start with some arbitrary policy $\sigma = \{\sigma_1, \sigma_2\}$.
- With assumption on the surge state providing higher potential earnings, replace σ_2 with a policy that provides higher earnings in state 2 than σ_1 does in state 1, without decreasing total reward.
- With theorems in B.3, replace σ_1 with policy of the appropriate form, without decreasing total reward.
- With theorems in B.3, replace σ_2 with policy of the appropriate form, without decreasing total reward.

THEOREM 4. *Consider pricing function $w = \{w_1, w_2\}$, where $i = 2$ is the surge state as defined. Then, there exists an optimal policy $\sigma = \{\sigma_1, \sigma_2\}$ that maximizes $R(w, \sigma)$, with the following properties.*

- *Non-surge state driver optimal policy σ_1:*

— *If* $w_1(\tau) = m_1\tau + a_1$*, for* $a_1 \geq 0$*, then* $\sigma_1 = (0, t_1)$*, for some* $t_1 \in [0, \infty) \cup \{\infty\}$*.*

— *If* $w_1(\tau) = m_1\tau - a_1$*, for* $a_1 > 0$*, then* $\sigma_1 = (t_2, t_3)$*, for some* $t_2, t_3 \in [0, \infty) \cup \{\infty\}$*.*

— *If* w_1 *such that* $\frac{\partial}{\partial u} R(w, \sigma' = \{\sigma'_1, \sigma'_2\}) \geq 0$ *for all* σ'*, where* u *is an upper endpoint of an interval that makes up* σ'_1*, then* $\sigma_1 = (0, \infty)$*.*

- *Surge state driver optimal policy* σ_2*:*

— *If* $w_2(\tau) = m_2\tau - a_2$*, for* $a_2 \geq 0$*, then* $\sigma_1 = (t_4, \infty)$*, for some* $t_4 \in [0, \infty)$*.*

— *If* $w_2(\tau) = m_2\tau + a_2$*, for* $a_2 > 0$*, then* $\sigma_1 = (0, t_5) \cup (t_6, \infty)$*, for some* $t_5, t_6 \in [0, \infty) \cup \{\infty\}$*.*

— *If* w_2 *such that* $\frac{\partial}{\partial u} R(w, \sigma' = \{\sigma'_1, \sigma'_2\}) \geq 0$ *for all* σ'*, where* u *is an upper endpoint of an interval that makes up* σ'_2*, then* $\sigma_2 = (0, \infty)$*.*

Proof. Let $\Delta(\sigma_i, \sigma_{-i}) = R_i(\sigma_i) - R_j(\sigma_{-i})$, where $\sigma_{-i} \triangleq \sigma_{3-i}$. Let $r(u, i, w, \sigma)$ be a function that has the same sign as $\frac{\partial}{\partial u} R(w, \sigma)$, where u is an upper endpoint of an interval that is part of σ_i. In B.1, we show

- (Remark 1). $\Delta(\sigma_i, \sigma_{-i}) > 0$ and $\frac{w_i(\tau)}{\tau}$ non-decreasing implies $r(u, i, w, \sigma)$ strictly increasing in $u \in \sigma_i$.
- (Remark 1). $\Delta(\sigma_i, \sigma_{-i}) < 0$ and $\frac{w_i(\tau)}{\tau}$ non-increasing implies $r(u, i, w, \sigma)$ strictly decreasing in $u \in \sigma_i$.
- (Lemma 5). $w(\tau) = m\tau + a$ for $m, a > 0$ and $\Delta(\sigma_i, \sigma_{-i}) > 0$ implies $r(u, i, w, \sigma)$ is strictly quasi-convex in $u \in \sigma_i$
- (Lemma 6). $w(\tau) = m\tau - a$ for $m, a > 0$ and $\Delta(\sigma_i, \sigma_{-i}) < 0$ implies $r(u, i, w, \sigma)$ is strictly quasi-concave in $u \in \sigma_i$

We need to show that there exists a σ of the appropriate form such that $R(w, \sigma) \geq R(w, \sigma')$, for all σ'.

Start with arbitrary $\sigma' = \{\sigma'_1, \sigma'_2\}$ where $\sigma'_1, \sigma'_2 \subseteq \mathbb{R}_+$ are open, measurable sets, but not of the correct form in the theorem statement. Invoking the theorems in Section B.3 as appropriate, we construct a sequence of changes to σ' such that the overall reward does not decrease with each change, and the sequence ends with a policy consistent with the theorem statement.

Step A First, we replace σ'_2 with a policy σ_2^A such that $R_2(\sigma_2^A) > R_1(\sigma_1), \forall \sigma_1$. This allows us to cite the appropriate theorems regarding the properties of the derivative of R, that only hold when the surge state provides higher earnings than the non-surge state.

Let σ_2^A be such that $\Delta(\sigma_2^A, \sigma''_1) > 0$, for all σ''_1 open and measurable, and $\sigma_2^A \geq \sigma'_2$. Such σ_2^A exists by our assumptions on F_i, w_i. Then, let $\sigma^A \triangleq \{\sigma_1^A = \sigma'_1, \sigma_2^A\}$. $R(w, \sigma^A) \geq R(w, \sigma')$: time spent earning reward at the rate of $R_2(w_2, \sigma'_2)$ is replaced by time spent earning at rate $R_1(w_1, \sigma'_1)$ or earning at rate $R_2(w_2, \sigma_2^A)$; time spent earning at $R_1(w_1, \sigma'_1)$ may be replaced by time earning at rate $R_2(w_2, \sigma_2^A)$.

Step B Now, we replace σ_1 with a policy that is of the appropriate form, citing one of the theorems we prove later on in the Appendix.

For $i = 1$: Let $\tilde{R}(\sigma_1) \triangleq R(w, \{\sigma_1, \sigma_2^A\})$.

By Theorem 5 (for derivative always positive), Theorem 7 (for $w_1(\tau)/\tau$ non-increasing), and Theorem 9 (for $w_1(\tau) = m\tau - a$ for $m, a > 0$) :

there exists σ_1^B such that $R(w, \{\sigma_1^B, \sigma_2^A\}) \geq R(w, \{\sigma_1^A, \sigma_2^A\})$, and σ_1^B is of the required form according to the table. Let $\sigma^B \triangleq \{\sigma_1^B, \sigma_2^B = \sigma_2^A\}$.

Note that all the assumptions of the associated theorems are met for the appropriate case: σ_2^B such that $\Delta(\sigma_2^B, \sigma_1') > 0$, $\forall \sigma_1'$, and so the scaled derivatives remain decreasing / strictly quasi-concave as necessary.

Step C For $i = 2$: Now, we replace σ_2 with a policy that is of the appropriate form, citing one of the theorems we prove later on in the Appendix.

Let $\tilde{R}(\sigma_2) \triangleq R(w, \{\sigma_1^B, \sigma_2\})$.

By Theorem 5 (for derivative always positive), Theorem 6 (for $w_2(\tau)/\tau$ non-decreasing), and Theorem 8 (for $w_2(\tau) = m\tau + a$ for $m, a > 0$): there exists σ_2^C such that $R(w, \{\sigma_1^B, \sigma_2^C\}) \geq R(w, \{\sigma_1^B, \sigma_2^B\})$, and σ_2^C is of the required form according to the table. Let $\sigma^C \triangleq \{\sigma_1^C = \sigma_1^B, \sigma_2^C\}$.

Note that, in this step, we need to confirm the assumption regarding $r(u, 2, w, \sigma)$ remaining strictly increasing / strictly quasi-convex in u for a fixed σ, for all σ such that $\tilde{R}(\sigma) \geq \tilde{R}(\sigma_2^B)$ or $F(\sigma_2^B \setminus \sigma \cup \sigma \setminus \sigma_2^B) < \delta$, for some $\delta > 0$. The chief concern is that because σ_2 is changing within the appropriate theorem, $\Delta(\sigma_2, \sigma_1^B)$ may not remain greater than 0, and so this condition might not be met. However this is not the case: $\Delta(\sigma_2, \sigma_1^B)$ remains positive – by continuity, σ_2 close to σ_2^B (by measure of set difference) implies $\Delta(\sigma_2, \sigma_1^B)$ positive. Furthermore

$$
\begin{aligned}
\tilde{R}(\sigma_2'') \geq \tilde{R}(\sigma_2^B) &\iff R(w, \{\sigma_1^B, \sigma_2''\}) \geq R(w, \{\sigma_1^B, \sigma_2^B\}) && \text{definition of } \tilde{R} \\
&\iff \pi_1 R_1(w_1, \sigma_1^B) + \pi_2 R_2(w_2, \sigma_2'') \geq \pi_a R_1(w_1, \sigma_1^B) + \pi_b R_2(w_2, \sigma_2^B) && (\pi_k \text{policy dependent}) \\
\Delta(\sigma_2, \sigma_1^B) \leq 0 &\implies \pi_1 R_1(w_1, \sigma_1^B) + \pi_2 R_2(w_2, \sigma_2'') \leq R_1(w_1, \sigma_1^B) \\
\Delta(\sigma_2^B, \sigma_1^B) > 0, \pi_b > 0 &\implies \pi_1 R_1(w_1, \sigma_1^B) + \pi_2 R_2(w_2, \sigma_2'') > R_1(w_1, \sigma_1^B)
\end{aligned}
$$

By **Step A**, $\Delta(\sigma_2^B, \sigma_1^B) > 0, \pi_b > 0$, and so $\Delta(\sigma_2, \sigma_1^B) \leq 0$ would be a contradiction for $R(w, \sigma) \geq \tilde{R}(\sigma_2^B)$.

Thus, we have constructed $\sigma^* = \{\sigma_1^* = \sigma_1^C, \sigma_2^* = \sigma_2^C\}$ such that σ_1^*, σ_2^* correspond to theorem statement for the appropriate cases, respectively, and $R(w, \sigma^*) \geq R(w, \sigma)$, for all $\sigma = \{\sigma_1, \sigma_2\}$ where $\sigma_1, \sigma_2 \subseteq \mathbb{R}_+$ are open, measurable sets.

□

Theorem 2 immediately follows.

A.3.2. Proof of Theorem 3

THEOREM 3. *Let R_1 and R_2 be target earning rates during non-surged and surge states, respectively. There exist prices $w = \{w_1, w_2\}$ of the form*

$$w_i(\tau) = m_i \tau + z_i q_{i \to j}(\tau),$$

where $m_1, m_2, z_2 \geq 0$ (but z_1 may be either positive or negative), such that the optimal driver policy is to accept every trip in the surge state and all trips up to a certain length in the non-surge state. Furthermore, for $\frac{R_1}{R_2} \in [C, 1]$, there exist fully incentive compatible prices of this form, where

$$C = 1 - \frac{1}{T_1} \frac{Q_2(\lambda_{12} T_1 - Q_1) + Q_1(T_2 \lambda_{1\to 2} + Q_2)}{Q_2(\lambda_{1\to 2} T_1 - Q_1) + \lambda_{1\to 2}(T_2 \lambda_{1\to 2} + Q_2)} \in [0, 1),$$

and $T_i = \lambda_i F_i(\sigma_i) T_i((0, \infty))$, and $Q_i = Q_i((0, \infty))$.

Proof. Note that in the theorem statement we defined Q_i, T_i as what we call $\bar{Q}_i, \bar{T}_i$ in the helper lemmas in Appendix Section B.2, i.e. they refer to their respective values when every trip is accepted.

Let $w_2(\tau) = m_2\tau + z_2 q_{2\to 1}(\tau)$, and $w_1(\tau) = m_1\tau + z_1 q_{1\to 2}(\tau)$.

From Lemmas 7 and 8 in Appendix Section B.2, the following constraints are sufficient for these prices to have always positive derivatives, with respect to upper endpoints u of the intervals that compose either σ_1 or σ_2:

$$\frac{T_1(\lambda_{2\to 1}T_2 - Q_2) - (Q_1 + T_1\lambda_{2\to 1})}{(Q_1(\lambda_{2\to 1}T_2 - Q_2) + \lambda_{2\to 1}(Q_1 + T_1\lambda_{2\to 1}))} \le \frac{z_2}{m_2 - R_1} \le \frac{Q_2 T_1 + Q_1}{Q_1(Q_2 - \lambda_{2\to 1})}$$

$$m_1 = R_2$$

$$-\frac{(T_2\lambda_{1\to 2} + Q_2)}{Q_2(\lambda_{1\to 2}T_1 - Q_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2)} \le \frac{z_1}{R_2} \le \frac{1}{(Q_1 - \lambda_{1\to 2})}$$

Now, applying Theorem 4, the policy that accepts everything, $\sigma = \{(0,\infty),(0,\infty)\}$, is optimal, given these constraints are satisfied, as the derivative is always positive.

Resulting constraints on R_1, R_2. These constraints limit R_1, R_2 with respect to each other. From Remark 2,

$$W_2 = m_2(T_2 - 1) + z_2(Q_2 - \lambda_{2\to 1})$$

$$W_1 = m_1(T_1 - 1) + z_1(Q_1 - \lambda_{1\to 2})$$

Given R_2, what's the range R_1 can be to still have IC in state 1?

$$\begin{aligned}
& W_1 \le R_2\left[T_1 - 1 + \left[\frac{1}{(Q_1 - \lambda_{1\to 2})}(Q_1 - \lambda_{1\to 2})\right]\right] \\
\implies & \frac{R_1}{R_2} \le 1 \\
& W_1 \ge R_2\left[T_1 - 1 - \left[\frac{(T_2\lambda_{1\to 2} + Q_2)}{Q_2(\lambda_{1\to 2}T_1 - Q_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2)}\right](Q_1 - \lambda_{1\to 2})\right] \\
\implies & \frac{R_1}{R_2} = \frac{W_1}{T_1}\frac{1}{R_2} \\
& \ge \frac{1}{T_1}\left[T_1 - 1 - \left[\frac{(T_2\lambda_{1\to 2} + Q_2)}{Q_2(\lambda_{1\to 2}T_1 - Q_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2)}\right](Q_1 - \lambda_{1\to 2})\right] \\
& = 1 - \frac{1}{T_1}\left[1 + \frac{(T_2\lambda_{1\to 2} + Q_2)(Q_1 - \lambda_{1\to 2})}{Q_2(\lambda_{1\to 2}T_1 - Q_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2)}\right] \\
& = 1 - \frac{1}{T_1}\frac{Q_2(\lambda_{1\to 2}T_1 - Q_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2) + (T_2\lambda_{1\to 2} + Q_2)(Q_1 - \lambda_{1\to 2})}{Q_2(\lambda_{1\to 2}T_1 - Q_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2)} \\
& = 1 - \frac{1}{T_1}\frac{Q_2(\lambda_{1\to 2}T_1 - Q_1) + Q_1(T_2\lambda_{1\to 2} + Q_2)}{Q_2(\lambda_{1\to 2}T_1 - Q_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2)} = C
\end{aligned}$$

What about IC in state 2? If only care about that state, can support any ratio of payments:

$$\begin{aligned}
\text{Let } \; z_2 &= \frac{Q_2T_1 + Q_1}{Q_1(Q_2 - \lambda_{2\to 1})}(m_2 - R_1) \triangleq c(m_2 - R_1) \\
R_2 &= \frac{1}{T_2}\left[m_2(T_2 - 1) + z_2(Q_2 - \lambda_{2\to 1})\right] \\
\implies \frac{R_2}{R_1} &= \frac{1}{R_1T_2}\left[m_2(T_2 - 1) + (m_2 - R_1)c(Q_2 - \lambda_{2\to 1})\right] \\
& \to 1 - \frac{1}{T_2} \le 1 \text{ as } m_2 \to R_1 \\
& \to \infty \text{ as } m_2 \to \infty
\end{aligned}$$

Thus, we can make the surge state IC for any ratio of payments $\frac{R_2}{R_1} \geq 1$, i.e. $\frac{R_1}{R_2} \leq 1$.

Now, suppose we want to achieve R_1, R_2 such that $\frac{R_1}{R_2} \in [0, C]$. From the previous line, we can still set w_2 such that every trip in state 2 is accepted (the derivative with respect to the surge policy is positive everywhere). Then, setting $z_1 = 0$, and m_1 to meet R_1, all trips up to a certain length will be accepted in the non-surge state: By Remark 1, $\frac{\partial}{\partial u} R(w, \sigma)$ is positive up to a certain value and then negative after that, where u is an upper endpoint of σ_1.Thus, by Theorem 4, the optimal policy is of the form
$\sigma = \{(0, t_1), (0, \infty)\}$. $\square$

A.4. Main text lemmas proofs

A.4.1. Single-state model Recall that $R(w, \sigma, t)$ is the total earnings from jobs accepted up from time 0 to time t, i.e. $R(w, \sigma, t) = \mathbb{E}\left[\sum_{k=1}^{N(t)} w(\tau_i)\right]$, where τ_i is the length of the ith job the driver accepts, e_i is the wait time to that job, and $N(t) = |\{i : 0 \leq \tau_i + e_i \leq t\}|$ is the number of accepted jobs up to time t.

As mentioned using the renewal reward theorem in the main text,

$$R(w, \sigma) \triangleq \liminf_{t \to \infty} \frac{R(w, \sigma, t)}{t} = \frac{\text{Expected cycle payment given } \sigma}{\text{Expected cycle length given } \sigma} = \frac{\frac{1}{F(\sigma)} \int_{\tau \in \sigma} w(\tau) dF(\tau)}{\frac{1}{F(\sigma)\lambda} + \frac{1}{F(\sigma)} \int_{\tau \in \sigma} \tau dF(\tau)}$$

The $\frac{1}{\lambda F(\sigma)}$ term is the expected value of a exponential random variable with rate $\lambda F(\sigma)$, which is the rate at which a driver receives ride requests that she accepts.

A.4.2. Dynamic model

LEMMA 1. *The overall earning rate can be decomposed into the earnings rate $R_i(w_i, \sigma_i)$ and fraction of time $\mu_i(\sigma)$ spent in state i. The following equality holds with probability 1:*

$$R(w, \sigma) \triangleq \liminf_{t \to \infty} \frac{R(w, \sigma, t)}{t} = \mu_1(\sigma) R_1(w_1, \sigma_1) + \mu_2(\sigma) R_2(w_2, \sigma_2).$$

As in the single-state model, $R_i(w_i, \sigma_i) = \frac{W_i(\sigma_i)}{T_i(\sigma_i)}$, where

$$W_i(\sigma_i) = \frac{1}{F_i(\sigma_i)} \int_{\tau \in \sigma_i} w_i(\tau) dF_i(\tau), \qquad T_i(\sigma_i) = \frac{1}{\lambda_i F_i(\sigma_i)} + \frac{1}{F_i(\sigma_i)} \int_{\tau \in \sigma_i} \tau dF_i(\tau)$$

Proof. Consider the renewal process (with cycles and sub-cycles) defined in the main text.

Let $M(t)$ be the total number of cycles that have been completed up to time t. Let $N_j(M)$ be the number of sub-cycles in state j in the Mth cycle – i.e., in the Mth cycle of the single renewal process described above, the number of times that the driver is open in state j (after transitioning from the other state, or finishing a trip that started in the same state j). Let $S_j(k, M)$ be the length of the kth such sub-cycle in the Mth cycle, with expected length $S_j(\sigma_j)$. Let $p_{ji}(\sigma_j)$ be the probability that the current sub-cycle is the last in state j for the current cycle – as the next cycle starts in the other state.

Finally, let $R_j(w_j, \sigma_j, M)$ be the total amount earned in state j after M such cycles. Then:

$$\begin{aligned}
R_j(w_j, \sigma_j, M(t)) &= \sum_{M=1}^{M(t)} \sum_{k=1}^{N_j(M)} W_j(k, M) \\
\lim_{t \to \infty} \frac{R_j(w_j, \sigma_j, M(t))}{M(t)} &= \lim_{t \to \infty} \frac{1}{M(t)} \left[\sum_{M=1}^{M(t)} \frac{W_j(k, M)}{p_{ji}}\right] \\
&= \frac{1}{p_{ji}(\sigma_j)} W_j(\sigma_j) \qquad \text{almost surely}
\end{aligned}$$

by the mean of a geometric random variable and the basic law of large numbers for renewal processes. Similarly, we know that $\frac{M(t)}{t}$ converges to its mean almost surely as $t\to\infty$, where the mean is based on the lengths of in each state in each cycle. Let $\tilde{S}(\sigma)$ be the expected length of one of these cycles. Then:

$$\begin{aligned}
\lim_{t\to\infty}\frac{M(t)}{t} &= \frac{1}{\tilde{S}(\sigma)} \\
\tilde{S}(\sigma) &= \mathbb{E}\left[\sum_{k=1}^{N_1(1)} S_1(k,1)\right] + \mathbb{E}\left[\sum_{k=1}^{N_2(1)} S_2(k,1)\right] \\
&= \mathbb{E}[N_1(1)]\mathbb{E}[S_1(k,1)] + \mathbb{E}[N_2(1)]\mathbb{E}[S_2(k,1)] && \text{Wald's identity} \\
&= \frac{1}{p_{12}(\sigma_1)}S_1(\sigma_1) + \frac{1}{p_{21}(\sigma_2)}S_2(\sigma_2) \\
\implies \lim_{t\to\infty}\frac{M(t)}{t} &= \frac{p_{21}(\sigma_2)p_{12}(\sigma_1)}{p_{21}(\sigma_2)S_1(\sigma_1)+p_{12}(\sigma_1)S_2(\sigma_2)}
\end{aligned}$$

Then, by standard algebra on multiplication with almost sure convergence

$$\begin{aligned}
\lim_{t\to\infty}\frac{R_j(w_j,\sigma_j,M(t))}{t} &= \lim_{t\to\infty}\frac{R_j(w_j,\sigma_j,M(t))}{M(t)}\frac{M(t)}{t} \\
&= \frac{1}{p_{ji}(\sigma_j)}W_j(w_j,\sigma_j)\left[\frac{p_{21}(\sigma_2)p_{12}(\sigma_1)}{p_{21}(\sigma_2)S_1(\sigma_1)+p_{12}(\sigma_1)S_2(\sigma_2)}\right] \\
&= \left[\frac{p_{ij}(\sigma_i)S_j(\sigma_j)}{p_{21}(\sigma_2)S_1(\sigma_1)+p_{12}(\sigma_1)S_2(\sigma_2)}\right]R_j(w_j,\sigma_j)
\end{aligned}$$

Let $\mu_j(\sigma)\triangleq\frac{p_{ij}(\sigma_i)S_j(\sigma_j)}{p_{21}(\sigma_2)S_1(\sigma_1)+p_{12}(\sigma_1)S_2(\sigma_2)}$. Putting everything together:

$$\begin{aligned}
\liminf_{t\to\infty}\frac{R(w,\sigma,t)}{t} &= \liminf_{t\to\infty}\frac{R_1(w_1,\sigma_1,M(t))}{t} + \liminf_{t\to\infty}\frac{R_2(w_2,\sigma_2,M(t))}{t} \\
&= \mu_1(\sigma)R_1(w_1,\sigma_1) + \mu_2(\sigma)R_2(w_2,\sigma_2)
\end{aligned}$$

□

LEMMA 2. *Suppose the world is in state i at time t. Let $q_{i\to j}(s)$ denote the probability that the world will be in state $j\neq i$ at time $t+s$. Then,*

$$q_{i\to j}(s) = \frac{\lambda_{i\to j}}{\lambda_{i\to j}+\lambda_{j\to i}}\left[1-e^{-(\lambda_{i\to j}+\lambda_{j\to i})s}\right]$$

Proof. Given the state dynamics in the model, $q_{i\to j}(s)$ is determined by the evolution of a CTMC in time s, given that the current state is i. We can use standard CTMC results here. Let Q denote the Q-matrix for the world state CTMC. From the model definition,

$$Q = \begin{bmatrix} -\lambda_{1\to2} & \lambda_{1\to2} \\ \lambda_{2\to1} & -\lambda_{2\to1} \end{bmatrix}$$

Recall that the state transition matrix after time t is then given by the matrix exponential e^{Qt}, which is equal to the inverse of the Laplace transform of the inverse of the resolvent of Q:

$$\begin{aligned}
q_{i\to j}(\tau) &= (e^{Q\tau})_{ij} \\
&= \mathcal{L}^{-1}((wI-Q)^{-1}_{ij})(\tau) && w \text{ is a Laplace transform parameter} \\
&= \frac{\lambda_{i\to j}}{\lambda_{i\to j}+\lambda_{j\to i}}\left[1-e^{-(\lambda_{i\to j}+\lambda_{j\to i})\tau}\right]
\end{aligned}$$

where the closed form in the last line emerges only in a 2 state model. □

LEMMA 3. *Let $T_i(\sigma_i)$ be as defined in Lemma 1. The fraction of time a driver following strategy $\sigma = \{\sigma_1, \sigma_2\}$ spends either open in state i or on a trip started in state i is*

$$\mu_i(\sigma) = \frac{\lambda_i F_i(\sigma_i) T_i(\sigma_i) Q_j(\sigma_j)}{\lambda_j F_j(\sigma_j) T_j(\sigma_j) Q_i(\sigma_i) + \lambda_i F_i(\sigma_i) T_i(\sigma_i) Q_j(\sigma_j)}$$

$$\textit{where} \quad Q_i(\sigma_i) = \lambda_{i\to j} + \lambda_i \int_{\tau\in\sigma_i} q_{i\to j}(\tau) dF_i(\tau)$$

Proof. From the proof of Lemma 1, we have

$$\mu_i(\sigma) = \frac{p_{ji}(\sigma_j) S_i(\sigma_i)}{p_{21}(\sigma_2) S_1(\sigma_1) + p_{12}(\sigma_1) S_2(\sigma_2)}$$

where $S_i(\sigma_i)$ is the expected length of the time between being open in a state i to being open again, either after a state transition or after finishing a job; and $p_{ij}(\sigma_i)$ is the probability that the driver is next open in state j given she is currently open in state i. These are:

$$\begin{aligned} S_i(\sigma_i) &= \frac{1}{\lambda_i F_i(\sigma_i) + \lambda_{i\to j}} + \frac{\lambda_i F_i(\sigma_i)}{\lambda_i F_i(\sigma_i) + \lambda_{i\to j}} \int_{\tau\in\sigma_i} \tau \frac{f_i(\tau)}{F_i(\sigma_i)} d\tau \\ &= \frac{1}{\lambda_i F_i(\sigma_i) + \lambda_{i\to j}} \left[1 + \lambda_i \int_{\tau\in\sigma_i} \tau dF_i(\tau)\right] \\ &= \left[\frac{\lambda_i F_i(\sigma_i)}{\lambda_i F_i(\sigma_i) + \lambda_{i\to j}}\right] T_i(\sigma_i) \end{aligned}$$

The first part of the sum $\frac{1}{\lambda_i F_i(\sigma_i) + \lambda_{i\to j}}$ is the expected time until either the driver receives a request that she accepts, or the world state transitions to the other state. This form emerges because there are two competing independent exponential clocks – that for a request and that for the world state changing. The second part of the sum is the probability of receiving an accepted trip request before a state transition, times the expected length of an accepted trip.

The next step is to find an expression for $p_{ij}(\sigma_i)$, the probability that the next renewal cycle is at state j, given the current one is at state i. We find it for $j \neq i$, and then $p_{ii} = 1 - p_{ij}$.

$$\begin{aligned} p_{ij}(\sigma) &= \frac{\lambda_{i\to j}}{\lambda_i F_i(\sigma_i) + \lambda_{i\to j}} + \frac{\lambda_i F_i(\sigma_i)}{\lambda_i F_i(\sigma_i) + \lambda_{i\to j}} \frac{1}{F_i(\sigma_i)} \int_{\sigma_i} q_{i\to j}(\tau) dF_i(\tau) \\ &= \left[\frac{1}{\lambda_i F_i(\sigma_i) + \lambda_{i\to j}}\right] Q_i(\sigma_i) \end{aligned}$$

The first part of the summation is the probability that the world state transitions to state j before the driver accepts a trip request. The second part is the probability that the driver accepts a trip request before the state transitions, times the probability $q_{i\to j}(\sigma_i) = \frac{1}{F_i(\sigma_i)} \int_{\sigma_i} q_{i\to j}(\tau) dF_i(\tau)$ that the world will be in state j when the driver's trip ends. The result follows. $\square$

A.5. Uniqueness of optimal policy for single-state model

LEMMA 4. *Consider the single-state model, and optimal policy σ^* of the form $\sigma^* = \{\tau : \frac{w(\tau)}{\tau} \geq c^*\}$. Then, $R(\sigma^*) = c^*$. Further, other policies of the form $\sigma_c = \{\tau : \frac{w(\tau)}{\tau} \geq c\}$ are not optimal unless $\sigma_c = \sigma^*$ (up to sets of measure 0).*

Proof. By Theorem 1, there exists an optimal policy of the form $\sigma^* = \{\tau : \frac{w(\tau)}{\tau} \geq c^*\}$, for some c^*. Here, we show (1) that $R(\sigma^*) = c^*$, and (2) this is the unique optimal policy of the form $\sigma^* = \{\tau : \frac{w(\tau)}{\tau} \geq c\}$.

1. $R(\sigma^*) = c^*$. The proof is identical to lines (1), (2).

   Suppose $R(\sigma^*) > c^*$. Then, consider $c = R(\sigma^*)$, $\sigma_c = \{\tau : \frac{w(\tau)}{\tau} \geq c\}$. If $F(\sigma^* \setminus \sigma_c) > 0$:

$$R(\sigma_c) = \frac{\lambda \int_{\tau \in \sigma^*} w(\tau) dF(\tau) - \lambda \int_{\tau \in \sigma^* \setminus \sigma_c} w(\tau) dF(\tau)}{1 + \lambda \int_{\tau \in \sigma^*} \tau dF(\tau) - \lambda \int_{\tau \in \sigma^* \setminus \sigma_c} \tau dF(\tau)}$$
$$> R(\sigma^*)$$

   Which follows from $\frac{\lambda \int_{\tau \in \sigma^* \setminus \sigma_c} w(\tau) dF(\tau)}{\lambda \int_{\tau \in \sigma^* \setminus \sigma_c} \tau dF(\tau)} < c = R(\sigma^*) = \frac{\lambda \int_{\tau \in \sigma^*} w(\tau) dF(\tau)}{1 + \lambda \int_{\tau \in \sigma^*} \tau dF(\tau)}$, and $\frac{x}{z} < \frac{w}{y} \implies \frac{w-x}{y-z} > \frac{w}{y}$. This contradicts that σ^* is optimal.

   Similarly, suppose $R(\sigma^*) < c^*$. Then, consider $c = R(\sigma^*)$, $\sigma_c = \{\tau : \frac{w(\tau)}{\tau} \geq c\}$. If $F(\sigma_c \setminus \sigma^*) > 0$:

$$R(\sigma_c) = \frac{\lambda \int_{\tau \in \sigma^*} w(\tau) dF(\tau) + \lambda \int_{\tau \in \sigma_c \setminus \sigma^*} w(\tau) dF(\tau)}{1 + \lambda \int_{\tau \in \sigma^*} \tau dF(\tau) + \lambda \int_{\tau \in \sigma_c \setminus \sigma^*} \tau dF(\tau)}$$
$$> R(\sigma^*)$$

   Which follows from $\frac{\lambda \int_{\tau \in \sigma_c \setminus \sigma^*} w(\tau) dF(\tau)}{\lambda \int_{\tau \in \sigma_c \setminus \sigma^*} \tau dF(\tau)} > c = R(\sigma^*) = \frac{\lambda \int_{\tau \in \sigma^*} w(\tau) dF(\tau)}{1 + \lambda \int_{\tau \in \sigma^*} \tau dF(\tau)}$, and $\frac{x}{z} > \frac{w}{y} \implies \frac{w+x}{y+z} > \frac{w}{y}$. This contradicts that σ^* is optimal.

2. Suppose there were two optimal $\sigma_{c_1}, \sigma_{c_2}$ of the appropriate form. Without loss of generality, let $c_1 < c_2$. Suppose $F(\sigma_{c_1} \setminus \sigma_{c_2}) > 0$, and so $\sigma_{c_1} \subsetneq \sigma_{c_2}$. By the previous part, $R(\sigma_{c_1}) = c_1, R(\sigma_{c_2}) = c_2$. Then

$$R(\sigma_{c_2}) = \frac{\lambda \int_{\tau \in \sigma_{c_1}} w(\tau) dF(\tau) - \lambda \int_{\tau \in \sigma_{c_1} \setminus \sigma_{c_2}} w(\tau) dF(\tau)}{1 + \lambda \int_{\tau \in \sigma_{c_1}} \tau dF(\tau) - \lambda \int_{\tau \in \sigma_{c_1} \setminus \sigma_{c_2}} \tau dF(\tau)}$$
$$< R(\sigma_{c_1})$$

   Which follows from $\frac{\lambda \int_{\tau \in \sigma_{c_1} \setminus \sigma_{c_2}} w(\tau) dF(\tau)}{\lambda \int_{\tau \in \sigma_{c_1} \setminus \sigma_{c_2}} \tau dF(\tau)} > c_1 = R(c_1) = \frac{\lambda \int_{\tau \in \sigma_{c_1}} w(\tau) dF(\tau)}{1 + \lambda \int_{\tau \in \sigma_{c_1}} \tau dF(\tau)}$, and $\frac{x}{z} > \frac{w}{y} \implies \frac{w}{y} > \frac{w-x}{y-z}$, contradicting the supposition that $R(\sigma_{c_1}) = c_1 < c_2 = R(\sigma_{c_2})$.

   □

Appendix B: Helper theorems and lemmas

B.1. Derivatives of reward and properties in dynamic model

Here we take derivatives of the reward function $R(w, \sigma)$ and its components in the dynamic model. These will be useful in proving both the incentive compatible pricing structure and the optimal policy structure with multiplicative or affine pricing.

Recall in the dynamic model that we constrain σ_i to be measurable, *open*, subsets of the $\mathbb{R}_+$. Then, σ_i can be written as a countable union of disjoint subsets of $\mathbb{R}_+$, i.e. $\sigma_i = \cup_{k=0}^{\infty} (\ell_k, u_k)$. We further assume that $u_k \neq \ell_m$, for any k, m; we can do so without loss of generality by making a measure 0 change to σ_i, by adding $u_k = \ell_m$ to σ_i.

Suppose u is an upper-endpoint of σ_i, ie. $\exists k$ such that $u = u_k$. Then, we use $\frac{\partial}{\partial u} H(\sigma_i)$ to denote the derivative of the set function H with respect to u at σ_i. Similarly, $\frac{\partial}{\partial \ell} H(\sigma_i)$ is the derivative of H at σ_i with respect to a lower-endpoint of σ_i.

Note that we also derive $\frac{\partial}{\partial u}R(w,\{\sigma_1,\sigma_2\})$, $\frac{\partial}{\partial \ell}R(w,\{\sigma_1,\sigma_2\})$. We will make it clear in each instance whether u or ℓ is an endpoint of σ_1 or σ_2. For all the derivatives in this subsection $\frac{\partial}{\partial u}$ refers to the derivative with respect to an upper endpoint in σ_i.

Throughout, we use the $\propto$ symbol to denote *has the same sign as*. For notational convenience, for a fixed σ_i, let

$$\begin{aligned}
Q_i &\triangleq Q_i(\sigma_i) = \lambda_{i\to j} + \lambda_i \int_{\sigma_i} q_{i\to j}(\tau)dF_i(\tau)\\
T_i &\triangleq \lambda_i F_i(\sigma_i)T_i(\sigma_i) = 1 + \lambda_i \int_{\tau\in\sigma_i} \tau dF_i(\tau)\\
W_i &\triangleq \lambda_i F_i(\sigma_i)W_i(\sigma_i) = \lambda_i \int_{\tau\in\sigma_i} w_i(\tau)dF_i(\tau)
\end{aligned}$$

Then,

$$\begin{aligned}
\mu_i(\{\sigma_j,\sigma_2\}) &= \frac{Q_jT_i}{Q_jT_i+Q_iT_j}\\
R(w,\sigma) &= \mu_1(\sigma)R_1(w_1,\sigma_1) + \mu_2(\sigma)R_2(w_2,\sigma_2)\\
&= \left[\frac{1}{Q_2T_1+Q_1T_2}\right][Q_2W_1+Q_1W_2]\\
R_i(\sigma_i) &= \frac{W_i}{T_i}\\
\frac{\partial}{\partial u}Q_i &= \frac{\partial}{\partial u}\left[\lambda_{i\to j} + \lambda_i\int_{\tau\in\sigma_i} q_{i\to j}(\tau)dF_i(\tau)\right] = \lambda_i q_{i\to j}(u)f_i(u)\\
\frac{\partial}{\partial u}W_i &= \frac{\partial}{\partial u}\left[\lambda_i\int_{\tau\in\sigma_i} w_i(\tau)dF_i(\tau)\right] = \lambda_i w_i(u)f_i(u)\\
\frac{\partial}{\partial u}T_i &= \lambda_i f_i(u)u
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial u}R(w,\sigma) &= \left[\frac{\lambda_i f_i(u)}{Q_iT_j+Q_jT_i}\right][[q_{i\to j}(u)W_j + Q_jw_i(u)] - R(w,\sigma)(uQ_j + q_{i\to j}(u)T_j)]\\
&\propto [q_{i\to j}(u)W_j + Q_jw_i(u)] - R(w,\sigma)(uQ_j + q_{i\to j}(u)T_j)\\
&\propto [q_{i\to j}(u)W_j + Q_jw_i(u)]\,(Q_iT_j\\
&\quad + Q_jT_i) - (Q_iW_j + Q_jW_i)(uQ_j + q_{i\to j}(u)T_j)\\
&= q_{i\to j}(u)W_j(Q_iT_j+Q_jT_i) + Q_jw_i(u)(Q_iT_j+Q_jT_i)\\
&\quad - uQ_j(Q_iW_j+Q_jW_i) - q_{i\to j}(u)T_j(Q_iW_j+Q_jW_i)\\
&\propto q_{i\to j}(u)W_jT_i + w_i(u)(Q_iT_j+Q_jT_i) - u(Q_iW_j+Q_jW_i) - q_{i\to j}(u)T_jW_i\\
&= q_{i\to j}(u)\,[W_jT_i - T_jW_i] + w_i(u)(Q_iT_j+Q_jT_i) - u(Q_iW_j+Q_jW_i)\\
&= uT_iT_j\left[\frac{q_{i\to j}(u)}{u}(R_j-R_i) + \frac{w_i(u)}{u}\left(\frac{Q_i}{T_i}+\frac{Q_j}{T_j}\right) - \left(\frac{Q_i}{T_i}R_j + \frac{Q_j}{T_j}R_i\right)\right]\\
&\propto \frac{q_{i\to j}(u)}{u}\Delta_{ji} + \frac{w_i(u)}{u}\left(\frac{Q_i}{T_i}+\frac{Q_j}{T_j}\right) - \left(\frac{Q_i}{T_i}R_j + \frac{Q_j}{T_j}R_i\right) \qquad \Delta_{ji} = R_j - R_i\\
&\triangleq r(u,i,w,\sigma)
\end{aligned}$$

In other words, $r(u,i,w,\sigma)$ has the *same sign* as the derivative of the overall reward with respect to u (an upper endpoint of σ_i) at w, σ, but it is not necessarily monotonic with it.

REMARK 1. Given assumptions on F_i, w_i:

- $R_i(\sigma), R(\sigma), \mu_i$ are continuous in σ
- $\frac{\partial}{\partial u} R(w,\sigma), r(u,i,w,\sigma)$ are both continuous u (for fixed σ), and continuous in σ.
- $\frac{q_{i\to j}(u)}{u}$ is strictly decreasing in u.
- If $\Delta_{ji} < 0$ (i.e. $i = 2$ the surge state) and $\frac{w_i(u)}{u}$ is non-decreasing in u, then $r(u,i,w,\sigma)$ is strictly increasing in u for a fixed σ. Thus, $\frac{\partial}{\partial u} R(w,\sigma)$ is negative up to a certain point $U \in (0,\infty) \cup \{\infty\}$ and then positive thereafter.
- If $\Delta_{ji} > 0$ (i.e. $i = 1$ the non-surge state) and $\frac{w_i(u)}{u}$ is non-increasing in u, then $r(u,i,w,\sigma)$ is strictly decreasing in u for a fixed σ. Thus, $\frac{\partial}{\partial u} R(w,\sigma)$ is positive up to a certain point $U \in (0,\infty) \cup \{\infty\}$ and then negative thereafter.
- $\frac{\partial}{\partial \ell} R(w,\sigma)$ at a lower endpoint of σ_i is just the negative of the derivative at the same place if a lower endpoint.

LEMMA 5. *Suppose $w_i(\tau) = m\tau + a$, where $m, a > 0$, and $\Delta_{ji} < 0$. Then, $r(u,i,w,\sigma)$ is strictly quasi-convex in u, for a fixed σ.*

Proof.

$$r(u,i,w,\sigma) = \frac{c_1 a - c_2 q_{i\to j}(u)}{u} + c_3 \qquad\qquad c_1, c_2 \geq 0; c_3 \text{ can be negative}$$

Then, $\frac{\partial}{\partial u} r(u,i,w,\sigma)$

$$\begin{aligned}
&= \frac{\partial}{\partial u}\left[\frac{c_1 - c_2 q_{i\to j}(u)}{u} + c_3\right] \\
&= \frac{1}{u^2}\left[-uc_2 \frac{\partial}{\partial u} q_{i\to j}(u) - [c_1 - c_2 q_{i\to j}(u)]\right] \\
&= \frac{1}{u^2}\left[-uc_2 \frac{\partial}{\partial u}\left[\frac{\alpha}{\alpha+\beta}\left[1 - e^{-(\alpha+\beta)u}\right]\right] - \left[c_1 - c_2\left[\frac{\alpha}{\alpha+\beta}\left[1 - e^{-(\alpha+\beta)u}\right]\right]\right]\right] \\
&= \frac{1}{u^2}\left[-uc_2\left[\alpha e^{-(\alpha+\beta)u}\right] + c_2\left[\frac{\alpha}{\alpha+\beta}\left[1 - e^{-(\alpha+\beta)u}\right]\right] - c_1\right] \\
&= \frac{1}{u^2}\left[-uc_2\left[\alpha\left[\sum_{n=0}^{\infty} \frac{u^n(-1)^n(\alpha+\beta)^n}{n!}\right]\right] + c_2\left[\frac{\alpha}{\alpha+\beta}\left[1 - \left[\sum_{n=0}^{\infty} \frac{u^n(-1)^n(\alpha+\beta)^n}{n!}\right]\right]\right] - c_1\right] \\
&= \frac{1}{u^2}\left[\frac{c_2\alpha}{\alpha+\beta}\left[\sum_{n=0}^{\infty} \frac{(-1)^{n+1}u^{n+1}(\alpha+\beta)^{n+1}}{n!} + 1 + \sum_{n=0}^{\infty} \frac{u^n(-1)^{n+1}(\alpha+\beta)^n}{n!}\right] - c_1\right] \\
&= \frac{1}{u^2}\left[\frac{c_2\alpha}{\alpha+\beta}\left[\sum_{n'=1}^{\infty} \frac{(-1)^{n'}u^{n'}(\alpha+\beta)^{n'}}{(n'-1)!} + \sum_{n=1}^{\infty} \frac{u^n(-1)^{n+1}(\alpha+\beta)^n}{n!}\right] - c_1\right] \qquad n' = n+1 \\
&= \frac{1}{u^2}\left[\frac{c_2\alpha}{\alpha+\beta}\left[\sum_{n=2}^{\infty}(-1)^n u^n(\alpha+\beta)^n\left[\frac{1}{(n-1)!} - \frac{1}{n!}\right]\right] - c_1\right]
\end{aligned}$$

Where last line follows because first $(n = 1)$ term of summation is zero.

Thus, $r(u,i,w,\sigma)$ is strictly quasi-convex if $\frac{c_2\alpha}{\alpha+\beta}\left[\sum_{n=2}^{\infty}(-1)^n u^n(\alpha+\beta)^n\left[\frac{1}{(n-1)!} - \frac{1}{n!}\right]\right] - c_1$ is strictly increasing (derivative is strictly negative up to a point, and then strictly positive above that point u, for a fixed σ.)

$$\frac{\partial}{\partial u}\left[\frac{c_2\alpha}{\alpha+\beta}\left[\sum_{n=2}^{\infty}(-1)^n u^n(\alpha+\beta)^n\left[\frac{1}{(n-1)!} - \frac{1}{n!}\right]\right] - c_1\right]$$

$$
\begin{aligned}
&= \frac{c_2\alpha}{\alpha+\beta}\left[\sum_{n=2}^{\infty}(-1)^n u^{n-1}(\alpha+\beta)^n\left[\frac{n}{(n-1)!}-\frac{n}{n!}\right]\right] \\
&= \frac{c_2\alpha}{\alpha+\beta}\left[\sum_{n=2}^{\infty}(-1)^n u^{n-1}(\alpha+\beta)^n\frac{1}{(n-2)!}\right] = \frac{c_2\alpha}{\alpha+\beta}\left[\sum_{n'=0}^{\infty}(-1)^{n'+2}u^{n'+1}(\alpha+\beta)^{n'+2}\frac{1}{n'!}\right] \quad n'=n-2 \\
&= c_2\alpha u(\alpha+\beta)\left[\sum_{n=0}^{\infty}(-1)^n u^n(\alpha+\beta)^n\frac{1}{n!}\right] = c_2\alpha u(\alpha+\beta)e^{-(\alpha+\beta)u} > 0
\end{aligned}
$$

□

LEMMA 6. *Suppose $w_i(\tau) = m\tau + a$, where $m > 0, a < 0$ and $\Delta_{ji} > 0$. Then, $r(u,i,w,\sigma)$ is strictly quasi-concave.*

Proof. Corollary of Lemma 5. $r(u,i,w,\sigma)$ is the negative of the previous case, modulo constants that do not affect quasi-concavity. □

B.2. Lemmas for Incentive Compatible policy

REMARK 2.

$$
\begin{aligned}
\text{Let} \quad & w_i(u) = mu + zq_{i\to j}(u) \\
\text{Then} \quad & W_i = m(T_i - 1) + z(Q_i - \lambda_{i\to j}) \\
& \frac{\partial}{\partial u}R(w,\sigma) \propto q_{i\to j}(u)\left[(R_j - m)T_jT_i + mT_j + zQ_jT_i + zT_j\lambda_{i\to j}\right] \\
& \qquad\qquad + u\left[Q_iT_j(m - R_j) + Q_j(m - zQ_i + z\lambda_{i\to j})\right]
\end{aligned}
$$

Proof.

$$
\begin{aligned}
w_i(u) &= mu + zq_{i\to j}(u) \qquad\qquad m, z \geq 0 \\
W_i &= \lambda_i\int_{\tau\in\sigma_i} w_i(\tau)dF_i(\tau) = \lambda_i\int_{\tau\in\sigma_i}\left[m\tau + zq_{i\to j}(\tau)\right]dF_i(\tau) = m(T_i - 1) + z(Q_i - \lambda_{i\to j})
\end{aligned}
$$

Then

$$
\begin{aligned}
W_jT_i - T_jW_i &= R_jT_jT_i - mT_j(T_i - 1) - zT_j(Q_i - \lambda_{i\to j}) \\
w_i(u)(Q_iT_j + Q_jT_i) &= (mu + zq_{i\to j}(u))(Q_iT_j + Q_jT_i) \\
&= q_{i\to j}(u)(zQ_iT_j + zQ_jT_i) + u(mQ_iT_j + mQ_jT_i) \\
\frac{\partial}{\partial u}R(w,\sigma) &\propto q_{i\to j}(u)\left[W_jT_i - T_jW_i\right] + w_i(u)(Q_iT_j + Q_jT_i) - u(Q_iW_j + Q_jW_i) \\
&= q_{i\to j}(u)\left[R_jT_jT_i - mT_j(T_i - 1) - zT_j(Q_i - \lambda_{i\to j}) + zQ_iT_j + zQ_jT_i\right] \\
&\quad + u\left[mQ_iT_j + mQ_jT_i - Q_iR_jT_j - Q_j(m(T_i - 1) + z(Q_i - \lambda_{i\to j}))\right] \\
&= q_{i\to j}(u)\left[(R_j - m)T_jT_i + mT_j + zQ_jT_i + zT_j\lambda_{i\to j}\right] \\
&\quad + u\left[Q_iT_j(m - R_j) + Q_j(m - zQ_i + z\lambda_{i\to j})\right]
\end{aligned}
$$

□

REMARK 3. $\lim_{u\to 0}\frac{q_{i\to j}(u)}{u} = \lambda_{i\to j}$.

Proof. Simple application of L'Hopital's rule.

$$\lim_{u\to 0}\frac{q_{i\to j}(u)}{u}=\lim_{u\to 0}\frac{\partial}{\partial u}q_{i\to j}(u)=\lim_{u\to 0}\frac{\partial}{\partial u}\frac{\lambda_{i\to j}}{\lambda_{i\to j}+\lambda_{j\to i}}\left[1-e^{-(\lambda_{i\to j}+\lambda_{j\to i})u}\right]=\lambda_{i\to j}$$

□

Remark 4. $\lambda_{i\to j}T_i-Q_i\geq 0$ and maximized when $\sigma_i=(0,\infty)$. Similarly, $Q_i\geq 0$ and maximized when $\sigma_i=(0,\infty)$.

Proof.

$$\begin{aligned}\lambda_{i\to j}T_i-Q_i&=\lambda_{i\to j}\left[1+\lambda_i\int_{\tau\in\sigma_i}\tau dF_i(\tau)\right]-\lambda_{i\to j}-\lambda_i\int_{\sigma_i}q_{i\to j}(\tau)dF_i(\tau)\\&=\lambda_i\int_{\tau\in\sigma_i}\left[\lambda_{i\to j}\tau-q_{i\to j}(\tau)\right]dF_i(\tau)\end{aligned}$$

$\lambda_{i\to j}\tau-q_{i\to j}(\tau)$ is increasing in τ:

$$\frac{\partial}{\partial\tau}\left[\lambda_{i\to j}\tau-q_{i\to j}(\tau)\right]=\lambda_{i\to j}-\left[\lambda_{i\to j}e^{-(\lambda_{i\to j}+\lambda_{j\to i})\tau}\right]\geq 0$$

and $\lambda_{i\to j}*0-q_{i\to j}(0)=0$. Thus, the function being integrated is positive, and so $\lambda_{i\to j}T_i-Q_i>0$ and maximized when $\sigma_i=(0,\infty)$. Identical proof holds for Q_i.

□

In the next lemma, we consider u an upper endpoint of σ_2, and so $\frac{\partial}{\partial u}R(w=\{w_1,w_2\},\sigma=\{\sigma_1,\sigma_2\})$ is a derivative with respect to an upper endpoint of σ_2.

Lemma 7. *Fix arbitrary σ_1, and thus Q_1,T_1,R_1. Let $\bar{Q}_2,\bar{T}_2$ be the respective values of Q_2,T_2 at $\sigma_2=(0,\infty)$. Let $w_2(\tau)=m\tau+zq_{2\to 1}(\tau)$, where $m>R_1$.*

If

$$\frac{T_1(\lambda_{2\to 1}\bar{T}_2-\bar{Q}_2)-(Q_1+T_1\lambda_{2\to 1})}{\left(Q_1(\lambda_{2\to 1}\bar{T}_2-\bar{Q}_2)+\lambda_{2\to 1}(Q_1+T_1\lambda_{2\to 1})\right)}\leq\frac{z}{m-R_1}\leq\frac{\bar{Q}_2T_1+Q_1}{Q_1(\bar{Q}_2-\lambda_{2\to 1})}$$

Then $\frac{\partial}{\partial u}R(w,\sigma)\geq 0$, for all u,σ_2. Furthermore, the constraint set is feasible regardless of the primitives.

Proof.

Suppose we have $w_2(u)=mu+zq_{2\to 1}(u)$, for some $m>R_1,z\geq 0$.

From Remark 2,

$$\begin{aligned}\frac{\partial}{\partial u}R(w,\sigma)\propto\ &u\left[\frac{q_{2\to 1}(u)}{u}\left[(R_1-m)T_1T_2+mT_1+zQ_1T_2+zT_1\lambda_{2\to 1}\right]\right]\\&+u\left[Q_2T_1(m-R_1)+Q_1(m-zQ_2+z\lambda_{2\to 1})\right]\end{aligned}$$

T_2,Q_2 are functions of σ_2.

As $u\to\infty$, the term in brackets in the first term goes to 0, and thus the first necessary condition is to have the second term always positive.

If the second term is always positive, then the first term may be negative as long as it has a smaller absolute value than the second term. As $u\to 0$, the ratio between (absolute value of) the first and second terms is maximized. Thus, the second necessary (and sufficient) condition is to have the entire value positive when we take the limit of $\frac{q_{2\to1}(u)}{u}$ as $u \to 0$.

These two conditions are sufficient for $\frac{\partial}{\partial u} R(w,\sigma) \ge 0$, for all u, σ_2.

From the first condition, we need m, z such that:

$$\begin{aligned} & Q_2 T_1 (m - R_1) + Q_1 (m - z Q_2 + z\lambda_{2\to1}) \ge 0 && \forall T_1, Q_1, Q_2, R_1 \\ \iff & \frac{z}{m - R_1} \le \frac{Q_2 T_1 + \frac{m}{m-R_1} Q_1}{Q_1 (Q_2 - \lambda_{2\to1})} \end{aligned}$$

From the second condition, and using Remark 3 we need:

$$\begin{aligned} & \lambda_{2\to1} \left[(R_1 - m) T_1 T_2 + m T_1 + z Q_1 T_2 + z T_1 \lambda_{2\to1} \right] + \left[Q_2 T_1 (m - R_1) + Q_1 (m - z Q_2 + z \lambda_{2\to1}) \right] \ge 0 \\ \iff & (m - R_1) T_1 (Q_2 - \lambda_{2\to1} T_2) + m (Q_1 + \lambda_{2\to1} T_1) + z Q_1 (\lambda_{2\to1} T_2 - Q_2 + \lambda_{2\to1}) + z T_1 \lambda_{2\to1}^2 \ge 0 \\ \iff & z \left(Q_1 (\lambda_{2\to1} T_2 - Q_2) + \lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1}) \right) \ge (m - R_1) T_1 (\lambda_{2\to1} T_2 - Q_2) - m (Q_1 + T_1 \lambda_{2\to1}) \\ \iff & \frac{z}{m - R_1} \ge \frac{T_1 (\lambda_{2\to1} T_2 - Q_2) - \frac{m}{m-R_1} (Q_1 + T_1 \lambda_{2\to1})}{(Q_1 (\lambda_{2\to1} T_2 - Q_2) + \lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1}))} \end{aligned}$$

Putting the conditions together, we need, for all T_i, Q_i, R_i:

$$\frac{T_1 (\lambda_{2\to1} T_2 - Q_2) - \frac{m}{m-R_1} (Q_1 + T_1 \lambda_{2\to1})}{(Q_1 (\lambda_{2\to1} T_2 - Q_2) + \lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1}))} \le \frac{z}{m - R_1} \le \frac{Q_2 T_1 + \frac{m}{m-R_1} Q_1}{Q_1 (Q_2 - \lambda_{2\to1})}$$

$m > R_1$ by supposition, and so $\frac{m}{m-R_1} > 1$. Thus, the following is sufficient as the constraints become tighter:

$$\begin{aligned} \frac{T_1 (\lambda_{2\to1} T_2 - Q_2) - (Q_1 + T_1 \lambda_{2\to1})}{(Q_1 (\lambda_{2\to1} T_2 - Q_2) + \lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1}))} & \le \frac{z}{m - R_1} \le \frac{Q_2 T_1 + Q_1}{Q_1 (Q_2 - \lambda_{2\to1})} \\ \iff \frac{T_1 - \frac{Q_1 + T_1 \lambda_{2\to1}}{(\lambda_{2\to1} T_2 - Q_2)}}{Q_1 + \frac{\lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1})}{(\lambda_{2\to1} T_2 - Q_2)}} & \le \frac{z}{m - R_1} \le \frac{T_1 + \frac{Q_1}{Q_2}}{Q_1 \left(1 - \frac{\lambda_{2\to1}}{Q_2}\right)} \end{aligned}$$

It turns out that both constraints are tightest when $\sigma_2 = (0, \infty)$. In the left constraint, the numerator is increasing and the denominator is decreasing with $\lambda_{2\to1} T_2 - Q_2$, and so the constraint becomes tighter as $\lambda_{2\to1} T_2 - Q_2$ increases. By Remark 4, $\lambda_{2\to1} T_2 - Q_2$ is always positive, and maximized when $\sigma_2 = (0, \infty)$. Similarly, in the right constraint, the numerator decreases and the denominator increases with Q_2.

Thus, it is sufficient for the two constraints to be feasible for $\sigma_2 = (0, \infty)$. Then, they are satisfied for all σ_2'. For feasibility, we need

$$\begin{aligned} & \frac{T_1 (\lambda_{2\to1} T_2 - Q_2) - (Q_1 + T_1 \lambda_{2\to1})}{(Q_1 (\lambda_{2\to1} T_2 - Q_2) + \lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1}))} \le \frac{Q_2 T_1 + Q_1}{Q_1 (Q_2 - \lambda_{2\to1})} \\ \iff & (T_1 (\lambda_{2\to1} T_2 - Q_2) - (Q_1 + T_1 \lambda_{2\to1})) Q_1 (Q_2 - \lambda_{2\to1}) \\ & \le (Q_2 T_1 + Q_1)(Q_1 (\lambda_{2\to1} T_2 - Q_2) + \lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1})) \\ \iff & Q_1 Q_2 (T_1 (\lambda_{2\to1} T_2 - Q_2) - (Q_1 + T_1 \lambda_{2\to1})) - Q_1 \lambda_{2\to1} (T_1 (\lambda_{2\to1} T_2 - Q_2) - (Q_1 + T_1 \lambda_{2\to1})) \\ & \le Q_2 T_1 (Q_1 (\lambda_{2\to1} T_2 - Q_2) + \lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1})) + Q_1 (Q_1 (\lambda_{2\to1} T_2 - Q_2) + \lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1})) \\ \iff & Q_1 Q_2 (-(Q_1 + T_1 \lambda_{2\to1})) - Q_1 \lambda_{2\to1} (T_1 (\lambda_{2\to1} T_2 - Q_2)) \\ & \le Q_2 T_1 (\lambda_{2\to1} (Q_1 + T_1 \lambda_{2\to1})) + Q_1 (Q_1 (\lambda_{2\to1} T_2 - Q_2)) \end{aligned}$$

For any valid Q_i, T_i, the left hand side of the final line is always non-positive, and the right hand side is always non-negative, and thus there exists feasible ratios $\frac{z}{m-R_1}$.

□

We can now do the same thing for the first state, assuming that $w_1(\tau)$ is of the form $w_1(\tau) = m\tau + zq_{1\to 2}(\tau)$, where now $z \leq$ and $m = R_2$. In the next lemma, we consider u an upper endpoint of σ_1, and so $\frac{\partial}{\partial u} R(w = \{w_1, w_2\}, \sigma = \{\sigma_1, \sigma_2\})$ is a derivative with respect to an upper endpoint of σ_1. Then,

LEMMA 8. *Fix arbitrary σ_2, and thus Q_2, T_2, R_2. Let $\bar{Q}_1, \bar{T}_1$ be the respective values of Q_1, T_1 at $\sigma_1 = (0, \infty)$. Let $w_1(\tau) = m\tau + zq_{1\to 2}(\tau)$, where $m = R_2$.*

If

$$-\frac{(T_2\lambda_{1\to 2} + Q_2)}{Q_2(\lambda_{1\to 2}\bar{T}_1 - \bar{Q}_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2)} \leq \frac{z}{R_2} \leq \frac{1}{(\bar{Q}_1 - \lambda_{1\to 2})}$$

Then $\frac{\partial}{\partial u} R(w, \sigma) \geq 0$, for all u, σ_1. Furthermore, the constraint set is feasible regardless of the primitives.

Proof.

Similar to previous proof. Suppose we have $w_1(u) = mu + zq_{1\to 2}(u)$, for some $m = R_2, z \leq 0$.

From Remark 2,

$$\begin{aligned}
\frac{\partial}{\partial u} R(w, \sigma) =& u \left[\frac{q_{1\to 2}(u)}{u} \left[(R_2 - m)T_1T_2 + mT_2 + zQ_2T_1 + zT_2\lambda_{1\to 2} \right] \right] \\
&+ u \left[Q_1T_2(m - R_2) + Q_2(m - zQ_1 + z\lambda_{1\to 2}) \right] \\
=& u \left[\frac{q_{1\to 2}(u)}{u} \left[R_2T_2 + zQ_2T_1 + zT_2\lambda_{1\to 2} \right] + \left[Q_2(R_2 - zQ_1 + z\lambda_{1\to 2}) \right] \right]
\end{aligned}$$

As before, we have two necessary and sufficient conditions for $\frac{\partial}{\partial u} R(w, \sigma) \geq 0$, for all u, σ_1.

From the first condition, we need m, z such that:

$$\begin{aligned}
& Q_2(R_2 - z(Q_1 - \lambda_{1\to 2})) \geq 0 \qquad \forall T_2, Q_2, Q_1, R_2 \\
\iff & \frac{z}{R_2} \leq \frac{1}{(Q_1 - \lambda_{1\to 2})}
\end{aligned}$$

This condition is trivially met when $z \leq 0$.

Similarly, the second condition becomes

$$\begin{aligned}
& \lambda_{1\to 2} \left[R_2T_2 + zQ_2T_1 + zT_2\lambda_{1\to 2} \right] + \left[Q_2(R_2 - zQ_1 + z\lambda_{1\to 2}) \right] \geq 0 \\
\iff & \lambda_{1\to 2}R_2T_2 + Q_2R_2 \geq -zQ_2(\lambda_{1\to 2}T_1 - Q_1) - z\lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2) \\
\iff & \frac{z}{R_2} \geq -\frac{(T_2\lambda_{1\to 2} + Q_2)}{Q_2(\lambda_{1\to 2}T_1 - Q_1) + \lambda_{1\to 2}(T_2\lambda_{1\to 2} + Q_2)}
\end{aligned}$$

Both constraints are tightest when $\sigma_1 = (0, \infty)$. By Remark 4, $\lambda_{1\to 2}T_1 - Q_1$ is always positive, and maximized when $\sigma_1 = (0, \infty)$.

As one is non-negative and the other is non-positive, the constraints are feasible.

□

B.3. Optimal policies as depend on derivatives

Here we prove how the optimal policies depend on the derivative of the reward function (whether they are always positive, strictly increasing, strictly decreasing, strictly quasi-convex, or strictly quasi-concave). The following five proofs, one for each case, are extremely similar. We start with an arbitrary open set σ', and then make a sequence of changes to the policy that result in a set σ of the appropriate form. The set-up in each proof is the same; only the exact changes made to improve the policy differ. These changes depend on the structure of the derivative of the reward with respect to the endpoints of the sets that make up the policy, $\frac{\partial}{\partial u}R(w,\sigma), \frac{\partial}{\partial \ell}R(w,\sigma)$. The idea is that as long as the derivative can be shown to be non-negative for some u that is an endpoint of σ_i, that policy can be locally modified to accept more trips while not decreasing the overall reward function.

THEOREM 5. *Consider a reward function $\tilde{R}(\sigma)$ that maps open, measurable subsets $\sigma = \cup_k^\infty(\ell_k, u_k) \subseteq (0,\infty)$ to the non-negative reals, and probability measure F such that F is continuous, i.e. f is bounded. Further consider an open subset $\sigma' \subseteq (0,\infty)$. Suppose*

1. *$F(\sigma') > 0$, and $\tilde{R}(\sigma') > \tilde{R}(\emptyset)$.*
2. *$\tilde{R}(\sigma)$ is continuous in σ.*
3. *Let $\sigma = \cup_k(\ell_k, u_k)$. Then, let $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ denote the partial derivative of $\tilde{R}$ with respect to an upper end-point u_k of the intervals that make up σ. Suppose $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ exists, for all σ and its endpoints u_k, and is continuous in u for a fixed σ.*
4. *Further suppose that $\frac{\partial}{\partial u_k}\tilde{R}(\cup_m(\ell_m, u_m))$ is continuous in u_k, i.e. $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ is continuous in σ.*
5. *Suppose $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ is **always non-negative** in u (for a fixed σ), for all σ.*

Then, the policy $\sigma^* = (0,\infty)$ such that $\tilde{R}(\sigma^*) \geq \tilde{R}(\sigma')$.

Proof.

We prove the result as follows: start at any arbitrary open measurable subset $\sigma' \subseteq (0,\infty) = \cup_k^\infty(\ell_k, u_k) = \cup_k^\infty \zeta_k$, where the intervals are disjoint and $\zeta_k = (\ell_k, u_k)$ denotes the kth interval. It is well known that any open subset of $\mathbb{R}$ can be uniquely written as the countable union of such disjoint intervals.

Starting with this set σ', we create a sequence $\sigma'_\delta \to \sigma'$ (as $\delta \to 0$), where, for each δ, $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$. Then, we show that there exists a $\sigma^* = (\ell^*, \infty)$, for some $\ell \in \mathbb{R}_+$, such that $\tilde{R}(\sigma'_\delta) \leq R(\sigma^*), \forall \delta$. By continuity of the set function $\tilde{R}$, this implies that $\tilde{R}(\sigma') \leq R(\sigma^*)$.

Sequence σ'_δ. Each σ'_δ will be of the form $\sigma'_\delta = (0,L)\,(\cup_{k=1}^K(\ell_k, u_k)) \cup (B,\infty)$, for some K, B, L that depend on δ. We construct a σ'_δ such that $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$ as follows:

- F is a finite (probability) measure, and so there exists K such that $F(\cup_{k=K+1}^\infty(\ell_k, u_k)) < \delta/2$. (Since $F(\sigma') \leq 1$, it follows by the Cauchy condition).
- Let $B \in \mathbb{R}$ s.t. $F((B,\infty)) < \delta/4$. Let $L \in \mathbb{R}$ s.t. $F((0,L)) < \delta/4$. Such B, L exist by condition on F.
- Set $\sigma'_\delta = (0,L) \cup (\cup_{k=1}^K(\ell_k, u_k)) \cup (B,\infty)$.

- For convenience, we re-index the disjoint intervals $\{\zeta_k\}_{k=1}^{K+2}$ such that they are in increasing order, i.e. $u_k > \ell_k \geq u_{k-1}, \forall k > 1$, starting at $(0, L)$, with the last interval (B, ∞) If there exist any intervals such that $\ell_k = u_{k-1}$, replace them with the combined interval (ℓ_{k-1}, u_k). If $\{B, \infty\}$ overlaps with the last interval, combine them.

Starting with this set σ', we create a sequence $\sigma'_\delta \to \sigma'$ (as $\delta = \frac{1}{N}, N = 1, 2, 3, \dots$), where, for each δ, $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$.

Showing that $\tilde{R}(\sigma^*) \geq \tilde{R}(\sigma')$, where $\sigma^* = (0, \infty)$.

Now, starting at $\sigma = \sigma'_\delta$, we describe a sequence of modifications to σ, such that each modification does not reduce the reward $\tilde{R}(\sigma)$. The limit of this sequence of modifications is the policy $\sigma^* = (0, \infty)$, regardless of the starting σ'_δ. This shows that $\tilde{R}(\sigma^*) \geq \tilde{R}(\sigma'_\delta)$.

Let $\sigma = \cup_{k=1}^K (\ell_k, u_k)$, and note that $\ell_1 = 0$ from above. By supposition that $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ is **always non-negative** in u, we can increase u_1 (merging with other intervals) without decreasing $\tilde{R}(\sigma)$.

Thus, we can keep increasing u_1, and $u_1 \to B$, and so $R((0, \infty)) \geq R(\sigma')$.

□

THEOREM 6. *Consider a reward function $\tilde{R}(\sigma)$ that maps open, measurable subsets $\sigma = \cup_k^\infty (\ell_k, u_k) \subseteq (0, \infty)$ to the non-negative reals, and probability measure F such that F is continuous, i.e. f is bounded. Further consider an open subset $\sigma' \subseteq (0, \infty)$.*

Suppose

1. *$F(\sigma') > 0$, and $\tilde{R}(\sigma') > \tilde{R}(\emptyset)$.*
2. *$\tilde{R}(\sigma)$ is continuous in σ.*
3. *Let $\sigma = \cup_k (\ell_k, u_k)$. Then, let $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ denote the partial derivative of $\tilde{R}$ with respect to an upper end-point u_k of the intervals that make up σ. Suppose $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ exists, for all σ and its endpoints u_k, and is continuous in u for a fixed σ.*
4. *Further suppose that $\frac{\partial}{\partial u_k}\tilde{R}(\cup_m (\ell_m, u_m))$ is continuous in u_k, i.e. $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ is continuous in σ.*
5. *Suppose $\exists \epsilon > 0$ s.t. $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ has the same sign as a function $r(u, \sigma)$ that is **strictly increasing** in u (for a fixed σ), for all σ such that $\tilde{R}(\sigma) \geq \tilde{R}(\sigma') - \epsilon$.*

Then, there exists a value $\ell^* \in \mathbb{R}_+ \cup \{\infty\}$ such that the policy $\sigma^* = (\ell^*, \infty)$ such that $\tilde{R}(\sigma^*) \geq \tilde{R}(\sigma')$.

Proof.

We prove the result as follows: start at any arbitrary open measurable subset $\sigma' \subseteq (0, \infty) = \cup_k^\infty (\ell_k, u_k) = \cup_k^\infty \zeta_k$, where the intervals are disjoint and $\zeta_k = (\ell_k, u_k)$ denotes the kth interval. It is well known that any open subset of $\mathbb{R}$ can be uniquely written as the countable union of such disjoint intervals.

Starting with this set σ', we create a sequence $\sigma'_\delta \to \sigma'$ (as $\delta \to 0$), where, for each δ, $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$. Then, we show that there exists a $\sigma^* = (\ell^*, \infty)$, for some $\ell \in \mathbb{R}_+$, such that $\tilde{R}(\sigma'_\delta) \leq R(\sigma^*), \forall \delta$. By continuity of the set function $\tilde{R}$, this implies that $\tilde{R}(\sigma') \leq R(\sigma^*)$.

Sequence σ'_δ. Each σ'_δ will be of the form $\sigma'_\delta = (\cup_{k=1}^K (\ell_k, u_k)) \cup (B, \infty)$, for some K, B that depends on δ. We construct a σ'_δ such that $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$ as follows:

- F is a finite (probability) measure, and so there exists K such that $F(\cup_{k=K+1}^\infty (\ell_k, u_k)) < \delta/2$. (Since $F(\sigma') \le 1$, it follows by the Cauchy condition).
- Let $B \in \mathbb{R}$ s.t. $F((B, \infty)) < \delta/2$. Such a B exists by condition on F.
- For convenience, we re-index the disjoint intervals $\{\zeta_k\}_{k=1}^K$ such that they are in increasing order, i.e. $u_k > \ell_k \ge u_{k-1}, \forall k > 1$. If there exist any intervals such that $\ell_k = u_{k-1}$, replace them with the combined interval (ℓ_{k-1}, u_k). If $\{B, \infty\}$ overlaps with the last interval, combine them.

Use the sequence: $\delta = \frac{1}{N}, N = 1, 2, 3, \ldots$. As constructed, $\sigma'_\delta \to \sigma'$ as $\delta \to 0$.

Showing that $\exists \ell^*$ such that $\exists \delta_0 : \forall \delta < \delta_0, \tilde{R}(\sigma'_\delta) \le \tilde{R}((\ell^*, \infty))$.

By suppositions, $\exists \delta_0$ small enough such that $r(u, \sigma)$ is strictly increasing for all σ such that $\tilde{R}\sigma \ge \tilde{R}(\sigma'_\delta)$, $\forall \delta < \delta_0$. Suppose $\delta < \delta_0$.

Now, starting at $\sigma = \sigma'_\delta$, we describe a sequence of modifications to σ, such that each modification does not reduce the reward $\tilde{R}(\sigma)$. The limit of this sequence of modifications is a policy $\sigma^* = (\ell^*, \infty)$, regardless of the starting σ'_δ. This shows that $\tilde{R}(\sigma^*) \ge \tilde{R}(\sigma'_\delta)$.

We continue our abuse of notation with $r(\ell, \sigma)$ indicating a function that has the same sign as the derivative of a lower endpoint of σ_i.

By the supposition that $r(u, \sigma)$ strictly increasing in u, we have:

$$
\begin{aligned}
& r(\ell, \sigma) \text{ strictly decreasing} & \\
& r(\ell_1, \sigma) \le 0 \implies r(u_1, \sigma) > 0 & \ell_1 < u_1 \\
\equiv\ & \frac{\partial}{\partial \ell_1} \tilde{R}(\sigma) \le 0 \implies \frac{\partial}{\partial u_1} \tilde{R}(\sigma) > 0 & x f(x) \ge 0 \\
& r(\ell_1, \sigma) > 0 \impliedby r(u_1, \sigma) \le 0 & \\
\equiv\ & \frac{\partial}{\partial \ell_1} \tilde{R}(\sigma) > 0 \impliedby \frac{\partial}{\partial u_1} \tilde{R}(\sigma) \le 0 &
\end{aligned}
$$

Case 1: $\exists \zeta_1, \zeta_2 \subset \sigma$ such that $\ell_2 > u_1, |\zeta_1|, |\zeta_2|$, i.e. there is more than one interval that makes up σ, and ζ_1, ζ_2 are the first and second such intervals, respectively, with positive mass.

Then we make the following sequence of changes (forming new σ), depending on $\frac{\partial}{\partial \ell_1} \tilde{R}(\sigma), \frac{\partial}{\partial u_1} \tilde{R}(\sigma)$:

Subcase 1A, $\frac{\partial}{\partial u_1} \tilde{R}(\sigma) > 0$: Increase u_1 until $u_1 = \ell_2$ (exit Case 1), or $\frac{\partial}{\partial u_1} \tilde{R}(\sigma) \le 0$ (go to Case 1B).

Sub-subcase 1AA, $\frac{\partial}{\partial \ell_1} \tilde{R}(\sigma) < 0$, $\ell_1 > 0$: Simultaneously, decrease ℓ_1.

Sub-subcase 1AB, $\frac{\partial}{\partial \ell_1} \tilde{R}(\sigma) \ge 0$ or $\ell_1 = 0$: Hold ℓ_1 fixed.

Subcase 1B, $\frac{\partial}{\partial u_1} \tilde{R}(\sigma) \le 0 \implies \frac{\partial}{\partial \ell_1} \tilde{R}(\sigma) > 0$: Increase ℓ_1 until $\ell_1 = u_1$ (exit Case 1), or $\frac{\partial}{\partial \ell_1} \tilde{R}(\sigma) \le 0$ (which implies $\frac{\partial}{\partial u_1} \tilde{R}(\sigma) > 0$, i.e. go to Case 1A).

Each of these changes cannot decrease $\tilde{R}(\sigma)$, due to the direction of the changes in u_1, ℓ_1 and the corresponding derivatives (and thus the scaled gradient remains strictly increasing by supposition). Note that these subcases are mutually-exclusive, and one is true as long as $\exists \zeta_1, \zeta_2 \subset \sigma, \ell_2 > u_1$. Further, note that u_1 is increasing in Subcase 1A and constant in Subcase 1B. Thus, with ℓ_2 fixed and bounded, eventually:

- $\ell_1 \to u_1$, in Subcase 1B (i.e. the first interval collapses to mass 0). OR
- $u_1 \to \ell_2$, in Subcase 1A (i.e. the first interval merges with the second).

Thus, this sequence of changes cannot decrease the reward, and results in there being one fewer interval then before (after combining the bottom 2 intervals by adding the point $u_1 = \ell_2$ of 0 measure). Case 1 can be iteratively applied until there is just a single interval $\sigma = (\ell', \infty)$.

Case 2: $\sigma = (\ell', \infty)$, i.e. there is a single interval that makes up σ

By supposition, $\tilde{R}(\sigma') > \tilde{R}(\emptyset)$ and so $\tilde{R}((\ell', \infty)) > \tilde{R}(\emptyset)$. Further $\tilde{R}((\ell, \infty))$ is a continuous function in ℓ. Thus, there exists L such that $\forall \ell > L$, $\tilde{R}((\ell', \infty)) > \tilde{R}((\ell, \infty))$.

Thus, there exists $\ell^* \in [0, L]$ such that $\tilde{R}((\ell^*, \infty)) \geq \tilde{R}((\ell, \infty)), \forall \ell \in \mathbb{R}_+ \cup \{\infty\}$ (continuous functions in a compact domain have a maximum). $\square$

THEOREM 7. *Consider a reward function $\tilde{R}(\sigma)$ that maps open, measurable subsets $\sigma = \cup_k^\infty (\ell_k, u_k) \subseteq (0, \infty)$ to the non-negative reals, and probability measure F such that F is continuous, i.e. f is bounded. Further consider an open subset $\sigma' \subseteq (0, \infty)$.*

Suppose

1. $F(\sigma') > 0$
2. *$\tilde{R}(\sigma)$ is continuous in σ.*
3. *Let $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ denote the partial derivative of $\tilde{R}$ with respect to an upper end-point u_k of the intervals that make up σ. Suppose $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ exists, for all σ and its endpoints u_k, and is continuous in u for a fixed σ.*
4. *Further suppose that $\frac{\partial}{\partial u_k}\tilde{R}(\cup_m(\ell_m, u_m))$ is continuous in u_k, i.e. $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ is continuous in σ.*
5. *Suppose $\exists \epsilon > 0$ s.t. $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ has the same sign as a function $r(u, \sigma)$ that is **strictly decreasing** in u (for a fixed σ), for all σ such that $\tilde{R}(\sigma) \geq \tilde{R}(\sigma') - \epsilon$.*

Then, there exists a value $u^* \in [0, \infty) \cup \{\infty\}$ such that the policy $\sigma^* = (0, u^*)$ such that $\tilde{R}(\sigma^*) \geq \tilde{R}(\sigma')$.

Proof. The proof is extremely similar to that of Theorem 6.

However, we now need to modify the starting σ' so it *does not* contain an interval (B, ∞), and the other cases have movement in different directions.

As before: start at any arbitrary open measurable subset $\sigma' \subseteq (0, \infty) = \cup_k^\infty (\ell_k, u_k) = \cup_k^\infty \zeta_k$, where the intervals are disjoint and $\zeta_k = (\ell_k, u_k)$ denotes the kth interval. Starting with this set σ', we create a sequence $\sigma'_\delta \to \sigma'$ (as $\delta \to 0$), where, for each δ, $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$. Then, we show that there exists a $\sigma^* = (\ell^*, \infty)$, for some $\ell \in \mathbb{R}_+$, such that $\tilde{R}(\sigma'_\delta) \leq R(\sigma^*), \forall \delta$. By continuity of the set function $\tilde{R}$, this implies that $\tilde{R}(\sigma') \leq R(\sigma^*)$.

Sequence σ'_δ. Each σ'_δ will be of the form $\sigma'_\delta = (0,L)\cup(\cup_{k=1}^K(\ell_k,u_k))$, for some K,L that depends on δ and $\ell_k,u_k<\infty$. We construct a σ'_δ such that $F(\sigma'\setminus\sigma'_\delta\cup\sigma'_\delta\setminus\sigma')<\delta$ as follows:

- F is a finite (probability) measure, and so there exists K such that $F(\cup_{k=K+1}^\infty(\ell_k,u_k))<\delta/2$. (Since $F(\sigma')\le 1$, it follows by the Cauchy condition).
- Let $B\in\mathbb{R}$ s.t. $F((B,\infty))<\delta/2$. Let $L\in\mathbb{R}$ s.t. $F((0,L))<\delta/2$. Such B,L exist by condition on F.
- Set $\sigma'_\delta=(0,L)\cup(\cup_{k=1}^K(\ell_k,u_k))\setminus(B,\infty)$.
- For convenience, we re-index the disjoint intervals $\{\zeta_k\}_{k=1}^{K+1}$ (or less, if some eliminated when removing B; include $(0,L)$ as the first interval) such that they are in increasing order, i.e. $u_k>\ell_k\ge u_{k-1},\forall k>1$. If there exist any intervals such that $\ell_k=u_{k-1}$, replace them with the combined interval (ℓ_{k-1},u_k).

Use the sequence: $\delta=\frac{1}{N}, N=1,2,3,\dots$. As constructed, $\sigma'_\delta\to\sigma'$ as $\delta\to 0$.

Showing that $\exists u^*$ such that $\exists\delta_0:\forall\delta<\delta_0,\tilde{R}(\sigma'_\delta)\le\tilde{R}((0,u^*))$.

By suppositions, $\exists\delta_0$ small enough such that $r(u,\sigma)$ is strictly decreasing for all σ such that $\tilde{R}\sigma\ge\tilde{R}(\sigma'_\delta)$, $\forall\delta<\delta_0$. Suppose $\delta<\delta_0$.

Now, starting at $\sigma=\sigma'_\delta$, we describe a sequence of modifications to σ, such that each modification does not reduce the reward $\tilde{R}(\sigma)$. The limit of this sequence of modifications is a policy $\sigma^*=(0,u^*)$, regardless of the starting σ'_δ. This shows that $\tilde{R}(\sigma^*)\ge\tilde{R}(\sigma'_\delta)$.

We continue our abuse of notation with $r(\ell,\sigma)$ indicating a function that has the same sign as the derivative of a lower endpoint of σ_i. By the supposition that $r(u,\sigma)$ strictly decreasing in u, we have:

$$
\begin{aligned}
& r(\ell,\sigma)\text{ strictly increasing} \\
& r(u_K,\sigma)\ge 0 \implies r(\ell_K,\sigma)<0 && \ell_1<u_1 \\
\equiv\ & \frac{\partial}{\partial u_K}\tilde{R}(\sigma)\ge 0 \implies \frac{\partial}{\partial \ell_K}\tilde{R}(\sigma)<0 && xf(x)\ge 0 \\
& r(u_K,\sigma)<0 \impliedby r(\ell_K,\sigma)\ge 0 \\
\equiv\ & \frac{\partial}{\partial u_K}\tilde{R}(\sigma)<0 \impliedby \frac{\partial}{\partial \ell_K}\tilde{R}(\sigma)\ge 0
\end{aligned}
$$

Case 1: $\exists\zeta_{K-1},\zeta_K\subset\sigma$ such that $\ell_K>u_{K-1},|\zeta_K|,|\zeta_{K-1}|$, i.e. there is more than one interval that makes up σ, and ζ_K,ζ_{K-1} are the last two such intervals, respectively, with positive mass.

Then we make the following sequence of changes (forming new σ), depending on $\frac{\partial}{\partial\ell_K}\tilde{R}(\sigma)$, $\frac{\partial}{\partial u_K}\tilde{R}(\sigma)$:

Subcase 1A, $\frac{\partial}{\partial\ell_K}\tilde{R}(\sigma)<0$: Decreasing ℓ_K until $\ell_K=u_{K-1}$ (exit Case 1), or $\frac{\partial}{\partial\ell_K}\tilde{R}(\sigma)\ge 0$ (go to Case 1B).

Sub-subcase 1AA, $\frac{\partial}{\partial u_K}\tilde{R}(\sigma)>0$: Simultaneously, increase u_K.

Sub-subcase 1AB, $\frac{\partial}{\partial u_K}\tilde{R}(\sigma)\le 0$: Hold u_K fixed.

Subcase 1B, $\frac{\partial}{\partial\ell_K}\tilde{R}(\sigma)\ge 0\implies\frac{\partial}{\partial u_K}\tilde{R}(\sigma)<0$: Decrease u_K until $u_K=\ell_K$ (exit Case 1), or $\frac{\partial}{\partial u_K}\tilde{R}(\sigma)\ge 0$ (which implies $\frac{\partial}{\partial\ell_K}\tilde{R}(\sigma)<0$, i.e. go to Case 1A).

Each of these changes cannot decrease $\tilde{R}(\sigma)$, due to the direction of the changes in ℓ_K, u_K and the corresponding derivatives (and thus the scaled gradient remains strictly increasing by supposition). Note that these subcases are mutually-exclusive, and one is true as long as $\exists$ such $\zeta_K, \zeta_{K-1} \subset \sigma$. Further, note that ℓ_K is decreasing in Subcase 1A and constant in Subcase 1B. Thus, eventually:

- $u_K \to \ell_K$, in Subcase 1B (i.e. the last interval collapses to mass 0). OR
- $\ell_K \to u_{K-1}$, in Subcase 1A (i.e. the last interval merges with the second to last).

Thus, this sequence of changes cannot decrease the reward, and results in there being one fewer interval then before Case 1 can be iteratively applied until there is just a single interval $\sigma = (0, u')$.

Case 2: $\sigma = (0, u')$**, i.e. there is a single interval that makes up** σ By supposition, $\tilde{R}((0,u))$ is a continuous function for $u \in [0,\infty) \cup \{\infty\}$. Further, $\frac{\partial}{\partial u}\tilde{R}((0,u))$ is strictly decreasing for all u such that $\tilde{R}((0,u)) \geq \tilde{R}((0,u'))$. If $\frac{\partial}{\partial u}\tilde{R}((0,u)) > 0, \forall u$, then $u^* = \infty$ is optimal. Otherwise if $\frac{\partial}{\partial u}\tilde{R}((0,u)) < 0 \forall u$, then u^* is optimal. Otherwise $\exists u^* \in (0,\infty)$ such that $\tilde{R}(\sigma^*) \geq \tilde{R}(\sigma')$.

□

Now, we show the structure of the optimal policies in the surge state when the derivative itself has the same sign as a quasi-concave or quasi-convex function.

THEOREM 8. *Consider a reward function* $\tilde{R}(\sigma)$ *that maps open, measurable subsets* $\sigma = \cup_k^\infty (\ell_k, u_k) \subseteq (0,\infty)$ *to the non-negative reals, and probability measure* F *such that* F *is continuous, i.e.* f *is bounded. Further consider an open subset* $\sigma' \subseteq (0,\infty)$.

Suppose

1. $F(\sigma') > 0$
2. $\tilde{R}(\sigma)$ *is continuous in* σ.
3. *Let* $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ *denote the partial derivative of* $\tilde{R}$ *with respect to an upper end-point* u_k *of the intervals that make up* σ. *Suppose* $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ *exists, for all* σ *and its endpoints* u_k, *and is continuous in* u *for a fixed* σ.
4. *Further suppose that* $\frac{\partial}{\partial u_k}\tilde{R}(\cup_m(\ell_m, u_m))$ *is continuous in* u_k, *i.e.* $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ *is continuous in* σ.
5. *Suppose* $\exists \epsilon > 0$ *s.t.* $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ *has the same sign as a function* $r(u,\sigma)$ *that is* ***strictly quasi-convex*** *in* u *(for a fixed* σ*), for all* σ *such that* $\tilde{R}(\sigma) \geq \tilde{R}(\sigma') - \epsilon$.

Then, there exists exist $\ell^*, u^* \in \mathbb{R}_+ \cup \{\infty\}$ *such that the policy* $\sigma^* = (0,\ell^*) \cup (u^*, \infty)$ *such that* $\tilde{R}(\sigma^*) \geq \tilde{R}(\sigma')$, *and it is not the case that both* $\ell^* = 0, u^* = \infty$.

Proof. Proof is similar to that of Theorem 6: start at any open subset $\sigma' \subseteq (0,\infty) = \cup_k^\infty (\ell_k, u_k)$.

Starting with this set σ', we create a sequence $\sigma'_\delta \to \sigma'$ (as $\delta \to 0$), where, for each δ, $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$.

Sequence σ'_δ. Each σ'_δ will be of the form $\sigma'_\delta = (0, L)\,(\cup_{k=1}^K (\ell_k, u_k)) \cup (B, \infty)$, for some K, B, L that depend on δ. We construct a σ'_δ such that $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$ as follows:

- F is a finite (probability) measure, and so there exists K such that $F(\cup_{k=K+1}^{\infty}(\ell_k,u_k)) < \delta/2$. (Since $F(\sigma') \le 1$, it follows by the Cauchy condition).
- Let $B \in \mathbb{R}$ s.t. $F((B,\infty)) < \delta/4$. Let $L \in \mathbb{R}$ s.t. $F((0,L)) < \delta/4$. Such B, L exist by condition on F.
- Set $\sigma'_\delta = (0,L) \cup (\cup_{k=1}^{K}(\ell_k,u_k)) \cup (B,\infty)$.
- For convenience, we re-index the disjoint intervals $\{\zeta_k\}_{k=1}^{K+2}$ such that they are in increasing order, i.e. $u_k > \ell_k \ge u_{k-1}, \forall k > 1$, starting at $(0,L)$, with the last interval (B,∞) If there exist any intervals such that $\ell_k = u_{k-1}$, replace them with the combined interval (ℓ_{k-1}, u_k). If $\{B,\infty\}$ overlaps with the last interval, combine them.

Then, we show that there exists a $\sigma^* = (0,\ell^*) \cup (u^*,\infty)$, for some $u^*, l^* \in \mathbb{R}_+$, such that
$\tilde{R}(\sigma'_\delta) \le R(\sigma^*), \forall\delta$. By continuity of the set function $\tilde{R}$, this implies that $\tilde{R}(\sigma') \le R(\sigma^*)$.

Showing that $\exists \ell^*, u^* \in \mathbb{R}_+ \cup \{\infty\}$ **such that** $\exists \delta_0 : \forall \delta < \delta_0, \tilde{R}(\sigma'_\delta) \le \tilde{R}((0,\ell^*) \cup (u^*,\infty))$.

By suppositions, $\exists\delta_0$ small enough such that $r(u,\sigma)$ is strictly quasi-convex for all σ such that
$\tilde{R}\sigma \ge \tilde{R}(\sigma'_\delta)$, $\forall \delta < \delta_0$. Suppose $\delta < \delta_0$.

Now, starting at $\sigma = \sigma'_\delta$, we describe a sequence of modifications to σ, such that each modification does not reduce the reward $\tilde{R}(\sigma)$. The limit of this sequence of modifications is the policy $\sigma^* = (0,\ell^*) \cup (u^*,\infty)$, regardless of the starting σ'_δ. This shows that $\tilde{R}(\sigma^*) \ge \tilde{R}(\sigma'_\delta)$.

Then, let $\sigma'_\delta = \cup_{k=1}^{K}(\ell_k,u_k)$, where $\zeta_k \triangleq (\ell_k,u_k)$.

The key step is noting that quasi-convexity of the transformed derivative implies that any σ with three intervals ζ_1,ζ_2,ζ_3 can be improved by eliminating the middle interval (or joining it with one of the others).

Case 1: $\exists$ **disjoint** $\zeta_1 = (0,u_1), \zeta_2 = (\ell_2,u_2), \zeta_3 = (\ell_3,u_3)$**, s.t.** $|\zeta_1|,|\zeta_2|,|\zeta_3| > 0$, i.e. σ is composed of at least three intervals, and ζ_1,ζ_2,ζ_3 are the first three such intervals with positive mass. (u_3 may be ∞).

By supposition, the transformed derivative with respect to any of the upper end-points u_k, $r(u_k,\sigma)$, is strictly quasi-convex in u. Then, the transformed derivative with respect to any of the lower end-points ℓ_k, $r(\ell_k,\sigma)$, is strictly quasi-concave in u, and further is the negative of $r(u,\sigma)$ when $u = \ell_k$.

Then, we have:

$$\frac{\partial}{\partial u_1}\tilde{R}(\sigma) \le 0 \text{ and } \frac{\partial}{\partial \ell_3}\tilde{R}(\sigma) \ge 0 \implies \frac{\partial}{\partial \ell_2}\tilde{R}(\sigma) > 0 \text{ and } \frac{\partial}{\partial u_2}\tilde{R}(\sigma) < 0$$
$$\frac{\partial}{\partial \ell_2}\tilde{R}(\sigma) \le 0 \text{ or } \frac{\partial}{\partial u_2}\tilde{R}(\sigma) \ge 0 \implies \frac{\partial}{\partial u_1}\tilde{R}(\sigma) > 0 \text{ or } \frac{\partial}{\partial \ell_3}\tilde{R}(\sigma) < 0$$

Then we make the following sequence of changes (forming new σ):

Subcase 1A, $\frac{\partial}{\partial u_1}\tilde{R}(\sigma) \le 0$ *and* $\frac{\partial}{\partial \ell_3}\tilde{R}(\sigma) \ge 0 \implies \frac{\partial}{\partial \ell_2}\tilde{R}(\sigma) > 0$ *and* $\frac{\partial}{\partial u_2}\tilde{R}(\sigma) < 0$*:* Increase ℓ_2 and decrease u_2 simultaneously until $\ell_2 = u_2$ (exit Case 1), $\frac{\partial}{\partial u_2}\tilde{R}(\sigma) \ge 0$, or $\frac{\partial}{\partial \ell_2}\tilde{R}(\sigma) \le 0$ (go to **1B** or **1C**).

Subcase 1B, $\frac{\partial}{\partial u_1}\tilde{R}(\sigma) > 0$*:* Increase u_1 until $u_1 = \ell_2$ (exit Case 1), or $\frac{\partial}{\partial u_1}\tilde{R}(\sigma) \le 0$ (go to **1A** or **1C**)

Subcase 1C, $\frac{\partial}{\partial \ell_3}\tilde{R}(\sigma) < 0$*:* Decrease ℓ_3 until $u_2 = \ell_3$ (exit Case 1), or $\frac{\partial}{\partial \ell_3}\tilde{R}(\sigma) \ge 0$ (go to **1B** or **1A**).

Each of these changes strictly increase $R\tilde{(\sigma)}$. **1B** and **1C** may both be true, in which case arbitrarily decide between them. At least one of the three subcases is true as long as the Case 1 condition holds. Thus, eventually:

- $\ell_2 = u_2$, in Subcase 1A (i.e. the middle interval collapses to mass 0). OR
- $u_1 = \ell_2$, in Subcase 1B (i.e. the first interval merges with the second). OR
- $u_2 = \ell_3$, in Subcase 1C (i.e. the third interval merges with the second).

Thus, this sequence of changes cannot decrease the reward, and result in there being one fewer interval then before. Case 1 can be iteratively applied until there are just two intervals $\sigma = (0, t_1) \cup (t_2, \infty)$.

Case 2: $\sigma = (0, t_1) \cup (t_2, \infty)$.

By supposition, $\tilde{R}(\sigma') > \tilde{R}(\emptyset)$ and so $\tilde{R}((0, t_1) \cup (t_2, \infty)) > \tilde{R}(\emptyset)$ for $t_1 > 0$ or $t_2 < \infty$.

Further $\tilde{R}((0, t_1) \cup (t_2, \infty))$ is a continuous function in t_1, t_2. Thus $\tilde{R}((0, t_1) \cup (t_2, \infty)) \to \tilde{R}(\emptyset)$ as $t_1 \to 0, t_2 \to \infty$ together.

Further, $\tilde{R}((0, t_1) \cup (t_2, \infty)) \to \tilde{R}((0, \infty))$ as $t_1 \to \infty$, regardless of how t_2 behaves. Similarly, fixing t_1, $\tilde{R}((0, t_1) \cup (t_2, \infty)) \to \tilde{R}((0, t_1))$ as $t_2 \to \infty$.

- If $\tilde{R}((0, t_1) \cup (t_2^*(t_1), \infty))$ is increasing for $t_1 > T_1$, for however $t_2^*(t_1)$ behaves as a function of t_1 then $\tilde{R}((0, \infty)) \geq \tilde{R}((0, t_1) \cup (t_2, \infty)), \forall t_1 > T_1, t_2$.
- For any fixed t_1, if $\tilde{R}((0, t_1) \cup (t_2, \infty))$ is increasing for $t_2 > T_2$, then $\tilde{R}((0, t_1)) \geq \tilde{R}((0, t_1) \cup (t_2, \infty)), \forall t_2$.

These limiting values eliminate the possible cases where t_1 or t_2 increasing to infinity, but the asymptotic values at ∞ produce lower rewards, which would have implied that the maximum is not achieved. Thus, either

1. $\exists t_1^* \in (0, \infty) : \tilde{R}((0, t_1^*)) \geq \tilde{R}((0, t_1) \cup (t_2, \infty)), \forall t_1, t_2$
2. $\exists t_1^*, t_2^* \in [0, \infty) : \tilde{R}((0, t_1^*) \cup (t_2^*, \infty)) \geq \tilde{R}((0, t_1) \cup (t_2, \infty)), \forall t_1, t_2$

$\square$

THEOREM 9. *Consider a reward function $\tilde{R}(\sigma)$ that maps open, measurable subsets $\sigma = \cup_k^\infty (\ell_k, u_k) \subseteq (0, \infty)$ to the non-negative reals, and probability measure F such that F is continuous, i.e. f is bounded. Further consider an open subset $\sigma' \subseteq (0, \infty)$.*

Suppose

1. $F(\sigma') > 0$
2. $\tilde{R}(\sigma)$ *is continuous in* σ.
3. *Let* $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ *denote the partial derivative of* $\tilde{R}$ *with respect to an upper end-point* u_k *of the intervals that make up* σ. *Suppose* $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ *exists, for all* σ *and its endpoints* u_k, *and is continuous in* u *for a fixed* σ.
4. *Further suppose that* $\frac{\partial}{\partial u_k}\tilde{R}(\cup_m(\ell_m, u_m))$ *is continuous in* u_k, *i.e.* $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ *is continuous in* σ.
5. *Suppose* $\exists \epsilon > 0$ *s.t.* $\frac{\partial}{\partial u}\tilde{R}(\sigma)$ *has the same sign as a function* $r(u, \sigma)$ *that is* ***strictly quasi-concave*** *in* u *(for a fixed* σ*), for all* σ *such that* $\tilde{R}(\sigma) \geq \tilde{R}(\sigma') - \epsilon$.

Then, there exists exist $\ell^*, u^* \in \mathbb{R}_+ \cup \{\infty\}$ *such that the policy* $\sigma^* = (\ell^*, u^*)$ *such that* $\tilde{R}(\sigma^*) \geq \tilde{R}(\sigma')$.

Proof.

Proof is similar to that of above theorems: start at any open subset $\sigma' \subseteq (0,\infty) = \cup_k^\infty (\ell_k, u_k)$.

Starting with this set σ', we create a sequence $\sigma'_\delta \to \sigma'$ (as $\delta \to 0$), where, for each δ, $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$, as above.

Then, we show that there exists a $\sigma^* = (0,\ell^*) \cup (u^*,\infty)$, for some $u^*, l^* \in \mathbb{R}_+$, such that $\tilde{R}(\sigma'_\delta) \leq R(\sigma^*), \forall \delta$. By continuity of the set function $\tilde{R}$, this implies that $\tilde{R}(\sigma') \leq R(\sigma^*)$.

Sequence σ'_δ. Each σ'_δ will be of the form $\sigma'_\delta = (\cup_{k=1}^K (\ell_k, u_k)) \setminus (0,L) \setminus (B,\infty)$, for some K, B, L that depend on δ. We construct a σ'_δ such that $F(\sigma' \setminus \sigma'_\delta \cup \sigma'_\delta \setminus \sigma') < \delta$ as follows:

- F is a finite (probability) measure, and so there exists K such that $F(\cup_{k=K+1}^\infty (\ell_k, u_k)) < \delta/2$. (Since $F(\sigma') \leq 1$, it follows by the Cauchy condition).
- Let $B \in \mathbb{R}$ s.t. $F((B,\infty)) < \delta/4$. Let $L \in \mathbb{R}$ s.t. $F((0,L)) < \delta/4$. Such B, L exist by condition on F.
- Set $\sigma'_\delta = (\cup_{k=1}^K (\ell_k, u_k)) \setminus (0,L) \setminus (B,\infty)$.
- For convenience, we re-index the disjoint intervals $\{\zeta_k\}_{k=1}^K$ such that they are in increasing order, i.e. $u_k > \ell_k \geq u_{k-1}, \forall k > 1$. If there exist any intervals such that $\ell_k = u_{k-1}$, replace them with the combined interval (ℓ_{k-1}, u_k).

Showing that $\exists \ell^*, u^* \in \mathbb{R}_+ \cup \{\infty\}$ **such that** $\exists \delta_0 : \forall \delta < \delta_0, \tilde{R}(\sigma'_\delta) \leq \tilde{R}((\ell^*, u^*))$.

By suppositions, $\exists \delta_0$ small enough such that $r(u,\sigma)$ is strictly quasi-concave for all σ such that $\tilde{R}\sigma \geq \tilde{R}(\sigma'_\delta)$, $\forall \delta < \delta_0$. Suppose $\delta < \delta_0$.

Now, starting at $\sigma = \sigma'_\delta$, we describe a sequence of modifications to σ, such that each modification does not reduce the reward $\tilde{R}(\sigma)$. The limit of this sequence of modifications is the policy $\sigma^* = (\ell^*, u^*)$, regardless of the starting σ'_δ. This shows that $\tilde{R}(\sigma^*) \geq \tilde{R}(\sigma'_\delta)$.

Then, let $\sigma'_\delta = \cup_{k=1}^K (\ell_k, u_k)$, where $\zeta_k \triangleq (\ell_k, u_k)$.

The key step is noting that quasi-concavity of the transformed derivative implies that any σ with two intervals ζ_1, ζ_2 can be improved by eliminating one (or joining the two).

Case 1: $\exists$ **disjoint** $\zeta_1 = (\ell_1, u_1), \zeta_2 = (\ell_2, u_2)$**, s.t.** $|\zeta_1|, |\zeta_2| > 0$, i.e. σ is composed of at least two intervals with positive mass, and ζ_1, ζ_2 are the first two such intervals.

By supposition, the transformed derivative with respect to any of the upper end-points u_k, $r(u_k, \sigma)$, is strictly quasi-concave in u. Then, the transformed derivative with respect to any of the lower end-points ℓ_k, $r(\ell_k, \sigma)$, is strictly quasi-convex in u, and further is the negative of $r(u,\sigma)$ when $u = \ell_k$.

Then, we have:

$$\frac{\partial}{\partial \ell_1}\tilde{R}(\sigma) \leq 0 \text{ and } \frac{\partial}{\partial \ell_2}\tilde{R}(\sigma) \leq 0 \implies \frac{\partial}{\partial u_1}\tilde{R}(\sigma) > 0$$
$$\frac{\partial}{\partial u_1}\tilde{R}(\sigma) \leq 0 \implies \frac{\partial}{\partial \ell_1}\tilde{R}(\sigma) > 0 \text{ or } \frac{\partial}{\partial \ell_2}\tilde{R}(\sigma) > 0$$

Then we make the following sequence of changes (forming new σ):

Subcase 1A, $\frac{\partial}{\partial \ell_1}\tilde{R}(\sigma) \leq 0$ *and* $\frac{\partial}{\partial \ell_2}\tilde{R}(\sigma) \leq 0 \Longrightarrow \frac{\partial}{\partial u_1}\tilde{R}(\sigma) > 0$*:* Increase u_1 until $u_1 = \ell_2$ (exit Case 1) or $\frac{\partial}{\partial u_1}\tilde{R}(\sigma) \leq 0$ (go to **1B** or **1C**).

Subcase 1B, $\frac{\partial}{\partial \ell_1}\tilde{R}(\sigma) > 0$*:* Increase ℓ_1 until $u_1 = \ell_1$ (exit Case 1), or $\frac{\partial}{\partial \ell_1}\tilde{R}(\sigma) \leq 0$ (go to **1A** or **1C**)

Subcase 1C, $\frac{\partial}{\partial \ell_2}\tilde{R}(\sigma) > 0$*:* Increase ℓ_2 until $u_2 = \ell_2$ (exit Case 1), or $\frac{\partial}{\partial \ell_2}\tilde{R}(\sigma) \leq 0$ (go to **1B** or **1A**).

Each of these changes strictly increase $R(\sigma)$. **1B** and **1C** may both be true, in which case arbitrarily decide between them. At least one of the three subcases is true as long as the Case 1 condition holds. Thus, eventually:

- $\ell_2 = u_1$, in Subcase 1A (i.e. the intervals combine). OR
- $u_1 = \ell_1$, in Subcase 1B (i.e. the first interval collapses to mass 0). OR
- $u_2 = \ell_2$, in Subcase 1C (i.e. the second interval collapses to mass 0).

Thus, this sequence of changes cannot decrease the reward, and result in there being one fewer interval then before. Case 1 can be iteratively applied until there is just one interval $\sigma_i = (t_1, t_2)$.

Case 2: $\sigma_i = (t_1, t_2)$**.** Similar to the same case in the previous theorem.

Further $\tilde{R}((t_1, t_2))$ is a continuous function in t_1, t_2. Thus $\tilde{R}((t_1, t_2)) \to \tilde{R}(\emptyset)$ as $t_1 \to t_2$.

Further, $\tilde{R}((t_1, t_2)) \to \tilde{R}(\emptyset)$ as $t_1 \to \infty$, regardless of how $t_2 \geq t_1$ behaves. Similarly, fixing t_1, $\tilde{R}((t_1, t_2)) \to \tilde{R}((t_1, \infty))$ as $t_2 \to \infty$.

- If $\tilde{R}((t_1, t_2^*(t_1)))$ is increasing for $t_1 > T_1$, for however $t_2^*(t_1)$ behaves as a function of t_1 then $\tilde{R}(\emptyset) \geq \tilde{R}((t_1, t_2)), \forall t_1 > T_1, t_2$.
- For any fixed t_1, if $\tilde{R}((t_1, t_2))$ is increasing for $t_2 > T_2$, then $\tilde{R}((t_1, \infty)) \geq \tilde{R}((t_1, t_2)), \forall t_2 > T_2$.

These limiting values eliminate the possible cases where t_1 or t_2 increasing to infinity, but the asymptotic values at ∞ produce lower rewards, which would have implied that the maximum is not achieved. Thus, either

1. $\tilde{R}(\emptyset) \geq \tilde{R}((t_1, t_2)), \forall t_1, t_2$
2. $\exists t_1^* \in [0, \infty) : \tilde{R}((t_1^*, \infty)) \geq \tilde{R}((t_1, t_2)), \forall t_1, t_2$
3. $\exists t_1^*, t_2^* \in [0, \infty) : \tilde{R}(t_1^*, t_2^*) \geq \tilde{R}((t_1, t_2)), \forall t_1, t_2$

$\square$