Informal and Privatized Transit: Incentives, Efficiency and Coordination

Devansh Jalota
Georgia Tech
devansh.jalota@isye.gatech.edu

Matthew Tsao
Lyft
mattxtsao@gmail.com

August 22, 2026

Abstract

Informal and privatized transit services, such as minibuses and shared auto-rickshaws, are integral to daily travel in large urban metropolises, particularly in developing nations, providing affordable commutes where formal public transport is inadequate and other options are unaffordable. A defining feature of these systems is their decentralized, market-based organization, with drivers providing service in response to rider demand and earning opportunities. While this structure helps fill critical mobility gaps, it can also generate inefficient service patterns when profit-driven driver route choices do not align with system-wide mobility goals.

We develop an analytically tractable game-theoretic framework to study incentives underlying informal and privatized transit systems with a fixed menu of routes, quantify efficiency losses from decentralized driver route choice, and design incentive mechanisms to mitigate these inefficiencies. Here, profit-maximizing informal operators (*drivers*) decide where to provide service and cost-minimizing commuters (*riders*) decide whether to use these services. Within this framework, we establish *tight* price of anarchy bounds showing that decentralized, profit-maximizing driver behavior can lead to bounded yet substantial losses in cumulative driver profit and rider demand served and that these losses can be mitigated through targeted interventions: *budget-balanced cross-subsidization*, which uses route-specific tolls/subsidies to shape driver payoffs, and *fare optimization*, which changes rider demand and driver margins through centrally regulated route-level fares. Finally, numerical experiments based on a real-world informal transit system in Nalasopara, India, reinforce these findings.

1 Introduction

Governments in many cities, particularly in developing nations, often struggle to provide adequate public transport to meet rising travel demands due to chronic fiscal constraints. Although on-demand mobility services, such as Uber and Lyft, have partially filled this gap, they remain unaffordable for most low and middle-income users in these cities. Thus, commuters face a stark mobility gap: on one side, low-frequency, overcrowded, and geographically limited public transit; on the other, high-cost individualized services that are unaffordable for daily travel. To fill this gap, informal transport (or intermediate public transport) services, such as minibuses, shared auto-rickshaws, matatus, jeepneys, and dollar vans, have become integral to daily travel, providing affordable commutes where a formal public transport system is absent or inadequate and other options are unaffordable (Cervero 2000). Operated largely by private providers, these services form

the backbone of urban mobility in many developing nation cities, often accounting for 50–100% of all trips (Conwell 2025), e.g., in Lagos, Nigeria, minibuses alone serve roughly 62% of daily commutes (Björkegren et al. 2025).

Informal or otherwise privatized fixed-route services are not limited to developing nations, with similar services filling gaps even in mature public transit systems in the developed world. In New York, for example, a range of private operators have emerged, from long-standing jitney services in outer-borough corridors (Goldwyn 2020) to Uber's recent airport shuttle offerings (Uber 2025), which pool riders along fixed routes to offer a lower cost shared alternative to individualized ride-hailing. Together, these examples highlight a broader global pattern: privatized and, often, informal shared transit systems are crucial to meeting the mobility needs of hundreds of millions of daily travelers worldwide that formal public transit leaves unmet.

A defining feature of these services that enables them to fill these mobility gaps, often at low or no fiscal cost to governments, is their decentralized, market-based organization, wherein private drivers or operators serve particular routes in response to rider demand and earning opportunities. While this decentralized structure allows these systems to extend service where formal public transit is absent, inadequate, or costly to expand, it can also produce inefficient service patterns, with operators concentrating on high-demand or lucrative corridors while low-income or peripheral neighborhoods with lower returns remain underserved. Thus, although these systems often resemble public transit operationally, with shared vehicles running along fixed routes, these profit-driven decentralized incentive patterns fundamentally shape where service is provided, creating inefficiencies that can significantly compromise system efficiency and equity (Klopp and Cavoli 2019). These inefficiencies raise a central policy question: *how large are the losses from decentralized driver route choices, and how should the incentives of informal and privatized transit operators be shaped to mitigate them?* Quantifying these losses is important for governments, driver associations, and platform-mediated shared transit providers seeking to understand the value of coordination, especially as informal and privatized transit systems are increasingly being integrated, professionalized, and corporatized, i.e., as informal operators are organized into formal businesses (SSATP 2021). At the same time, the practical policy problem is not only whether coordination is valuable, but how much of that value can be recovered through the design of incentive mechanisms.

Our Contributions: This work addresses the above question by developing a framework for analyzing the incentives underlying informal and privatized transit systems and quantifying the performance loss from decentralized driver behavior. We further study two natural monetary mechanisms to shape incentives in these systems, characterize their optimal design, and identify the limits of these mechanisms, i.e., when losses from selfish driver behavior remain unavoidable.

In Section 3, we introduce a novel game-theoretic model of a privatized informal transit system with a fixed set of routes. On these routes, cost-minimizing *riders* choose between using informal transit to complete their trip or an outside option, such as walking or staying at home, while profit-maximizing informal or privatized operators (*drivers*) decide which routes to serve. Inspired by the practical operation and global prevalence of fixed-route informal transit systems, our framework captures two key features shaping incentives and equilibrium outcomes in such settings: (i) competition among drivers for profitable routes, which leads to over-provision on some routes and under-provision in others, and (ii) rider queuing and waiting delays, a defining feature of informal transit during morning and evening peaks that shapes riders' willingness to use these services.

To capture driver competition and its effect on their route choices, we use a non-atomic congestion game in which congestion reflects the impact of service over or under-provision on rider

and driver payoffs, rather than physical road congestion. To capture rider queuing and waiting delays, we adapt Vickrey's seminal bottleneck model of peak-period congestion (Vickrey 1969) to our informal transit setting. A key departure from the standard bottleneck model is that a route's service capacity is not fixed exogenously, but emerges endogenously from drivers' profit-maximizing route choices. In this sense, the rider-side problem of our framework can be viewed as a *bottleneck model with endogenous service capacities*. By jointly capturing both driver competition and rider delays in a unified framework, our model is sufficiently rich to capture the key operational features of informal and privatized transit systems, yet tractable enough to admit closed-form solutions and isolate the core forces at play. We thus view the *formulation* of our model as a central contribution.

For this model, in Section 4, we derive price of anarchy (PoA) guarantees under both cumulative driver profit and rider welfare objectives, where rider welfare is measured by the total number of riders served by informal transit, a central policy objective in transit planning. Our PoA bounds quantify the worst-case ratio (over problem instances) between the optimal and equilibrium outcomes, thereby capturing the performance loss due to drivers' selfish route choices. In particular, we establish *tight* PoA bounds of 2 for cumulative driver profit and $1+\frac{p_{\max}}{p_{\min}}$ for rider welfare, where $p_{\min}$ and $p_{\max}$ are the minimum and maximum per-rider profits across routes. To our knowledge, these constitute the first PoA bounds for informal transit systems and demonstrate that selfish driver decisions can result in *bounded yet substantial inefficiencies* in both objectives.

To mitigate these inefficiencies, we study two mechanisms to steer informal transit operators to improved system outcomes: *cross-subsidization* (Section 5) and *fare optimization* (Section 6).

First, we study cross-subsidization, in which a central planner uses route-specific tolls or subsidies to encourage drivers to operate on particular routes. We show that for any informal transit instance, there exists a budget-balanced (i.e., requiring no net spending) cross-subsidy scheme that aligns driver incentives with any desired system objective, and can be computed in polynomial time for both cumulative driver profit and rider welfare objectives. Thus, by appropriately accounting for incentives, central planners can eliminate inefficiencies from selfish driver behavior.

Despite its efficacy, cross-subsidies can be difficult to implement in informal transit systems, as monitoring driver behavior and enforcing compliance can be challenging. We thus study fare optimization, in which route-level fares can be directly set by a central planner and which is closer to how informal and shared privatized transit systems are regulated by driver associations and governments in practice. In this setting, we define a fare-optimized PoA notion that compares two benchmarks: a centralized benchmark, where the planner jointly optimizes fares and driver allocations to maximize either objective, and an equilibrium benchmark, where the planner chooses fares but drivers choose routes selfishly in response. Unlike cross-subsidies, fare optimization has *asymmetric power* across objectives: it can recover the centralized rider welfare benchmark arbitrarily closely, but cannot improve the factor-two worst-case bound for cumulative profit. Moreover, despite the non-convexity of the joint fare-setting and driver allocation problems, we develop polynomial-time algorithms to compute the fare-optimized centralized and equilibrium benchmarks under both objectives.

We complement our theory with numerical experiments in Section 7 based on a real-world informal transit system in Nalasopara, India. Our results under the existing fares in Nalasopara show that, although real-world inefficiencies do not reach the worst-case levels given by our PoA bounds, the corresponding fractions of the optimal cumulative driver profit and rider welfare remain meaningfully bounded away from 1, implying substantial inefficiencies in practice. We further examine a setting in which a central planner can adjust route-level fares, showing that fare optimization is

an operationally viable tool for improving outcomes in informal transit systems, but its efficacy depends on how fare adjustments are paired with driver earning objectives or constraints.

The appendix contains additional related literature, proofs omitted from the main text, additional theoretical results, and numerical implementation details. While framed in the context of informal transit, our model may be applicable to other settings, such as ride-sharing, in which the forces we study, including driver competition for service and rider queuing delays, play a central role.

2 Related Literature

A large literature documents the key role of informal transit in urban mobility and its defining aspects, including institutional arrangements, regulatory challenges, and service patterns. Despite taking many forms across contexts, "informal transport is, stripped to the basics, about profit-seeking operators serving consumer demands" (Cervero 2000). Motivated by this core feature, we study incentive interactions in fixed-route informal transit, common in cities worldwide. However, unlike the largely qualitative and case study based informal transit literature rooted in urban planning and political science (Klopp 2021, Kerzhner 2022), we develop a formal framework that complements these diagnostic studies with prescriptive tools for incentive design in such systems.

By modeling the incentives of informal transit systems and the effects of selfish driver behavior on system performance, our work connects to the literature on the efficiency of privatized transportation provision. Prior work studies competition among private service providers (Rosaia 2025) and emphasizes search frictions (Brancaccio et al. 2023, Buchholz 2015), where spatial mismatch leads to matching costs or delays. In the context of informal transit, Mbonu and Eaglin (2024) document how territorial segmentation among minibus associations affects service provision, while Mittal et al. (2024) show that informal transit routes often deviate less from shortest paths than formal public transit. Closest to our work, Conwell (2025) develops a model of privatized shared transit with profit-maximizing drivers and rider queuing delays, and studies the efficiency gains from reorganizing privately operated minibuses. However, unlike the discrete stochastic queuing model of Conwell (2025), we adopt a deterministic continuum queuing framework inspired by Vickrey (1969) and leverage ideas from congestion games, enabling a closed-form equilibrium analysis.

Our work also extends the PoA literature on efficiency losses from selfish behavior (Roughgarden 2005), typically focused on a single class of strategic agents (e.g., commuters in routing games), to informal transit systems, where service provision and demand are jointly shaped by the strategic actions of both drivers and riders. Related work on informal transit systems considers two-sided incentives but relies primarily on numerical comparisons rather than PoA bounds (Chavis and Daganzo 2013, Sangveraphunsiri et al. 2022), while others focus on equilibrium computation without efficiency comparisons (Fernandez and Marcotte 1992). In contrast, we derive explicit PoA bounds, which, to our knowledge, are the first such results for informal transit systems.

A key component of our framework is Vickrey's bottleneck model (Vickrey 1969), which we adapt to study the rider-side decision of our framework, where we treat waiting time for informal transit akin to queuing delay at a bottleneck (Kraus and Yoshida 2002). Moreover, in line with prior extensions of the bottleneck model, we incorporate outside options (Arnott et al. 1993, Gonzales and Daganzo 2012, Jalota et al. 2026), allowing riders to forgo transit when queuing delays become large. Despite these similarities, unlike existing bottleneck-based models that focus exclusively on rider behavior under fixed service capacities, we embed bottleneck-style rider behavior within

a two-sided informal transit system in which service capacities and outcomes are endogenously determined by the strategic interaction between riders and profit-maximizing drivers.

By studying the allocation of drivers to serve riders in mobility systems, our work relates to the ride-sharing and shared rides literature. Queuing models have been used to study pricing in ride-hailing (Banerjee et al. 2015, 2016), typically in stochastic and stationary settings, unlike Vickrey's deterministic and non-stationary bottleneck model. Shared rides services have been studied from the lens of online stochastic optimization (Aouad and Sarıtaç 2022, Ashlagi et al. 2019) and dynamic programming (Yan et al. 2026, Dogan and Jacquillat 2026). These works treat rider arrival as exogenous, and do not model the option for riders to strategically time their entry into the market, which we incorporate in our model. These works study exclusively matching pairs of riders together, with the exception of Alonso-Mora et al. (2017), which considers higher-capacity shared rides but abstracts from strategic rider and driver behavior and pricing. Overall, these works study discrete models while we use a fluid model, though fluid approximations are often useful even in discrete models for designing algorithms with guarantees (Aouad and Sarıtaç 2022, Feng et al. 2024).

Our work is also related to the design of mechanisms to mitigate inefficiencies in decentralized mobility systems, the design of public transit systems, and the economic impacts of public transportation infrastructure, which we review in Appendix A.

3 Model

This section introduces a stylized model of informal and privatized transit with a fixed menu of routes. On these routes, *riders* seek to make trips, and service is provided by profit-seeking informal or privatized operators, whom we refer to as *minibus drivers* (or simply *drivers*). Motivated by the practical operations of these systems, our framework captures two key features shaping incentives and equilibrium outcomes: (i) driver route choice and competition for profitable routes, which endogenously determine service capacities across routes, and (ii) rider queuing and waiting delays, which influence their willingness to use informal transit.

In the following, we first introduce notation and describe the features of the informal transit environment we study (Section 3.1) and specify rider demand for informal transit as a function of driver supply (Section 3.2). Then, we define driver payoffs and our equilibrium notion in Section 3.3. Next, we introduce the performance metrics of cumulative driver profit and rider welfare and present the price of anarchy notion in Section 3.4. Finally, we discuss modeling assumptions in Section 3.5.

3.1 Preliminaries: Notation and Features of Informal Transit Environment

We consider an informal transit system during the morning or evening peak period, defined by the interval $[t_1, t_2]$, with a fixed menu of n routes. In this system, a continuum of D *minibus drivers* choose which routes to serve, where each minibus has a fixed capacity F, denoting the mass of riders that can be served in a single trip. Consistent with observed practice during peak periods, a driver serving a route operates at full capacity, picking up riders at the origin, transporting them to their destination, and then returning empty to the origin to continue service (see left of Figure 1).

Each route $i \in [n]$ is characterized by a vector $\Theta_i = (\bar{p}_i, l_i, c_i)$, where $\bar{p}_i$ is the per-trip fare paid by minibus riders, l_i is the fixed one-way minibus travel time (so the round-trip travel time is $2l_i$), and c_i is the driver's per-trip operating cost. We first study a fixed-fare setting in this section and in Sections 4 and 5, in which the route-level fares $\bar{p}_i$ are exogenously given. This captures many

informal transit systems, including Nalasopara in Section 7, where fares remain stable over short- to medium-run horizons, and allows us to focus on equilibrium behavior under decentralized driver route choice. Section 6 generalizes the model to allow fares to be optimized by a central planner.

As in Vickrey's seminal bottleneck model (Vickrey 1969), let Λ_i be the mass of riders seeking to travel on route i with desired arrival times at the destination uniformly distributed over the peak-period interval $[t_1, t_2]$, and let $\lambda_i = \frac{\Lambda_i}{t_2 - t_1}$ be the deterministic constant desired arrival rate. Then, riders on each route choose between the minibus (M) and an outside option (O), e.g., walking or a local public transit option, based on their travel costs defined in Section 3.2 (see right of Figure 1).

A central primitive of our model is the mapping from driver supply on a route to the rider demand served by minibuses, as only a fraction of riders may be served when service capacity is limited. Let $\mathbf{x} = (x_i)_{i \in [n]}$ denote the driver allocation across routes, where $x_i \geq 0$ is the mass of drivers serving route i, and let $\Omega = \{\mathbf{x} \in \mathbb{R}^n_{\geq 0} : \sum_{i \in [n]} x_i = D\}$ denote the feasible set of driver allocations. All results extend naturally to the case when $\sum_{i \in [n]} x_i \leq D$, and we focus on the equality-constrained case for analytical clarity. Given a driver allocation $\mathbf{x}$, we define $\mathbf{\Lambda}^M(\mathbf{x}) = (\Lambda_i^M(x_i))_{i \in [n]}$ as the vector of rider demands served by the minibuses, where $\Lambda_i^M(x_i)$ is the number of riders served when x_i drivers operate on that route. The remaining $\Lambda_i - \Lambda_i^M(x_i)$ riders take the outside option.

We let $\mathcal{L}$ denote the class of admissible rider demand functions $\mathbf{\Lambda}^M(\cdot)$, induced under particular modeling choices for rider costs and queuing behavior, which we specify in Section 3.2. Then, we have the following description of an instance of an informal transit system.

Definition 1 (The Informal Transit System)**.** An instance of an informal transit system is specified by the tuple $I = (n, F, \overline{\mathbf{p}}, \mathbf{l}, \mathbf{c}, t_1, t_2, D, \mathbf{\Lambda}^M)$, where n is the number of routes, F is the vehicle capacity, and $\overline{\mathbf{p}}, \mathbf{l}, \mathbf{c} \in \mathbb{R}^n$ denote the vectors of per-trip fares, travel times, and driver costs for the n routes. $[t_1, t_2]$ is the peak-period interval over which riders' desired destination arrival times are uniformly distributed, D is the total mass of minibus drivers, and $\mathbf{\Lambda}^M : \mathbb{R}^n \to \mathbb{R}^n$ specifies the rider demand for each route as a function of the driver supply on that route. Then, under a rider demand function family $\mathcal{L}$, the set of all instances of the informal transit system is given by:

$$\mathcal{I}_{\mathcal{L}} := \{(n, F, \overline{\mathbf{p}}, \mathbf{l}, \mathbf{c}, t_1, t_2, D, \mathbf{\Lambda}^M) : n \geq 1, \overline{\mathbf{p}}, \mathbf{l}, \mathbf{c} \in \mathbb{R}^n_{>0}, F, t_1, t_2, D \in \mathbb{R}_{>0}, \mathbf{\Lambda}^M \in \mathcal{L}\}.$$

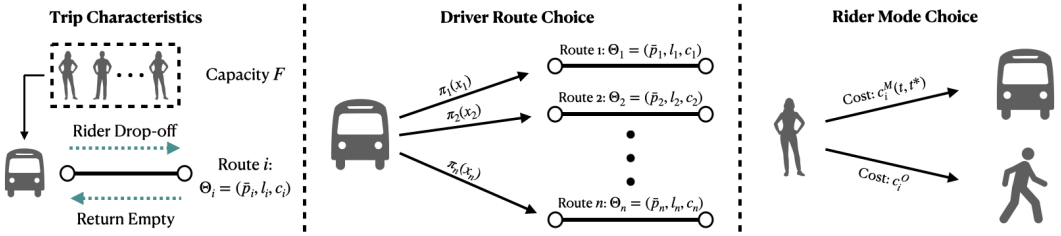


Figure 1: Depiction of the characteristics of the informal transit system on which minibus drivers service riders on a fixed menu of routes. The left panel depicts the trip characteristics, where, on each trip, a minibus transports F riders from the route's origin to its destination and then returns empty to the origin to serve additional riders. The middle panel illustrates driver route choice, with each driver selecting one of n routes to operate on based on profitability. The right panel shows the mode choice decision faced by riders who either travel by minibus or take the outside option.

3.2 Minibus Rider Demand Function

This section specifies the class $\mathcal{L}$ of admissible demand functions $\mathbf{\Lambda}^M(\cdot)$ studied in this work. We consider demand functions arising endogenously from equilibrium rider behavior, for each level of

driver supply, under an adaptation of Vickrey's bottleneck model of peak-period congestion, in which rider costs reflect queuing delays, a central feature of informal transit during morning or evening peaks. While we focus on this class of demand functions, our formulation is sufficiently general as it subsumes an important special case of capacity-constrained demand, closely related to formulations commonly studied in adjacent domains such as ride-sharing, where rider queuing delays are abstracted away (Bimpikis et al. 2019), and central to our results (see Propositions 2, 3, and Theorem 5). That said, extending our analysis to alternate demand classes arising from different modeling choices of rider costs and queuing behavior is an important direction for future work.

Demand Function Under Rider Queuing: To specify the rider demand function class we study, we first describe rider behavior under an adaptation of Vickrey's bottleneck model (Vickrey 1969, Gonzales and Daganzo 2012). On each route $i \in [n]$, cost-minimizing riders choose either an outside option (e.g., walking) with a fixed cost c_i^O or a minibus service. Given x_i drivers operating on route i, the cost of using the minibus depends on its service capacity $\mu_i(x_i) = \frac{x_i F}{2l_i}$, where F is the vehicle capacity and $2l_i$ is the round-trip travel time. In the *over-supplied* regime, when the service capacity exceeds the desired rider arrival rate, i.e., $\mu_i(x_i) \geq \lambda_i$, all riders can arrive at the destination at their desired arrival times when using the minibus without queuing or schedule delays. In this case, the cost of using the minibus is a constant given by $c_i^M = \bar{p}_i + \eta_T l_i$, where η_T represents the common value of time across riders. Without loss of generality, we assume that the minibus absent rider queuing and schedule delays is weakly more attractive than the outside option, i.e., $S_i := c_i^O - c_i^M \geq 0$ for all routes i; otherwise, such routes can be removed from consideration, as riders would always prefer the outside option regardless of minibus supply.

In the *under-supplied* regime when $\mu_i(x_i) < \lambda_i$, rider queuing and schedule delays may arise, as is common during morning or evening peaks in informal transit systems. In this case, riders, beyond choosing their mode (i.e., outside option or minibus), select their arrival time when using the minibus. Accordingly, in addition to the fixed cost c_i^M, the cost of using the minibus consists of (i) queuing (or waiting) delays $w(t)$, representing the time spent waiting for the minibus by a rider whose trip is completed at time t, and (ii) schedule delays $|t^* - t|$, capturing the discrepancy between a rider's desired destination arrival time t^* and actual arrival time t. Let η_W be the penalty for incurring a unit of waiting time delay, and η_E and η_L be the schedule delay penalties for arriving early or late, respectively, with $0 < \eta_E < \eta_W$ and $\eta_L > 0$, as is standard in the literature. Then, normalizing the waiting time penalty $\eta_W = 1$ (with all other parameters scaled accordingly), the minibus travel cost for a user with a desired destination arrival time t^* and actual arrival time t is:

$$c_i^M(t, t^*) = \bar{p}_i + \eta_T l_i + w(t) + \eta_E(t^* - t)_+ + \eta_L(t - t^*)_+. \tag{1}$$

Under this cost structure, for a given driver supply x_i on a route i, a *rider equilibrium* arises when all Λ_i riders choose a mode (minibus or outside option) and, if using the minibus, an arrival time, to minimize travel costs, with no rider having an incentive to deviate. For a fixed driver supply x_i, the resulting equilibrium is akin to that in Vickrey's bottleneck model with an outside option (Gonzales and Daganzo 2012), and as x_i varies, traces out the minibus rider demand function $\Lambda_i^M(\cdot)$ depicted on the left of Figure 2. We provide a closed-form characterization of this demand function in Equation (3) in Proposition 1 (see Section 4.1), along with a discussion of the regimes that arise. Given this characterization, we define the class $\mathcal{L}_V$ as the set of all rider demand functions in Equation (3) for arbitrary choices of strictly positive primitives $F, \overline{\mathbf{p}}, \mathbf{l}, \mathbf{c}, t_2, t_1, \eta_E, \eta_L, (\Lambda_i)_{i\in[n]}$ and non-negative values of $(S_i)_{i\in[n]}$, where, recall that $S_i = c_i^O - c_i^M$ captures the relative attractiveness

of the minibus absent queuing and schedule delays. Note that if $S_i < 0$, the rider demand is zero.

Capacity-Constrained Demand Function: We now consider an important special case of the above demand function, corresponding to $S_i = 0$ on route i. In this case, no queuing or schedule delays arise at equilibrium, as any such delays would make the minibus strictly more costly than the outside option and induce riders to switch. Hence, the demand function reduces to Equation (4) (see Section 4.2), exhibiting a piecewise-linear relationship shown on the right of Figure 2. Specifically, served demand on route i increases linearly in the mass of allocated drivers up to a threshold $k_i^* = \frac{2l_i\lambda_i}{F}$ at which $\mu_i(x_i) = \lambda_i$, beyond which the service capacity is sufficient to serve all riders.

The capacity-constrained demand function mirrors standard equilibrium relations where the served demand is limited by the minimum of the demand arrival rate and service capacity, with full demand capture occurring only once the saturation condition $\mu_i(x_i) = \lambda_i$ holds. While both demand functions in Figure 2 coincide in the over-supplied regime (i.e., $\mu_i(x_i) \geq \lambda_i$) when driver supply exceeds k_i^* and all riders are served, they diverge in the under-supplied regime. In this regime, unlike the $S_i = 0$ case corresponding to the capacity-constrained demand function, when $S_i > 0$, riders are willing to incur queuing or schedule delays, allowing the minibus to capture additional demand even when $\mu_i(x_i) < \lambda_i$. In this case, we show that the demand function is concave quadratic below a threshold $\tilde{k}_i^*$, which need not coincide with the saturation threshold at which $\mu_i(x_i) = \lambda_i$, i.e., $k_i^* \neq \tilde{k}_i^*$; in particular, there are regimes where $\mu_i(x_i) < \lambda_i$, yet all arriving riders are served by the minibus at equilibrium. For a further discussion on these demand functions, see Section 4.1.

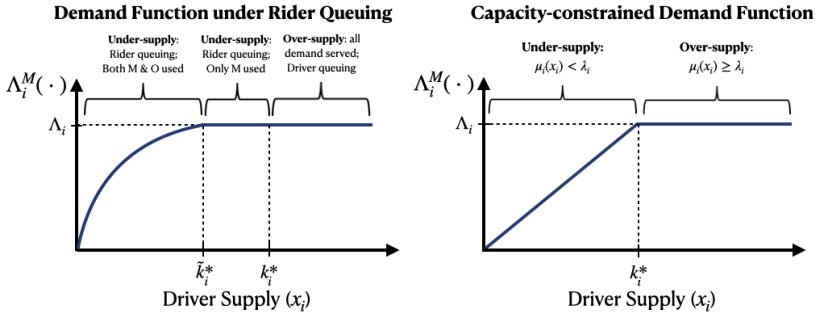


Figure 2: Depiction of the minibus demand function when incorporating rider queuing (left) and under a capacity-constrained formulation without rider queuing (right). When rider queuing is incorporated via a Vickrey-style bottleneck model, the minibus rider demand is concave quadratic below a threshold $\tilde{k}_i^* \leq k_i^*$, and remains flat thereafter. When the cost difference $S_i = 0$, this demand function reduces to the capacity-constrained formulation, where the served minibus demand increases linearly with driver supply until k_i^* at which $\mu_i(x_i) = \lambda_i$, and remains flat thereafter.

3.3 Driver Payoffs and Equilibrium Notion

We now describe the supply-side of the model and how drivers allocate themselves across routes. Motivated by the profit motive of informal transit operators, we assume drivers choose routes to maximize individual profits (see middle of Figure 1). Let $p_i = \bar{p}_i - \frac{c_i}{F} \geq 0$ denote the per-rider profit on route i, equal to the trip fare minus per-rider operating cost. Then, given x_i drivers serving $\Lambda_i^M(x_i)$ riders on route i, the total route profit is $p_i\Lambda_i^M(x_i)$, which we assume is shared equally among drivers, yielding a per-driver profit $\pi_i(x_i) = \frac{\Lambda_i^M(x_i)p_i}{x_i}$. Under the demand functions in Equation (3), $\pi_i(\cdot)$ is well-defined and bounded for $x_i \geq 0$. Drivers select routes to maximize this payoff, taking the demand functions and the allocation of other drivers as given. This strategic interaction induces

a Wardrop-style (Wardrop 1952) equilibrium driver allocation, in which no driver can profitably deviate. We formalize this notion below and let Ω^{Eq} denote the set of equilibrium driver allocations.

Definition 2 (Equilibrium Driver Allocation)**.** Consider a driver allocation $\mathbf{x}$, with $\Lambda_i^M(\cdot)$ denoting the minibus rider demand function on each route i. Then, $\mathbf{x}$ is an equilibrium driver allocation if for any route i with $x_i > 0$, it holds that $\pi_i(x_i) \geq \pi_j(x_j)$ for all routes $j \in [n]$.

3.4 Price of Anarchy

Since drivers selfishly maximize individual profits, the resulting equilibrium allocation of drivers across routes may not be optimal from the perspective of a central planner (e.g., government or driver association). In this work, we evaluate the equilibrium outcomes for an informal transit instance $I = (n, F, \overline{\mathbf{p}}, \mathbf{l}, \mathbf{c}, t_1, t_2, D, \mathbf{\Lambda}^M)$ using two metrics of direct relevance to a central planner: (i) *cumulative driver profits*, defined as $P_I(\mathbf{x}) = \sum_{i=1}^n p_i \Lambda_i^M(x_i)$, an aggregate notion of driver welfare, and (ii) *cumulative rider demand served* by the minibus system, given by $R_I(\mathbf{x}) = \sum_{i=1}^n \Lambda_i^M(x_i)$, an aggregate notion of rider welfare commonly used in transit planning, where maximizing cumulative demand served is often a central policy objective (Bertsimas et al. 2021, Yu et al. 2012).

To quantify the performance loss under these metrics due to the decentralized profit-maximizing route choices of drivers, we adopt the classical notion of the *price of anarchy* (PoA), which quantifies the worst-case ratio (over problem instances) between the equilibrium system objective and optimal system objective obtained under a centralized authority coordinating the drivers' actions.

Definition 3 (Price of Anarchy)**.** For a given family of demand functions $\mathcal{L}$, the profit PoA, denoted P-PoA($\mathcal{L}$), is defined to be the worst case profit ratio over informal transit instances (see Definition 1) of the optimum cumulative driver profit to that achieved by an equilibrium driver allocation. The rider welfare PoA, denoted R-PoA($\mathcal{L}$), is defined similarly.

$$\text{P-PoA}(\mathcal{L}) = \sup_{I \in \mathcal{I}_{\mathcal{L}}} \sup_{\mathbf{x}^* \in \Omega} \sup_{\mathbf{x}^{Eq} \in \Omega^{Eq}} \frac{P_I(\mathbf{x}^*)}{P_I(\mathbf{x}^{Eq})} \qquad \text{R-PoA}(\mathcal{L}) = \sup_{I \in \mathcal{I}_{\mathcal{L}}} \sup_{\mathbf{x}^* \in \Omega} \sup_{\mathbf{x}^{Eq} \in \Omega^{Eq}} \frac{R_I(\mathbf{x}^*)}{R_I(\mathbf{x}^{Eq})} \tag{2}$$

In the remainder of this work, we focus on the demand function class $\mathcal{L}_V$ induced under an adaptation of Vickrey's bottleneck model (see Section 3.2). Hence, for notational simplicity, we will drop the dependency on $\mathcal{L}$ and use P-PoA, R-PoA in place of P-PoA($\mathcal{L}$), R-PoA($\mathcal{L}$) respectively. Similarly, we will use P, R instead of P_I, R_I when the choice of instance I is unambiguous and clear.

3.5 Additional Discussion of Modeling Assumptions

We provide additional discussion on our modeling assumptions. First, in modeling a minibus trip (left of Figure 1), we assume that after dropping off riders, drivers return empty to the origin and continue serving the same route, reflecting the operation of informal transit systems during morning or evening peaks when demand is highly unidirectional (Björkegren et al. 2025). While we assume a round-trip travel time of $2l_i$, our results extend naturally when the two legs of the trip have asymmetric travel times. Further, consistent with practice, where over 96% of minibuses often operate at full capacity (Björkegren et al. 2025), we assume all minibuses operate at capacity F.

Next, we abstract from physical road congestion and assume fixed travel times l_i independent of the number of drivers on a route, consistent with prior work (Björkegren et al. 2025, Conwell 2025). Rather than physical road congestion, our congestion game captures *competition among*

drivers for riders, reflecting how service over or under-provision affects payoffs in informal transit systems.

Further, in line with the operational realities and institutional norms of informal transit systems, we assume that drivers commit to a single route to provide service for the entire horizon (e.g., the duration of the morning or evening peak) and do not switch routes between trips. We further show in Appendix D.1 that gains from route switching are limited at equilibria in our framework.

Finally, our assumption of a fixed menu of routes reflects the operation of many informal transit systems, where the set of routes remain stable over short to medium horizons (e.g., several months), as in the Nalasopara system studied in Section 7. Given a set of routes, our analysis focuses on the incentive and equilibrium effects in informal transit systems, where changes to route structures are often infeasible or of limited relevance due to regulatory constraints. In this sense, our work serves as a natural starting point for studying higher-level planning decisions such as route design, particularly when new informal transit systems are planned or old ones substantially redesigned.

Overall, while real-world informal and privatized transit systems are complex, our model abstracts from some operational details to isolate the core forces at play, most notably, driver competition for profitable routes and rider queuing delays (Cervero 2000), while remaining faithful to how these systems operate in practice. These abstractions enable a tractable framework that yields clear insights into the incentives and equilibrium behavior in informal transit systems, and extending the model to incorporate additional operational features is a valuable future research direction.

4 Price of Anarchy Bounds

We begin the study of our informal transit system through a price of anarchy (PoA) analysis under both cumulative driver profit and rider welfare metrics introduced in Section 3.4. To this end, under the rider behavior model in Section 3.2, we first derive the minibus rider demand as a function of driver supply and analyze its key properties in Section 4.1. Leveraging this characterization, we establish tight PoA bounds for cumulative driver profit and rider welfare in Sections 4.2 and 4.3, respectively. Our results show that decentralized, self-interested driver decisions can lead to bounded yet substantial inefficiencies in both objectives, with comparatively larger losses in rider welfare.

4.1 Derivation of Minibus Demand Function Under Rider Queuing

This section derives the rider demand served by the minibus as a function of driver supply on a route and studies its properties, which underpin our PoA analysis in Sections 4.2 and 4.3.

We first characterize, for any driver supply x_i on a route, the total demand served by the minibus at equilibrium under the costs defined in Equation (1). In presenting this result, we focus on the regime when the cost difference between the outside option and the minibus without rider queuing and schedule delays, given by $S_i = c_i^O - c_i^M$, is non-negative. Recall that when $S_i < 0$, regardless of the driver supply, the outside option is always preferred to the minibus, i.e., $\Lambda_i^M(x_i) = 0$ for all x_i.

Proposition 1 (Minibus Demand Under Rider Queuing)**.** *Consider an informal transit system with a fixed menu of routes, where riders on each route $i \in [n]$ choose either the minibus and incur a cost as defined in Equation* (1) *or an outside option with a fixed cost c_i^O. Further, suppose that $S_i = c_i^O - c_i^M \geq 0$, and define the driver allocation threshold $\tilde{k}_i^* = \min\left\{\frac{2l_i\lambda_i}{F}, \frac{2l_i\Lambda_i\eta_E\eta_L}{FS_i(\eta_E+\eta_L)}\right\}$. Then:*

$$\Lambda_i^M(x_i) = \begin{cases} \frac{F}{2l_i}\left(t_2 - t_1 + S_i\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\right)x_i - \left(\left(\frac{F}{2l_i}\right)^2 S_i \frac{\eta_E+\eta_L}{\eta_E\eta_L}\frac{t_2-t_1}{\Lambda_i}\right)x_i^2, & \text{if } x_i \in [0, \tilde{k}_i^*) \\ \Lambda_i, & \text{if } x_i \geq \tilde{k}_i^* \end{cases} \quad (3)$$

Proof. Proof Sketch. Fix a route i with $S_i \geq 0$ and x_i drivers, resulting in a service rate of $\mu_i(x_i) = \frac{x_i F}{2l_i}$. Then, defining the time horizon over which drivers service the route as $T_i(x_i)$, the total mass of riders served is: $\Lambda_i^M(x_i) = \min\{\mu_i(x_i), \lambda_i\}T_i(x_i)$. Note $T_i(x_i) \geq \Delta := t_2 - t_1$, as riders may be willing to arrive earlier or later than their desired destination arrival times under the costs in Equation (1).

To establish an expression for $T_i(x_i)$, we first characterize the equilibrium of riders' mode choice (minibus versus outside option) and arrival-time decisions on a route i with x_i drivers by adapting the equilibrium characterization in Vickrey's bottleneck model with an outside option from Gonzales and Daganzo (2012) to our setting (see Proposition 4 in Appendix B.1). Combining this equilibrium characterization with the property that rider's desired destination arrival times are uniformly distributed, we derive a closed-form expression for the total service time. When the system is under-supplied (i.e., $\mu_i(x_i) < \lambda_i$), $T_i(x_i) = \Delta + \min\{S_i, \bar{S}(x_i)\}\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\left(1 - \frac{\Delta x_i F}{2l_i\Lambda_i}\right)$, where $\bar{S}(x_i) = \frac{\Lambda_i\eta_E\eta_L}{\mu_i(x_i)(\eta_E+\eta_L)}$ (see Corollary 4 in Appendix B.1). When the system is over-supplied, $T_i(x_i) = \Delta$.

Finally, we evaluate the rider demand $\Lambda_i^M(x_i) = \mu_i(x_i)T_i(x_i)$ by substituting the derived service time relation in the different regimes for driver supply, yielding the expression in Equation (3). □

For a complete proof, see Appendix B.1. Proposition 1 implies that the rider demand function is concave quadratic up to a driver supply threshold $\tilde{k}_i^*$, beyond which all rider demand is served. The resulting demand function in Equation (3), shown on the left of Figure 2, highlights a key departure of our setting from standard equilibrium models, in which full demand capture occurs only once the saturation condition $\mu_i(x_i) = \lambda_i$ is met, as with the capacity-constrained demand function on the right of Figure 2. Specifically, Proposition 1 implies that all rider demand is served at a threshold $\tilde{k}_i^*$, which may be strictly smaller than the saturation threshold $k_i^* = \frac{2l_i\lambda_i}{F}$ at which $\mu_i(x_i) = \lambda_i$. Thus, all rider demand can be served even in an under-supplied system ($\mu_i(x_i) < \lambda_i$) if the outside option is sufficiently unattractive, albeit while inducing rider queuing. Once the driver supply exceeds the saturation threshold $k_i^* = \frac{2l_i\lambda_i}{F}$, rider queuing is eliminated at equilibrium.

We now leverage Proposition 1 to establish key continuity and monotonicity properties.

Corollary 1 (Continuity and Monotonicity of Minibus Rider Demand)**.** *The minibus rider demand function $\Lambda_i^M(\cdot)$ given by Equation* (3) *for any route i is continuous and monotonically non-decreasing in the driver allocation x_i on that route.*

Corollary 2 (Continuity and Monotonicity of Per-Driver Profits)**.** *Consider a route i with a demand function $\Lambda_i^M(\cdot)$ in Equation* (3)*. Then, the per-driver profit $\pi_i(x_i) = \frac{\Lambda_i^M(x_i)p_i}{x_i}$ is continuous and non-increasing in the driver allocation x_i on route i.*

For proofs of Corollaries 1 and 2, see Appendices B.2 and B.3, respectively. Corollary 1 implies a natural monotonicity condition that the served rider demand is non-decreasing in the driver supply, implying that the cumulative profit $p_i\Lambda_i^M(x_i)$ on a route is also non-decreasing. However, Corollary 2 shows that per-driver profit decreases in driver supply, a consequence of the concavity of the rider demand function in Equation (3) and, hence, of the profit function $p_i\Lambda_i^M(x_i)$.

4.2 Price of Anarchy for Cumulative Driver Profit

Building on the rider demand function characterization in Proposition 1 and the monotonicity properties in Corollaries 1 and 2, we now analyze the PoA in our informal transit system with respect to cumulative driver profits. We show that the profit PoA is at most 2 and that this bound is tight: for any $\varepsilon > 0$, there exists an instance $I \in \mathcal{I}_{\mathcal{L}_V}$ for which the profit ratio is at least $2 - \varepsilon$.

Theorem 1 (PoA Upper Bound for Cumulative Driver Profit)**.** *Consider an informal transit system with a class $\mathcal{L}_V$ of rider demand functions $(\Lambda_i^M(\cdot))_{i\in[n]}$ in Equation (3). Then, P-PoA ≤ 2.*

Proposition 2 (Tightness of P-PoA Bound)**.** *For any $\varepsilon > 0$, there exists an informal transit instance $I \in \mathcal{I}_{\mathcal{L}_V}$ such that the profit ratio exceeds $2 - \varepsilon$.*

These results show that in the worst case, selfish driver behavior can reduce cumulative driver profits to half of the centralized optimum, implying a substantial efficiency loss. Yet, unlike settings with unbounded PoAs (Kannan et al. 2013), the resulting loss in cumulative driver profits remains bounded due to the concavity of the demand function in Equation (3). We now prove these results.

Proof. Proof of Theorem 1. Fix an informal transit instance I with rider demand functions given by Equation (3) for all routes $i \in [n]$. Since our proof applies to all feasible instances I, we drop the dependency on I in our notation. For this instance, let $\mathbf{x}^*$ be the cumulative profit-maximizing allocation, i.e., $\mathbf{x}^* \in \arg\max_{\mathbf{x}\in\Omega} P(\mathbf{x}) = \sum_{i=1}^n p_i \Lambda_i^M(x_i)$, and let $\mathbf{x}^{Eq}$ be any equilibrium allocation. Moreover, let Q be the set of routes for which the the cost difference $S_i \geq 0$. Note for any route $i \in [n]\backslash Q$, the minibus serves no rider demand and generates no profit regardless of the allocation.

At any equilibrium $\mathbf{x}^{Eq}$, all routes i with $x_i^{Eq} > 0$ have the same per-driver profit $\pi^{Eq} = \frac{\Lambda_i^M(x_i^{Eq})p_i}{x_i^{Eq}}$. Further, by Definition 2, $\pi^{Eq} \geq \pi_j(x_j^{Eq})$ for all $j \in [n]$. Thus, the total profit under $\mathbf{x}^{Eq}$ can be re-expressed as $P(\mathbf{x}^{Eq}) = \sum_{i\in Q} p_i \Lambda_i^M(x_i^{Eq}) \overset{(a)}{=} \sum_{i\in Q} \pi_i(x_i^{Eq}) x_i^{Eq} \overset{(b)}{=} \sum_{i\in Q} \pi^{Eq} x_i^{Eq}$, where (a) follows from the definition of $\pi_i(\cdot)$ and (b) follows as $\pi^{Eq} = \pi_i(x_i^{Eq})$ for all routes i with $x_i^{Eq} > 0$. Analogously, letting $\pi_i^* = \pi_i(x_i^*)$ be the per-driver profit on each route $i \in Q$ at $\mathbf{x}^*$, the optimal cumulative profit is $P(\mathbf{x}^*) = \sum_{i\in Q} \pi_i^* x_i^*$. Then, defining $L_1 = \{i \in Q : x_i^{Eq} \geq x_i^*\}$ and $L_2 = \{i \in Q : x_i^{Eq} < x_i^*\}$:

$$P(\mathbf{x}^*) = \sum_{i\in Q} \pi_i^* x_i^* = \sum_{i\in L_1} \pi_i^* x_i^* + \sum_{i\in L_2} \pi_i^* x_i^* \overset{(a)}{\leq} \sum_{i\in L_1} \pi^{Eq} x_i^{Eq} + \sum_{i\in L_2} \pi_i^* x_i^* \overset{(b)}{\leq} P(\mathbf{x}^{Eq}) + \sum_{i\in L_2} \pi_i^* x_i^*,$$
$$\overset{(c)}{\leq} P(\mathbf{x}^{Eq}) + \sum_{i\in L_2} \pi^{Eq} x_i^* \overset{(d)}{\leq} P(\mathbf{x}^{Eq}) + P(\mathbf{x}^{Eq}) = 2P(\mathbf{x}^{Eq})$$

where (a) follows from Corollary 1, as $x_i^{Eq} \geq x_i^*$ for all $i \in L_1$, (b) follows as $\sum_{i\in L_1} \pi^{Eq} x_i^{Eq} \leq \sum_{i\in Q} \pi^{Eq} x_i^{Eq}$ as $L_1 \subseteq Q$, (c) follows as $\pi^{Eq} \geq \pi_i^*$ for all $i \in L_2$ by Corollary 2, and (d) follows as $\sum_{i\in L_2} \pi^{Eq} x_i^* \leq \sum_{i\in Q} \pi^{Eq} x_i^* = \pi^{Eq} \sum_{i\in Q} x_i^{Eq}$ as $L_2 \subseteq Q$ and $\sum_{i\in Q} x_i^* = \sum_{i\in Q} x_i^{Eq} = D$. The above analysis holds for any equilibrium allocation $\mathbf{x}^{Eq}$ and instance I, thus establishing our claim. □

Proof. Proof of Proposition 2. Consider an instance with $n = 2$ routes and a driver supply $D = 1$. Moreover, suppose that $S_i = 0$ on both routes. Then, the minibus rider demand function on both routes reduces to the capacity-constrained demand function shown on the right of Figure 2:

$$\Lambda_i^M(x_i) = \begin{cases} \frac{F(t_2-t_1)}{2l_i} x_i, & \text{if } x_i \in [0, \tilde{k}_i^*] \\ \Lambda_i, & \text{if } x_i > \tilde{k}_i^*, \end{cases} \tag{4}$$

where $\tilde{k}_i^* = k_i^* = \frac{2l_i\lambda_i}{F}$. Given Equation (4), consider an instance where $\tilde{k}_1^* + \tilde{k}_2^* = D = 1$. Then, the cumulative profit-maximizing allocation is $\mathbf{x}^* = (\tilde{k}_1^*, \tilde{k}_2^*)$, resulting in a total profit $P(\mathbf{x}^*) = p_1\Lambda_1 + p_2\Lambda_2$. Next, suppose that the per-driver profit on the two routes satisfies $\pi_1(1) = \pi_2(0)$, i.e., the system supports an equilibrium $\mathbf{x}^{Eq} = (1, 0)$. Then, the following equalities must hold: $p_1\Lambda_1 = \pi_1(1) = \pi_2(0) = p_2\frac{F(t_2-t_1)}{2l_2} = p_2\Lambda_2\frac{F}{2l_2\lambda_2} = p_2\Lambda_2\frac{1}{\tilde{k}_2^*}$. Using this relation, the profit ratio between the optimal and equilibrium allocation is given by: $\frac{P(\mathbf{x}^*)}{P(\mathbf{x}^{Eq})} = \frac{p_1\Lambda_1+p_2\Lambda_2}{p_1\Lambda_1} = 1+\frac{p_2\Lambda_2}{p_1\Lambda_1} = 1+\tilde{k}_2^*$. Finally, taking the limit as $\tilde{k}_2^* \to 1$ and $\tilde{k}_1^* \to 0$ while satisfying $\tilde{k}_1^* + \tilde{k}_2^* = D = 1$, the above analysis implies that the profit ratio approaches two, establishing our claim. □

These proofs rely on key properties of our informal transit system established in Section 4.1. The profit PoA upper bound in Theorem 1 follows from the monotonicity of the minibus demand and per-driver profits (Corollaries 1 and 2) and concavity of the demand function in Equation (3). For tightness, Proposition 2 constructs an instance with $S_i = 0$ for all routes i. In this regime, the demand function reduces to a piecewise-linear form shown on the right of Figure 2, implying that the capacity-constrained demand function corresponds to the worst-case profit PoA instance.

4.3 Price of Anarchy for Rider Welfare

This section extends the PoA analysis from the previous section to the rider welfare metric. Letting $p_{\max}$ and $p_{\min}$ denote the maximum and minimum per-rider minibus profit across the n routes, we show that the PoA with respect to cumulative rider welfare is bounded above by $1+\frac{p_{\max}}{p_{\min}}$. Moreover, we show that this bound is tight akin to Proposition 2. These results are formalized below.

Theorem 2 (PoA Upper Bound for Rider Welfare)**.** *Consider an informal transit system with a class $\mathcal{L}_V$ of rider demand functions $(\Lambda_i^M(\cdot))_{i\in[n]}$ in Equation (3). Then, R-PoA $\leq 1 + \frac{p_{\max}}{p_{\min}}$.*

Proposition 3 (Tightness of R-PoA Bound)**.** *For any $\varepsilon > 0$, there exists an informal transit instance $I \in \mathcal{I}_{\mathcal{L}_V}$ such that rider welfare ratio exceeds $1 + \frac{p_{\max}}{p_{\min}} - \varepsilon$.*

These results highlight that decentralized, selfish driver decisions can substantially reduce the cumulative rider demand served to as little as a $1/(1 + \frac{p_{\max}}{p_{\min}})$ fraction of the demand that can be served under a centralized optimum that coordinates driver actions. Since $\frac{p_{\max}}{p_{\min}} \geq 1$, in the worst case, selfish driver behavior is more detrimental to rider welfare than to cumulative driver profits (see Section 4.2). That said, in practical settings, the ratio $\frac{p_{\max}}{p_{\min}}$ is typically small, as suggested by our numerical experiments (see Section 7); for instance, using data from Nalasopara, India, this ratio equals three, resulting in a R-PoA of four. While the proof of Proposition 3 follows analogously to the proof of Proposition 2 (see Appendix B.5), we emphasize that Theorem 2 does not follow directly from the profit PoA bound in Theorem 1. In particular, a naive application of the profit PoA bound would yield a rider welfare PoA of $2\frac{p_{\max}}{p_{\min}}$. However, establishing the sharper bound in Theorem 2 requires additional arguments and a different set of inequalities (see Appendix B.4).

5 Mechanism I: Cross-subsidization

To mitigate the efficiency losses from selfish driver behavior, we study mechanisms through which a public authority can steer informal transit operators toward improved system outcomes in terms of cumulative driver profit and rider welfare. This section studies *cross-subsidization*, in which the public authority sets route-specific tolls and subsidies to influence driver route choices and shape service patterns. The key idea is that tolls can deter excessive driver entry on highly lucrative routes,

with the resulting revenue used to subsidize service on less profitable routes. Such mechanisms have practical antecedents in adjacent domains such as ride-hailing (Uber Blog 2022).

In this section, we show that for any instance of our informal transit system, there exists a budget-balanced (i.e., zero net expenditure) cross-subsidization scheme, which can be derived in closed form, that aligns individual driver incentives with any desired system objective, and can be computed in polynomial time for both cumulative driver profit and rider welfare objectives. These results demonstrate that by appropriately accounting for driver incentives, a public authority can eliminate the PoA inefficiencies identified in Section 4 without incurring any net fiscal costs.

We first introduce cross-subsidization and define the notion of a *budget-balanced* cross-subsidy scheme. In our setting with cross-subsidies, a public authority assigns each route $i \in [n]$ a transfer $\tau_i \in \mathbb{R}$, which influences driver payoffs. Negative transfers represent tolls paid by drivers to operate on a route, while positive transfers represent subsidies. Then, under a driver allocation x_i on route i, drivers earn transfer-adjusted profits $\tilde{\pi}(x_i) = \pi_i(x_i) + \tau_i$. With these modified payoffs, the equilibrium driver allocation under cross-subsidies can be defined analogously to Definition 2, with the two notions coinciding when the transfer vector $\boldsymbol{\tau} = (\tau_i)_{i\in[n]} = \mathbf{0}$. Crucially, cross-subsidies affect only driver payoffs and have no direct impact on rider payoffs or the rider demand function, with riders being influenced only through induced changes in driver allocations across routes. We focus, in particular, on *budget-balanced* cross-subsidy schemes, under which the induced equilibrium driver allocation $\mathbf{x}$ satisfies $\sum_{i=1}^{n} \tau_i x_i = 0$, so that the collected revenues finance driver subsidies.

We now characterize and establish the existence of a budget-balanced cross-subsidy scheme that induces any target driver allocation $\mathbf{x}^*$, e.g., those maximizing cumulative driver profits or rider welfare, as an equilibrium. Moreover, the corresponding cross-subsidy transfer vector yields per-driver profits that are a weighted combination of their profits at $\mathbf{x}^*$ without cross-subsidies.

Theorem 3 (Optimal Budget-Balanced Cross-Subsidization Scheme)**.** *For any informal transit instance with a class $\mathcal{L}_V$ of demand functions in Equation* (3)*, let $\mathbf{x}^* \in \Omega$ be a target driver allocation with $x_j^* > 0$ for at least one route j. Then, there exists a budget-balanced cross-subsidization scheme $\boldsymbol{\tau} = (\tau_i)_{i\in[n]}$, which induces $\mathbf{x}^*$ as an equilibrium. Moreover, $\boldsymbol{\tau}$ can be derived in closed-form and induces equilibrium per-driver profits $\tilde{\pi}^{Eq} = \frac{\sum_{i\in[n]} \pi_i^* x_i^*}{\sum_{i\in[n]} x_i^*}$, where $\pi_i^* = \pi_i(x_i^*)$.*

Proof. Proof Sketch. For each route i, define the per-driver profit at $\mathbf{x}^*$ (without cross-subsidies) as $\pi_i^* = \pi_i(x_i^*) = \frac{\Lambda_i^M(x_i^*)p_i}{x_i^*}$. To induce $\mathbf{x}^*$ as an equilibrium under a cross-subsidy scheme $\boldsymbol{\tau}$, for any route i with $x_i^* > 0$, its transfer-adjusted per-driver profits are at least that of any other route. Thus, we construct $\boldsymbol{\tau}$ that satisfies this equilibrium condition with equality for all routes, i.e., $\pi_i^* + \tau_i = \pi_j^* + \tau_j$ for all i, j. Combining these relations with the budget-balance condition $\sum_{i\in[n]} \tau_i x_i^* = 0$ yields a system of n (unique) equations with n unknowns (see Equation (8)), corresponding to the entries of $\boldsymbol{\tau}$. Solving this linear system yields an expression for $\boldsymbol{\tau}$, under which per-driver profits are $\tilde{\pi}^{Eq}$. □

For a complete proof, see Appendix B.6. Theorem 3 shows that cross-subsidies can eliminate PoA inefficiencies arising from selfish driver behavior in informal transit systems by aligning driver incentives with any target driver allocation. In this sense, cross-subsidies play a role akin to marginal-cost pricing in congestion games, which restores efficiency under selfish routing (Roughgarden 2005); however, unlike marginal-cost pricing, our cross-subsidy scheme is budget-balanced.

The constructive proof of Theorem 3 implies a natural algorithm to implement any desired driver allocation as an equilibrium. First, a central planner computes a target allocation $\mathbf{x}^*$ that

maximizes a chosen objective. It then computes the per-driver profits $\pi_i^* = \pi_i(x_i^*) = \frac{\Lambda_i^M(x_i^*)p_i}{x_i^*}$ using the demand function in Equation (3) for each route i. Finally, it sets the transfer vector $\boldsymbol{\tau}$ that induces $\mathbf{x}^*$ as an equilibrium by solving the linear system in Theorem 3's proof (see Equation (8)). This procedure implies that the efficacy of cross-subsidies is limited only by the central planner's ability to compute $\mathbf{x}^*$. As a corollary, the cumulative driver profit and rider welfare maximizing allocations can be implemented in polynomial time as equilibria via a budget-balanced cross-subsidy scheme.

Corollary 3. *For any informal transit instance with a class $\mathcal{L}_V$ of demand functions, let $\mathbf{x}^*$ be the cumulative driver profit or rider welfare-maximizing allocation. Then, there exists a budget-balanced cross-subsidy scheme, computable in polynomial time, that induces $\mathbf{x}^*$ as an equilibrium.*

The proof of this result follows by combining Theorem 3 with the fact that the rider demand functions in Equation (3) are concave quadratic and the feasible set Ω of driver allocations is convex, enabling the cumulative driver profit and rider welfare maximization problems to be solved in polynomial time. Beyond the above algorithmic and computational results, we provide additional properties of the budget-balanced cross-subsidy scheme derived in Theorem 3 in Appendix D.2, and discuss multiplicity and uniqueness of induced equilibrium outcomes in Appendices D.3 and D.4.

6 Mechanism II: Fare Optimization

Despite its efficacy, cross-subsidization can be difficult to implement in informal transit systems, where monitoring route choices and revenues across drivers is challenging and compliance is limited. This section therefore studies a more directly enforceable policy lever that is closer to how informal and shared privatized transit systems are regulated in practice: *fare optimization*. Unlike the fixed-fare model studied thus far, we now consider a setting where route-level fares can be set and regulated by a central authority, such as a transit agency or association. This focus on centralized fare setting institutions reflects the institutional reality and status quo of many informal and privatized transit markets, as individual drivers typically do not freely set fares, and, instead, route-level fares are regulated by public authorities and driver associations. For example, jeepney fares in Philippines are approved by the Land Transportation Franchising and Regulatory Board (Raymond Carl Dela Cruz 2026), trotro and shared-taxi fares in Ghana are periodically adjusted through transport unions (Baah 2025), and minibus-taxi associations in South Africa retain substantial authority over fare adjustments (Ngoepe 2026). These examples also highlight that the fare-setting objective may differ by institution, as public regulators may prioritize rider welfare and affordability, while driver unions may place greater weight on driver profitability.

Although cross-subsidies and fare adjustments are both monetary levers, they differ in a fundamental way. As noted in Section 5, cross-subsidies directly affect driver payoffs while leaving rider payoffs and the rider demand function unchanged, with riders affected only indirectly through the induced reallocation of drivers across routes. Fare adjustments, in contrast, directly affect both sides of the market, changing the margin drivers earn per rider and the cost riders incur from taking the minibus, thereby influencing both equilibrium driver allocations and realized demand.

This distinction between the policies crucially shapes the efficacy of fare optimization as a tool for mitigating inefficiencies from selfish driver behavior. Specifically, we show that, unlike cross-subsidies, fare optimization has *asymmetric power* across objectives. Comparing the best fare-optimized centralized outcome to the best fare-optimized equilibrium outcome, formalized

through a fare-optimized PoA notion introduced in Section 6.1, we show that optimized fares can recover the fare-optimal centralized benchmark for rider welfare arbitrarily closely (Section 6.2), but cannot improve the tight factor-two PoA bound for cumulative driver profit (Section 6.3). In addition to these PoA results, in Section 6.4, we develop polynomial-time algorithms to compute the fare-optimized centralized and equilibrium benchmarks under both objectives despite the bi-level and non-convex structure of the joint fare-setting and driver allocation problems.

6.1 Fare Optimization Setup

This section extends the model of an informal transit system described in Section 3 to the setting where the route-level fare vector $\bar{\mathbf{p}}$ is chosen by the central planner to optimize a desired objective, cumulative driver profit or rider welfare, and is not fixed apriori. In the following, we define the fare-dependent minibus rider demand function and the associated PoA notions to compare the fare-optimized centralized and equilibrium benchmarks under both objectives.

Price Feasibility Set and Notation: For each route $i \in [n]$, the fare $\bar{p}_i$ is chosen from the interval $\mathcal{P}_i := \left[\frac{c_i}{F},\ c_i^O - \eta_T l_i\right]$, which we assume is non-empty. The lower bound ensures non-negative per-rider profits, while the upper bound ensures that, absent queuing and schedule delays, the minibus is weakly preferred to the outside option. Let $\mathcal{P} := \prod_{i=1}^n \mathcal{P}_i$ denote the set of feasible fare vectors.

Given a fare $\bar{p}_i$, the per-rider profit on a route can be defined as $p_i(\bar{p}_i) := \bar{p}_i - \frac{c_i}{F}$ and the cost advantage of the minibus, without rider queuing and waiting delays, to the outside option is $S_i(\bar{p}_i) := c_i^O - \eta_T l_i - \bar{p}_i$. Both these quantities now depend on the chosen fares, where increasing $\bar{p}_i$ raises the driver margin $p_i(\bar{p}_i)$ but lowers the rider-side minibus cost advantage $S_i(\bar{p}_i)$.

Fare-Dependent Minibus Rider Demand: Since fares are no longer exogenously fixed, we express the minibus rider demand function, derived in Proposition 1, as a function of both the driver allocation and the fare. Specifically, the minibus rider demand on route i is:

$$\Lambda_i^M(x_i, \bar{p}_i) = \begin{cases} \frac{F}{2l_i}\left(t_2 - t_1 + S_i(\bar{p}_i)\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\right)x_i - \left(\left(\frac{F}{2l_i}\right)^2 S_i(\bar{p}_i)\frac{\eta_E+\eta_L}{\eta_E\eta_L}\frac{t_2-t_1}{\Lambda_i}\right)x_i^2, & \text{if } x_i \in [0, \tilde{k}_i^*(\bar{p}_i)) \\ \Lambda_i, & \text{if } x_i \geq \tilde{k}_i^*(\bar{p}_i) \end{cases} \tag{5}$$

which is obtained by substituting $S_i = S_i(\bar{p}_i)$ in Equation (3) . Note here that the threshold $\tilde{k}_i^*(\bar{p}_i)$ also is fare-dependent, where $\tilde{k}_i^*(\bar{p}_i) := \min\left\{\frac{2l_i\Lambda_i}{F(t_2-t_1)}, \frac{2l_i\Lambda_i\eta_E\eta_L}{FS_i(\bar{p}_i)(\eta_E+\eta_L)}\right\}$. Given a fare-dependent rider demand function, define the per-driver profit function on a route i as $\pi_i(x_i, \bar{p}_i) := \frac{p_i(\bar{p}_i)\Lambda_i^M(x_i,\bar{p}_i)}{x_i}$. Then, for any fare vector $\bar{\mathbf{p}} \in \mathcal{P}$, let $\Omega^{Eq}(\bar{\mathbf{p}}) \subseteq \Omega$ denote the set of induced equilibrium driver allocations, defined analogously to Definition 2. That is, a driver allocation $\mathbf{x} \in \Omega^{Eq}(\bar{\mathbf{p}})$ is an equilibrium if every route i with $x_i > 0$ yields weakly higher per-driver profit than any other route.

Fare-Optimized Benchmarks and PoA: We now define the fare-optimized cumulative driver profit and rider welfare benchmarks and the associated PoA notions. Throughout this section, an instance I consists of the model primitives in Definition 1, except that the route-level fare vector $\bar{\mathbf{p}}$ is not fixed and is instead chosen from $\mathcal{P}$. For any instance I with a driver allocation $\mathbf{x}$ and fare vector $\bar{\mathbf{p}} \in \mathcal{P}$, the cumulative driver profit is $P_I(\mathbf{x}, \bar{\mathbf{p}}) := \sum_{i=1}^n p_i(\bar{p}_i)\Lambda_i^M(x_i, \bar{p}_i)$ and rider welfare is $R_I(\mathbf{x}, \bar{\mathbf{p}}) := \sum_{i=1}^n \Lambda_i^M(x_i, \bar{p}_i)$. The corresponding fare-optimized centralized benchmarks are:

$$P_I^{\text{fare,opt}} := \sup_{\bar{\mathbf{p}}\in\mathcal{P}} \sup_{\mathbf{x}\in\Omega} P_I(\mathbf{x}, \bar{\mathbf{p}}), \qquad R_I^{\text{fare,opt}} := \sup_{\bar{\mathbf{p}}\in\mathcal{P}} \sup_{\mathbf{x}\in\Omega} R_I(\mathbf{x}, \bar{\mathbf{p}}).$$

Similarly, the fare-optimized equilibrium cumulative driver profit and rider welfare are:

$$P_I^{\text{fare,eq}} := \sup_{\bar{\mathbf{p}} \in \mathcal{P}} \inf_{\mathbf{x} \in \Omega^{\text{Eq}}(\bar{\mathbf{p}})} P_I(\mathbf{x}, \bar{\mathbf{p}}), \qquad R_I^{\text{fare,eq}} := \sup_{\bar{\mathbf{p}} \in \mathcal{P}} \inf_{\mathbf{x} \in \Omega^{\text{Eq}}(\bar{\mathbf{p}})} R_I(\mathbf{x}, \bar{\mathbf{p}}).$$

Here, for any fixed fare vector, the infimum evaluates the worst equilibrium outcome induced by those fares. The outer supremum over fare vectors captures a central planner's ability to choose fares to optimize a desired objective, a common feature of informal transit systems, and Section 6.4 shows that the corresponding fares under both objectives can be computed efficiently. We then define the fare-optimized PoAs by comparing the centralized and equilibrium benchmarks:

$$\text{P-PoA}^{\text{fare}} := \sup_{I \in \mathcal{I}} \frac{P_I^{\text{fare,opt}}}{P_I^{\text{fare,eq}}}, \qquad \text{R-PoA}^{\text{fare}} := \sup_{I \in \mathcal{I}} \frac{R_I^{\text{fare,opt}}}{R_I^{\text{fare,eq}}}.$$

A few comments about the above PoA notions are in order. First, they are the fare-optimization analogues of the PoA notions in Definition 3, with the difference that the planner may choose fares to optimize either objective under both the centralized and equilibrium benchmarks. Next, the fare-optimized PoAs are upper bounded by their fixed-fare counterparts, i.e., $\text{P-PoA}^{\text{fare}} \leq \text{P-PoA}$ and $\text{R-PoA}^{\text{fare}} \leq \text{R-PoA}$, as the planner can recover the fixed-fare PoAs by choosing the centralized optimum fare vector. In the remainder of this section, we characterize the fare-optimized PoA under both objectives and develop algorithms for computing the centralized and equilibrium benchmarks.

6.2 Fare-Optimized PoA for Rider Welfare

This section analyzes the fare-optimized PoA for rider welfare. Our main result is that fare optimization can essentially eliminate the rider welfare loss from decentralized driver behavior. In particular, Theorem 4 establishes that optimized fares can induce an equilibrium whose rider welfare is arbitrarily close to the fare-optimized centralized benchmark. This guarantee is constructive, and the fares needed to realize it can be computed efficiently, as discussed in Appendix C.

Theorem 4 (Fare-optimized Rider Welfare PoA)**.** *Suppose $\frac{c_i}{F} < c_i^O - \eta_T l_i$ for all routes $i \in [n]$. Then, for every instance I and every $\varepsilon > 0$, there exists a fare vector $\bar{\mathbf{p}}^\varepsilon \in \mathcal{P}$ such that the induced driver equilibrium $\mathbf{x}^{\text{Eq}}$ is unique and $R_I(\mathbf{x}^{\text{Eq}}, \bar{\mathbf{p}}^\varepsilon) \geq \frac{1}{1+\varepsilon} R_I^{\text{fare,opt}}$. Consequently,* $\text{R-PoA}^{\text{fare}} = 1$.

The proof of this result relies on the fact that the minibus rider demand is weakly decreasing in the fare, so the fare-optimized rider welfare centralized benchmark is achieved at the minimum feasible fare vector $\bar{\mathbf{p}}^{\min} := \left(\frac{c_1}{F}, \ldots, \frac{c_n}{F}\right)$. While minimum fares maximize rider welfare, they also eliminate driver margins and leave drivers indifferent across routes. Then, the key to proving Theorem 4 lies in showing that vanishingly small, route-specific fare perturbations can resolve this driver indifference and steer drivers toward a rider-welfare maximizing allocation while preserving nearly all of the centralized rider welfare. For a complete proof of Theorem 4, see Appendix B.7.

Remark 1 (Driver Reservation Wages)**.** The construction in Theorem 4 uses vanishingly small fares close to their minimum feasible levels, and therefore drives per-driver profits arbitrarily close to zero. This is useful for characterizing the fare-optimized PoA, but may be less suitable as a practical policy when drivers require a positive reservation wage $W > 0$. To that end, in Appendix C.2, we extend Theorem 4 to a setting in which both the centralized and equilibrium fare-optimized rider welfare benchmarks must satisfy the reservation wage constraint $\pi_i(x_i, \bar{p}_i) \geq W$ on every used route. This extension is technically more involved because, when $W > 0$, the welfare-maximizing fare for a fixed allocation is no longer simply the minimum feasible fare; the planner must choose the lowest

fare satisfying the reservation wage constraint, which depends on the allocation x_i. Appendix C.2 shows that, despite this tighter coupling between fares and allocations, the PoA guarantee of one from Theorem 4 continues to hold.

6.3 Fare-Optimized PoA for Cumulative Driver Profit

This section extends the PoA analysis to the cumulative driver profit metric. In contrast to the rider welfare result in Section 6.2, we show that fare optimization cannot eliminate equilibrium inefficiency for the cumulative profit objective. Theorem 5 shows that even though the planner can optimize fares, the worst-case fare-optimized profit PoA remains exactly two, matching the corresponding bound in the setting with fixed fares (see Theorem 1).

Theorem 5 (Fare-optimized Cumulative Driver Profit PoA)**.** *For every instance I, $P_I^{\text{fare,opt}} \leq 2P_I^{\text{fare,eq}}$, and for all $\varepsilon > 0$ there exists an instance I_ε with $\frac{P_{I_\varepsilon}^{\text{fare,opt}}}{P_{I_\varepsilon}^{\text{fare,eq}}} > 2 - \varepsilon$. Thus,* $\text{P-PoA}^{\text{fare}} = 2$.

For a proof of Theorem 5, see Appendix B.8. Theorem 5 shows that fare optimization has limited power for the cumulative profit objective. Unlike rider welfare, where optimized fares can essentially eliminate the inefficiencies from selfish driver behavior, fare optimization cannot improve the factor-two PoA guarantee for cumulative profit. The key feature driving these different guarantees is monotonicity: rider welfare is weakly decreasing in fares for a fixed driver allocation, whereas cumulative driver profit need not be, since higher fares increase driver margins but may reduce rider participation. Thus, fare regulation is an effective lever for mitigating welfare inefficiencies, but it cannot by itself overcome the worst-case profit inefficiency caused by selfish driver behavior. In this sense, fare optimization is strictly less powerful than cross-subsidization for profit maximization: cross-subsidies can correct driver incentives without changing rider payoffs or demand, whereas fare adjustments alter driver margins by changing the prices faced by riders, creating a tradeoff between higher driver margins and lower rider participation. That said, fare optimization can be paired with cross-subsidies to recover the centralized cumulative driver profit benchmark.

6.4 Computing Fare-Optimized Outcomes

While the fare-optimized PoAs in Sections 6.2 and 6.3 characterize the power of fares in mitigating inefficiencies from selfish driver behavior, they are operationally meaningful only if the corresponding fare-optimized outcomes can be efficiently computed. Unlike the fixed fare setting, the centralized benchmarks require jointly optimizing over fares and driver allocations, and the equilibrium benchmarks have a bi-level structure, since the planner chooses fares while driver allocations are determined by the induced equilibrium. Despite the general difficulty of solving such joint pricing-allocation and bi-level optimization problems, we exploit the structure of our model to show that the centralized and equilibrium benchmarks introduced in Section 6.1 can be computed efficiently for both rider welfare and cumulative driver profit objectives. In this section, we focus on computing the fare-optimized centralized and equilibrium benchmarks for cumulative driver profit; the corresponding computational results for rider welfare are presented in Appendix C.

Fare-Optimized Equilibrium Cumulative Driver Profit: We first show that the fare-optimized equilibrium profit benchmark $P_I^{\text{fare,eq}}$ can be computed efficiently for any feasible instance I. For each route i, define the largest per-driver profit route i can generate with $x_i \in [0, D]$ drivers when its fare is optimized as $q_i(x_i) := \max_{\bar{p}_i \in \mathcal{P}_i} \pi_i(x_i, \bar{p}_i)$. Moreover, for each candidate common per-driver profit level $\rho \geq 0$, define $u_i(\rho) := \max\{x_i \in [0, D] : q_i(x_i) \geq \rho\}$, with $u_i(\rho) = 0$ when

$\rho > q_i(0)$. Thus, $u_i(\rho)$ is the largest mass of drivers that route i can support while still achieving per-driver profit at least ρ, after optimizing its fare. We now show that for any $\varepsilon > 0$, $P_I^{\text{fare,eq}}$ can be computed in time $O(n \log^2(\frac{1}{\varepsilon}))$ by reducing the problem into a one-dimensional binary search over ρ.

Theorem 6. *Fix an instance I and define $\rho_I^* := \sup\{\rho \geq 0 : \sum_{i=1}^n u_i(\rho) \geq D\}$. Then, $P_I^{\text{fare,eq}} = D\rho_I^*$ can be computed to additive accuracy $\varepsilon > 0$ in time $O\big(n \log^2 \frac{1}{\varepsilon}\big)$.*

Fare-Optimized Centralized Cumulative Driver Profit: Next, we show that the fare-optimized centralized profit benchmark $P_I^{\text{fare,opt}}$ can also be computed efficiently for any feasible instance I. The optimal fare for each route admits a semi-closed form as a function of the allocation, due to the objective being quadratic in the fare for any fixed allocation. Hence, the profit of each route is determined completely by the number of drivers assigned to it. With centralized control of the drivers, the problem becomes optimal allocation of drivers across routes. We approximately solve the allocation problem by constructing an ε-net for the feasible set, running dynamic programming on the ε-net, and bounding suboptimality using Lipschitzness of the objective function. Crucially, the profit of one route does not depend on the prices or allocations to other routes, meaning that the state space for the dynamic program is only linear in the number of routes, rather than exponential.

Theorem 7 (Polynomial Time Computability of Fare-Optimized Centralized Profit)**.** *For every $\varepsilon > 0$, there is an algorithm running in $O(\frac{n}{\varepsilon^2})$ time that returns a feasible allocation $\hat{\mathbf{x}}$ and route fares $\hat{p}_i \in \arg\max_{\bar{p}_i \in \mathcal{P}_i} P_i(\hat{x}_i, \bar{p}_i) := p_i(\bar{p}_i)\Lambda_i^M(\hat{x}_i, \bar{p}_i)$, such that $\sum_{i=1}^n P_i(\hat{x}_i, \hat{p}_i) \geq P_I^{\text{fare,opt}} - \varepsilon$.*

For complete proofs of Theorems 6 and 7, see Appendices B.9 and B.10, respectively.

7 Numerical Experiments

This section presents numerical experiments based on a real-world case study of an informal transit system in Nalasopara, India, where shared auto-rickshaws serve nearly 90,000 riders daily. Our results under the existing route-level fares in Nalasopara's system demonstrate that, although real-world inefficiencies in cumulative driver profit and rider welfare do not reach the worst-case levels given by our PoA bounds, the profit and welfare ratios under operational data from this system remain meaningfully bounded away from 1. We further examine a setting in which the central planner can adjust route-level fares and show that fare optimization has asymmetric effects across objectives, consistent with Section 6. In particular, prioritizing rider welfare can eliminate welfare losses from decentralized route choice, but at the expense of low cumulative driver profits. Incorporating cumulative driver profit into the fare-setting objective or modest reservation wage constraints substantially improves driver earnings with only modest rider welfare losses, although equilibrium outcomes can remain meaningfully away from the centralized benchmark, as in the fixed-fare setting. In the following, we describe the methodology used to calibrate our model parameters from data on Nalasopara's system in Section 7.1 and present our results in Section 7.2. All the code for this section is available at the following link.

7.1 Experimental Setup and Model Calibration

In Nalasopara, shared auto-rickshaws serve riders on $n = 18$ routes (see Appendix E.1) connecting residential neighborhoods to a railway station. For each route i, we obtained trip times (l_i) and

fares ($\bar{p}_i$) from a local NGO, and estimated the per-trip operating costs (c_i) using expenditure statistics, which indicate that auto expenditures constitute roughly 80% of drivers' daily earnings. Accordingly, we set $c_i = 0.8F\bar{p}_i$, yielding a per-rider profit $p_i = \bar{p}_i - \frac{c_i}{F} = 0.2\bar{p}_i$, where $F = 4$ is the fixed capacity of the auto-rickshaws. We study this system during the evening peak period from 5 PM to midnight, when demand in Nalasopara is highest. Unlike typical evening peaks between 3-8 PM, Nalasopara exhibits a late-night demand surge, driven by long rail commutes as riders return home from major employment hubs and substantial nighttime commercial activity near the station. During this evening peak period, we estimate the total rider demand Λ_i for each route i using population data from the most recent Indian Census. For details of the demand calibration procedure, see Appendix E.2. The resulting estimates imply approximately 90,000 daily travelers across the eighteen routes, consistent with current estimates of local train ridership in Nalasopara after accounting for population growth since the last census (Mumbai Live Team 2025).

Finally, we calibrate the rider cost function parameters. We set the value of time to $\eta_T =$ Rs. 2.5 per minute, corresponding to an hourly wage of Rs. 150, reflecting average worker earnings in India in 2026 (ERI Economic Research Institute 2026). For schedule delay penalties, we follow the estimates in Small (1982) and set the earliness parameter to $\eta_E = 0.61$ and the lateness parameter $\eta_L = 2.4$. Finally, we calibrate the outside option cost c_i^O as the time required to walk from the origin to the destination stop on route i, using a walking speed of 1.3m/s. To keep this cost consistent with plausible fare levels, we cap it at the cost of taking the minibus, excluding waiting and schedule delays, when the route fare is 25% above its current value. This reflects the fact that riders are unlikely to use the minibus service once fares become sufficiently high, and will instead switch to an outside option (e.g., not taking the trip). While we use this specification as our baseline, our results are qualitatively robust to alternate specifications of c_i^O.

7.2 Results

We first quantify inefficiencies from decentralized driver behavior in this real-world informal transit setting under the current fares in Nalasopara. We measure these inefficiencies through driver profit and rider welfare ratios, defined as the ratio of the optimal cumulative driver profit or rider welfare to that achieved under an equilibrium allocation under the calibrated model parameters. We then turn to fare optimization, where route-level fares can be chosen by a central planner while drivers continue to choose routes selfishly in response. In this setting, we examine how optimized fares mediate the tradeoff between rider welfare and cumulative driver profit by solving fare-optimization problems for convex combinations of the two objectives, indexed by $\alpha \in [0, 1]$, where $\alpha = 0$ corresponds to rider welfare and $\alpha = 1$ corresponds to the cumulative driver profit. We also study how equilibrium rider welfare and cumulative driver profit change as the reservation wage requirement varies, which captures how minimum driver earning guarantees affect the performance of fare optimization.

Profit and Welfare Ratios: Figure 3 depicts the variation in the profit ratio, rider welfare ratio, and equilibrium per-driver profits as the number of drivers D is varied from 100 to 2,000, while holding route-level fares fixed at their current values in Nalasopara's system. We find that the profit ratio ranges between 1.1-1.2 and the rider welfare ratio ranges between 1.1-1.4 when $D \leq 1300$, a range consistent with practice. While the number of drivers is not observed in the data, it is likely to lie between 750-1250, as this corresponds to equilibrium per-driver profits of around Rs. 250 ($\approx$ \$3) in Figure 3 (right), consistent with average daily driver profits in Nalasopara. Once the number of drivers exceeds 1300, both ratios converge to one as profit-maximizing and rider welfare maximizing

allocations already absorb all rider demand; hence, additional drivers increase the corresponding objective under the equilibrium allocation without improving the optimal benchmarks.

While the empirically observed profit and rider welfare ratios do not reach the worst-case levels given by our PoA bounds, which equal 2 for profit and $1 + \frac{p_{\max}}{p_{\min}} = 4$ for rider welfare, where $\frac{p_{\max}}{p_{\min}} = 3$ in our data, they imply substantial inefficiencies in the practically relevant regime when $D \leq 1300$. Specifically, our observed profit and rider welfare ratios correspond to a 9–17% reduction in total profits and a 9–29% reduction in rider demand served. These losses translate into roughly 8,000–26,000 fewer riders served and average income losses of Rs. 22.5–42.5 per driver per day (about \$0.27–\$0.51), a substantial amount for daily-wage workers whose net income is around Rs. 250 (\$3) daily. These results highlight the significant value of cross-subsidization in mitigating the inefficiencies of selfish driver behavior in such high-stakes informal transit settings (see Section 5).

Finally, note from the right of Figure 3 that the equilibrium profit per driver is monotonically decreasing in the number of drivers. This pattern is consistent with Corollary 2 and reflects the increasing competition for riders as driver supply grows.

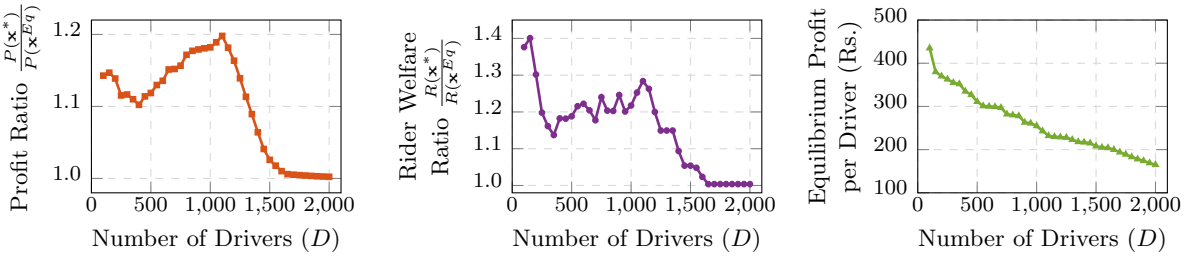


Figure 3: Depiction of the profit ratio (left), rider welfare ratio (center), and equilibrium profit per driver (right) as the number of drivers in the system is varied.

Fare optimization: While the results in Figure 3 fix route-level fares to their current values in Nalasopara, we now consider a setting in which a central planner can optimize fares. To keep the resulting fares practically meaningful, we restrict the fare on each route to be no more than 25% above its current value. In this setting, the fare-optimized profit ratio closely mirrors the fixed-fare profit ratio in Figure 3, while the fare-optimized rider welfare ratio is approximately one when the reservation wage $W = 0$ and exactly one when $W > 0$ (see Section 6.2). Thus, we focus on the tradeoff between the two objectives under fare optimization. Specifically, we optimize a convex combination of the two objectives indexed by $\alpha \in [0, 1]$, where $\alpha = 0$ corresponds to rider welfare and $\alpha = 1$ to cumulative driver profit, and compare the resulting centralized and equilibrium outcomes on both objectives. We note that for any α, the corresponding objective reduces to a special case of the cumulative driver profit objective, given by $\sum_i \tilde{p}_i(\bar{p}_i)\Lambda_i^M(x_i, \bar{p}_i)$, where $\tilde{p}_i(\bar{p}_i) = \alpha + (1 - \alpha)p_i(\bar{p}_i)$, so the computational procedures from Section 6.4 apply in computing the corresponding centralized and equilibrium benchmarks. In these experiments, we set the reservation wage under the rider welfare objective to $W =$ Rs. 250, consistent with average daily driver profits in Nalasopara.

Figure 4 depicts the profit and rider welfare ratios as functions of α for the centralized and equilibrium fare-optimized problems. The reported profit ratio is measured relative to the fare-optimized centralized profit benchmark at $\alpha = 1$, and the rider welfare ratio relative to the fare-optimized centralized welfare benchmark at $\alpha = 0$; hence, ratios close to one indicate that the corresponding single-objective benchmark is nearly attained. Our results show that moving from

the pure rider welfare objective ($\alpha = 0$) to even a small positive weight on profit substantially improves cumulative driver profit with only a small loss in rider welfare. This sharp improvement arises because the pure rider welfare objective sets fares as low as possible subject to the reservation wage constraint, maximizing ridership but leaving little margin for driver profits. Once a small profit weight is introduced, fares rise enough to improve driver margins substantially, but not so much as to deter many riders; thus, the profit ratio falls sharply while the rider welfare ratio remains close to one. In fact, under the centralized solution, even the profit-maximizing fare-optimized solution achieves nearly the rider welfare optimum, indicating that the welfare cost of incorporating driver profits is small. Moreover, as in the fixed-fare setting, the equilibrium solution remains bounded away from the centralized benchmark on both metrics, incurring roughly a 20–30% loss in cumulative driver profits and rider welfare for most α, reflecting the loss from decentralized driver route choice even when fares are optimized. Note that, at the extremes, the rider welfare ratio of the equilibrium solution is equal to one at $\alpha = 0$, while the profit ratio at $\alpha = 1$ is strictly larger than one but below two, consistent with our guarantees in Section 6.

The preceding results show that incorporating driver profits into the fare-optimization objective can improve cumulative profits with little loss in rider welfare. However, in practice, driver earnings are often protected not through the objective being optimized but through minimum earning requirements. We thus examine how the fare-optimized rider welfare equilibrium changes with the reservation wage. Figure 5 shows that rider welfare is largely insensitive to the reservation wage up to $W = \text{Rs. } 450$, indicating that moderate earning guarantees can be accommodated with little loss in ridership. Over this range, total driver profit is given by $P_I^{\text{fare,eq}} = DW$ and rises linearly with W. Beyond this point, however, the reservation wage constraint becomes costly: both rider welfare and cumulative driver profit fall because the system can no longer allocate all drivers while satisfying the higher reservation wage requirement. Thus, moderate reservation wage guarantees are compatible with high rider welfare and cumulative driver profits under optimized fares, but aggressive wage floors can become counterproductive by forcing fare and allocation changes that reduce demand and limit driver participation.

Overall, these results suggest that fare optimization is an operationally viable tool for improving outcomes in informal transit systems, but its effectiveness depends on how it is paired with driver earning objectives or constraints. When fares are optimized purely for rider welfare, the equilibrium can attain the centralized welfare benchmark, but driver profits remain low. Adding even a small weight on profit substantially improves driver earnings with little loss in rider welfare, although decentralized route choice still creates a nontrivial gap relative to the centralized benchmark. Similarly, moderate reservation wage guarantees can be accommodated under optimized fares without meaningfully reducing ridership, but aggressive earning requirements can become counterproductive by reducing realized demand and limiting driver participation. Thus, fare adjustments can help balance rider affordability and driver earnings, but are most effective when earning targets are moderate and policymakers account for the inefficiency from decentralized driver allocation.

8 Conclusion and Future Work

This work developed a framework for analyzing the incentives in informal and privatized transit systems and proposed incentive mechanisms to mitigate inefficiencies from decentralized driver behavior. We showed, through PoA bounds, that selfish driver behavior can result in bounded yet substantial losses in cumulative driver profit and rider welfare. However, these losses can be mit-

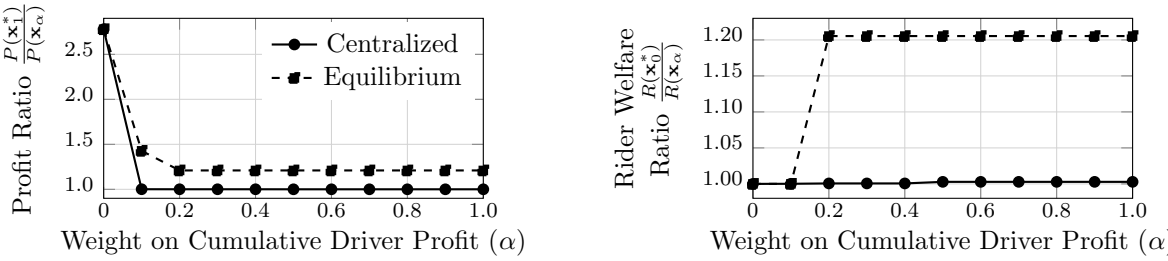


Figure 4: Fare-optimized profit and rider welfare ratios as functions of the objective weight α, where $\alpha = 0$ corresponds to rider welfare and $\alpha = 1$ to cumulative driver profit. For each $\alpha \in [0, 1]$, we optimize a convex combination of the two objectives and report the resulting ratios for both the fare-optimized centralized and equilibrium solutions $\mathbf{x}_\alpha$. Profit ratios are computed relative to the centralized profit benchmark at $\alpha = 1$, denoted $\mathbf{x}_1^*$, while rider welfare ratios are computed relative to the centralized rider welfare benchmark at $\alpha = 0$, denoted $\mathbf{x}_0^*$.

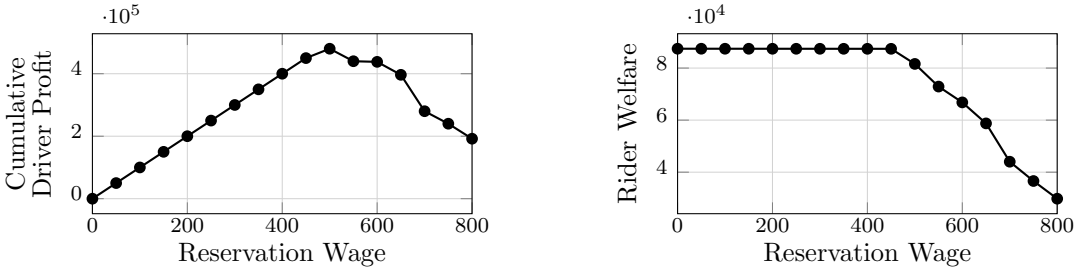


Figure 5: Cumulative driver profit and rider welfare accrued under the fare-optimized equilibrium rider welfare maximizing solution as the reservation wage W is varied.

igated through targeted interventions, including cross-subsidization and fare optimization, though their efficacy depends on the objective. While cross-subsidization can mitigate inefficiencies across objectives, fare optimization, which is easier to implement and enforce, has asymmetric power: it can recover the centralized rider welfare benchmark arbitrarily well, but cannot improve the tight PoA bound of two for cumulative driver profit. We further reinforced these findings through numerical experiments, which, in particular, highlighted that fare optimization is an operationally viable tool for mitigating inefficiencies in informal transit systems, but its efficacy depends on how rider welfare is balanced against driver earnings. Overall, our work highlights both the promise and the limits of incentive design in informal transit systems, showing how practical monetary levers can align operator incentives with system objectives while identifying when losses from decentralized behavior remain unavoidable. In doing so, we introduce a new, relatively understudied, application domain, particularly relevant in developing nation contexts, to the operations community.

Several directions merit future study. First, our model can be extended to incorporate heterogeneous driver route preferences and *physical* road congestion, where travel times depend on the number of drivers on a route. For the rider-side problem in our framework, it would be valuable to explore standard relaxations of Vickrey's bottleneck model, such as heterogeneous values of time or non-uniform desired arrival time distributions. Further, future work could examine whether the computational complexity of our algorithms for fare-optimized equilibrium and centralized benchmarks is tight or can be improved, and whether fare optimization can mitigate efficiency losses

under alternative objectives. Finally, while this paper focuses on fully privatized informal transit systems, the framework naturally extends to settings in which public and informal transit coexist. A fuller treatment of this extension, omitted here for space and to preserve the focus of the present paper, would study how public transit service and incentive mechanisms for decentralized private operators should be jointly designed to improve system-wide outcomes.

Acknowledgments

This work was supported by the Data Science Institute Postdoctoral Fellowship at Columbia University. We also thank Dr. Jacqueline Klopp for several insightful discussions and the team at the Society for the Promotion of Area Resource Centers (SPARC), including Alex Mohan Kandathil, Anreerudha Paul, Shreya Baoni, and Kshitija Akre for their support in providing the data on Nalasopara's informal transit system. This data collection was enabled through the Mumbai Living Lab as part of PRISM (Partnership for Research on Informal and Shared Mobility).

References

Robert Cervero. *Informal transport in the developing world*. UN-HABITAT, 2000.

Lucas Conwell. Privatized provision of public transit. SSRN Scholarly Paper No. 4489658, revised April 26, 2025, 2025.

Daniel Björkegren, Alice Duhaut, Geetika Nagpal, and Nick Tsivanidis. Public and private transit: Evidence from lagos. Technical report, National Bureau of Economic Research, 2025.

Eric Goldwyn. Anatomy of a new dollar van route: Informal transport and planning in new york city. *Journal of Transport Geography*, 88:102309, 2020.

Uber. Jfk airport uber shuttle: our new low-cost shared-ride option. Online, 2025. Accessed: 2025-12-05.

Jacqueline M. Klopp and Clemence Cavoli. Mapping minibuses in maputo and nairobi: engaging paratransit in transportation planning in african cities. *Transport Reviews*, 39(5):657–676, 2019.

SSATP. From complete 'formalization' encompassing bus rapid transit (BRT) to corporatization of informal public transport owner/operators: Cape town, south africa. Technical report, July 2021.

William S. Vickrey. Congestion theory and transport investment. *The American Economic Review*, 59(2): 251–260, 1969.

Jacqueline M. Klopp. Chapter seven - from "para-transit" to transit? power, politics and popular transport. In Rafael H.M. Pereira and Geneviève Boisjoly, editors, *Social Issues in Transport Planning*, volume 8 of *Advances in Transport Policy and Planning*, pages 191–209. Academic Press, 2021.

Tamara Kerzhner. Is informal transport flexible? *Journal of Transport and Land Use*, 15(1):671–689, Nov. 2022.

Nicola Rosaia. Competing platforms and transport equilibrium. *Econometrica*, 93(6):2235–2271, 2025.

Giulia Brancaccio, Myrto Kalouptsidi, Theodore Papageorgiou, and Nicola Rosaia. Search frictions and efficiency in decentralized transport markets. *The Quarterly Journal of Economics*, 138(4):2451–2503, 2023.

Nicholas Buchholz. Spatial equilibrium, search frictions and efficient regulation in the taxi industry. *Working paper*, 2015.

Oluchi Mbonu and F Christopher Eaglin. Market segmentation and coordination costs: Evidence from johannesburg's minibus networks. Technical report, Working Paper, 2024.

Kush Mohan Mittal, Marc Timme, and Malte Schröder. Efficient self-organization of informal public transport networks. *Nature Communications*, 15(1):4910, 2024.

Tim Roughgarden. *Selfish routing and the price of anarchy*. MIT press, 2005.

Celeste Chavis and Carlos F. Daganzo. Analyzing the structure of informal transit: The evening commute problem. *Research in Transportation Economics*, 39(1):277–284, 2013.

Tawit Sangveraphunsiri, Michael J. Cassidy, and Carlos F. Daganzo. Jitney-lite: a flexible-route feeder service for developing countries. *Transportation Research Part B: Methodological*, 156:1–13, 2022.

Enrique Fernandez and Patrice Marcotte. Operators-users equilibrium model in a partially regulated transit system. *Transportation Science*, 26(2):93–105, 1992.

Marvin Kraus and Yuichiro Yoshida. The commuter's time-of-use decision and optimal pricing and service in urban mass transit. *Journal of Urban Economics*, 51(1):170–195, 2002.

Richard Arnott, André de Palma, and Robin Lindsey. A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *The American Economic Review*, 83(1):161–179, 1993.

Eric J. Gonzales and Carlos F. Daganzo. Morning commute with competing modes and distributed demand: User equilibrium, system optimum, and pricing. *Transportation Research Part B: Methodological*, 46 (10):1519–1534, 2012.

Devansh Jalota, Xuan Di, and Adam N. Elmachtoub. Simple vs. optimal congestion pricing. *arXiv:2602.21495*, 2026.

Siddhartha Banerjee, Ramesh Johari, and Carlos Riquelme. Pricing in ride-sharing platforms: A queueing-theoretic approach. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC '15, page 639, 2015.

Siddhartha Banerjee, Ramesh Johari, and Carlos Riquelme. Dynamic pricing in ridesharing platforms. *SIGecom Exch.*, 15(1):65–70, September 2016.

Ali Aouad and Ömer Sarıtaç. Dynamic stochastic matching under limited time. *Operations Research*, 70(4): 2349–2383, 2022.

Itai Ashlagi, Maximilien Burq, Chinmoy Dutta, Patrick Jaillet, Amin Saberi, and Chris Sholley. Edge weighted online windowed matching. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, EC '19, page 729–742, 2019.

Chiwei Yan, Julia Yan, and Yifan Shen. Pricing shared rides. *Operations Research*, 0(0):null, 2026.

Mustafa Dogan and Alexandre Jacquillat. On-demand service sharing via collective dynamic pricing. *Manufacturing & Service Operations Management*, 28(2):400–420, 2026.

Javier Alonso-Mora, Samitha Samaranayake, Alex Wallar, Emilio Frazzoli, and Daniela Rus. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences*, 114(3):462–467, 2017.

Yiding Feng, Rad Niazadeh, and Amin Saberi. Two-stage stochastic matching and pricing with applications to ride hailing. *Operations Research*, 72(4):1574–1594, 2024.

Kostas Bimpikis, Ozan Candogan, and Daniela Saban. Spatial pricing in ride-sharing networks. *Operations Research*, 67(3):744–769, 2019.

John Glen Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1(3):325–362, 1952.

Dimitris Bertsimas, Yee Sian Ng, and Julia Yan. Data-driven transit network design at scale. *Operations Research*, 69(4):1118–1133, 2021.

Bin Yu, Zhong-Zhen Yang, Peng-Huan Jin, Shan-Hua Wu, and Bao-Zhen Yao. Transit route network design-maximizing direct and transfer demand density. *Transportation Research Part C: Emerging Technologies*, 22:58–75, 2012.

Rajgopal Kannan, Costas Busch, and Paul Spirakis. The price of anarchy is unbounded for congestion games with superpolynomial latency costs. *arXiv:1308.4101*, 2013.

Uber Blog. Introducing boost+. Online, 2022. Accessed: 2026-02-04.

Raymond Carl Dela Cruz. LTFRB Announces PUV Fare Hikes for Land Transport. Philippine News Agency, March 2026. URL https://www.pna.gov.ph/articles/1271201. Accessed May 3, 2026.

Elizabeth Larkwor Baah. Nungua Residents Commend GPRTU for Reducing Fares. Online, May 2025. Accessed May 3, 2026.

Karabo Ngoepe. Taxi Associations Evaluate Fare Adjustments amid Rising Operational Costs. The Post, April 2026. URL https://tinyurl.com/5b2ts6ta. Accessed May 3, 2026.

Mumbai Live Team. Mumbai local news: Nalasopara station gets elevated deck aimed to reduce congestion. Mumbai Live, 2025. URL https://tinyurl.com/3f5ja8k4. Accessed: 2026-01-21.

ERI Economic Research Institute. General worker salary in India. Online, 2026. Accessed: 2026-01-20.

Kenneth A. Small. The scheduling of consumer activities: Work trips. *American Economic Review*, 72(3): 467–479, 1982.

Zheng Zhu, Jintao Ke, and Hai Wang. A mean-field markov decision process model for spatial-temporal subsidies in ride-sourcing markets. *Transportation Research Part B: Methodological*, 150:540–565, 2021.

Yi-Neng Wang, Xi Lin, Fang He, Zhengtian Xu, and Zuo-Jun Max Shen. Efficiency and interventions of strategic driver relocation for ride-hailing platforms. *Available at SSRN 4454177*, 2023.

Devansh Jalota, Kiril Solovey, Karthik Gopalakrishnan, Stephen Zoepf, Hamsa Balakrishnan, and Marco Pavone. When efficiency meets equity in congestion pricing and revenue refunding schemes. *IEEE Transactions on Control of Network Systems*, 11(2):1127–1138, 2024.

Anita Schöbel. Line planning in public transportation: models and methods. *OR spectrum*, 34(3):491–510, 2012.

Gabriel Kreindler, Arya Gaduh, Tilman Graff, Rema Hanna, and Benjamin A Olken. Optimal public transportation networks: Evidence from the world's largest bus rapid transit system in jakarta. Technical report, National Bureau of Economic Research, 2023.

Siddhartha Banerjee, Chamsi Hssaine, Qi Luo, and Samitha Samaranayake. Plan your system and price for free: Fast algorithms for multimodal transit operations. *Transportation Science*, 59(1):13–27, 2025.

Noémie Périvier, Chamsi Hssaine, Samitha Samaranayake, and Siddhartha Banerjee. Real-time approximate routing for smart transit systems. *SIGMETRICS Perform. Eval. Rev.*, 49(1):73–74, June 2022.

Pascal Van Hentenryck, Connor Riley, Anthony Trasatti, Hongzhao Guan, Tejas Santanam, Jorge A Huertas, Kevin Dalmeijer, Kari Watkins, Juwon Drake, and Samson Baskin. Marta reach: Piloting an on-demand multimodal transit system in atlanta. *arXiv:2308.02681*, 2023.

Kayla Cummings, Alexandre Jacquillat, and Vikrant Vaze. Activated benders decomposition for day-ahead paratransit itinerary planning. *INFORMS Journal on Computing*, 38(1):126–149, 2026.

Yizhen Gu, Chang Jiang, Junfu Zhang, and Ben Zou. Subways and road congestion. *American Economic Journal: Applied Economics*, 13(2):83–115, April 2021.

Nick Tsivanidis. Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio. *Unpublished manuscript*, 18, 2022.

A Additional Related Work

This section surveys additional works in addition to those covered in Section 2 that are related to this work.

Our work is also related to the design of mechanisms to mitigate inefficiencies in decentralized mobility systems. The cross-subsidy mechanisms we study use route-specific tolls or subsidies to shape service patterns of informal or privatized transit drivers, akin to targeted driver subsidies in ride-sharing (Zhu et al. 2021, Wang et al. 2023). In contrast to these works, we show that a *budget-balanced* cross-subsidy mechanism, requiring no net government expenditure, can align operator incentives in the informal transit setting we study, echoing revenue-neutral congestion pricing with revenue-refunding schemes developed for traffic routing contexts (Jalota et al. 2024).

Beyond cross-subsidies, we also study fare optimization as a mechanism to mitigate inefficiencies from selfish driver behavior, relating both to congestion pricing in selfish routing and to platform pricing in two-sided mobility markets. As with congestion pricing, route-level fares can improve decentralized outcomes (Roughgarden 2005); however, unlike classical tolls, which only price agents generating congestion externalities, fares in our setting influence both rider costs and driver margins. This two-sided feature of fare adjustments connects our work to pricing models in ride-hailing platforms (Banerjee et al. 2015, 2016, Bimpikis et al. 2019). Our setting differs, however, in that drivers choose which routes to serve rather than being matched, dispatched, or repositioned by a platform. Thus, in our setting, fare optimization does not simply balance rider demand and driver supply; it also reshapes the route-level earnings that determine decentralized service patterns.

Our work also contributes to the literature on public transit design, including network and line planning (Schöbel 2012, Kreindler et al. 2023), multi-modal integration (Banerjee et al. 2025, Périvier et al. 2022), and micro-transit design (Van Hentenryck et al. 2023, Cummings et al. 2026). Whereas these works study centrally coordinated transit systems, we adopt a complementary approach by studying equilibria in decentralized transit systems with a predefined menu of routes. Finally, our work relates to studies on the economic impacts of public transportation infrastructure, which have mainly focused on formal transit systems, including subways (Gu et al. 2021) and BRT lines (Tsivanidis 2022). In contrast, our analysis centers on informal and privatized transit, highlighting how their incentives shape outcomes in settings where formal public transit is limited.

B Proofs

B.1 Proof of Proposition 1

Fix a route i with $S_i \geq 0$ and x_i minibus drivers operating on that route, resulting in a service rate of $\mu_i(x_i) = \frac{x_i F}{2l_i}$. Then, defining the total time horizon over which the drivers service the route as $T_i(x_i)$, the total mass of riders served by the minibus is given by: $\Lambda_i^M(x_i) = \min\{\mu_i(x_i), \lambda_i\}T_i(x_i)$. Note that $T_i(x_i) \geq \Delta = t_2 - t_1$, as riders may be willing to arrive earlier or later than their desired arrival times at the destination under the cost functions in Equation (1). Thus, the key to proving our result is to establish a relation for the total time $T_i(x_i)$ over which riders are serviced on a route i, given an allocation of x_i minibus drivers on that route.

Note that if the system is over-supplied with $\mu_i(x_i) \geq \lambda_i$, since $S_i \geq 0$, there will be no rider queuing or waiting delays and all riders can arrive at their destinations at their desired times by the minibus; hence, $T_i(x_i) = \Delta$ in this regime.

Thus, consider the under-supplied regime when $\mu_i(x_i) < \lambda_i$. In this regime, to derive a relation for the total service time, we first characterize riders' equilibrium mode and arrival time decisions on a given route i with x_i drivers operating on that route. We do so by adapting the associated equilibrium characterization in Vickrey's bottleneck model with an outside option from Gonzales and Daganzo (2012) to our informal transit setting. In particular, defining the threshold $\bar{S}(x_i) = \frac{\Lambda_i \eta_E \eta_L}{\mu_i(x_i)(\eta_E+\eta_L)}$ given an allocation of x_i drivers to route i, we characterize riders' equilibrium mode and arrival time decisions in two cases: (i) $S_i \geq \bar{S}(x_i)$, and (ii) $S_i \in [0, \bar{S}(x_i))$.

Case (i): In this setting, the outside option is never cost-effective relative to traveling by the minibus, even for the user experiencing the highest travel cost with a wait time of $\bar{S}(x_i)$. The resulting equilibrium waiting time profile of drivers reduces to the classical Vickrey bottleneck model without an outside option and is depicted on the left of Figure 6, with riders being serviced over the interval $[t'_A, t'_B]$. In this case, since all riders are serviced by the minibus at equilibrium, the total service time $T_i(x_i) = t'_B - t'_A = \frac{\Lambda_i}{\mu_i(x_i)}$.

Case (ii): In this regime, only a fraction of the total rider demand is served by the minibus. Adapting Proposition 1 from Gonzales and Daganzo (2012) to our informal transit setting, we obtain the following equilibrium characterization in the regime when $S_i \in [0, \bar{S}(x_i))$.

Proposition 4 (Rider Equilibrium Characterization (Gonzales and Daganzo 2012))**.** *Suppose x_i minibus drivers operate on route i with $\mu_i(x_i) < \lambda_i$, where users choose between the minibus and an outside option where the cost difference between the outside option and the minibus without rider queuing delays $S_i \in [0, \bar{S}(x_i))$. Then, assuming riders are serviced in the order of their desired destination arrival times, there exists a unique rider equilibrium satisfying (see right of Figure 6):*

- *The number of minibus riders that arrive at their destination earlier than their desired time is given by $N_E = \frac{\mu_i(x_i)S_i}{\eta_E}$ and they travel at the beginning of the rush between periods $[t_A, t_B]$.*
- *The number of minibus riders that arrive at their destination later than their desired time is given by $N_L = \frac{\mu_i(x_i)S_i}{\eta_L}$ and they travel at the end of the rush between periods $[t_C, t_D]$.*
- *The number of users that arrive exactly on time by the minibus and those that use the outside option are strictly decreasing functions of the cost difference S_i and they travel in the middle of the rush between $[t_B, t_C]$.*

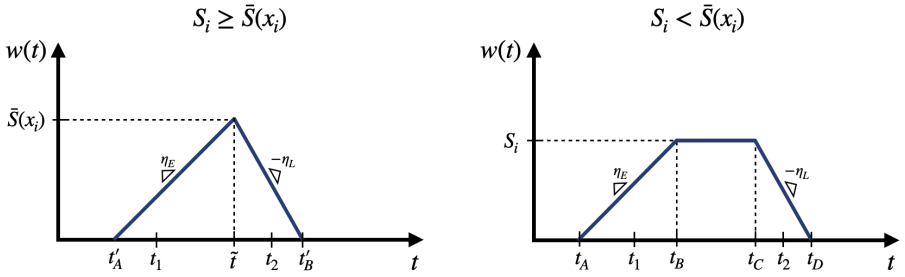


Figure 6: Depiction of the equilibrium rider waiting time profiles in the regime when the cost difference between the outside option and that of using the minibus without queuing delays satisfies $S_i \geq \bar{S}(x_i)$ (left) and $S_i < \bar{S}(x_i)$ (right).

While Gonzales and Daganzo (2012) characterize the above equilibrium for a broad class of desired destination arrival time distributions for riders, they do not provide a closed-form characterization of the total service time $T_i(x_i) = t_D - t_A$. In our setting, by focusing on uniformly

distributed desired destination arrival time distributions, which is consistent with the classical bottleneck model of Vickrey (1969) while also significantly more general than the single departure time formulations most commonly studied in the literature, we obtain a closed-form characterization of the service times in our setting.

Corollary 4 (Equilibrium Service Time)**.** *Suppose x_i minibus drivers operate on route i with $\mu_i(x_i) < \lambda_i$. Then, in the regime when the cost difference between the outside option and the minibus without rider queuing delays is non-negative, i.e., $S_i \geq 0$, the total service time $T_i(x_i) = \Delta + \min\{S_i, \bar{S}(x_i)\} \left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right) \left(1 - \frac{\Delta x_i F}{2l_i\Lambda_i}\right)$.*

Proof. Proof. First consider the regime when $S_i \geq \bar{S}(x_i)$. Then, substituting the expression for $\bar{S}(x_i)$ in the expression for the total service time in the statement of the corollary and simplifying, we obtain that $T_i(x_i) = \frac{\Lambda_i}{\mu_i(x_i)}$, consistent with the expression in the analysis of case (i) above.

Next, consider the regime when $S_i < \bar{S}(x_i)$. In this case, from the right of Figure 6, note that the service time $T_i(x_i) = t_D - t_A = (t_D - t_C) + (t_C - t_B) + (t_B - t_A)$. Next, from Proposition 4, since all riders are served by the minibus in the period $[t_A, t_B]$, it follows that $N_E = \frac{\mu_i(x_i)S_i}{\eta_E} = \mu_i(x_i)(t_B - t_A)$, which implies $t_B - t_A = \frac{S_i}{\eta_E}$. Analogously, it follows that $t_D - t_C = \frac{S_i}{\eta_L}$.

Finally, to derive the relation for $t_C - t_B$, note that the total fraction of the demand Λ_i that is served early or late during the periods $[t_A, t_B]$ and $[t_C, t_D]$ is $\frac{S_i}{\bar{S}(x_i)}$ by the linearity of the equilibrium waiting time function in Figure 6. Hence, it follows that $t_C - t_B = \left(1 - \frac{S_i}{\bar{S}(x_i)}\right)\Delta$. Then:

$$
\begin{aligned}
T_i(x_i) = t_D - t_A &= (t_D - t_C) + (t_C - t_B) + (t_B - t_A) = \frac{S_i}{\eta_L} + \frac{S_i}{\eta_E} + (t_2 - t_1)(1 - \frac{S_i}{\bar{S}(x_i)}), \\
&= \Delta + S_i \left(\frac{1}{\eta_L} + \frac{1}{\eta_E} - \frac{\Delta}{\bar{S}(x_i)}\right) = \Delta + S_i \left(\frac{\eta_E + \eta_L}{\eta_E\eta_L} - \frac{t_2 - t_1}{\frac{\Lambda_i\eta_E\eta_L}{\mu_i(x_i)(\eta_E+\eta_L)}}\right), \\
&= \Delta + S_i \left(\frac{\eta_E + \eta_L}{\eta_E\eta_L}\right) \left(1 - \frac{\Delta x_i F}{2l_i\Lambda_i}\right),
\end{aligned}
$$

where we substitute $\mu_i(x_i) = \frac{x_i F}{2l_i}$ to obtain the final equality. This establishes our total service time relation. □

Having established the relations for the total service time in both the under and over-supplied regimes, we now evaluate the minibus rider demand $\Lambda_i^M(x_i) = \min\{\mu_i(x_i), \lambda_i\}T_i(x_i)$ by substituting the derived expressions for the service time.

First consider the over-supplied regime when $\mu_i(x_i) \geq \lambda_i$, corresponding to the setting when $x_i \geq \frac{2l_i\lambda_i}{F}$. In this case, recalling that $T_i(x_i) = \Delta$, we have $\Lambda_i^M(x_i) = \min\{\mu_i(x_i), \lambda_i\}T_i(x_i) = \lambda_i\Delta = \Lambda_i$.

Next, in the under-supplied case, suppose that it further holds that $S_i \geq \bar{S}(x_i)$, corresponding to the setting when $x_i \geq \frac{2l_i\Lambda_i\eta_E\eta_L}{FS_i(\eta_E+\eta_L)}$. Then, substituting the expression for $\bar{S}(x_i)$ in the total service time expression in Corollary 4, it follows that $T_i(x_i) = \frac{\Lambda_i}{\mu_i(x_i)}$. Hence, in this regime, it again follows that $\Lambda_i^M(x_i) = \Lambda_i$. Together, the above relations imply that for any driver allocation $x_i \geq \min\{\frac{2l_i\lambda_i}{F}, \frac{2l_i\Lambda_i\eta_E\eta_L}{FS_i(\eta_E+\eta_L)}\} = \tilde{k}_i^*$, all the rider demand is served by the minibus, i.e., $\Lambda_i^M(x_i) = \Lambda_i$ for all $x_i \geq \tilde{k}_i^*$.

Finally, consider the regime when $x_i < \tilde{k}_i^*$. In this case, we leverage the relation for the total service time derived in Corollary 4 to derive the minibus rider demand as a function of the driver allocation on that route. In this regime:

$$\Lambda_i^M(x_i) = \min\{\mu_i(x_i), \lambda_i\} T_i(x_i) \overset{(a)}{=} \mu_i(x_i) T_i(x_i) \overset{(b)}{=} \frac{x_i F}{2l_i}\left(\Delta + S_i\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\left(1 - \frac{\Delta x_i F}{2l_i\Lambda_i}\right)\right),$$
$$= \frac{F}{2l_i}\left(\Delta + S_i\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\right)x_i - \left(\left(\frac{F}{2l_i}\right)^2 S_i \frac{\eta_E+\eta_L}{\eta_E\eta_L}\frac{\Delta}{\Lambda_i}\right)x_i^2$$

where (a) follows as $\mu_i(x_i) < \lambda_i$ in the regime when $x_i < \tilde{k}_i^*$ and (b) follows by substituting the relation for $\mu_i(x_i)$ and that for $T_i(x_i)$ derived in Corollary 4. This establishes our desired relation for the minibus rider demand in the regime when $x_i < \tilde{k}_i^*$, which proves our result.

B.2 Proof of Corollary 1

To establish continuity of the rider demand function, it suffices to show that $\Lambda_i^M(\tilde{k}_i^*) = \Lambda_i$. This follows directly by substituting the two candidate expressions for $\tilde{k}_i^*$, i.e., $\frac{2l_i\lambda_i}{F}$ and $\frac{2l_i\Lambda_i\eta_E\eta_L}{FS_i(\eta_E+\eta_L)}$, into the quadratic expression for $\Lambda_i^M(\cdot)$.

Next, to establish monotonicity, it suffices to show that $\Lambda_i^M(\cdot)$ is monotonically non-decreasing in the range $x_i \in [0, \tilde{k}_i^*]$. To see this, we consider the regime when the derivative of the minibus rider demand function is non-negative:

$$\frac{d\Lambda_i^M(x_i)}{dx_i} = \frac{F}{2l_i}\left(t_2 - t_1 + S_i\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\right) - 2\left(\left(\frac{F}{2l_i}\right)^2 S_i\frac{\eta_E+\eta_L}{\eta_E\eta_L}\frac{t_2-t_1}{\Lambda_i}\right) \geq 0.$$

Rearranging the above inequality and simplifying, we obtain the following condition on the driver allocation x_i for the non-negativity of the derivative of the minibus demand function $\Lambda_i^M(\cdot)$: $x_i \leq \frac{2l_i\lambda_i}{F} + \frac{2l_i\Lambda_i\eta_E\eta_L}{FS_i(\eta_E+\eta_L)}$. Since the expression on the right hand side of the this inequality is at least $\tilde{k}_i^* = \min\left\{\frac{2l_i\lambda_i}{F}, \frac{2l_i\Lambda_i\eta_E\eta_L}{FS_i(\eta_E+\eta_L)}\right\}$, it follows that $\Lambda_i^M(\cdot)$ is monotonically non-decreasing in the range $x_i \in [0, \tilde{k}_i^*]$.

B.3 Proof of Corollary 2

In the regime when $x_i > \tilde{k}_i^*$, the per-driver profit $\pi_i(x_i) = \frac{\Lambda_i p_i}{x_i}$. On the other hand, when $x_i \leq \tilde{k}_i^*$, substituting the relation for the demand function $\Lambda_i^M(\cdot)$, the per-driver profit is given by:

$$\pi_i(x_i) = \frac{\Lambda_i^M(x_i)p_i}{x_i} = p_i\frac{F}{2l_i}\left(t_2 - t_1 + S_i\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\right) - p_i\left(\left(\frac{F}{2l_i}\right)^2 S_i\frac{\eta_E+\eta_L}{\eta_E\eta_L}\frac{t_2-t_1}{\Lambda_i}\right)x_i. \quad (6)$$

From these relations and the continuity of the rider demand function $\Lambda_i^M(\cdot)$, the per-driver profit function is continuous and monotonically decreasing in the driver allocation x_i.

B.4 Proof of Theorem 2

Fix an informal transit instance I with associated rider demand functions belonging to the class $\mathcal{L}_V$, given by Equation (3) for all routes $i \in [n]$. Since our proof applies to all feasible instances I, we drop the dependency on I in our notation for the remainder of this proof.

For this instance, let $\mathbf{x}^*$ be the cumulative rider welfare allocation, i.e., $\mathbf{x}^* \in \arg\max_{\mathbf{x}\in\Omega} R(\mathbf{x}) = \sum_{i=1}^{n} \Lambda_i^M(x_i)$, and let $\mathbf{x}^{Eq}$ be any equilibrium driver allocation. Moreover, as in the proof of Theorem 1, it suffices to restrict attention to the set of routes Q for which the difference $S_i \geq 0$.

Then, defining the two sets: $L_1 = \{i \in Q : x_i^{Eq} \geq x_i^*\}$ and $L_2 = \{i \in Q : x_i^{Eq} < x_i^*\}$, we have:

$$\begin{aligned}
R(\mathbf{x}^*) &= \sum_{i\in Q} \Lambda_i^M(x_i^*) = \sum_{i\in L_1} \Lambda_i^M(x_i^*) + \sum_{i\in L_2} \Lambda_i^M(x_i^*) \overset{(a)}{\leq} \sum_{i\in L_1} \Lambda_i^M(x_i^{Eq}) + \sum_{i\in L_2} \Lambda_i^M(x_i^*), \\
&\overset{(b)}{\leq} \sum_{i\in L_1} \Lambda_i^M(x_i^{Eq}) + \frac{1}{p_{\min}} \sum_{i\in L_2} p_i \Lambda_i^M(x_i^*) \overset{(c)}{=} \sum_{i\in L_1} \Lambda_i^M(x_i^{Eq}) + \frac{1}{p_{\min}} \sum_{i\in L_2} \pi_i^* x_i^*, \\
&\overset{(d)}{\leq} \sum_{i\in L_1} \Lambda_i^M(x_i^{Eq}) + \frac{1}{p_{\min}} \pi^{Eq} \sum_{i\in L_2} x_i^* \overset{(e)}{\leq} \sum_{i\in L_1} \Lambda_i^M(x_i^{Eq}) + \frac{1}{p_{\min}} \pi^{Eq} \sum_{i\in Q} x_i^{Eq}, \\
&= \sum_{i\in L_1} \Lambda_i^M(x_i^{Eq}) + \frac{1}{p_{\min}} \sum_{i\in Q} p_i \Lambda_i^M(x_i^{Eq}) \leq \sum_{i\in L_1} \Lambda_i^M(x_i^{Eq}) + \frac{p_{\max}}{p_{\min}} \sum_{i\in Q} \Lambda_i^M(x_i^{Eq}), \\
&\overset{(f)}{\leq} \left(1 + \frac{p_{\max}}{p_{\min}}\right) R(\mathbf{x}^{Eq}),
\end{aligned}$$

where (a) follows by the monotonicity of the minibus rider demand in the driver allocation as established in Corollary 1, (b) follows as $p_{\min} \leq p_i$ for all routes i, and (c) follows by the definition of $\pi_i^* = \frac{p_i \Lambda_i^M(x_i^*)}{x_i^*}$. Moreover, (d) follows as $\pi^{Eq} \geq \pi_i^*$ for all $i \in L_2$ by the monotonicity of the per-driver profit function in Corollary 2, (e) follows as $\sum_{i\in L_2} x_i^* \leq \sum_{i\in Q} x_i^* = \sum_{i\in Q} x_i^{Eq} = D$, as the same number of drivers are allocated under the optimal and equilibrium allocations, and (f) follows as $\sum_{i\in L_1} \Lambda_i^M(x_i^{Eq}) \leq \sum_{i\in Q} \Lambda_i^M(x_i^{Eq}) = R(\mathbf{x}^{Eq})$, as $L_1 \subseteq Q$. The above analysis holds for any equilibrium allocation $\mathbf{x}^{Eq}$ and instance of our informal transit system, thus establishing our claim that the R-PoA is at most $1 + \frac{p_{\max}}{p_{\min}}$.

B.5 Proof of Proposition 3

We proceed in the vein of Proposition 2, and consider an informal transit instance with $n = 2$ routes, a normalized demand $D = 1$, and the cost difference $S_i = 0$ for both routes. Consequently, we obtain piece-wise linear rider demand functions for both routes, as in Equation (4) with $\tilde{k}_i^* = \frac{2l_i\lambda_i}{F}$ for $i \in \{1, 2\}$. Moreover, consider the setting where $\tilde{k}_1^* + \tilde{k}_2^* = D = 1$. In this case, the rider welfare maximizing allocation is $\mathbf{x}^* = (\tilde{k}_1^*, \tilde{k}_2^*)$, resulting in a total maximum achievable rider demand served of $R(\mathbf{x}^*) = \Lambda_1^M(\tilde{k}_1^*) + \Lambda_2^M(\tilde{k}_2^*) = \Lambda_1 + \Lambda_2$.

Next, suppose that profit per driver on the two routes satisfies $\pi_1(1) = \pi_2(0)$, i.e., the informal transit system supports an equilibrium driver allocation $\mathbf{x}^{Eq} = (1, 0)$. For this condition to hold, the following equalities must be satisfied:

$$p_1\Lambda_1 = \pi_1(1) = \pi_2(0) = p_2 \frac{F(t_2 - t_1)}{2l_2} = p_2\Lambda_2 \frac{F}{2l_2\lambda_2} = p_2\Lambda_2 \frac{1}{\tilde{k}_2^*}$$

Using this relation, the minibus rider demand ratio between the optimal and above-defined equilibrium allocation is:

$$\frac{R(\mathbf{x}^*)}{R(\mathbf{x}^{Eq})} = \frac{\Lambda_1 + \Lambda_2}{\Lambda_1} = 1 + \frac{\Lambda_2}{\Lambda_1} = 1 + \tilde{k}_2^* \frac{p_1}{p_2}.$$

Finally, taking the limit as $\tilde{k}_2^* \to 1$ and $\tilde{k}_1^* \to 0$ while satisfying $\tilde{k}_1^* + \tilde{k}_2^* = D = 1$, the above analysis implies that the minibus rider demand ratio approaches $1 + \frac{p_{\max}}{p_{\min}}$, thus establishing our claim.

B.6 Proof of Theorem 3

For brevity of notation, we define the per-driver profit on route i (without cross-subsidies) at the target allocation $\mathbf{x}^*$ as $\pi_i^* = \pi_i(x_i^*) = \frac{\Lambda_i^M(x_i^*)p_i}{x_i^*}$, where the rider demand function is given by Equation (3). Then, to induce $\mathbf{x}^*$ as an equilibrium driver allocation under a cross-subsidy scheme defined by $\boldsymbol{\tau}$, the following equilibrium condition must hold: for any route i with $x_i^* > 0$, the per-driver profits under cross-subsidy transfers is the at least that of any other route j, i.e.,

$$\tilde{\pi}_i(x_i^*) \geq \tilde{\pi}_j(x_i^*) \quad \implies \quad \pi_i^* + \tau_i \geq \pi_j^* + \tau_j, \tag{7}$$

where the inequality is met with an equality for all routes j with $x_j^* > 0$.

In the remainder of this proof, we construct a transfer vector $\boldsymbol{\tau}$ that satisfies the above equilibrium condition with equality for all routes, i.e., $\pi_i^* + \tau_i = \pi_j^* + \tau_j$ for all routes i, j, while satisfying budget balance, i.e., $\sum_{i\in[n]} \tau_i x_i^* = 0$. Thus, we seek to satisfy the following n (unique) equations with n unknowns, corresponding to the entries of the transfer vector $\boldsymbol{\tau}$: $\pi_1^* + \tau_1 = \pi_2^* + \tau_2$, $\pi_1^* + \tau_1 = \pi_3^* + \tau_3$, $\dots$ $\pi_1^* + \tau_1 = \pi_n^* + \tau_n$, $\sum_{i=1}^n \tau_i x_i^* = 0$. Rearranging the above equations, our goal is to find a vector $\boldsymbol{\tau}$ that solves the following system:

$$\begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \dots & 0 \\ & & & \vdots & \\ 1 & 0 & 0 & \dots & -1 \\ x_1^* & x_2^* & x_3^* & \dots & x_n^* \end{bmatrix} \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \vdots \\ \tau_n \end{bmatrix} = \begin{bmatrix} \pi_2^* - \pi_1^* \\ \pi_3^* - \pi_1^* \\ \vdots \\ \pi_n^* - \pi_1^* \\ 0 \end{bmatrix} \tag{8}$$

Note that the above matrix is full-rank and thus admits a unique solution if there is any route j such that $x_j^* > 0$. Then, note from the budget-balance constraint that:

$$\begin{aligned} 0 = \sum_{i=1}^n \tau_i x_i^* &= \sum_{i\in[n]:x_i^*>0} \tau_i x_i^* = \tau_j x_j^* + \sum_{i\in[n]:x_i^*>0, i\neq j} \tau_i x_i^* \overset{(a)}{=} \tau_j x_j^* + \sum_{i\in[n]:x_i^*>0, i\neq j} (\pi_j^* + \tau_j - \pi_i^*) x_i^* \\ &= \sum_{i\in[n]:x_i^*>0, i\neq j} (\pi_j^* - \pi_i^*) x_i^* + \tau_j \sum_{i\in[n]:x_i^*>0} x_i^*, \end{aligned}$$

where (a) follows from the equation $\pi_i^* + \tau_i = \pi_j^* + \tau_j$ for all pairs of routes i, j. Rearranging the above expression, we obtain $\tau_j = \frac{\sum_{i\in[n]:x_i^*>0, i\neq j} (\pi_i^* - \pi_j^*) x_i^*}{\sum_{i\in[n]:x_i^*>0} x_i^*}$. Next, for the remaining $i' \neq j$, we have:

$$\tau_{i'} = \pi_j^* - \pi_{i'}^* + \tau_j = \pi_j^* - \pi_{i'}^* + \frac{\sum_{i\in[n]:x_i^*>0, i\neq j} (\pi_i^* - \pi_j^*) x_i^*}{\sum_{i\in[n]:x_i^*>0} x_i^*}.$$

The resulting per-driver profit under these tolls, which is fixed across routes is given by:

$$\tilde{\pi}^{Eq} = \tilde{\pi}_j(x_j^*) = \pi_j^* + \tau_j = \pi_j^* + \frac{\sum_{i\in[n]:x_i^*>0, i\neq j} (\pi_i^* - \pi_j^*) x_i^*}{\sum_{i\in[n]:x_i^*>0} x_i^*} = \frac{\sum_{i\in[n]:x_i^*>0} \pi_i^* x_i^*}{\sum_{i\in[n]:x_i^*>0} x_i^*} = \frac{\sum_{i\in[n]} \pi_i^* x_i^*}{\sum_{i\in[n]} x_i^*},$$

which establishes our claim.

B.7 Proof of Theorem 4

To prove this claim, we first show the following lemma, which establishes the monotonicity of the minibus rider demand $\Lambda_i^M(x_i, \bar{p}_i)$ in the fare for any fixed driver allocation $x_i \geq 0$.

Lemma 1 (Monotonicity of Rider Demand in Route Fare)**.** *Fix a route i and a driver allocation $x_i \geq 0$. Then, the rider demand $\Lambda_i^M(x_i, \bar{p}_i)$ is weakly decreasing in the fare $\bar{p}_i$ on $\left[\frac{c_i}{F},\, c_i^O - \eta_T l_i\right]$.*

Proof. Proof. Fix route i and allocation $x_i \geq 0$, and let $\Delta := t_2 - t_1$. Moreover, we define $S_i(\bar{p}_i) := c_i^O - \eta_T l_i - \bar{p}_i$, so $S_i(\bar{p}_i) \geq 0$ on the feasible fare interval and $S_i(\bar{p}_i)$ is strictly decreasing in $\bar{p}_i$. Then, we have the following relation for the rider demand served by the minibus:

$$\Lambda_i^M(x_i, \bar{p}_i) = \begin{cases} \frac{F}{2l_i}\left(\Delta + \frac{\eta_E+\eta_L}{\eta_E\eta_L} S_i(\bar{p}_i)\right) x_i - \left(\frac{F}{2l_i}\right)^2 S_i(\bar{p}_i) \frac{\eta_E+\eta_L}{\eta_E\eta_L} \frac{\Delta}{\Lambda_i} x_i^2, & \text{if } x_i \in [0, \tilde{k}_i^*(\bar{p}_i)), \\ \Lambda_i, & \text{if } x_i \geq \tilde{k}_i^*(\bar{p}_i), \end{cases}$$

where $\tilde{k}_i^*(\bar{p}_i) = \min\left\{k_i^*,\, \frac{2l_i\Lambda_i\eta_E\eta_L}{F\, S_i(\bar{p}_i)(\eta_E+\eta_L)}\right\}$, and $k_i^* := \frac{2l_i\lambda_i}{F} = \frac{2l_i\Lambda_i}{F\Delta}$.

For the quadratic branch of the rider demand function, it is convenient to rewrite it as:

$$\Lambda_i^M(x_i, \bar{p}_i) = A_i(x_i) + B_i(x_i) S_i(\bar{p}_i), \qquad \text{for } x_i \in [0, \tilde{k}_i^*(\bar{p}_i)), \tag{9}$$

where $A_i(x_i) := \frac{F\Delta}{2l_i} x_i$ and $B_i(x_i) = \frac{\eta_E+\eta_L}{\eta_E\eta_L} \frac{F}{2l_i} x_i \left(1 - \frac{F\Delta}{2l_i\Lambda_i} x_i\right)$.

In the quadratic branch of the rider demand function, we have $x_i < \tilde{k}_i^*(\bar{p}_i) \leq k_i^* = \frac{2l_i\Lambda_i}{F\Delta}$; hence, $1 - \frac{F\Delta}{2l_i\Lambda_i} x_i \geq 0$, which implies that $B_i(x_i) \geq 0$. It then follows from (9) and the fact that $S_i(\bar{p}_i)$ is decreasing in $\bar{p}_i$ that, in the quadratic branch, $\Lambda_i^M(x_i, \bar{p}_i)$ is weakly decreasing in $\bar{p}_i$.

We now establish our monotonicity result for the two regimes corresponding to $\tilde{k}_i^*(\bar{p}_i)$.
Case 1: $\tilde{k}_i^*(\bar{p}_i) = k_i^*$. In this case, the threshold $\tilde{k}_i^*(\bar{p}_i)$ is independent of the fare $\bar{p}_i$. In this case, if $x_i < k_i^*$, then the rider demand is in the quadratic branch and, by the argument above, $\Lambda_i^M(x_i, \bar{p}_i)$ weakly decreases with $\bar{p}_i$. If $x_i \geq k_i^*$, then the non-quadratic branch applies and $\Lambda_i^M(x_i, \bar{p}_i) = \Lambda_i$, which is constant in $\bar{p}_i$.
Case 2: $\tilde{k}_i^*(\bar{p}_i) = \dfrac{2l_i\Lambda_i\eta_E\eta_L}{F\, S_i(\bar{p}_i)(\eta_E+\eta_L)}$. In this regime, since $S_i(\bar{p}_i)$ is decreasing in $\bar{p}_i$, the threshold $\tilde{k}_i^*(\bar{p}_i)$ is weakly increasing in $\bar{p}_i$ (until it possibly reaches the cap k_i^*). Let $\bar{p}_i' \geq \bar{p}_i$.

If $x_i < \tilde{k}_i^*(\bar{p}_i)$, then also $x_i < \tilde{k}_i^*(\bar{p}_i')$, so both rider demands lie in the quadratic branch under these prices. Since $B_i(x_i) \geq 0$ and $S_i(\bar{p}_i') \leq S_i(\bar{p}_i)$, we obtain

$$\Lambda_i^M(x_i, \bar{p}_i') = A_i(x_i) + B_i(x_i) S_i(\bar{p}_i') \leq A_i(x_i) + B_i(x_i) S_i(\bar{p}_i) = \Lambda_i^M(x_i, \bar{p}_i).$$

If instead $x_i \geq \tilde{k}_i^*(\bar{p}_i)$, then at fare $\bar{p}_i$, the rider demand is not in the quadratic branch; hence, $\Lambda_i^M(x_i, \bar{p}_i) = \Lambda_i$. For the higher fare $\bar{p}_i'$, either $x_i \geq \tilde{k}_i^*(\bar{p}_i')$, in which case the rider demand is still in the non-quadratic branch and the rider demand remains Λ_i, or $x_i < \tilde{k}_i^*(\bar{p}_i')$, in which case the rider demand moves to the quadratic branch. In the latter case, by the definition of $\tilde{k}_i^*(\bar{p}_i')$ as the threshold beyond which the full demand Λ_i is served, it must be the case that $\Lambda_i^M(x_i, \bar{p}_i') \leq \Lambda_i = \Lambda_i^M(x_i, \bar{p}_i)$. Thus, in all cases, $\Lambda_i^M(x_i, \bar{p}_i') \leq \Lambda_i^M(x_i, \bar{p}_i)$.

Combining the two cases analyzed above establishes that $\Lambda_i^M(x_i, \bar{p}_i)$ is weakly decreasing in $\bar{p}_i$ for every fixed x_i. □

We now leverage this monotonicity property to complete the proof of Theorem 4.

Proof. Proof of Theorem 4. Fix an instance I and $\varepsilon > 0$. For each route i, define $\bar{p}_i^{\min} := \frac{c_i}{F}$ and $\bar{\mathbf{p}}^{\min} := (\bar{p}_1^{\min}, \ldots, \bar{p}_n^{\min})$. By Lemma 1, for every feasible allocation $x \in \Omega_I$ the rider welfare $R_I(x, \bar{p}) := \sum_{i=1}^n \Lambda_i^M(x_i, \bar{p}_i)$ is weakly decreasing in each coordinate $\bar{p}_i$. Therefore, the fare vector that maximizes rider welfare is $\bar{\mathbf{p}}^{\min}$. Thus, we define $R_I^{F,\text{opt}} = \max_{x\in\Omega_I} R_I(x, \bar{p}^{\min})$ and let $x^* \in \arg\max_{x\in\Omega_I} R_I(x, \bar{p}^{\min})$ be any rider-welfare maximizing allocation under these prices. Note, however, that at the minimum fare vector, the per-driver profit on all routes is $p_i(\bar{p}_i) = p_i(\frac{c_i}{F}) = \frac{c_i}{F} - \frac{c_i}{F} = 0$, so drivers are indifferent across routes and every feasible allocation $\mathbf{x} \in \Omega$ is an equilibrium. Thus, although the minimum fares maximize rider welfare, $\inf_{\mathbf{x}\in\Omega(\bar{\mathbf{p}}^{\min})} R_I(\mathbf{x}, \bar{\mathbf{p}}^{\min})$ can be arbitrarily small.

To eliminate these low welfare equilibria, we introduce route-specific price perturbations of the minimum fare vector that induces $\mathbf{x}^*$ as the unique equilibrium driver allocation. To this end, define the sets $U := \{i \in [n] : x_i^* > 0\}$ and $Z := [n]\backslash U$, and for each route i, define $M_i := c_i^O - \eta_T l_i - \frac{c_i}{F} > 0$. Since $M_i > 0$, each route admits strictly interior fares of the form $\bar{p}_i = \frac{c_i}{F} + q_i$ with $q_i \in (0, M_i)$, and for such fares one has $S_i(\bar{p}_i) = M_i - q_i > 0$.

We first choose a fare vector in a small neighborhood of $\bar{p}^{\min}$ under which the rider welfare at x^* remains arbitrarily close to optimal. Since, for fixed x^*, each map $\bar{p}_i \mapsto \Lambda_i^M(x_i^*, \bar{p}_i)$ is continuous on $\mathcal{P}_I$, the function $\bar{\mathbf{p}} \mapsto R_I(x^*, \bar{p})$ is continuous at $\bar{\mathbf{p}}^{\min}$. Hence there exist $\delta_i \in (0, M_i)$ for all $i \in [n]$ such that whenever $\bar{p}_i \in \left[\frac{c_i}{F}, \frac{c_i}{F} + \delta_i\right]$ for all i, it follows that: $R_I(x^*, \bar{\mathbf{p}}) \geq \frac{1}{1+\varepsilon} R_I(x^*, \bar{\mathbf{p}}^{\min}) = \frac{1}{1+\varepsilon} R_I^{F,\text{opt}}$.

We now choose route-specific fare perturbations that make x^* the unique equilibrium. For each used route $i \in U$, define $f_i(q) := \pi_i\big(x_i^*; \frac{c_i}{F} + q\big) = \frac{\Lambda_i^M\left(x_i^*, \frac{c_i}{F}+q\right)}{x_i^*} q$, for $q \in [0, \delta_i]$. Each f_i is continuous, satisfies $f_i(0) = 0$, and is strictly positive for every $q > 0$. Therefore, define $\underline{\kappa} := \min_{i\in U} f_i(\delta_i) > 0$ and choose any $\kappa \in (0, \underline{\kappa})$. By the intermediate value theorem, for each $i \in U$ there exists $q_i^\varepsilon \in (0, \delta_i]$ such that $f_i(q_i^\varepsilon) = \kappa$, i.e., $\pi_i\big(x_i^*; \frac{c_i}{F} + q_i^\varepsilon\big) = \kappa$ for all $i \in U$.

For each unused route $j \in Z$, define $g_j(q) := \pi_j\big(0; \frac{c_j}{F} + q\big)$ for $q \in [0, \delta_j]$. Since g_j is continuous and $g_j(0) = 0$, we may choose $q_j^\varepsilon \in (0, \delta_j]$ sufficiently small so that $g_j(q_j^\varepsilon) < \kappa$ for all $j \in Z$.

Now define the fare vector $\bar{p}^\varepsilon$ route-wise by $\bar{p}_i^\varepsilon := \frac{c_i}{F} + q_i^\varepsilon$ for all $i \in [n]$. By construction, for every used route $i \in U$, $\pi_i(x_i^*; \bar{p}_i^\varepsilon) = \kappa$, while for every unused route $j \in Z$, $\pi_j(0; \bar{p}_j^\varepsilon) < \kappa$. Hence, at the allocation x^*, every used route yields the same per-driver profit and every unused route yields strictly lower per-driver profit. Therefore x^* is an equilibrium allocation under $\bar{\mathbf{p}}^\varepsilon$.

To see uniqueness, note that since $\bar{p}_i^\varepsilon < c_i^O - \eta_T l_i$ for all i, we have $S_i(\bar{p}_i^\varepsilon) > 0$ for all i, which implies the uniqueness of the equilibrium under $\bar{\mathbf{p}}^\varepsilon$. Hence the equilibrium induced by $\bar{\mathbf{p}}^\varepsilon$ is unique, and in fact $\mathcal{E}_I(\bar{\mathbf{p}}^\varepsilon) = \{x^*\}$.

Finally, since each $q_i^\varepsilon \leq \delta_i$, the fare vector $\bar{\mathbf{p}}^\varepsilon$ lies in the continuity neighborhood chosen above. Therefore, $R_I(x^*, \bar{p}^\varepsilon) \geq \frac{1}{1+\varepsilon} R_I^{F,\text{opt}}$. Since x^* is the unique equilibrium under $\bar{\mathbf{p}}^\varepsilon$, we obtain

$$R_I^{F,\text{eq}} \geq \min_{x\in\mathcal{E}_I(\bar{\mathbf{p}}^\varepsilon)} R_I(x, \bar{\mathbf{p}}^\varepsilon) = R_I(x^*, \bar{\mathbf{p}}^\varepsilon) \geq \frac{1}{1+\varepsilon} R_I^{F,\text{opt}}.$$

Equivalently, $\frac{R_I^{F,\text{opt}}}{R_I^{F,\text{eq}}} \leq 1 + \varepsilon$. Since the instance I was arbitrary, the same bound holds uniformly over all instances, which proves the claim. □

B.8 Proof of Theorem 5

We first prove the upper bound. Fix an instance I and a feasible fare vector $\bar{\mathbf{p}} \in \mathcal{P}$. Conditional on $\bar{\mathbf{p}}$, the per-rider margins $p_i(\bar{p}_i)$ and the minibus cost advantages $S_i(\bar{p}_i)$ are fixed, so the induced rider demand functions remain exactly of the form in Equation (3) in the fixed-fare model. Hence, the fixed-fare profit PoA guarantee applies to the induced instance: $\sup_{\mathbf{x}\in\Omega} P_I(\mathbf{x}, \bar{\mathbf{p}}) \leq 2 \inf_{\mathbf{x}\in\Omega^{\mathrm{Eq}}(\bar{\mathbf{p}})} P_I(\mathbf{x}, \bar{\mathbf{p}})$. Taking the supremum over $\bar{\mathbf{p}} \in \mathcal{P}$ gives $P_I^{\mathrm{fare,opt}} \leq 2P_I^{\mathrm{fare,eq}}$.

It remains to show tightness. To do so, for any $\varepsilon > 0$, we construct a two-route instance parametrized by $\rho \in (0, \varepsilon)$. For this instance, let $D = 1$, $k_1^* = \rho$, $k_2^* = 1 - \rho$, $\beta := \frac{\eta_E + \eta_L}{\eta_E \eta_L (t_2 - t_1)}$, and choose a constant $K > 0$ satisfying $\beta K < 1$. Next, choose $\Lambda_1 > 0$ arbitrarily, and set $\Lambda_2 := k_2^* \Lambda_1$. Finally, choose route parameters so that $k_i^* = \frac{2l_i\lambda_i}{F}$ and $c_i^O - \eta_T l_i - \frac{c_i}{F} = K$ for $i \in \{1, 2\}$. Thus, on both routes, the feasible per-rider profit margin interval is $p_i(\bar{p}_i) \in [0, K]$. For this instance, we establish the desired lower bound in two steps.

Step 1: Lower bounding the centralized fare-optimized profit. Set the maximum fare on both routes, so $p_i(\bar{p}_i) = K$ and $S_i(\bar{p}_i) = 0$. Then, the rider demand function reduces to the capacity-constrained form in Equation (4) and $\mathbf{x}^* = (k_1^*, k_2^*)$ serves the full demand on both routes. Then: $P_I^{\mathrm{fare,opt}} \geq K(\Lambda_1 + \Lambda_2) = K\Lambda_1(1 + k_2^*) = K\Lambda_1(2 - \rho)$.

Step 2: Upper bounding the best equilibrium profit over all fare vectors. Fix any feasible $\bar{\mathbf{p}}$, and let $\mathbf{x} \in \Omega^{\mathrm{Eq}}(\bar{\mathbf{p}})$ be any equilibrium under $\bar{\mathbf{p}}$. We show that $P_I(\mathbf{x}, \bar{\mathbf{p}}) \leq K\Lambda_1$ by considering two cases.

Case 1: $x_2 = 0$. Here, $x_1 = 1$ and at most Λ_1 riders are served; hence, $P_I(x, \bar{p}) \leq p_1(\bar{p}_1)\Lambda_1 \leq K\Lambda_1$.

Case 2: $x_2 > 0$. In this case, let π^{eq} be the common per-driver profit on the used routes. Since $D = 1$, $P_I(\mathbf{x}, \bar{\mathbf{p}}) = \pi^{\mathrm{eq}}$. Since the per-driver profit on a route is weakly decreasing in its own driver allocation $\pi^{\mathrm{eq}} = \pi_2(x_2, \bar{p}_2) \leq \pi_2(0, \bar{p}_2)$. Now, let $p_2 := p_2(\bar{p}_2) \in [0, K]$. Since $S_2(\bar{p}_2) = K - p_2$, the minibus rider demand function in Equation (3) implies that, for $x_2 < \tilde{k}_2^*(\bar{p}_2)$,

$$\Lambda_2^M(x_2, \bar{p}_2) = \frac{\Lambda_2}{k_2^*}\Big(1 + \beta(K - p_2)\Big)x_2 - \frac{\Lambda_2}{(k_2^*)^2}\beta(K - p_2)x_2^2.$$

Therefore, $\pi_2(0, \bar{p}_2) = \frac{\Lambda_2}{k_2^*}\, p_2\Big(1+\beta(K-p_2)\Big)$. Then, consider the scalar function $h(p) := p\Big(1+\beta(K-p)\Big)$ for $p \in [0, K]$. Its derivative is $h'(p) = 1 + \beta K - 2\beta p \geq 1 - \beta K > 0$, where the strict inequality uses $\beta K < 1$. Thus h is strictly increasing on $[0, K]$, so $\pi_2(0, \bar{p}_2) \leq \frac{\Lambda_2}{k_2^*} K = \frac{k_2^* \Lambda_1}{k_2^*} K = K\Lambda_1$. Consequently, $P_I(\mathbf{x}, \bar{\mathbf{p}}) = \pi^{\mathrm{eq}} \leq K\Lambda_1$. Since both cases yield the same bound, for every feasible fare vector and every equilibrium under that fare vector, $P_I(\mathbf{x}, \bar{\mathbf{p}}) \leq K\Lambda_1$. Hence, $P_I^{\mathrm{fare,eq}} \leq K\Lambda_1$.

Combining the lower bound on $P_I^{\mathrm{fare,opt}}$ with the upper bound on $P_I^{\mathrm{fare,eq}}$, we obtain

$$\frac{P_I^{\mathrm{fare,opt}}}{P_I^{\mathrm{fare,eq}}} \geq \frac{K\Lambda_1(2 - \rho)}{K\Lambda_1} = 2 - \rho > 2 - \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the factor two upper bound is tight, establishing that $\text{P-PoA}^{\mathrm{fare}} = 2$.

B.9 Proof of Theorem 6

We begin by introducing some notation. For an instance I, let $M_i := c_i^O - \eta_T l_i - \frac{c_i}{F}$, $\Delta := t_2 - t_1$, $a_i := \frac{F\Delta}{2l_i}$, $b_i := \frac{F(\eta_E+\eta_L)}{2l_i\eta_E\eta_L}$, $k_i^* := \frac{\Lambda_i}{a_i} = \frac{2l_i\Lambda_i}{F\Delta}$, and $H_i := \frac{\Lambda_i}{b_i} = \frac{2l_i\Lambda_i\eta_E\eta_L}{F(\eta_E+\eta_L)}$. Furthermore, let the per-

rider profit be $p_i := \bar{p}_i - \frac{c_i}{F} \in [0, M_i]$. Then, for $x_i < \tilde{k}_i^*(p_i)$, Equation (3) implies $\Lambda_i^M(x_i, p_i) = a_i x_i + b_i(M_i - p_i)x_i\left(1 - \frac{x_i}{k_i^*}\right)$, while for $x_i \geq \tilde{k}_i^*(p_i)$, $\Lambda_i^M(x_i, p_i) = \Lambda_i$. Accordingly, re-defining the per-driver profit as a function of the per-rider profit (rather than the fare), we have:

$$\pi_i(x_i, p_i) = \begin{cases} p_i\left(a_i + b_i(M_i - p_i)\left(1 - \dfrac{x_i}{k_i^*}\right)\right), & x_i < \tilde{k}_i^*(p_i), \\ \dfrac{p_i\Lambda_i}{x_i}, & x_i \geq \tilde{k}_i^*(p_i), \end{cases}$$

and the total profit on the route is $P_i(x_i, p_i) := p_i\Lambda_i^M(x_i, p_i) = x_i\,\pi_i(x_i, p_i)$.

Next, to prove our claim, fix a route i. For every fixed fare $\bar{p}_i$ (and, hence, fixed per-rider profit p_i), first note that the map $x \mapsto \pi_i(x, p_i)$ is weakly decreasing; hence, $q_i(\cdot)$, which is the point-wise maximum of continuous weakly decreasing functions, is also continuous and weakly decreasing.

We next show that $q_i(x)$ is computable in $O(1)$ time for each x. If $x \geq k_i^*$, the entire demand is served for every $p_i \in [0, M_i]$; hence, $q_i(x) = \max_{p_i \in [0, M_i]} \frac{p_i\Lambda_i}{x} = \frac{M_i\Lambda_i}{x}$. If, on the other hand, $x < k_i^*$, the entire demand is served when $p_i \leq M_i - \frac{H_i}{x}$. This relation implies that when $p_i \leq p_i^{\text{sat}}(x) := \max\left\{0,\, M_i - \frac{H_i}{x}\right\}$, it follows that $\pi_i(x, p_i) = \frac{p_i\Lambda_i}{x}$,; hence, the highest profit in this regime is achieved when $p_i = p_i^{\text{sat}}(x)$. When $p_i \geq p_i^{\text{sat}}(x)$, it follows that:

$$\pi_i(x, p_i) = p_i\left(a_i + b_i(M_i - p_i)\left(1 - \frac{x}{k_i^*}\right)\right),$$

which is a concave quadratic function of p_i. Hence, its maximizer is the projection of the unique stationary point $p_i^{\text{quad}}(x) = \frac{1}{2}\left(M_i + \frac{H_i}{k_i^* - x}\right)$ onto the interval $[p_i^{\text{sat}}(x), M_i]$. Thus, $q_i(x)$ is obtained by evaluating a constant number (three) of candidate prices.

Now fix $\rho \geq 0$. Since q_i is continuous and weakly decreasing, $u_i(\rho)$ is well-defined, and

$$q_i(x) \geq \rho \quad \Longleftrightarrow \quad x \in [0, u_i(\rho)].$$

Next, we show that an equilibrium with common profit ρ exists if and only if $\sum_{i=1}^n u_i(\rho) \geq D$. To this end, suppose first that $\sum_{i=1}^n u_i(\rho) \geq D$. Since each feasible interval starts at 0, we can choose $x_i \in [0, u_i(\rho)]$ for all i such that $\sum_i x_i = D$. For every route with $x_i > 0$, we have $q_i(x_i) \geq \rho$, i.e., there exists some price with per-driver profit at least ρ at the allocation x_i. Since $p_i \mapsto \pi_i(x_i, p_i)$ is continuous and satisfies $\pi_i(x_i, 0) = 0$, there exists $p_i \in [0, M_i]$ such that $\pi_i(x_i, p_i) = \rho$. For routes with $x_i = 0$, set $p_i = 0$. Under these fares, every used route yields a per-driver profit of exactly ρ, while every unused route yields profit $0 \leq \rho$. Since each route profit is weakly decreasing in its own allocation, no driver can profitably deviate. Hence, the resulting allocation is an equilibrium, and its cumulative profit is $\sum_{i=1}^n P_i(x_i, p_i) = \sum_{i=1}^n x_i\rho = D\rho$.

Conversely, suppose $\mathbf{x}$ is an equilibrium driver allocation, $\mathbf{p}$ is the vector of per-rider profits, with common per-driver profit ρ on all used routes. Then, for every used route i, $\rho = \pi_i(x_i, p_i) \leq q_i(x_i)$, so $x_i \leq u_i(\rho)$. Summing over all routes gives $D = \sum_{i=1}^n x_i \leq \sum_{i=1}^n u_i(\rho)$. Thus, we have established that an equilibrium allocation with common profit ρ exists if and only if $\sum_{i=1}^n u_i(\rho) \geq D$.

It follows that the highest equilibrium common profit is $\rho_I^* = \sup\{\rho \geq 0 : \sum_{i=1}^n u_i(\rho) \geq D\}$, and thus the highest equilibrium cumulative profit is $D\rho_I^*$. Finally, since every $u_i(\rho)$ is weakly decreasing in ρ for $i \in [n]$, the function $\sum_{i=1}^n u_i(\rho)$ is also weakly decreasing. Hence ρ_I^* can be

found by binary search on $[0, \overline{\pi}]$, where, for each candidate ρ, $u_i(\rho)$ is computed using binary search on $x \in [0, D]$. Here $\overline{\pi} = \max_i \max_{p_i \in \mathcal{P}_i} \pi_i(0, p_i)$. Finally, recalling here that $q_i(x)$ can be evaluated in $O(1)$ time, we obtain the desired complexity.

B.10 Proof of Theorem 7

In the following proof, we express all quantities as a function of the per-rider profits $p_i = \bar{p}_i - \frac{c_i}{F}$ for each route i. Define $G_i(x) := \max_{p_i \in [0, M_i]} P_i(x, p_i)$ for $x \in [0, D]$ and consider the same quantities a_i, b_i, M_i defined in the proof of Theorem 6. Then, for each fixed x, the maximization defining $G_i(x)$ is one-dimensional and can be solved in $O(1)$ time akin to the proof of Theorem 6. Next, for every fixed per-rider profit $p_i = \bar{p}_i - \frac{c_i}{F}$, the route profit $x \mapsto P_i(x, p_i)$ is weakly increasing and on the quadratic branch has derivative

$$\frac{\partial P_i(x, p_i)}{\partial x} = p_i \left(a_i + b_i(M_i - p_i) \left(1 - \frac{2x}{k_i^*} \right) \right) \leq p_i \big(a_i + b_i(M_i - p_i) \big) \leq L_i^P,$$

where $L_i^P := \max_{\bar{p}_i \in \mathcal{P}_i} p_i(\bar{p}_i) \big(a_i + b_i(M_i - p_i) \big)$. On the saturated branch the derivative is 0. Thus, each map $x \mapsto P_i(x, p_i)$ is L_i^P-Lipschitz, and so G_i, the pointwise maximum of such maps, is also L_i^P-Lipschitz. Moreover G_i is weakly increasing.

Since G_i is weakly increasing for all i, the profit maximization problem with constraints $\sum_i x_i = D$ and $\sum_i x_i \leq D$ have the same optimum value. Next, define $h := \frac{\varepsilon}{\sum_{i=1}^n L_i^P}$ and $m := \left\lfloor \frac{D}{h} \right\rfloor$. Further, for $i \in [n]$ and $s \in \{0, 1, \ldots, m\}$, define

$$V_i(s) := \max \left\{ \sum_{j=1}^{i} G_j(x_j) : x_j \in \{0, h, 2h, \ldots\},\ \sum_{j=1}^{i} x_j \leq sh \right\}.$$

Then $V_0(s) = 0$ and $V_i(s) = \max_{0 \leq t \leq s} \Big\{ V_{i-1}(s - t) + G_i(th) \Big\}$. This dynamic program has $n(m+1)$ states, and each state scans at most $m + 1$ transitions, so the running time is $O(nm^2)$.

Next, let $\mathbf{x}^*$ be an optimal solution of the (continuous) profit maximization problem, and define the rounded allocation $\tilde{x}_i := h \left\lfloor \frac{x_i^*}{h} \right\rfloor$ for all i. Then, $\sum_i \tilde{x}_i \leq D$, so $\tilde{\mathbf{x}}$ is feasible for the discretized problem with the inequality constraint $\sum_i \tilde{x}_i \leq D$. Since each G_i is L_i^P-Lipschitz, it follows that:

$$G_i(x_i^*) - G_i(\tilde{x}_i) \leq L_i^P (x_i^* - \tilde{x}_i) \leq L_i^P h.$$

Summing the above inequality over all routes gives $P_I^{\text{fare,opt}} - \sum_{i=1}^n G_i(\tilde{x}_i) \leq h \sum_{i=1}^n L_i^P = \varepsilon$. Since the dynamic program computes the best grid allocation:

$$V_n(m) \geq \sum_{i=1}^{n} G_i(\tilde{x}_i) \geq P_I^{\text{fare,opt}} - \varepsilon.$$

Finally, for the dynamic program solution $\hat{\mathbf{x}}$, choose $\hat{p}_i \in \arg\max_{p_i \in [0, M_i]} P_i(\hat{x}_i, p_i)$ for each route i, which can be performed in $O(1)$ time. Then, $\sum_{i=1}^n P_i(\hat{x}_i, \hat{p}_i) = \sum_{i=1}^n G_i(\hat{x}_i) \geq P_I^{\text{fare,opt}} - \varepsilon$, which proves our claim.

C Fare Optimization for Rider Welfare

C.1 Computational Tractability of Fare-optimized Rider Welfare Maximization with no Reservation Wage

The implementing perturbation in Theorem 4 is constructive and can be computed efficiently. First, the rider-welfare maximizing allocation $\mathbf{x}^*$ at the minimum fare vector $\bar{\mathbf{p}}^{\min}$ can be computed efficiently, since the resulting problem is a separable concave resource-allocation problem. Given $\mathbf{x}^*$, for any candidate common profit level κ, we compute the perturbation q_i^ε on each used route i by solving $f_i(q_i^\varepsilon) = \kappa$, where $f_i(q_i^\varepsilon) = \pi_i(x_i^*, \frac{c_i}{F} + q_i^\varepsilon)$ is either linear or quadratic in q_i^ε, and hence q_i^ε can be computed in closed form; for unused routes, we choose any sufficiently small q_i^ε so that $f_j(q_j^\varepsilon) < \kappa$. Moreover, for fixed $\mathbf{x}^*$, the following Lipshitzness relation can be established: the rider welfare loss induced by these perturbations is at most linear in κ. In other words, there exists a constant $C > 0$ such that $R_I(\mathbf{x}^*, \bar{\mathbf{p}}^{\min}) - R_I(\mathbf{x}^*, \bar{\mathbf{p}}^{\,\kappa}) \le C\kappa$. Therefore, choosing $\kappa = O(\varepsilon)$ suffices to obtain a $(1+\varepsilon)$-approximate implementation, and once $\mathbf{x}^*$ is known, the corresponding implementing fare vector can be computed in $O(n)$ time. Note that if one prefers to determine such a κ by search rather than by an explicit bound, this adds only $O(\log(1/\varepsilon))$ iterations, for an overall complexity of $O(n \log(1/\varepsilon))$ after computing $\mathbf{x}^*$.

Proof. Proof of Lipshitzness Relation. Fix the numbers $\delta_i > 0$ in the proof of Theorem 4, and define $\underline{\kappa} := \min_{i \in U} f_i(\delta_i) > 0$ and $m_i := \frac{\Lambda_i^M(x_i^*, \frac{c_i}{F} + \delta_i)}{x_i^*} > 0$, where $i \in U$. For each $i \in U$ and $\kappa \in (0, \underline{\kappa}]$, let $q_i(\kappa) \in (0, \delta_i]$ be any solution of $f_i(q) = \kappa$.

Next, define $r_i(q) := \Lambda_i^M\big(x_i^*, \frac{c_i}{F} + q\big)$. By Lemma 1, r_i is decreasing, so for every $i \in U$, $\kappa = f_i(q_i(\kappa)) = \frac{r_i(q_i(\kappa))}{x_i^*} q_i(\kappa) \ge m_i q_i(\kappa)$, and therefore $q_i(\kappa) \le \frac{\kappa}{m_i}$. For fixed x_i^*, the branch condition depends on q through

$$\tilde{k}_i^*\Big(\frac{c_i}{F} + q\Big) = \min\left\{k_i^*,\ \frac{2 l_i \Lambda_i \eta_E \eta_L}{F(M_i - q)(\eta_E + \eta_L)}\right\}, \qquad M_i := c_i^O - \eta_T l_i - \frac{c_i}{F}.$$

Hence $r_i(q) := \Lambda_i^M\big(x_i^*, \frac{c_i}{F} + q\big)$ is either (i) always on the quadratic branch, (ii) always equal to Λ_i, or (iii) equal to Λ_i up to the unique switching point $\hat{q}_i = M_i - \frac{2 l_i \Lambda_i \eta_E \eta_L}{F(\eta_E + \eta_L) x_i^*}$, and affine thereafter. On the quadratic branch, we have $r_i(q) = A_i(x_i^*) + B_i(x_i^*)(M_i - q)$, so its slope is $-B_i(x_i^*)$, while on the saturated branch its slope is 0. Therefore r_i is globally L_i-Lipschitz on $[0, \delta_i]$, with

$$L_i = \begin{cases} B_i(x_i^*), & x_i^* < k_i^*, \\ 0, & x_i^* \ge k_i^*. \end{cases}$$

Thus, for every $i \in U$, $0 \le r_i(0) - r_i(q_i(\kappa)) \le L_i q_i(\kappa) \le \frac{L_i}{m_i}\kappa$. Summing over $i \in U$ yields $R(\mathbf{x}^*, \bar{\mathbf{p}}^{\min}) - R_I(\mathbf{x}^*, \bar{\mathbf{p}}^{\,\kappa}) \le C\kappa$, where $C := \sum_{i \in U} \frac{L_i}{m_i}$. Therefore, any $\kappa \le \min\left\{\underline{\kappa}, \frac{\varepsilon}{1+\varepsilon}\frac{R_I^{\text{fare,opt}}}{C}\right\}$ guarantees $R_I(\mathbf{x}^*, \bar{\mathbf{p}}^{\,\kappa}) \ge \frac{1}{1+\varepsilon} R_I^{\text{fare,opt}}$. □

C.2 Fare-Optimized Rider Welfare PoA under Reservation Wages

Suppose drivers have a reservation wage $W \ge 0$, interpreted as an individual rationality constraint on per-driver profit. When $W = 0$, this reduces to the setting studied in Theorem 4; below, we

focus on the case $W > 0$. Under a reservation wage, the centralized fare-optimized rider-welfare benchmark is

$$R_I^{\text{fare,opt}}(W) := \max_{\substack{\mathbf{x}\in\Omega,\ \bar{\mathbf{p}}\in\mathcal{P} \\ \pi_i(x_i,\bar{p}_i)\geq W\ \forall i:\, x_i>0}} R_I(\mathbf{x},\bar{\mathbf{p}}).$$

For a fare vector $\bar{\mathbf{p}} \in \mathcal{P}$, define the set of reservation-wage-feasible equilibria as $\Omega_W^{\text{Eq}}(\bar{\mathbf{p}}) := \left\{\mathbf{x} \in \Omega^{\text{Eq}}(\bar{\mathbf{p}}) : \pi_i(x_i,\bar{p}_i) \geq W,\ \forall i \text{ with } x_i > 0\right\}$. The corresponding fare-optimized equilibrium benchmark is

$$R_I^{\text{fare,eq}}(W) := \sup_{\bar{\mathbf{p}}\in\mathcal{P}} \inf_{\mathbf{x}\in\Omega_W^{\text{Eq}}(\bar{\mathbf{p}})} R_I(\mathbf{x},\bar{\mathbf{p}}).$$

For each route i and allocation $x_i > 0$, define the smallest fare that meets the reservation wage as $\bar{p}_i^W(x_i) := \inf\left\{\bar{p}_i \in \left[\frac{c_i}{F},\, c_i^O - \eta_T l_i\right] : \pi_i(x_i,\bar{p}_i) \geq W\right\}$, whenever this set is nonempty. By the monotonicity of rider demand in fares, rider welfare is weakly decreasing in $\bar{p}_i$ for fixed x_i. Hence, for any fixed feasible driver allocation $\mathbf{x}$, the rider welfare maximizing fare on each used route is precisely $\bar{p}_i^W(x_i)$. Therefore, the reservation wage constrained rider welfare optimum can be written as

$$R_I^{\text{fare,opt}}(W) = \max_{\mathbf{x}\in\Omega_W} \sum_{i:\, x_i>0} \Lambda_i^M\big(x_i, \bar{p}_i^W(x_i)\big),$$

where Ω_W denotes the set of allocations for which $\bar{p}_i^W(x_i)$ exists on every used route.

Let $\mathbf{x}^{W,*}$ be an optimizer of the above problem, and set

$$\bar{p}_i^{W,*} := \begin{cases} \bar{p}_i^W(x_i^{W,*}), & \text{if } x_i^{W,*} > 0, \\ \frac{c_i}{F} + \delta_i, & \text{if } x_i^{W,*} = 0, \end{cases}$$

where $\delta_i > 0$ is chosen sufficiently small so that $\pi_i\big(0, \frac{c_i}{F} + \delta_i\big) < W$ for all i with $x_i^{W,*} = 0$. Such a choice is possible whenever $\frac{c_i}{F} < c_i^O - \eta_T l_i$ by continuity of $\pi_i(0,\bar{p}_i)$ and the fact that $\pi_i\big(0, \frac{c_i}{F}\big) = 0 < W$.

By construction, every used route i satisfies $\pi_i(x_i^{W,*}, \bar{p}_i^{W,*}) = W$, while every unused route j satisfies $\pi_j(0, \bar{p}_j^{W,*}) < W$. Hence, all used routes yield the same per-driver profit and every unused route yields strictly less. Therefore, $\mathbf{x}^{W,*} \in \Omega_W^{\text{Eq}}(\bar{\mathbf{p}}^{W,*})$.

Consequently, if in addition $\bar{p}_i^{W,*} < c_i^O - \eta_T l_i$ for all routes i with $x_i^{W,*} > 0$, then $S_i(\bar{p}_i^{W,*}) > 0$ on every used route. In this regime, the equilibrium under the fare vector $\bar{\mathbf{p}}^{W,*}$ is unique and thus the fare-optimized rider welfare PoA remains equal to one in the reservation wage constrained setting.

If, however, some used route satisfies $\bar{p}_i^{W,*} = c_i^O - \eta_T l_i$, then the strict uniqueness argument may fail because $S_i(\bar{p}_i^{W,*}) = 0$ on that route. In this case, there still exists an equilibrium attaining the reservation wage constrained rider welfare optimum, but the equilibrium need not be unique; consequently, other equilibria under the same fare vector may achieve lower rider welfare.

C.3 Computational Tractability of Fare-optimized Rider Welfare Maximization with Reservation Wage

Proposition 5 (Additive approximation for the reservation wage-constrained system optimum)**.** *Assume the reservation $W > 0$, which can be interpreted as the outside option for drivers, so the*

planner may leave some drivers unmatched. Then, the planner solves

$$\max_{\substack{x_i\geq 0,\ \bar{p}_i\in[\frac{c_i}{F},\, c_i^O-\eta_T l_i] \\ \sum_{i=1}^n x_i\leq D,\ \pi_i(x_i,\bar{p}_i)\geq W\ \ \forall i:\, x_i>0}} \sum_{i=1}^n \Lambda_i^M(x_i,\bar{p}_i).$$

For each route i, define $M_i := c_i^O - \eta_T l_i - \frac{c_i}{F}$, $a_i := \frac{F(t_2-t_1)}{2l_i}$, $b_i := \frac{F(\eta_E+\eta_L)}{2l_i\eta_E\eta_L}$, $L_i := a_i + b_i M_i$, and $U_i := \min\Big\{D, \frac{M_i\Lambda_i}{W}\Big\}$. Also, define the route-wise reservation wage welfare function

$$R_i^W(x) := \max\Big\{\Lambda_i^M(x,\bar{p}_i) : \bar{p}_i \in \Big[\frac{c_i}{F},\, c_i^O-\eta_T l_i\Big],\ \pi_i(x,\bar{p}_i)\geq W\Big\}, \qquad R_i^W(0) := 0.$$

Then, for every $\varepsilon > 0$, there is an algorithm that returns a feasible allocation $\hat{\mathbf{x}}$ with $\sum_{i=1}^n R_i^W(\hat{x}_i) \geq R_I^{\text{fare,opt}} - \varepsilon$. The algorithm runs in $O(nm^2)$, where $m =:= \left\lceil \frac{D\sum_{i=1}^n L_i}{\varepsilon} \right\rceil$, which is polynomial in n and $1/\varepsilon$.

Proof. Proof. By Lemma 2, for each fixed x_i, rider welfare is weakly decreasing in $\bar{p}_i$. Hence $R_i^W(x)$ is attained at the smallest fare satisfying the reservation wage constraint. Therefore, for any fixed x, each value $R_i^W(x)$ can be computed in $O(1)$ time by solving a single linear or quadratic equation in the profit per-rider $q_i := \bar{p}_i - \frac{c_i}{F}$: on the full-demand branch one has $q_i = \frac{Wx}{\Lambda_i}$, while on the quadratic branch q_i is the smaller root of the quadratic equation $q_i \frac{\Lambda_i^M(x,\frac{c_i}{F}+q_i)}{x} = W$.

Moreover, any feasible route allocation must satisfy $Wx_i \leq \Big(\bar{p}_i - \frac{c_i}{F}\Big)\Lambda_i^M(x_i,\bar{p}_i) \leq M_i\Lambda_i$; hence, $x_i \leq U_i$.

Lipschitz Bound: We next prove a one-sided Lipschitz bound for R_i^W. Fix $0 \leq y \leq x \leq U_i$, and let $\bar{p}_i^W(x)$ be a welfare-maximizing fare for route i at allocation x. Since the per-driver profit on a route is non-increasing in the route allocation, the same fare $\bar{p}_i^W(x)$ is also feasible at the smaller allocation y. Therefore, $R_i^W(y) \geq \Lambda_i^M\big(y,\bar{p}_i^W(x)\big)$ and $R_i^W(x) = \Lambda_i^M\big(x,\bar{p}_i^W(x)\big)$. Hence,

$$R_i^W(x) - R_i^W(y) \leq \Lambda_i^M\big(x,\bar{p}_i^W(x)\big) - \Lambda_i^M\big(y,\bar{p}_i^W(x)\big).$$

For any fixed fare, the function $z \mapsto \Lambda_i^M(z,\bar{p}_i)$ is continuous, has derivative 0 on the saturated branch, and on the quadratic branch has derivative at most $a_i + b_i M_i = L_i$. Thus, for every fixed fare, $\Lambda_i^M(x,\bar{p}_i) - \Lambda_i^M(y,\bar{p}_i) \leq L_i(x-y)$, which implies that: $R_i^W(x) - R_i^W(y) \leq L_i(x-y)$.

Dynamic Program: We now use a dynamic programming approach to establish our guarantee. To this end, choose the grid size $h := \frac{\varepsilon}{\sum_{i=1}^n L_i}$ and $m := \big\lfloor \frac{D}{h} \big\rfloor$. For each $i \in [n]$ and $s \in \{0,1,\ldots,m\}$, define the set

$$\mathcal{F}_i(s) := \left\{ (x_1,\ldots,x_i) : x_j \in \{0,h,2h,\ldots\},\ 0 \leq x_j \leq U_j,\ \sum_{j=1}^i x_j \leq sh \right\},$$

and define $V_i(s) := \max_{(x_1,\ldots,x_i)\in\mathcal{F}_i(s)} \sum_{j=1}^i R_j^W(x_j)$.

Thus, $V_i(s)$ is the maximum rider welfare attainable using only the first i routes and at most sh units of driver mass on the grid. Note that the boundary condition is $V_0(s) = 0$ for all s. We claim that $V_i(s)$ satisfies the recursion

$$V_i(s) = \max_{0\leq t\leq \min\{s,\lfloor U_i/h\rfloor\}} \Big\{ V_{i-1}(s-t) + R_i^W(th) \Big\}.$$

To prove this, fix i, s. First, let $(x_1,\ldots,x_i) \in \mathcal{F}_i(s)$ be any feasible grid allocation with $x_i = th$. Then, $0 \leq t \leq \min\{s, \lfloor U_i/h \rfloor\}$, and the vector $(x_1,\ldots,x_{i-1})$ belongs to $\mathcal{F}_{i-1}(s-t)$, since $\sum_{j=1}^{i-1} x_j \leq$

$sh - th = (s-t)h$. Thus,

$$\sum_{j=1}^{i} R_j^W(x_j) = \sum_{j=1}^{i-1} R_j^W(x_j) + R_i^W(th) \le V_{i-1}(s-t) + R_i^W(th).$$

Since this holds for every feasible allocation in $\mathcal{F}_i(s)$, we obtain

$$V_i(s) \le \max_{0 \le t \le \min\{s, \lfloor U_i/h \rfloor\}} \Big\{ V_{i-1}(s-t) + R_i^W(th) \Big\}.$$

For the reverse inequality, fix any integer $0 \le t \le \min\{s, \lfloor U_i/h \rfloor\}$, and let $(x_1, \ldots, x_{i-1}) \in \mathcal{F}_{i-1}(s-t)$ attain the value $V_{i-1}(s-t)$. Then setting $x_i := th$ yields a vector in $\mathcal{F}_i(s)$, because $th \le U_i$ and

$$\sum_{j=1}^{i} x_j = \sum_{j=1}^{i-1} x_j + th \le (s-t)h + th = sh.$$

Hence, $V_i(s) \ge V_{i-1}(s-t) + R_i^W(th)$, and taking the maximum over all feasible t yields the opposite inequality, and so the recursion holds.

By induction on i, the dynamic program computes the exact optimum over all grid allocations. The dynamic program has $n(m+1)$ states, and each state $V_i(s)$ is computed by scanning over at most $m+1$ possible values of t; resulting in a total running time of $O(nm^2)$ and a space complexity of $O(nm)$.

Concluding the Proof: Finally, consider a rounded vector $\tilde{\mathbf{x}}$ with $\tilde{x}_i := h \left\lfloor \frac{x_i^*}{h} \right\rfloor$, then $\sum_{i=1}^n \frac{\tilde{x}_i}{h} = \sum_{i=1}^n \left\lfloor \frac{x_i^*}{h} \right\rfloor \le \left\lfloor \sum_{i=1}^n \frac{x_i^*}{h} \right\rfloor \le \left\lfloor \frac{D}{h} \right\rfloor = m$, so the rounded solution $\tilde{\mathbf{x}}$ is feasible for the grid problem and is therefore dominated by the dynamic-programming optimum $V_n(m)$. Next, by the earlier obtained Lipschitz bound, it follows that:

$$R_i^W(x_i^*) - R_i^W(\tilde{x}_i) \le L_i(x_i^* - \tilde{x}_i) \le L_i h.$$

Summing over i gives $R_I^{\text{fare,opt}} - \sum_{i=1}^n R_i^W(\tilde{x}_i) \le h \sum_{i=1}^n L_i = \varepsilon$. Since the dynamic program computes the best grid allocation, it follows that:

$$V_n(m) = \max_{0 \le s \le m} V_n(s) \ge \sum_{i=1}^{n} R_i^W(\tilde{x}_i) \ge R_I^{\text{fare,opt}} - \varepsilon.$$

This proves the additive-ε guarantee. Note here that since $R_i^W(x) \ge 0$ for all i, x, the values $V_n(s)$ are non-decreasing in s, so the best grid solution is $V_n(m)$. □

Remark 2. In the above proof, we can retain the original equality constraint $\sum_i x_i = D$, with the same proof going through after adding a dummy outside option route that yields payoff W and rider welfare 0. Then, the planner may send any unused drivers to the dummy route, and the approximation algorithm above applies without change.

D Additional Discussions and Theoretical Results

D.1 Limited Gains from Route Switching

We show that, once all riders on a route have been served, the potential gains to a driver from switching routes are bounded and unlikely to outweigh the associated switching costs, such as the travel time required to move between geographically separated routes.

To this end, we first note from Corollary 4 that the total service time T_i on any route under our studied framework are bounded above by $\Delta + S_i \left(\frac{\eta_E + \eta_L}{\eta_E \eta_L} \right)$. Letting $\bar{S} = \max_i S_i$, it follows for

any two routes i, i' that the difference in the service times satisfies $|T_i - T_{i'}| \leq \bar{S}\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)$.

We now show that there exists a bounded switching cost under which drivers would not change routes if $T_i < T_{i'}$. In this case, we can define the driver profit as the sum of three terms: (i) Driver profit on route i, (ii) Driver profit on route i' during the period $T_{i'} - T_i$, and (iii) the negative of the cost to switch routes. Thus, for the drivers to not switch routes, we just require the switching cost to be high enough to cancel out the Driver profit on route i' during the period $T_{i'} - T_i$.

Now, the maximum driver profit on any route i' is bounded above by

$$p_{i'}(\Lambda_{i'}^M)'(0) = \frac{p_{i'}F}{2l_{i'}}\left(\Delta + S_{i'}\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\right) \leq \frac{p_{\max}F}{2\underline{l}}\left(\Delta + \bar{S}\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\right),$$

as the rider demand function Λ_i^M has a decreasing slope. Here, we take $p_{\max} = \max_i p_i$ and let $\underline{l} = \min_i l_i > 0$. Thus, for a bounded switching cost that is at least $\left(\Delta + \bar{S}\left(\frac{\eta_E+\eta_L}{\eta_E\eta_L}\right)\right)$, drivers will not seek to change routes even if $T_i < T_{i'}$.

D.2 Additional Properties of Budget-Balanced Cross-Subsidization Scheme

Beyond the algorithmic and computational properties highlighted in Section 5, we emphasize the generality of Theorem 3 by highlighting several other properties of the optimal budget-balanced cross-subsidization scheme.

First, if $\mathbf{x}^*$ is a solution to a system optimization problem beyond cumulative driver profit or rider welfare maximization that can only be solved approximately, up to an approximation factor β, then implementing this allocation via cross-subsidization results in an equilibrium that is likewise β-optimal. Thus, cross-subsidization preserves approximation guarantees when translating system-optimal driver allocations into equilibrium outcomes.

Next, since Theorem 3 provides a method to implement any feasible target driver allocation $\mathbf{x}^*$ with at least one strictly positive entry as an equilibrium, it applies not only to allocations optimizing different system objectives but also to allocations satisfying additional constraints, such as upper or lower bounds on driver supply across routes.

Furthermore, while the proof of Theorem 3 analyzed the setting when the equilibrium condition in Equation (7) holds with equality for all routes, it can be directly extended to settings when this equilibrium condition holds with a strict inequality for routes where the target allocation satisfies $x_j^* = 0$.

Finally, while our equilibrium condition assumes that drivers choose routes based solely on relative per-driver profits, selecting the route that maximizes earnings without imposing requirements on the absolute level of profits, our framework can be readily generalized to incorporate individual rationality constraints when drivers have an outside option with reservation wage W. In particular, if a target allocation $\mathbf{x}^*$ satisfies individual rationality for drivers, so that $\pi_j^* \geq W$ for all routes j with $x_j^* > 0$, then this condition is preserved under cross-subsidization:

$$\tilde{\pi}^{Eq} = \frac{\sum_{i\in[n]} \pi_i^* x_i^*}{\sum_{i\in[n]} x_i^*} \geq \frac{W\sum_{i\in[n]} x_i^*}{\sum_{i\in[n]} x_i^*} = W.$$

Thus, drivers are never made worse off under cross-subsidization relative to their reservation wage.

D.3 Uniqueness and Multiplicity of Equilibria under Cross-Subsidization

An equilibrium driver allocation induced by a cross-subsidy scheme may, in general, be non-unique. This not influence the validity of Theorem 3 and Corollary 3, as our cross-subsidization scheme guarantees that the target allocation is an equilibrium, irrespective of the existence of other equilibria. That said, in the regime when the cost difference $S_i > 0$ for all routes i, a condition that commonly holds in informal transit systems, where the cost of using the minibus without rider queuing and schedule delays is substantially lower than that of the outside option, the equilibrium driver allocation is guaranteed to be unique (see Appendix D.4). More generally, even when multiple equilibria arise, cumulative driver profits are identical across all equilibrium allocations, although rider welfare may vary across equilibrium allocations (see Appendix D.4). Thus, potential equilibrium multiplicity does not undermine the effectiveness of cross-subsidization in aligning driver incentives with the cumulative driver profit metric. While equilibrium multiplicity under rider welfare may lead to the realization of an equilibrium allocation with a lower rider welfare than the optimal, this issue is unlikely to arise in the empirically relevant regimes described above.

D.4 Uniqueness of Cumulative Driver Profits and Equilibrium Driver Allocation

In this section, we show that the cumulative driver profits are the same at any equilibrium allocation for the rider demand functions specified in Equation (3). Moreover, if the quadratic coefficient in Equation (3) is strictly positive (i.e., if the free-flow cost difference $S_i > 0$), we show that the resulting equilibrium driver allocation is unique.

To see this, consider two equilibrium driver allocations $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$, where $\mathbf{x}^{(1)} \neq \mathbf{x}^{(2)}$. Then, since $\sum_{i\in[n]} x_i^{(1)} = D = \sum_{i\in[n]} x_i^{(2)}$, it follows that there exist non-empty sets $L_1 = \{i : x_i^{(1)} > x_i^{(2)}\}$ and $L_2 = \{i : x_i^{(1)} < x_i^{(2)}\}$. Moreover, define $L_3 = \{i : x_i^{(1)} = x_i^{(2)}\}$ and let $\pi^{(1)}$ be the equilibrium per-driver profit under $\mathbf{x}^{(1)}$, where $\pi_i(x_i^{(1)}) = \pi^{(1)}$ for all routes i with $x_i^{(1)} > 0$. Analogously define $\pi^{(2)}$. Then, for some routes $i_1 \in L_1$ and $i_2 \in L_2$, we obtain the following relation for the per-driver profit:

$$\pi_{i_1}(x_{i_1}^{(2)}) \overset{(a)}{\geq} \pi_{i_1}(x_{i_1}^{(1)}) = \pi^{(1)} \overset{(b)}{\geq} \pi_{i_2}(x_{i_2}^{(1)}) \overset{(c)}{\geq} \pi_{i_2}(x_{i_2}^{(2)}) = \pi^{(2)} \overset{(d)}{\geq} \pi_{i_1}(x_{i_1}^{(2)}), \tag{10}$$

where (a) and (c) follow by the monotonicity of the per-driver profit (see Corollary 2), and (b) and (d) follow since $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ are equilibrium driver allocations. Since the left and right most terms in the above sequence of inequalities are the same, it follows that each of the above inequalities is met with an equality and, in particular, $\pi^{(1)} = \pi^{(2)}$.

Given this, we now compare the cumulative profits under the two equilibrium driver allocations. In particular:

$$P(\mathbf{x}^{(1)}) = \sum_{i=1}^{n} \pi_i(x_i^{(1)}) x_i^{(1)} = \pi^{(1)} \sum_{i=1}^{n} x_i^{(1)} \overset{(a)}{=} \pi^{(2)} \sum_{i=1}^{n} x_i^{(2)} = \sum_{i=1}^{n} \pi_i(x_i^{(2)}) x_i^{(2)} = P(\mathbf{x}^{(2)}),$$

where (a) follows as $\pi^{(1)} = \pi^{(2)}$ and the same number of drivers are allocated under both allocations. Thus, we have that the cumulative profits under any two equilibrium allocations are the same.

Next, we show that the equilibrium allocation is unique if the quadratic coefficient in Equation (3) is strictly positive. To see this, first note that under this condition, the per-driver profits is strictly monotonically decreasing in the driver allocation. Then, the only way for the inequalities

in Equation (10) to be held with equality is if $x_{i_1}^{(1)} = x_{i_1}^{(2)}$ for all $i \in L_1$ and $x_{i_2}^{(1)} = x_{i_2}^{(2)}$ for all $i \in L_2$, a contradiction. Hence, it must follow that $\mathbf{x}^{(1)} = \mathbf{x}^{(2)}$, which establishes our claim.

E Additional Details on Numerical Experiments

E.1 Details on Route Selection

Our original dataset on the informal transit system in Nalasopara contains information on twenty-one routes. However, some of these routes serve destinations in the same geographic area, for which only a single destination coordinate is available. In these cases, we aggregate routes by constructing weighted averages of route characteristics, including trip time, distance, and fare. In addition, geographic coordinates are unavailable for one route, which are required for demand estimation. Thus, for our experiments, we focus on a final set of eighteen routes.

E.2 Rider Demand Calibration Procedure

We calibrate the demand Λ_i for each route i using population statistics from the 2011 Indian Census, the most recent census conducted in India, which reports data at the level of administrative subdivisions referred to as *wards*. To this end, we first assign each route i to a ward based on the geographic coordinates of its destination stop. Since we study the evening commute period in our experiments, we take the Nalasopara railway station to be the common origin of all routes with the destination being the respective residential neighborhoods, reflecting that the majority of trips during the study period are home-bound.

Next, we estimate the population that can feasibly access each route by defining a pedestrian catchment area around its destination stop based on walking accessibility. Specifically, for routes with total length of at most 5 km, we assume users are willing to walk five minutes to access the route, while for longer routes we assume a walking radius of ten minutes. Using a walking speed of 1.3 m/s, we obtain circular catchment areas around each destination stop.

However, since these circular catchment areas may overlap across destination stops, we then generate disjoint catchments by assigning users to their nearest stop, ensuring that each user is associated with at most one route. To do so, we use Monte Carlo nearest-neighbor sampling. Specifically, for each circular catchment area, we repeatedly draw a large number of points uniformly at random, interpreting each draw as a representative residential location, and determine the fraction of those points that are geographically closest (in Euclidean distance) to that catchment area's destination stop. The fraction of points assigned to a given destination stop determines the share of the circular catchment area that can access that stop. Assuming that population density within the circular catchment area is equal to the ward-level population density of the ward to which the destination stop belongs, we compute the total population that can access each route by multiplying the total population within the circular catchment area by this fraction.

Finally, we assume that 40% of the population commute daily by train, consistent with standard mode share estimates in the Mumbai Metropolitan Region. Multiplying this commuting share with the above estimated population that can access each route yields the total mass of users Λ_i seeking to make trips on that route during our study period.

E.3 Implementation Details to Compute Equilibrium Driver Allocation

We now describe a procedure to compute an approximate equilibrium driver allocation, where we leverage the fact that, in equilibrium, all routes with a strictly positive mass of drivers have the same per-driver profit. Specifically, we search over a discretized set of possible equilibrium profit levels within a bounded interval $[\underline{\pi}, \bar{\pi}]$, where $\underline{\pi}, \bar{\pi}$ denote lower and upper bounds on the equilibrium per-driver profits. For each candidate per-driver profit level on the discretized grid, we compute the corresponding allocation of drivers across routes that would result in per-driver profits at that level and sum these allocations across the routes to obtain the total number of drivers willing to operate at that per-driver profit level. Then, we obtain an approximate equilibrium per-driver profit as the value on the discretized grid for which the resulting aggregate driver allocation is closest to the actual mass of drivers D, with the corresponding route-level allocation of drivers across routes taken as the approximate equilibrium driver allocation.