

On the Right Track? Designing Optimal Public Transit Contracts

Matías Navarro*
Cornell University

Job Market Paper | September 4, 2025

[Please click here for the latest version](#)

Abstract

Private transit provision under weak regulation creates two key inefficiencies: market power leading to underprovision, and uninternalized network effects where firms fail to consider how service on one route affects demand across the network. To address these market failures, governments increasingly use contracts with private operators featuring quality targets and route bundling. This paper studies how these contract instruments should be designed to maximize welfare. The key challenge is that underlying market failures push policy in opposite directions: increasing competition mitigates market power but can fragment service and weaken network coordination. I exploit unique quasi-experimental variation from Santiago, Chile's large-scale 2022 contract reform, which imposed stricter quality targets and rebundled routes among private operators. Using high-frequency GPS tracking data from 373 bus routes and an event-study difference-in-differences design, I find that stricter quality targets improve service regularity by 17.1% and increase ridership by 7.5%. I develop and estimate a structural model incorporating traveler demand and operator costs to study optimal contract design. Results show welfare losses from monopoly pricing outweigh network gains from single-operator coordination, with optimal market structure involving four to five competing firms. Route bundle composition plays a critical role in preserving network connectivity while balancing costs and network effect internalization.

Keywords: Public Transit. Private Provision. Contracts. Network Effects. JEL: D22, D62, L91, L92, R41, R42

*Acknowledge Shanjun Li, Todd Gerarden, Andrew Waxman, Stuart Rosenthal, Seba Tamblay, Beia Spiller. Thankful to seminar participants at Cornell University, and comments by XX. Email: min26@cornell.edu.

1 Introduction

Governments delegate a large share of public service delivery to private firms, with public procurement accounting for 12% of global GDP (Bosio *et al.*, 2022). This is especially true in sectors like energy distribution, broadband, and transportation, where service quality and access are critical. In the case of private provision of public transit, absence of regulation often leads to inefficiencies, including high prices, poor service quality, and gaps in network coverage (Conwell, 2023; Mbonu and Eaglin, 2024). Economic theory explains these inefficiencies through two market failures: market power, which leads to suboptimal service levels and markups, and uninternalized network effects, where independent operators fail to account for network complementarities.

In recent decades, many transit agencies have adopted contractual approaches to regulate service provision (see Figure XX for examples across cities). These contracts typically rely on two design choices. The first involves quality targets, which impose minimum service standards to curb underprovision, creating a trade-off between service quality and provision costs as firms face performance penalties when GPS monitoring reveals deviations. The second concerns route bundling, specifically, how many routes to group into each contract to balance competition and coordination. Smaller bundles foster price competition through more bidders but can fragment the network; larger bundles preserve coordination but may reduce competitive pressure. Despite their widespread use, we know little about how quality targets and route bundling interact through their respective trade-offs. This leaves unresolved how to design contracts that maximize welfare when market power and network effects pull in opposing directions.

In this paper, I study how quality targets and route bundling interact in addressing market power and network effects in public transit contracts. The answer to optimal contract design depends on the relative importance of each distortion for welfare: if consumers are more sensitive to prices, contracts should foster more competition by fragmenting the network into smaller bundles; if consumers place greater value on service quality and network redundancy, fewer and larger bundles may be preferred. I find that [main result will go here].

I study this question in the context of Santiago de Chile's public transit system, Transantiago, one of the first large-scale systems to adopt service contracts for bus provision. Transantiago incentivizes private provision of bus service through competitively awarded operating contracts. The system serves approximately 3.5 million trips per day, with several private firms operating the bus network under contract and a public firm running the subway. Contracts compensate firms for service provision based on kilometers traveled and passengers transported, and specify obligations on service quality such as frequency and regularity, typically over periods of five to seven years.

I combine a rich set of operational data from Transantiago, including a Household Travel Survey, smart card fare validations, and GPS records from all transit vehicles. These sources allow me to link travelers' mode and route choices to firm-level service quality—such as frequency and headway regularity—at fine spatial and temporal resolution. The 2023 redesign of Transantiago's contracts provides a unique empirical setting for studying how firms respond to changes in contract structure. The reform affected over 30 percent of the network, modifying both the size and composition of route bundles and the stringency of quality targets. Moreover, the insights from this setting are broadly applicable: Santiago's contractual framework is representative of other large cities that rely on competitive tendering to regulate privately operated transit networks.

I begin by analyzing the market failures that arise in public transit provision using a stylized framework adapted from Barwick *et al.* (2024), who study attribute-based subsidies in a standard price-attribute space. In my setting, a monopolist chooses both prices and service attributes—such as frequency and headway regularity—to maximize profits, while a social planner chooses the same variables to maximize welfare. A key assumption inherited from Barwick *et al.* is that utility is additively separable in prices and service attributes, which decouples pricing from quality decisions and simplifies the analysis. I extend the framework by allowing consumers' willingness to pay for service attributes on one route to depend on the service levels offered across the entire network, introducing demand-side network effects. Both the monopolist and the social planner internalize these network effects when optimizing over the full network. However, the monopolist underprovides service attributes because it does not account for the environmental externality of shifting travelers from cars to public transit. As a result, even though both planners internalize network effects, the monopoly's lower service quality levels reduce the strength of those spillovers in equilibrium. In addition, the monopolist charges a markup due to market power, further reducing welfare.

I use this framework to study two contractual instruments: quality targets and route bundling. Quality targets can be effective in restoring efficient service levels—if penalties are designed to reflect the environmental externality, the monopolist's chosen service attributes align with those of the social planner. However, because the penalty affects only quality choices, it does not eliminate the markup distortion. Route bundling addresses the pricing distortion by introducing competition: the network is divided into bundles, which are allocated via competitive tendering to reduce markups. Yet this comes at the cost of coordination. Once the network is fragmented, firms no longer internalize demand spillovers across bundles, weakening network effects and reintroducing inefficiencies in service quality even under quality targets. This creates a core policy trade-off: increasing competition reduces prices but undermines coordination. How to strike this balance—and how to design bundles to mitigate the loss of network internalization—are ultimately

empirical questions.

2 Theoretical Framework

In this section, I present a theoretical framework for optimal regulation of transit service provision in the presence of market power, externalities, and network effects. As a key departure from the literature on transportation regulation under market power, firms in this model respond to government policies by adjusting both prices and service attributes of their route networks. This framework enables me to characterize the welfare implications of two policy instruments: quality targets and competitive route bundling mechanisms.

The framework involves N differentiated transit routes, indexed by j , each characterized by a K -element service attribute vector $x_j = (x_j^1, x_j^2, \dots, x_j^K)$ (e.g., frequency, regularity) and price P_j . Route j generates external benefits $e_j(x_j) > 0$ due to reductions in pollution and traffic congestion externalities when travelers shift from more externality-intensive transportation modes. Both consumer willingness-to-pay $B_j(x)$ and the marginal cost of service provision $C_j(x_j)$ depend on service attributes, where $x = (x_1, \dots, x_N)$ captures network effects arising from the interdependence of route choices.

Throughout the theoretical analysis, I assume that consumer demand exhibits additive separability between price disutility and service attributes: $Q_j(P, x) = Q_j(P_j - B_j(x))$. This demand function is motivated by discrete choice models where utility depends on prices and attributes additively. Following Barwick *et al.* (2024), the additive separability makes firms' choices of prices and attributes independent and greatly simplifies the model. Most importantly, it enables me to characterize optimal policy design in the presence of market power and network externalities, which has not been done in the transit regulation literature. A limitation of the additivity assumption is that the marginal value of service attributes is the same across consumers, which rules out the Spence distortion in quality provision.

The theoretical analysis compares the privately and socially optimal outcomes and discusses the choice of regulatory instruments to rectify market failures in transit provision. To build intuition, I first analyze the baseline case of a social planner before examining monopoly provision without regulation, then extending to quality regulation and competitive bundling mechanisms.

2.1 Social Planner

The social planner maximizes social welfare that consists of consumer surplus, producer surplus, and externalities. Consider a social planner that chooses prices and service

attributes to maximize social welfare:

$$\max_{P,x} \quad SW(P,x) = \sum_{j=1}^N \left[\underbrace{\int_0^{Q_j(P,x)} \left(B_j(x) + Q_j^{-1}(s) - P_j \right) ds}_{\text{Consumer surplus}} \right. \\ \left. + \underbrace{(P_j - C_j(x_j)) Q_j(P,x)}_{\text{Producer surplus}} + \underbrace{\phi \cdot e_j(x_j) Q_j(P,x)}_{\text{Externality}} \right] \quad (1)$$

The socially optimal prices P_j^* and service attributes x_j^* satisfy the following first-order conditions:

$$[P_j] : \quad P_j^* - C_j(x_j^*) + \phi \cdot e_j(x_j^*) = 0 \quad (2)$$

$$[x_\ell^i] : \quad \underbrace{\left[\frac{\partial B_i(x^*)}{\partial x_\ell^i} - \frac{\partial C_i(x_i^*)}{\partial x_\ell^i} + \phi \cdot \frac{\partial e_i(x_i^*)}{\partial x_\ell^i} \right] Q_i}_{\text{Direct effect on route } i} + \underbrace{\sum_{j \neq i}^N \left[\frac{\partial B_j(x^*)}{\partial x_\ell^i} Q_j \right]}_{\text{Network effect}} = 0 \quad (3)$$

The first-order conditions reflect that service attributes are chosen to maximize per-unit social surplus, $B_j(x) - C_j(x_j) + \phi \cdot e_j(x_j)$, while prices reflect the social cost of service provision, $C_j(x_j^*) - \phi \cdot e_j(x_j^*)$. The socially optimal price P_j^* eliminates quantity distortions, while the socially optimal attributes x_j^* internalize both environmental externalities and network effects across routes.

2.2 Monopoly without Regulation

Consider a monopolist that controls the entire transit network and chooses prices and service attributes to maximize profit:

$$\max_{P,x} \quad \Pi(P,x) = \sum_{j=1}^N [(P_j - C_j(x_j)) Q_j(P,x)] \quad (4)$$

The privately optimal prices P_j^m and service attributes x_j^o satisfy:

$$[P_j] : \quad \frac{P_j^m - C_j(x_j^o)}{P_j^m} = \frac{1}{\varepsilon_{P_j}} \quad (5)$$

$$[x_\ell^i] : \underbrace{\left[\frac{\partial B_i(x^o)}{\partial x_\ell^i} - \frac{\partial C_i(x_\ell^o)}{\partial x_\ell^i} \right] Q_i}_{\text{Direct effect on route } i} + \underbrace{\sum_{j \neq i}^N \left[\frac{\partial B_j(x^o)}{\partial x_\ell^i} Q_j \right]}_{\text{Network effect}} = 0 \quad (6)$$

where ε_{P_j} is the price elasticity of demand for route j . The first-order conditions differ from Equations (2) and (3) in three important ways.

First, the monopolist sets $P_j^m > P_j^*$, resulting in underprovision of quantity relative to the social optimum due to markup pricing under downward-sloping demand. Second, the monopolist does not internalize external benefits $\phi \cdot e_j(x_j)$, leading to inefficient service quality choices where routes may be under-served in terms of environmentally beneficial attributes. Third, while the monopolist controls all routes and thus internalizes network effects through the term $\sum_{j \neq i} \frac{\partial B_j(x^o)}{\partial x_\ell^i} Q_j$, it chooses suboptimal attribute levels because it ignores externalities. Consequently, it does not generate the full social value of network coordination that would justify higher service levels.

2.3 Monopoly with Quality Regulation

Suppose the government introduces quality targets \bar{x}_j for each route j to address externality distortions. To incentivize compliance, the firm faces penalties when deviating from these targets:

$$\mathcal{P}_j(x_j, \bar{x}_j) = \tau \cdot s_j(x_j, \bar{x}_j),$$

where $\tau > 0$ is a penalty strength parameter and $s_j(\cdot)$ represents the penalty function. The regulated monopolist maximizes:

$$\max_{P, x} \Pi(P, x) = \sum_{j=1}^N [(P_j - C_j(x_j) - \tau \cdot s_j(x_j, \bar{x}_j)) Q_j(P, x)] \quad (7)$$

The regulated equilibrium prices P_j^z and service attributes x_j^z satisfy:

$$[P_j] : \frac{P_j^z - C_j(x_j^z) - \tau \cdot s_j(x_j^z, \bar{x}_j)}{P_j^z} = \frac{1}{\varepsilon_{P_j}} \quad (8)$$

$$[x_\ell^i] : \underbrace{\left[\frac{\partial B_i(x^z)}{\partial x_\ell^i} - \frac{\partial C_i(x_\ell^z)}{\partial x_\ell^i} - \tau \cdot \frac{\partial s_i(x_\ell^z, \bar{x}_i)}{\partial x_\ell^i} \right] Q_i}_{\text{Direct effect on route } i} + \underbrace{\sum_{j \neq i}^N \left[\frac{\partial B_j(x^z)}{\partial x_\ell^i} Q_j \right]}_{\text{Network effect}} = 0 \quad (9)$$

Quality regulation addresses externality distortions but leaves market power intact. The monopolist continues to set $P_j^z > P_j^*$, leading to underprovision of quantity relative to the social optimum, as the penalty function affects profit levels but not marginal pricing incentives. However, if the marginal penalty equals the marginal externality:

$$\tau \cdot \frac{\partial s_j(x_j, \bar{x}_j)}{\partial x_\ell^j} = \phi \cdot \frac{\partial e_j(x_j)}{\partial x_\ell^j} \quad \text{for all } j, \ell,$$

then the regulated monopolist chooses efficient service quality levels: $x_j^z = x_j^*$. Under this condition, the monopolist also internalizes the full network benefits that a social planner would, since it controls all routes and quality incentives are properly aligned.

2.4 Competitive Route Bundling

I now introduce a two-stage mechanism where competition occurs via auctions over packages of transit routes, followed by decentralized service provision by winning firms. This mechanism addresses market power distortions while potentially affecting the internalization of network effects.

The N transit routes are partitioned into B disjoint packages $\mathcal{R}_b \subset \{1, \dots, N\}$. In the first stage, firms $k \in \mathcal{K}$ submit bids P_{kb} representing the per-passenger price they would charge for operating bundle b . The regulator awards each bundle to the lowest bidder:

$$k(b) = \arg \min_{k \in \mathcal{K}} P_{kb}.$$

In the second stage, each winning firm chooses service attributes x_j for routes $j \in \mathcal{R}_b$ to maximize profit, taking prices as given from the auction outcome.

2.4.1 First-Stage Bidding

In the auction stage, firm k chooses bid P for bundle b to maximize expected profit. Assuming symmetric firms with i.i.d. rival bids following distribution $F(\cdot)$, the probability of winning with bid P is $(1 - F(P))^{n-1}$, where n is the number of bidders. Expected profit is:

$$\max_P \quad (1 - F(P))^{n-1} \cdot \pi_k(P),$$

where $\pi_k(P) = \sum_{j \in \mathcal{R}_b} \left(P - C_j(x_j^{auction}) \right) Q_j$.

The first-order condition for optimal bidding yields:

$$\sum_{j \in \mathcal{R}_b} \left[(P_b - C_j(x_j^{auction})) \cdot \frac{\partial Q_j}{\partial P} + Q_j \right] = (n-1) \cdot \pi_k(P_b) \cdot \frac{f(P_b)}{1 - F(P_b)} \quad (10)$$

This condition balances the marginal profit from raising the bid (left side) against the increased probability of losing the auction (right side). As competition intensifies (n increases), the equilibrium bid P_b decreases toward marginal cost, addressing the quantity distortion from market power.

2.4.2 Second-Stage Service Provision

After winning bundle b , firm k chooses service attributes to maximize profit given the auction-determined price P_b :

$$\max_{x_j; j \in \mathcal{R}_b} \sum_{j \in \mathcal{R}_b} [(P_b - C_j(x_j)) \cdot Q_j(P_b, x)].$$

The optimal service attributes satisfy:

$$[x_\ell^i] : \underbrace{\left[\frac{\partial B_i(x^{auction})}{\partial x_\ell^i} - \frac{\partial C_i(x^{auction})}{\partial x_\ell^i} \right] Q_i}_{\text{Direct effect on route } i} + \underbrace{\sum_{j \in \mathcal{R}_b, j \neq i} \left[\frac{\partial B_j(x^{auction})}{\partial x_\ell^i} Q_j \right]}_{\text{Within-bundle network effect}} = 0 \quad (11)$$

Comparing Equation (11) with (3) reveals that competitive bundling internalizes network effects only within each bundle \mathcal{R}_b , but not across bundles operated by different firms. The mechanism eliminates externality distortions only if combined with appropriate quality regulation, and the efficiency of network effect internalization depends critically on how routes are grouped into bundles.

2.5 Discussion

My theoretical framework illustrates the distinct roles of quality regulation and competitive bundling in addressing market failures in transit provision. Quality targets can eliminate externality distortions and restore efficient network coordination when properly calibrated, but leave market power intact. Competitive bundling addresses market power through price competition but may fragment network coordination depending on bundle design. The optimal regulatory approach depends on the relative importance of these distortions and the administrative feasibility of different policy instruments, questions I address empirically in the next sections.

3 Background and Institutional Setting

My empirical setting is Santiago's public transit system, Transantiago, which uses contract-based provision of urban bus services and it is administered by the *Directorio de Transporte*

Público Metropolitano (DTPM). Since its launch in 2007, Transantiago has delegated operation of its 373 bus routes—served by a fleet of 6,550 buses—to private firms through competitively tendered contracts, while the subway system remains publicly operated. The fare-integrated network serves 3.5 million daily trips across a metropolitan area of 7 million people, with annual budget exceeding US\$1.1 billion.¹ Transantiago is among the largest and most mature examples of bus contracting globally, making it an interesting case for studying how contract design affects service provision in regulated markets. Its relevance has grown as cities such as Singapore, Paris, and Hong Kong have adopted similar models, increasing interest in the design of incentives and oversight in transit systems.

3.1 Tendering

Unlike fixed-price or cost-plus contracts, DTPM allocates bus service contracts through a competitive scoring auction that creates ex-ante competition among private operators. Firms submit bids for contracts that specify bundles of bus routes, each linked to a designated set of bus depots controlled by DTPM. Depot proximity facilitates efficient operations, while capacity constraints limit the feasible combinations of routes that can be assigned to a given location.

Each bid includes two key decision variables: the per-kilometer price the firm is willing to accept and the fleet size it proposes to operate the bundle. These economic components are scored alongside technical criteria—such as the firm’s experience in urban transit, proposed fleet characteristics, and compliance with formal requirements—using a transparent scoring rule in which the economic score receives 80–90% of the total weight. At the time of bidding, firms observe historical ridership, the depot assignment for each bundle, the regulator-set per-passenger price, and the penalty structure that governs performance monitoring. These auction-stage choices shape both the firm’s expected profits and its operational incentives under the contract. In my structural model, I treat the price and fleet size as the firm’s decision variables at the tendering stage.

3.2 Operations

Beyond the auction stage, Transantiago’s contract structure creates ongoing incentives for firms to determine how to operate their assigned routes. Firms receive revenue from two sources: a per-passenger payment set by DTPM and a per-kilometer payment determined by their bid. To earn the latter, firms must actively deploy service—i.e., dispatch vehicles to cover the planned kilometers—making service frequency a choice variable. In addition to frequency, firms also choose how evenly to space vehicle departures, as regulators monitor

¹Add reference

the regularity of headways on each route. Both frequency and regularity affect traveler waiting times and are central to the system's service quality targets.

DTPM monitors these attributes using GPS data transmitted from each bus every 30 seconds. Monitoring is conducted within predefined periods (e.g., 30 min, 60 min, 120 min), and financial penalties are applied when firms deviate from route-level frequency or regularity targets. These penalties are determined by publicly known functions written into each contract and directly reduce firms' revenues. As a result, the auction outcomes—specifically, the per-kilometer price and fleet size—interact with operational choices made during the contract period. In my structural model, I explicitly treat frequency and regularity as firm-level decision variables shaped by these contractual incentives.

4 Data

My empirical analysis combines several administrative and survey datasets to study how transit contract design affects service provision, bidding behavior, and travel decisions. This section describes each dataset and how I use it in the analysis.

4.1 Mode and Route Choices

I combine individual travel choices from the 2012–2013 Household Travel Survey with reconstructed trip attributes based on operational data from the same period. The survey records all daily trips made by approximately 60,000 individuals across 18,000 randomly sampled households in the Santiago metropolitan area. The sample is representative at a fine geographic level across 866 origin-destination zones averaging 1 km² in size. I restrict attention to the 700 urban zones, which account for nearly 50,000 urban trips and cover 80% of work-related trips and 83% of the city's residential population.

For each trip, I observe origin and destination coordinates, traveler demographics (income, car ownership, age, gender, education), and the chosen mode (car, public transit, walking, or other). Figure C.2a shows the distribution of mode choices by income group, revealing substantial heterogeneity in travel patterns: high-income travelers use cars for 68% of trips compared to only 21.8% for low-income travelers. For car trips, I compute monetary costs using fuel prices and maintenance costs from the *Comisión Nacional de Energía de Chile*, assuming an average fuel economy of 8.3 km/liter.² I obtain travel times and distances using the OSRM routing engine based on OpenStreetMap data, using the exact origin and destination coordinates from the survey.

²Fuel economy source: https://energia.gob.cl/sites/default/files/documentos/20240304_informe_final_estandar_-_vehiculos_medianos_vf.pdf

For public transit trips, the survey identifies the exact bus and metro routes used in each leg. However, it does not report leg-level attributes such as fare, travel time, distance, or frequency. To recover these, I use DTPM's trip reconstruction algorithm, developed by Munizaga and Palma (2012), which combines GPS and smartcard data to reconstruct complete transit itineraries and their characteristics.³ This allows me to merge observed survey trips with corresponding travel time, distance, and frequency for representative weeks in 2012 and 2013. I obtain fares separately from administrative fare tables published by DTPM for the corresponding period.

I successfully **match 23,000 out of 27,000 public transit trips** in the survey to operational data, enabling the construction of a rich dataset of individual travel choices linked to detailed service characteristics across all main transport modes. The resulting data reveals clear trade-offs in route choice, as illustrated in Figure C.2b, where faster travel speeds are associated with longer wait times, particularly distinguishing between direct transit routes and those requiring transfers. I use this dataset to estimate travel preference parameters in the structural model.

4.2 Bidding Choices in Tendering

I use administrative records from the 2022 and 2024 tendering processes, in which DTPM awarded bus service contracts through scoring auctions. In 2022, six route bundles were auctioned; in 2024, five additional bundles were tendered using the same mechanism. I observe complete bidding data for all eleven auctions, covering 100 bids in total. For each bid, I observe the identity of the bidding firm, the per-kilometer price bid, the proposed fleet size, and the resulting economic and final scores used to determine contract awards. In all cases, the final score combines a technical and economic component, with the economic score receiving 80–90% weight.

Although the actual scoring rule includes additional components—such as depot infrastructure costs and vehicle technology—I focus on the two primary choice variables: the per-kilometer price and the proposed fleet size. These two components explain roughly 80% of the variation in economic scores across bids and form the core of firms' decision problem in my structural model of the Tendering stage.

This data has two limitations. First, while I observe the full set of submitted bids, the sample size (100 bids) limits the feasibility of a rich nonparametric estimation. I therefore estimate a stylized bidding model that captures the main trade-offs firms face under the

³The algorithm uses all transit card transaction data to reconstruct individual trips. For each card, it matches tap-in times with GPS vehicle locations to infer trip origin, route legs, leg-level travel times, and service attributes such as frequency and regularity. Since users do not tap out, destinations are imputed based on weekly travel patterns, assuming reciprocal morning and evening origins. DTPM reports that the method matches approximately 25 million trips per week, consistent with system-wide tap-in volumes.

observed scoring rule. Second, the 2024 contracts had not yet entered into operation during my observation window. As a result, I estimate cost parameters using realized operations only for the 2022 tenders, which I link to awarded bids and depot assignments.

4.3 Frequency and Headway Regularity Choices

I use GPS data from DTPM covering the entire Transantiago bus network between August 2022 and August 2023. Each vehicle transmits its location every 30 seconds, and the dataset includes the route, operator, vehicle ID, timestamp, and coordinates. These data are used by DTPM for contract monitoring and penalty enforcement, and achieve full coverage across all firms, routes, and days. I am not aware of any gaps or missing vehicles during the observation window.

I construct two key service attributes that enter the firm's decision problem in my structural model of the Operations stage. First, frequency measures the number of vehicle dispatches per route within a monitoring period. Monitoring periods vary in length across contracts but generally span between 30 minutes and 3 hours. Second, regularity captures the consistency of headways between consecutive buses, computed as the coefficient of variation within each monitoring period. Both dimensions shape user waiting times and are directly incentivized by contract enforcement rules. Figure C.3 illustrates these concepts using representative GPS trajectories and shows how firms make strategic trade-offs between frequency and regularity.

For estimation, I aggregate these measures to the route-day level, excluding night service (12am–5am) and dispatches flagged by DTPM as exempt from performance evaluation. This yields a panel of 549,000 route-day observations, balancing data consistency with institutional variation in monitoring periods across contracts.

4.4 Equilibrium Outcomes: Travel Time, Traffic Flow, and Ridership

First, I measure public transit ridership using smartcard transaction data that record all system tap-ins. These data provide route-level boardings in 30-minute intervals for every day in the observation period, covering approximately 3.5 million trips per weekday. I use these data to examine whether more stringent quality targets increase ridership at the route level.

Second, I use vehicle count data from 70 automatic traffic sensors distributed across major corridors in Santiago. These sensors record vehicle flows in 15-minute intervals and provide citywide coverage of key traffic arteries. I complement these data with travel speed information from Google Maps API, matched to the same locations and timestamps. These traffic and speed data span the period from August 1 to September 17, 2022, and were

originally collected by Bordeu (2023) using sensors maintained by the Chilean Ministry of Transportation.

Figure C.4 illustrates relationships in these data: Panel (a) shows the traffic flow-speed relationship that underlies the road technology, while Panel (b) demonstrates the association between transit service reliability and ridership demand.

I use this combination of traffic flow and speed data to estimate a road technology model that maps vehicle flows to travel times. This relationship allows me to predict how equilibrium travel speeds adjust when transit service quality or car usage changes endogenously in the model.

5 Descriptive Evidence

I exploit a major contract retendering process launched by DTPM at the end of 2022, which reassigned more than 40% of the routes due to contract expirations. These expirations followed the original 10–12 year contract cycle, making the selection of routes for retendering plausibly orthogonal to contemporaneous performance. As part of the retendering, DTPM introduced two changes to contract design. First, it split three large route bundles into six smaller ones to foster competition and improve operational flexibility. Second, it made quality targets more stringent by standardizing monitoring periods from uneven 2–3 hour intervals to uniform 30-minute windows.

These changes affected incumbent operators asymmetrically: some firms lost nearly all routes due to contract expiration, others lost only a subset, and one firm retained all existing operations. This variation generates three empirical margins that I use to identify different effects of contract design.

First, I study the effect of stricter quality targets on service outcomes by comparing routes that transitioned to the new monitoring regime with those that did not, holding fixed route, depot, firm, and time-specific factors, as well as controlling for time-variant route characteristics. Second, I isolate the impact of depot load on performance by exploiting changes in the number of routes operated from a depot—among firms that experienced partial route loss—while keeping route-level contracts and quality targets constant. Third, I explore how the change in bundle size influenced bidding behavior by comparing per-kilometer price bids across the two tendering regimes.

5.1 Effect of Stricter Quality Targets on Service Attributes

Figure 1

I quantify the effects of the policy on service outcomes by using an event study

difference-in-differences design. I estimate the following regression:

$$\log(y_{rdkt}) = \sum_{l \neq -1} \beta_l \text{Treated}_r \mathbb{1}\{t = l + 1\} + \mu_r + \lambda_k + \delta_t + X'_{rt} \phi + \varepsilon_{rdkt} \quad (12)$$

where y_{rdkt} represents service outcomes for route r operated by firm k on day t : (a) frequency and (b) coefficient of variation of headways. Treated_r is an indicator for whether route r transitions to stricter quality targets. X_{rt} includes time-varying route characteristics such as route length and average speed. The specification includes route fixed effects (μ_r), firm fixed effects (λ_k), and time fixed effects (δ_t). Standard errors are clustered at the firm level.

Figure 2 displays the results of estimating Equation (12). In the pre-period, the coefficients are small and not significantly different from zero, confirming the parallel trends assumption. After the stricter quality targets were implemented, the effects vary across outcomes. Panel (a) shows that frequency remains largely unaffected, with coefficients close to zero throughout the post-treatment period. Panel (b) demonstrates that service regularity (measured by the coefficient of variation) improves significantly, with a reduction of approximately 17.1% that persists throughout the sample period. Panel (c) reveals that ridership increases, with coefficients rising to around 7.5% in the post-treatment period. The impacts appear immediately after implementation and remain fairly stable over time, suggesting that stricter quality targets led to sustained improvements in service regularity and ridership, while leaving service frequency unchanged.

5.2 Effect of Increased Fragmentation on Transit Network Attractiveness

Figure 3

5.3 Effect of Increased Auction Competition on Bid Prices

Observed prices per kilometer paid to operators include both structural and institutional components: infrastructure payments, per-passenger transfers, and features specific to the contract regime (e.g., penalties and bundling rules). To isolate the component of prices that reflects firm-level behavior and potentially strategic pricing, I construct a clean price measure that simulates what prices would have looked like under a common, pre-reform institutional framework.

I begin by estimating the following equation at the firm-month level:

$$\log(p_{kt}^{\text{dist}}) = \alpha_k + \lambda_t + \beta_1 \log(\text{Fleet Pay}_{kt} + 1) + \beta_2 \log(\text{Infra Pay}_{kt} + 1) + \beta_3 \log(p_{kt}^{\text{pax}}) \quad (13)$$

$$+ \beta_4 \log(\text{Passengers}_{kt}) + \beta_5 \log(\text{Kilometers}_{kt}) + \beta_6 \text{Contract Type}_k + \varepsilon_{kt}, \quad (14)$$

where p_{kt}^{dist} is the observed per-kilometer price to firm f in month t , Fleet Pay $_{kt}$ and Infra Pay $_{kt}$ capture capital payments, p_{kt}^{pax} is the per-passenger price, and fixed effects α_k and λ_t absorb firm fixed effects and month fixed effects, respectively.

I then predict counterfactual prices for each firm-month under a common “old contract” regime by setting Fleet Pay = Infra Pay = 0, Contract Type = 0, and assigning p_{kt}^{pax} the average value observed among firms that did not retender. The resulting predicted value, $\log(\hat{p}_{kt}^{\text{dist}})$, reflects what each firm’s price would have been in the absence of structural reforms, holding bundle size and ridership fixed.

To estimate the effect of the reform on cleaned prices, I implement a difference-in-differences design focusing on “stable firms”—those who operated both before and after the retendering but under different contractual terms.

I estimate:

$$\log(\hat{p}_{kt}^{\text{dist}}) = \delta \cdot \text{Treated}_k \cdot \text{Post}_t + \alpha_k + \lambda_t + \varepsilon_{kt}, \quad (15)$$

where $\text{Treated}_k = 1$ for the retendered firms, and $\text{Post}_t = 1$ after their respective contract transitions. The coefficient δ captures the change in cleaned prices attributable to the reform, net of contract structure and input payments.

I find that, once institutional features are neutralized, the estimated effect of treatment on cleaned prices is small but statistically significant: $\delta = -0.024$ (s.e. = 0.008). This suggests that retendering into smaller bundles under the new regime is associated with approximately a 2.4% decrease in per-kilometer prices, consistent with increased competitive pressure.

5.4 Demand Responses

6 Empirical Model

The reduced-form analysis confirms that contract parameter changes significantly affect service quality, ridership patterns, and operator bidding behavior in transit markets. To evaluate the welfare implications of different contract designs and to quantify the underlying mechanisms driving these effects, I now turn to an equilibrium model that features three key economic agents: travelers making mode and route choices, private transit operators competing for route bundles and making operational decisions, and a transit agency designing contract parameters.

The model characterizes the complex interactions between travel demand and transit service provision in a regulated oligopolistic environment. On the one hand, travelers’ mode and route choices determine ridership patterns, which affect operators’ revenues and hence their incentives for service provision. On the other hand, operators’ choices on service attributes—particularly frequency and headway regularity—directly affect the

attractiveness of different travel options through travelers' wait times, and therefore consumer surplus. The equilibrium nature of my model allows counterfactual simulations of different contract designs and provides direct comparative statics of service quality, ridership, traffic congestion, and social welfare across policy alternatives.

My approach addresses a gap in the existing literature on transit regulation, which has primarily focused on either completely decentralized markets (e.g., minibuses in African cities) or public monopolistic provision (e.g., transit systems in US cities or BRTs in developing countries). No previous study has examined a regulated oligopolistic environment where service provision is delegated to private firms operating under performance contracts with quality targets and competitive route bundling mechanisms. [There is one exception, which is Marra and Oswald (2024), talk about this]

The model assumes that travelers' origins and destinations are determined *ex ante* and examines mode and route choices given these fixed trip patterns. This assumption is motivated by three considerations. First, for most urban trips, origins and destinations reflect longer-term residential and employment decisions that respond slowly to transportation policy changes. Second, the contract parameter variation I exploit occurred over a relatively short time horizon, making it unlikely that fundamental location patterns adjusted significantly. Third, incorporating joint location-transportation choices would substantially complicate the empirical analysis given the rich individual-level preference heterogeneity I incorporate into the model.

My approach offers several methodological advantages over existing studies of transportation policy. First, my rich operational data on service provision decisions allows me to estimate cost parameters directly from observed operator choices rather than relying solely on auction bids, which is important given the limited bidding data available in my setting. Second, the discrete route choice framework enables me to decompose travelers' expected utility from transit into direct effects (utility from the chosen route) and network effects (utility from having other routes as alternatives), providing novel insights into the welfare consequences of network fragmentation under competitive bundling. [I think the third methodological advance is the source of variation I use to identify the parameters, which is quasi-experimental and not that many structural papers have this]

The model has several limitations that I acknowledge. I abstract from dynamic considerations and treat the analysis as a static equilibrium. I assume travelers' origins, destinations, and the physical route network are fixed, focusing on how existing routes can be reallocated across different operators rather than network redesign. I model road congestion affecting car travel times but abstract from crowding effects on transit vehicles. Finally, I focus on the three dominant transportation modes (car, public transit, walking) and abstract from smaller-share alternatives like taxis and ride-sharing services.

The empirical model proceeds in two stages. Travelers first choose among transporta-

tion modes (car, public transit, walking) based on the expected utility from each option, then those choosing transit select among available routes. Transit operators participate in auctions for route bundles by submitting bids that maximize expected profits, then winning operators choose service attributes (frequency and headway regularity) for each route in their bundle to maximize actual profits given the auction-determined contract terms. The equilibrium is characterized by market clearing conditions where travel demand depends on prices and service attributes, and travel times depend on traffic flows determined by the joint distribution of mode and route choices. This framework allows me to simulate counterfactual scenarios examining optimal quality target design and the welfare-maximizing number and composition of route bundles under competitive tendering.

6.1 Travelers

I specify a nested discrete choice model where travelers first choose among transportation modes and then, conditional on selecting public transit, choose among available routes. This nested structure allows me to capture both the direct utility from chosen transportation options and the network effects arising from the availability of alternative routes within the transit system.

6.1.1 Stage 1: Mode Choice

For a given origin-destination pair (market m), traveler i 's utility from choosing mode j is given by:

$$u_{ijm} = \theta_{ij} + v_{jm} + \chi_{jm} + \epsilon_{ijm} \quad (16)$$

where θ_{ij} is a mode-specific random coefficient that varies across individuals, v_{jm} is the deterministic utility component, χ_{jm} represents observable mode-market characteristics, and ϵ_{ijm} is an idiosyncratic error term assumed to follow a Type I extreme value distribution.

The choice set includes three transportation modes: car, public transit, and walking (the outside option). Mode availability depends on individual circumstances and infrastructure access. For car availability, the Household Travel Survey data indicates whether individuals have access to a vehicle. For transit accessibility, I identify all bus stops and subway stations within 1,000 meters of both trip origins and destinations, determining the set of transit routes serving these access points.

The deterministic utility component v_{jm} varies by transportation mode:

$$v_{jm} = \begin{cases} \mathbb{E} \max_{r \in \mathcal{R}_m} u_{rm}, & \text{if } j = \text{transit} \\ \alpha_{\text{price}} P_{jm} + \alpha_{\text{veh}} T_{jm}^{\text{veh}}, & \text{if } j = \text{car} \end{cases} \quad (17)$$

For car travel, the deterministic utility depends on trip cost P_{jm} and travel time T_{jm}^{veh} . Trip costs include both fuel expenses and maintenance costs.⁴ For public transit, the deterministic utility equals the expected maximum utility across all available routes \mathcal{R}_m in market m , which I discuss in Section 6.1.2, where I provide details of the route choice model. Walking serves as the outside option with utility normalized to zero, ensuring model identification and providing a baseline for comparing other transportation modes.

The observable characteristics χ_{jm} include three sets of fixed effects interacted with mode indicators. First, I include mode-specific fixed effects to capture average preferences across transportation options. Second, I incorporate mode-trip characteristic interactions, where trip variables include distance, purpose, time period, and indicators for whether trips originate or terminate in the central business district. Third, I include mode-demographic interactions with traveler characteristics including education, age, and gender.

Individual heterogeneity enters through the mode-specific random coefficient θ_{ij} , which allows baseline preferences for each transportation mode to vary across travelers. This specification maintains tractability while capturing unobserved preference heterogeneity that affects mode choice decisions.

Given the Type I extreme value distribution assumption for ϵ_{ijm} , the probability that traveler i chooses mode j in market m follows a standard logit specification:

$$\mathbb{P}_{ijm} = \frac{\exp(\theta_{ij} + v_{jm} + \chi_{jm})}{\sum_{k \in J_m} \exp(\theta_{ik} + v_{km} + \chi_{km})} \quad (18)$$

where J_m represents the set of available modes in market m .

6.1.2 Stage 2: Transit Route Choice

Building on the model developed by Kreindler *et al.* (2023), consider a traveler choosing among transit route options $h \in \mathcal{H}_m$ in market m (origin-destination pair). Each option h consists of either a direct route r or a combination of two routes r_1 and r_2 connected through a transfer. The utility from option h depends on deterministic components and a random wait time component.

⁴Fuel expenses are calculated as the product of distance, fuel efficiency, and gasoline prices. Maintenance costs are calculated as the product of distance and by per-kilometer maintenance rates

This model assumes that bus arrivals on route r follow a Poisson process with arrival rate λ_r . This generates exponentially distributed wait times with the property that $\Pr(T_r^{\text{wait}} > w) = \exp(-\lambda_r w)$. While Kreindler *et al.* (2023) focus on network expansion effects in a setting with fixed service levels, I extend their model to examine how operators' frequency and regularity choices affect the arrival rates that travelers experience. In my model, the arrival rate λ_r represents the "effective frequency" that travelers experience, which depends on operators' service provision decisions as detailed below.

The utility from option h can be decomposed into deterministic and random wait time components:

$$u_h = v_h + \alpha_{\text{wait}} T_h^{\text{wait}} \quad (19)$$

where the deterministic utility component v_h is given by:

$$v_h = \begin{cases} \alpha_{\text{price}} P_h + \alpha_{\text{veh}} T_h^{\text{veh}}, & \text{if } h = \text{Direct} \\ \alpha_{\text{price}} P_h + \alpha_{\text{veh}} T_{r_1}^{\text{veh}} \\ \quad + \mathbb{E} \max_{r_2} [\alpha_{\text{veh}} T_{r_2}^{\text{veh}} + \alpha_{\text{wait}} T_{r_2}^{\text{wait}}] + \mu_{\text{transfer}}, & \text{if } h = \text{Transfer} \end{cases} \quad (20)$$

where P_h is the fare for option h , T_h^{veh} is in-vehicle travel time, T_h^{wait} is wait time governed by the Poisson arrival process, and μ_{transfer} captures the pure disutility of making a transfer.

For direct routes, utility depends on the price, in-vehicle time, and realized wait time for that route. For transfer options, the traveler experiences utility from the first leg (including its wait time) plus the expected utility from optimally choosing among available second-leg routes at the transfer station. This expected utility formulation captures the option value from having multiple connections available at transfer points.

The fare structure P_h varies across options within markets. While bus-to-bus transfers incur no additional cost, bus-to-subway transfers require an additional reduced fare, generating price variation that helps identify travel cost sensitivity.

Transit Choice Set. I determine route choice sets using detailed origin-destination information from the Household Travel Survey combined with complete network topology data. For each trip, I identify all bus stops and subway stations within 1,000 meters of both origin and destination points. The choice set \mathcal{H}_m for market m includes all direct routes and single-transfer combinations that connect the accessible origin and destination stations. Choice set sizes vary substantially across markets, with a median of 8 options, ranging from a minimum of 1 to a maximum of 32 available combinations.

Choice Probabilities. The exponential wait time assumption yields tractable expres-

sions for choice probabilities and expected utility. Following the results in Kreindler *et al.* (2023), the probability of choosing option h among alternatives ranked by deterministic utility $v_1 \leq v_2 \leq \dots \leq v_H$ is:

$$\lambda_h^{-1} \pi_h = \sum_{i=1}^h e^{-\alpha_{\text{wait}}^{-1} M_i} \frac{e^{v_i \alpha_{\text{wait}}^{-1} \Lambda_i} - e^{v_{i-1} \alpha_{\text{wait}}^{-1} \Lambda_i}}{\Lambda_i} \quad (21)$$

where $\Lambda_i = \sum_{j=i}^H \lambda_j$ and $M_i = \sum_{j=i}^H v_j \lambda_j$, with $v_0 = -\infty$ by convention.

This framework ensures computational tractability and avoids the "red bus, blue bus" problem that affects standard logit models, as combining identical routes with split frequencies yields identical choice probabilities and expected utilities.

Expected Utility. The expected utility from choosing optimally among available routes is:

$$\mathbb{E} \max_{h \in \mathcal{H}_m} u_h = v_{h^*} - \pi_{h^*} \frac{\alpha_{\text{wait}}}{\lambda_{h^*}} \quad (22)$$

where h^* denotes the option with highest deterministic utility v_{h^*} , π_{h^*} is the probability of choosing option h^* , and λ_{h^*} is its effective arrival rate. The influence of alternative options $h \neq h^*$ on expected utility is captured through the choice probability π_{h^*} —when more attractive alternatives are available, the probability of choosing any single option decreases, increasing the expected utility from the entire choice set.

Network Effects. I can separate total expected utility into two components:

$$\mathbb{E} \max_{h \in \mathcal{H}_m} u_h = \underbrace{v_{h^*} - \frac{\alpha_{\text{wait}}}{\lambda_{h^*}}}_{\text{Direct Effect}} + \underbrace{(1 - \pi_{h^*}) \frac{\alpha_{\text{wait}}}{\lambda_{h^*}}}_{\text{Network Effect}} \quad (23)$$

The direct effect represents the utility a traveler would obtain if only their chosen route h^* were available, consisting of the deterministic utility plus expected wait time cost when that route has choice probability equal to one. The network effect captures the additional utility from having alternative routes available, measured by how the presence of other options reduces the probability of choosing any single route ($1 - \pi_{h^*} > 0$).

For policy analysis, this decomposition allows me to evaluate how regulatory changes affect welfare through different channels. When route r is improved (higher frequency, better regularity), travelers who use that route experience direct utility gains. Simultaneously, all travelers who have route r in their choice set—including those who ultimately choose other routes—benefit from improved option value, as reflected in lower choice probabilities π_{h^*} for their chosen alternatives. These network effects are particularly important for evaluating policies that change route bundling and coordination, as fragmenting route operations across different firms may reduce the internalization of these cross-route

externalities.

Linking Supply Decisions to Arrival Rates. The main departure from Kreindler *et al.* (2023)'s model is connecting operators' service provision decisions to the arrival rates λ_r , that determine traveler wait times. Transit operators choose both frequency (buses per hour) and headway regularity (measured by the coefficient of variation of headways) for each route r in their bundle.

I model the relationship between these operational decisions and travelers' experienced service quality using the following engineering relationship:

$$\lambda_r = \frac{f_r}{1 + CV_r^2} \quad (24)$$

where f_r is the number of dispatched buses per interval of time and CV_r is the coefficient of variation of headways. When service is perfectly regular ($CV_r = 0$), travelers experience the full dispatched frequency. As service becomes more irregular, the effective frequency decreases, reflecting longer average wait times due to service bunching and gaps.

This specification captures an important operational trade-off: operators can increase service quality either by running more buses (higher frequency) or by improving service reliability (lower coefficient of variation). The regulatory mechanisms I examine—quality targets and competitive bundling—affect operators' incentives along both dimensions.

6.2 Private Transit Operators

I model private transit operators as making decisions in two sequential stages: first bidding competitively for bundles of routes in government tenders, then choosing service attributes (frequency and headway regularity) for each route in their awarded bundle. This two-stage structure captures the key trade-offs operators face between service quality and costs, and between regulatory compliance and profit maximization. Crucially, operators make bundle-level decisions that account for both shared depot resources and demand spillovers across routes, enabling them to internalize network effects within their bundle while potentially fragmenting coordination across bundles operated by different firms.

6.2.1 Stage 1: Tendering

In the tendering stage, private transit firms compete for the right to operate route bundles through a competitive scoring auction. Each firm k submits a two-part bid $\mathbf{b}_k = (p_k^{\text{dist}}, \mathbf{x}_k)$ consisting of a price per kilometer p_k^{dist} and a contract vector \mathbf{x}_k that includes fleet specifications. Firms observe the regulatory environment before bidding, including quality targets for each route, the price per passenger p_k^{pax} that the regulator will pay operators for

each passenger transported (distinct from passenger fares that travelers pay), and bundle characteristics.

The firm's bidding problem involves maximizing expected profit:

$$\max_{\mathbf{b}_k} \mathbb{E}[\Pi_k] = H(s_k) \cdot \Pi_k(\mathbf{b}_k) \quad (25)$$

where $s_k = s(\mathbf{b}_k)$ is the score associated with firm k 's bid, $H(s_k)$ is the probability of winning given the score, and Π_k represents the operational stage profit conditional on winning the tender.

This structure captures the fundamental trade-off in auction participation: firms can increase their winning probability by submitting more attractive bids (lower prices, better fleet specifications), but this reduces the profitability of operations conditional on winning. The optimal bid balances these competing incentives.

Scoring Rule and Selection Mechanism. The regulator awards bundles using a scoring auction where the highest-scoring bid wins. The scoring rule combines multiple bid components to evaluate overall proposal attractiveness. While the actual scoring mechanism is complex, involving fleet specifications by vehicle type and technology, depot infrastructure costs, and operational pricing, I model the essential structure as:

$$s_k = \beta_1 + \beta_2 p_k^{\text{dist}} + \beta_3 F_k + \mathbf{X}_{\text{bundle}}^\top \boldsymbol{\beta}_4 + \epsilon_k \quad (26)$$

where F_k represents fleet cost components, $\mathbf{X}_{\text{bundle}}$ captures bundle characteristics that affect scoring, and ϵ_k reflects unobserved bid attributes. The parameters β_2 and β_3 determine how operational pricing and fleet investments translate into competitive advantage.

This simplified scoring rule captures the key insight that firms face trade-offs between different bid dimensions. Lower per-kilometer prices increase scores but reduce operational margins, while fleet investments may improve scores but require upfront capital commitments.

Profit Expectations and Strategic Considerations. Firms form profit expectations $\Pi_k(\mathbf{b}_k)$ by solving their anticipated operational stage problem conditional on winning with bid \mathbf{b}_k . Since quality targets, price per passenger, and bundle compositions are announced before bidding, firms can accurately forecast the regulatory environment they would face as operators.

The operational profit function incorporates the revenue structure established by the

bid:

$$\pi_k(\mathbf{b}_k) = p_k^{\text{pax}} \sum_{r \in \mathcal{R}_k} q_r(\lambda_r, \lambda_{-r}) + p_k^{\text{dist}} \sum_{r \in \mathcal{R}_k} f_r s_r - \mathcal{P}_k - C_k - FC_k(\mathbf{x}_k) \quad (27)$$

where p^{pax} is the regulatory-set passenger fare, \mathcal{P}_k represents expected regulatory penalties, C_k captures variable operational costs, and $FC_k(\mathbf{x}_k)$ reflects fixed costs from fleet and infrastructure commitments in the bid.

The win probability function $H(s_k)$ depends on the distribution of rival bids and the competitive environment. Firms with superior operational efficiency, better fleet availability, or local market knowledge may systematically achieve higher scores for given bid parameters, affecting their bidding strategies.

Bundle Characteristics and Selective Participation. Route bundles vary in size, geographic coverage, and operational complexity. Bundle characteristics affect both the scoring mechanism (through $\mathbf{X}_{\text{bundle}}$) and operational profitability through depot assignments, route complementarities, and economies of scale. Larger bundles may offer greater scope for efficiency gains but require more substantial fleet commitments and operational capacity.

Not all firms participate in all bundle auctions. Selective participation reflects firms' capacity constraints, geographic specialization, or strategic focus on particular market segments. This participation pattern affects the competitive intensity across different bundle types and influences the regulator's bundle design incentives.

Equilibrium Bidding Conditions. The first-order condition for optimal per-kilometer pricing reveals the trade-off between competitive positioning and operational margins:

$$\frac{\partial H(s_k)}{\partial s_k} \cdot \frac{\partial s_k}{\partial p_k^{\text{dist}}} \cdot \Pi_k + H(s_k) \cdot \frac{\partial \Pi_k}{\partial p_k^{\text{dist}}} = 0 \quad (28)$$

This condition shows that firms set per-kilometer prices where the marginal benefit from improved winning probability (first term) equals the marginal cost from reduced operational profit (second term). Using the envelope theorem, the profit derivative simplifies to the total service kilometers, linking pricing decisions directly to operational scale.

Similarly, the first-order condition for fleet and infrastructure investments \mathbf{x}_k balances scoring advantages against fixed cost commitments:

$$\frac{\partial H(s_k)}{\partial s_k} \cdot \frac{\partial s_k}{\partial x_{kj}} \cdot \Pi_k = H(s_k) \cdot \frac{\partial FC_k}{\partial x_{kj}} \quad (29)$$

These equilibrium conditions demonstrate how competitive pressure disciplines both

operational pricing and fleet investment decisions, creating incentives for cost efficiency while maintaining service quality standards through the regulatory scoring mechanism.

6.2.2 Stage 2: Operations

After winning route bundles in the competitive tender, transit firms make operational decisions for each route in their bundle to maximize profits subject to regulatory quality targets. Each transit firm k operates a bundle of routes \mathcal{R}_k across multiple depots. To provide service on route r , firms choose two variables: service frequency f_{rdkt} (quantity) and service regularity measured by the coefficient of variation of headways CV_{rdkt} (quality). These decisions jointly determine the passenger experience through wait times and service reliability.

The firm's profit can be expressed as:

$$\Pi_k = \text{Revenue}_k - \text{Penalties}_k - \text{Operational Costs}_k - \text{Fixed Costs}_k \quad (30)$$

This structure captures the central trade-offs between service quality, regulatory compliance, and cost minimization. I now detail each component of this profit function.

Production Technology and Cost Structure. The technological relationship in transit operations dictates that vehicle-hour requirements depend on both frequency and route characteristics:

$$\text{Vehicle-hours}_{rdkt} = f_{rdkt} \cdot \frac{S_r}{v_{rt}} \quad (31)$$

where S_r is the total route length including deadhead distance from depot to route origin, and v_{rt} is average route speed. This total distance differs from the in-service distance s_r for which operators receive distance-based revenue, making depot location a crucial determinant of operational costs.

Service quality, captured by headway regularity, is not a byproduct of quantity but requires dedicated labor-intensive effort. Achieving smooth, regular intervals between buses demands active dispatching, real-time monitoring, and contingency planning—activities that increase operational complexity and cost.

Transit firms face a multi-dimensional production problem where quantity and quality are jointly produced using composite labor inputs. I model the production technology as:

$$L_{rdkt} = \left(f_{rdkt} \cdot \frac{S_r}{v_{rt}} \right)^\gamma \cdot g(CV_{rdkt}) \quad (32)$$

where L_{rdkt} represents composite labor requirements, γ is a technological parameter governing how vehicle-hours translate into labor needs, and $g(CV_{rdkt})$ captures the additional labor required to achieve better regularity.

The function $g(\cdot)$ reflects the operational reality that maintaining consistent headways requires increasingly intensive management effort. Following standard assumptions in the operations literature, I specify $g(CV_{rdkt}) = CV_{rdkt}^{-\phi}$ where $\phi > 0$, such that lower coefficients of variation (better regularity) require exponentially more labor input.

The route-level cost function combines input prices with input quantities:

$$C_{rdkt} = w_k \cdot \left(f_{rdkt} \cdot \frac{S_r}{v_{rt}} \right)^\gamma \cdot CV_{rdkt}^{-\phi} \cdot \varepsilon_{rdkt} \quad (33)$$

where w_k is the firm-specific wage rate and ε_{rdkt} captures route-specific productivity shocks.

Transit operations exhibit scale effects at the depot level. Routes operated from the same depot can share resources such as maintenance facilities, spare vehicles, and supervisory staff. I model depot-level costs as:

$$C_{dkt} = \left(\sum_{r \in \mathcal{R}_{dkt}} C_{rdkt} \right) \cdot |\mathcal{R}_{dkt}|^\rho \quad (34)$$

where $|\mathcal{R}_{dkt}|$ is the number of routes operated from depot d , and ρ captures scale economies ($\rho < 0$) or diseconomies ($\rho > 0$) across routes within a depot.

Revenue Structure. Firm revenue comes from two sources reflecting the dual nature of transit service provision. The demand-based component captures passenger revenue:

$$R_{kt}^{\text{pax}} = p_k^{\text{pax}} \sum_{r \in \mathcal{R}_k} q_{rt}(\lambda_{rt}, \lambda_{-rt}) \quad (35)$$

where ridership q_r depends on the effective arrival rate of route r and the effective arrival rates of other routes λ_{-rt} , capturing demand spillovers within and across route bundles.

The service-based component provides revenue for kilometers of service supplied:

$$R_{kt}^{\text{dist}} = p_k^{\text{dist}} \sum_{r \in \mathcal{R}_k} f_{rdkt} s_r \quad (36)$$

This dual revenue structure creates incentives for both ridership maximization and service provision, with the relative weights determined by the prices p_k^{pax} and p_k^{dist} established in the competitive tender.

Penalties and Quality Targets. The regulator establishes quality targets for frequency \bar{f}_r and service regularity \bar{CV}_r for each route, with penalties for deviations:

$$\mathcal{P}^f(f_r, \bar{f}_r) = \max\{0; \tau^f \cdot (\bar{f}_r - f_r)\} \quad (37)$$

$$\mathcal{P}^{\text{wait}}(f_r, CV_r, \bar{f}_r, \bar{CV}_r) = \max\{0; \tau^w \cdot (W_r - \bar{W}_r)\} \quad (38)$$

where W_r is the realized average wait time and \bar{W}_r is the target wait time. The relationship between operational choices and wait times follows the engineering formula:

$$W_r = \frac{1}{2f_r} \cdot (1 + CV_r^2) \quad (39)$$

This specification captures how both frequency and regularity affect passenger wait times, providing the link between supply-side operational decisions and demand-side service quality.

Equilibrium Conditions. The first-order conditions for optimal frequency and regularity reveal the trade-offs firms face:

$$\frac{\partial C_k}{\partial f_r} = p_k^{\text{pax}} \left[\frac{\partial q_r}{\partial \lambda_r} \frac{\partial \lambda_r}{\partial f_r} + \sum_{u \neq r} \frac{\partial q_u}{\partial \lambda_u} \frac{\partial \lambda_u}{\partial f_r} \right] + p_k^{\text{dist}} s_r - \frac{\partial \mathcal{P}^f}{\partial f_r} - \frac{\partial \mathcal{P}^{\text{wait}}}{\partial f_r} \quad (40)$$

$$\frac{\partial C_k}{\partial CV_r} = p_k^{\text{pax}} \left[\frac{\partial q_r}{\partial \lambda_r} \frac{\partial \lambda_r}{\partial CV_r} + \sum_{u \neq r} \frac{\partial q_u}{\partial \lambda_u} \frac{\partial \lambda_u}{\partial CV_r} \right] - \frac{\partial \mathcal{P}^{\text{wait}}}{\partial CV_r} \quad (41)$$

These conditions show that optimal operational decisions balance marginal costs against three sources of marginal benefit: direct ridership effects on the route, demand spillovers to other routes in the bundle, distance-based revenue, and the avoidance of regulatory penalties. The presence of demand spillovers $\sum_{u \neq r} \frac{\partial q_u}{\partial \lambda_u} \frac{\partial \lambda_u}{\partial f_r}$ in the first-order conditions demonstrates how bundle-level coordination enables firms to internalize demand interactions across routes within their portfolio, a key mechanism for evaluating the welfare effects of alternative bundling designs.

6.3 Road Technology

7 Estimation

7.1 Identification

7.2 Travel Preference Parameters

The parameters for mode and route choice are estimated jointly via simulated maximum likelihood estimation (MLE) using the Household Travel Survey data. The key parameters of interest include preferences for in-vehicle travel time, wait time, monetary costs, transfer penalties, and mode-specific random coefficients for car and public transit. These parameters are crucial for evaluating how regulatory changes affect travel behavior and welfare.

The identifying variation for travel preference parameters comes from the rich variation in choice alternative characteristics and observed choices made by travelers. The cost, travel time, wait time, and transfer penalty sensitivities are informed by how differences in these attributes across choice alternatives affect the relative odds of choosing different modes and routes. With choice sets ranging from 1 to 32 transit route options (median of 8), there is substantial cross-sectional variation in the attractiveness of transit relative to car travel and in the composition of available route alternatives.

For specifications with random coefficients, since there is no closed form for choice probabilities when integrating out the random preference distribution, I simulate the choice probabilities using Monte Carlo integration with a sequence of 100 Halton draws. This simulation approach maintains computational tractability while allowing for unobserved preference heterogeneity that is essential for welfare analysis.

I include an extensive set of fixed effects and control variables to address potential sources of bias in preference parameter estimation. The specification includes mode-specific fixed effects, mode-trip characteristic interactions (including trip distance, purpose, time period, and CBD origin/destination indicators), and mode-demographic interactions with traveler characteristics including education, age, and gender. These interactions control for a rich set of time-varying and location-specific unobservables by travel mode.

7.3 Cost Parameters

Estimation

To estimate the supply model, we need to recover three key technological parameters: (a) the returns to scale in vehicle-hours, γ , (b) the quality-cost elasticity, ϕ , and (c) the depot-level scale parameter, ρ . We also need to identify firm-specific wage rates w_k and route-specific productivity shocks ε_{rdkt} .

I recover these parameters by exploiting the first-order conditions from firms' cost minimization. Taking logarithms of the marginal cost expressions yields:

$$\begin{aligned} \log(MC_{rdkt}^f) = & \log(w_k) + \log(\gamma) + (\gamma - 1) \log(f_{rdkt}) + \gamma \log\left(\frac{l_r}{v_{rt}}\right) \\ & + \phi \log(CV_{rdkt}) + \rho \log(|\mathcal{R}_{dkt}|) + \log(\varepsilon_{rdkt}) \end{aligned} \quad (42)$$

$$\begin{aligned} \log(MC_{rdkt}^{CV}) = & \log(w_k) + \log(\phi) + \gamma \log(f_{rdkt}) + \gamma \log\left(\frac{l_r}{v_{rt}}\right) \\ & + (\phi - 1) \log(CV_{rdkt}) + \rho \log(|\mathcal{R}_{dkt}|) + \log(\varepsilon_{rdkt}) \end{aligned} \quad (43)$$

Identification Strategy

My identification relies on four sources of exogenous variation:

Frequency instruments (Z_{rdkt}^f): Unexpected operational incidents such as vehicle breakdowns, security incidents, or route blockages create exogenous shocks to frequency decisions without directly affecting the fundamental cost structure.

Regularity instruments (Z_{rdkt}^{CV}): Changes in monitoring intensity induced by contract renewals provide variation in regularity requirements that is orthogonal to underlying route productivity.

Scale instruments (Z_{dkt}^{scale}): Reallocation of routes across depots during contract renewals creates exogenous variation in depot-level scale without affecting route-specific characteristics.

Speed instruments (Z_{rdkt}^v): Free-flow speeds during off-peak hours capture the technological speed potential of routes, providing variation in the vehicle-hour requirement that is independent of operational decisions.

I decompose the productivity shock as $\log(\varepsilon_{rdkt}) = \alpha_k + \beta_d + \eta_r + u_{rdkt}$, where α_k , β_d , and η_r represent firm, depot, and route fixed effects, respectively, and u_{rdkt} is the structural error term.

GMM Estimation

I estimate the parameters $\theta = (\gamma, \phi, \rho)$ using the generalized method of moments (GMM). After absorbing fixed effects, my moment conditions require that the instruments be orthogonal to the structural residuals:

$$\mathbb{E}[Z_{rdkt} \cdot u_{rdkt}] = 0 \quad (44)$$

The system is exactly identified with eight moment conditions (four instruments \times two marginal cost equations) and three parameters plus fixed effects.

8 Counterfactuals

9 Conclusions

References

- BARWICK, P. J., KWON, H.-S. and LI, S. (2024). *Attribute-based Subsidies and Market Power: an Application to Electric Vehicles*. Working Paper 32264, National Bureau of Economic Research.
- BORDEU, O. (2023). Commuting infrastructure in fragmented cities, working Paper.
- BOSIO, E., DJANKOV, S., GLAESER, E. and SHLEIFER, A. (2022). Public procurement in law and practice. *American Economic Review*, **112** (4), 1091–1117.
- CONWELL, L. (2023). Subways or minibuses? privatized provision of public transit, working Paper.
- KREINDLER, G., GADUH, A., GRAFF, T., HANNA, R. and OLKEN, B. A. (2023). *Optimal Public Transportation Networks: Evidence from the World's Largest Bus Rapid Transit System in Jakarta*. Working Paper 31369, National Bureau of Economic Research.
- MBONU, O. and EAGLIN, F. C. (2024). Market segmentation and coordination costs: Evidence from johannesburg's minibus networks, working Paper.
- MUNIZAGA, M. A. and PALMA, C. (2012). Estimation of a disaggregate multimodal public transport origin-destination matrix from passive smartcard data from santiago, chile. *Transportation Research Part C: Emerging Technologies*, **24**, 9–18.

Tables and Figures

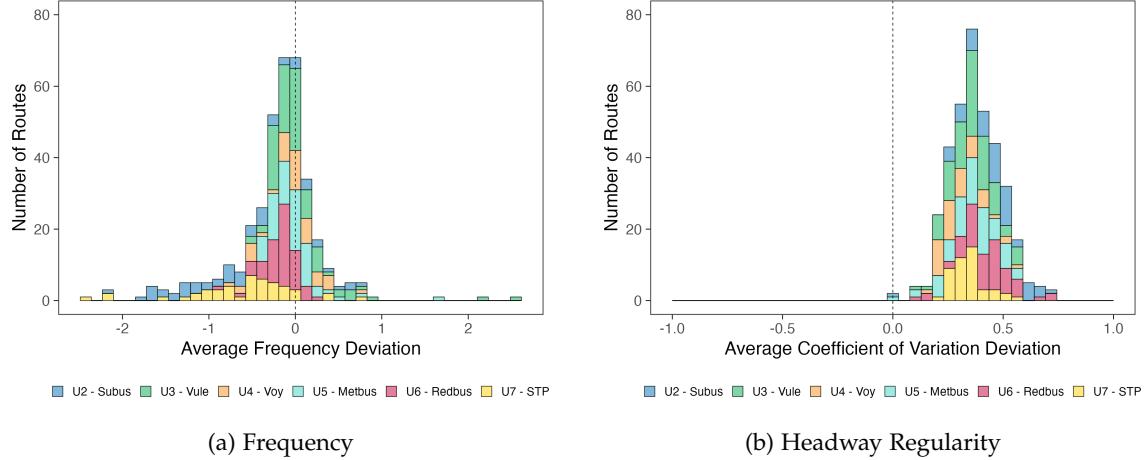


Figure 1: Distribution of Deviations from Planned Transit Service Attributes (Pre-Reform Period).

Notes: Panel (a) shows the distribution of average frequency deviations (observed minus planned frequency) across transit routes during the pre-reform period. Panel (b) displays the distribution of coefficient of variation deviations for headway regularity (observed minus planned variability). The vertical dashed line at zero represents perfect adherence to planned schedules. Colors distinguish routes operated by different firms.

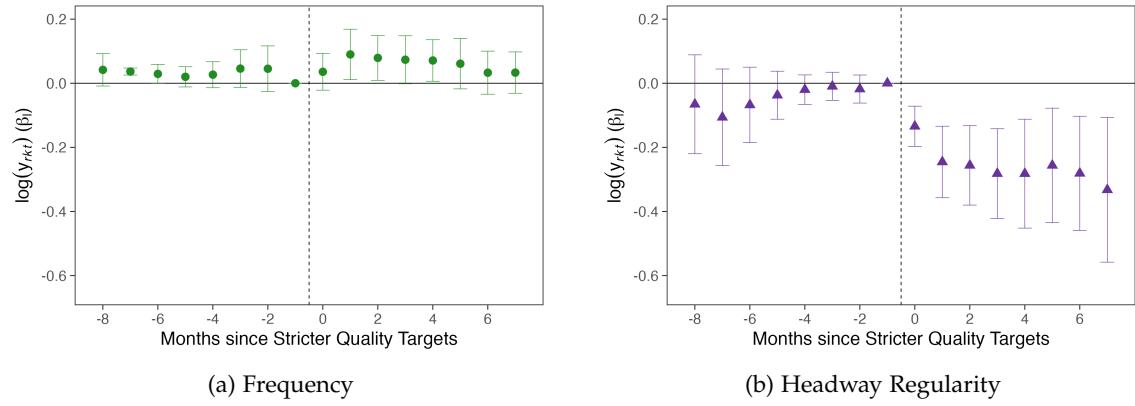


Figure 2: Effect of Stricter Quality Targets on Transit Service Attributes.

Notes: The figures show event study estimates of the impact of stricter quality targets on transit service outcomes, with time measured in months relative to policy implementation. Panel (a) displays the effects on service frequency, showing relatively stable coefficients in the pre-treatment period followed by a slight increase in frequency levels after the policy implementation. Panel (b) shows the effects on headway regularity (measured in coefficient of variation), with pre-treatment estimates fluctuating around zero and post-treatment effects indicating improved headway regularity (negative coefficients suggest lower variability in headways). Error bars represent 95% confidence intervals, and estimates control for firm, route, and time fixed effects.

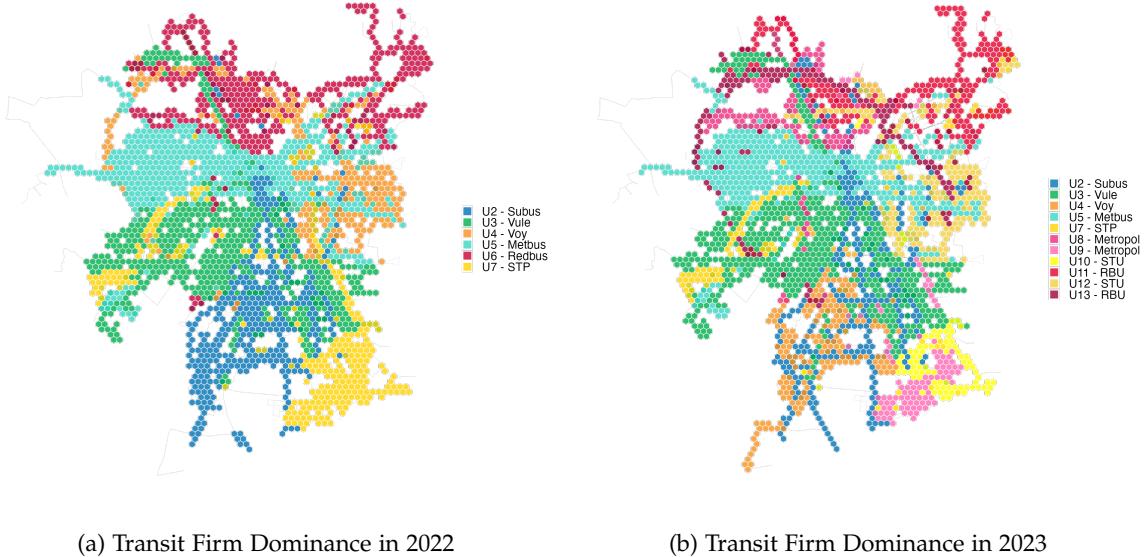


Figure 3: Network Fragmentation Before and After the Reform.

Notes: The figures display transit network fragmentation through geographic dominance of transit firms using a hexagonal grid overlay. Panel (a) shows the transit network in 2022 before the rebundling and tendering process, while Panel (b) shows the transit network in 2023. Each hexagon represents a spatial area where a single transit firm operates the plurality of routes and holds more than 25% of the local route share, with a minimum threshold of 3 routes per hexagon. Colors correspond to different transit operators as shown in the legend. Areas without hexagons indicate either insufficient route density or high fragmentation where no single firm achieves dominance.

Appendix A Data sources

Santiago Household Travel Survey (SHTS) Data

The sample size was determined using the Smith method to establish the number of surveys required in each of the municipality, ensuring estimation of trip generation rates, modal shares, and car ownership rates. Within each municipality, blocks were selected through probability proportional to size (PPS) sampling with replacement. A minimum of 160 surveys were conducted in each municipality—100 on weekdays and 60 on weekends.

The total sample includes 18,264 households, with 11,246 surveyed on weekdays during the normal season and 7,018 surveyed on weekends, covering both normal and summer seasons. Surveys conducted during the normal season took place between July 2012 and November 2013, while summer season surveys were carried out in January and February 2013.

Data collection was conducted in person using mobile devices and involved two visits. During the first visit, surveyors introduced the study, gathered household-level information, and assigned a random travel day for each household member. Participants received a travel diary to record their trips on the assigned day. The second visit involved collecting trip data through in-person interviews with each household member. If any household member was unavailable, the surveyor returned until all trip data were collected. This methodology substantially reduces the underreport issue of surveys based on collecting the trips taken during the preceding 24 hours.

The SHTS includes detailed individual (e.g., age, gender, number of trips, driver license, education level, occupation, and income) and household demographics (e.g., georeferenced location, household size, number of vehicles, rent or mortgage, household income) and trip characteristics (georeferenced origin and destination, purpose, mode choice, travel time, period of day).

I restrict my sample to the municipalities within the urban area of Santiago which involves 13,696 households, 31,735 travelers and 74,166 trips.

I drop the observations that take a mode of transportation different than car, public transit, or walking (15.73%) I drop the observations with purpose other than work, study or other (0.38%). I drop the observations with implausible trip distance and travel time (0.43%). I drop the observations with missing income (0.44%). I drop observations with incomplete cases (0.01%). After doing this, my final trip dataset contains 12,668 households, 27,111 travelers, and 61,574 trips.

The monetary cost for walking is zero. The public transit fare is flat for buses over the day at 1.1 USD and varies by period (peak and off-peak) for subway –1.3 or 1.2 USD–. Transfers are free within a period of 1.5 hours except for transfers between bus and subway, in which the traveler pays a small fee to cover the subway fare. High school students don't pay and college students have a flat fare of 0.4 USD. Fuel cost is a major component of the monetary cost associated with driving. Based on the average fuel economy reported by XXX, I use 0.11 liter/km (9 km/liter). Gasoline prices are XXX USD/liter.

Appendix B Model estimation

Appendix C Figures and tables (for online publication)

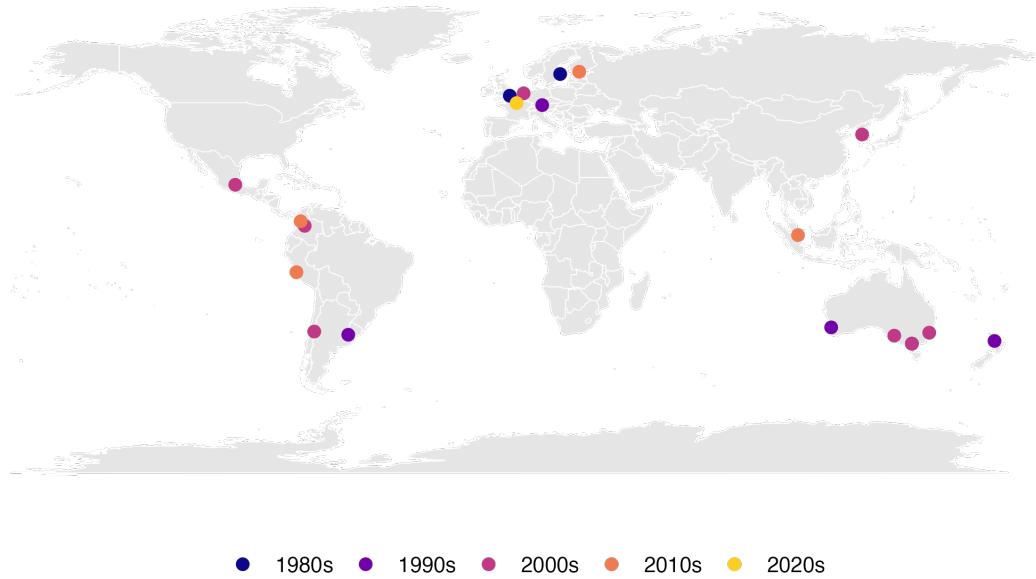


Figure C.1: Geographic distribution of cities adopting competitive tendering for public transit, by decade.

Table C.1: Mode Choice Summary Statistics

	N	Mean	SD
Panel A: Household Characteristics			
Household size	11,612	3.40	1.56
Vehicle ownership	11,612	0.47	0.50
Household number of vehicles	11,612	0.58	0.73
Income: < USD 10k	11,612	0.31	0.46
Income: [USD 10k, 40k)	11,612	0.60	0.49
Income: > USD 40k	11,612	0.09	0.28
Panel B: Traveler Characteristics			
Age (in years)	22,653	38.39	20.39
Female (= 1)	22,653	0.52	0.50
Education: Less than High School	22,653	0.24	0.42
Education: High School	22,653	0.43	0.49
Education: Associate degree	22,653	0.09	0.29
Education: College degree or higher	22,653	0.25	0.43
Panel C: Trip Characteristics			
Car available (= 1)	47,622	0.44	0.50
Origin within CBD (= 1)	47,622	0.31	0.46
Destination within CBD (= 1)	47,622	0.31	0.46
Driving (= 1)	47,622	0.39	0.49
Public Transit (= 1)	47,622	0.37	0.48
Walking (= 1)	47,622	0.25	0.43
Distance: [0 km, 2 km)	47,622	0.39	0.49
Distance: [2 km, 5 km)	47,622	0.23	0.42
Distance: > 5 km	47,622	0.38	0.48
Purpose: Work	47,622	0.32	0.47
Purpose: Study	47,622	0.15	0.36
Purpose: Other	47,622	0.53	0.50

Notes: This table presents summary statistics for transportation mode choice decisions. Sample includes all observations from the travel preference parameters estimation sample covering the period 2012-2013. All monetary values are in 2013 USD.

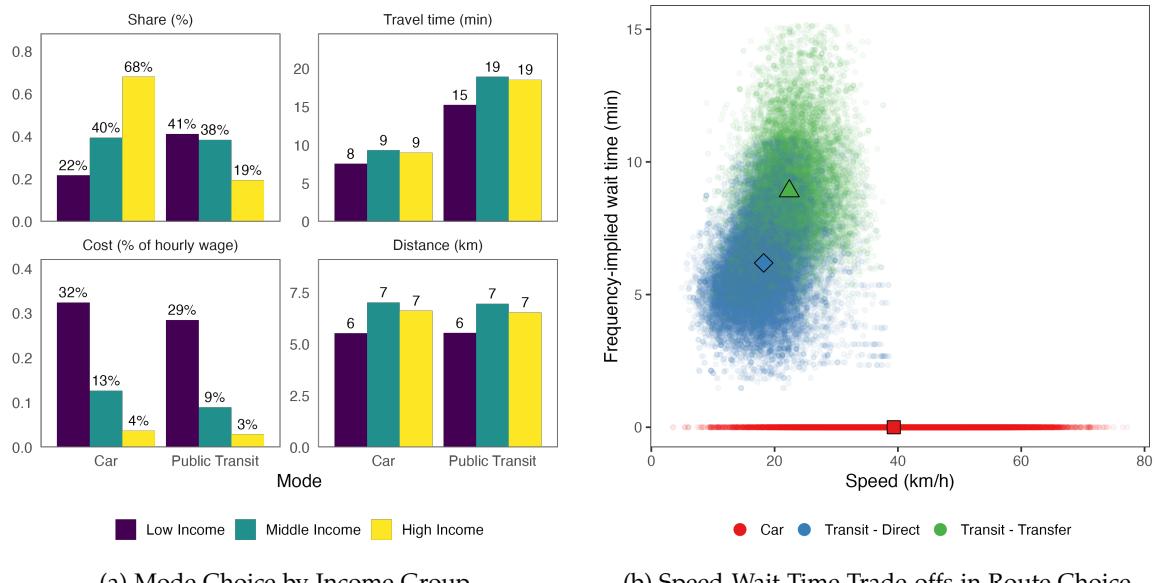


Figure C.2: Transportation Mode and Route Choice Patterns.

Notes: Panel (a) displays mode choice distributions across income groups, showing share, travel time, cost (as percentage of hourly wage), and distance for car versus public transit users. The displayed shares represent only car and public transit usage, with walking serving as the outside option to bring total mode shares to 100%. High-income travelers show greater car usage (68% vs 21.8% for low-income), while public transit users face longer travel times but lower relative costs. Panel (b) illustrates the speed-frequency trade-off in route choice, plotting frequency-implied wait times against travel speeds. The scatter plot reveals distinct clusters for different route types: cars, direct transit routes, and transfer-based transit options. Squared, Triangular and diamond markers indicate mean values for different route categories.

Table C.2: Route Choice Summary Statistics

	Direct		Transfer	
	Mean	SD	Mean	SD
Panel A: Off-peak				
Number of options	3.13	3.07	5.38	6.95
At least one option (= 1)	0.83	0.37	0.73	0.45
Subway (= 1)	0.11	0.31	0.43	0.49
Fare (USD)	1.10	0.30	1.11	0.31
Travel time (min)	12.97	11.05	19.72	13.21
Wait time (min)	3.93	1.01	7.57	1.46
Panel B: Peak				
Number of options	2.34	2.47	4.31	6.06
At least one option (= 1)	0.77	0.42	0.66	0.47
Subway (= 1)	0.12	0.32	0.46	0.50
Fare (USD)	1.09	0.33	1.15	0.34
Travel time (min)	16.08	13.07	27.06	14.56
Wait time (min)	3.61	1.10	7.01	1.61

Notes: This table presents summary statistics for route choice decisions. Sample includes all observations from travel preference parameters estimation sample covering the period 2012-2013. All monetary values are in 2013 USD.

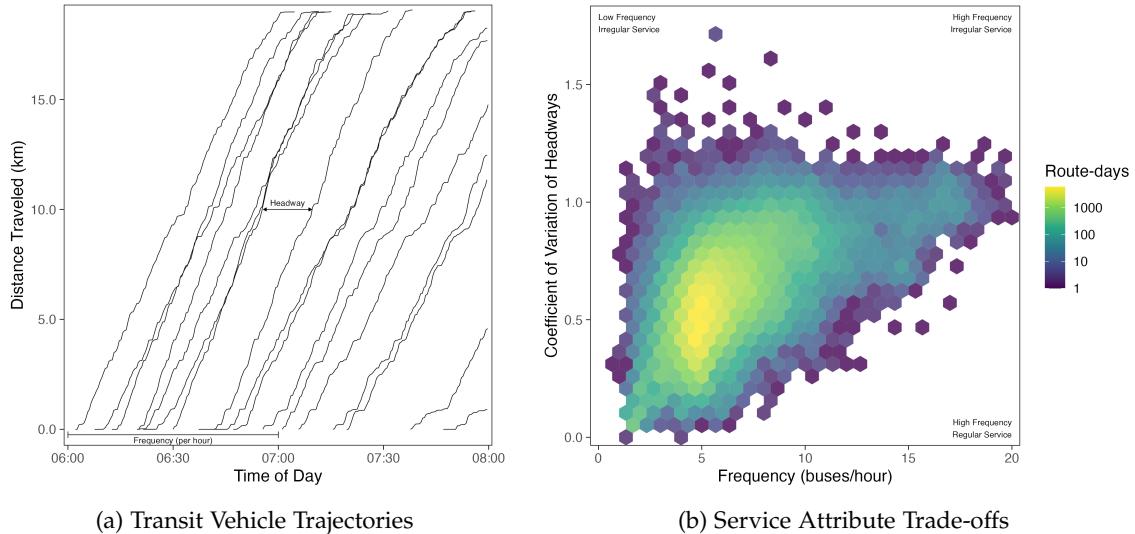


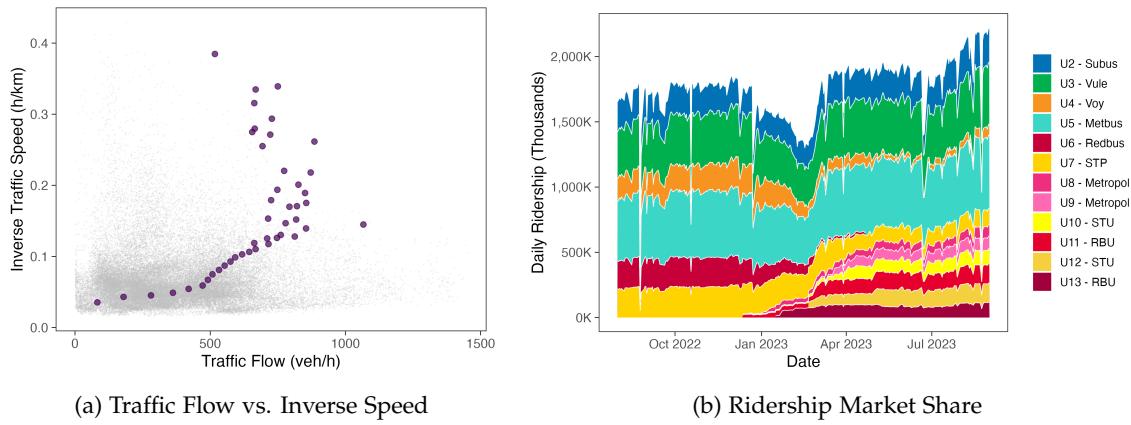
Figure C.3: Frequency and Headway Regularity Choice Patterns.

Notes: Panel (a) shows transit vehicle trajectories over time, illustrating frequency and headway concepts through dispatch timing and distance traveled. Panel (b) reveals the strategic trade-off between service frequency and observed reliability (coefficient of variation of headways), with density patterns showing common operational strategies.

Table C.3: Public Transit Supply Summary Statistics

	Mean	SD	Min	Max
Panel A: System-level				
Number of Bundles	8.50	2.53	6.00	11.00
Number of Firms	7.00	1.01	6.00	8.00
Number of Depots	65.84	2.05	62.00	68.00
Number of Routes	354.86	10.86	295.00	363.00
Daily Ridership (millions)	1.86	0.32	0.43	2.21
Panel B: Bundle-level				
Number of Depots	8.10	6.03	2.00	19.00
Number of Routes	41.75	23.67	11.00	89.00
Daily Ridership (thousands)	218.72	145.45	49.48	562.91
Panel C: Route-level				
Frequency (bus/h)	5.82	1.96	1.06	20.27
Headway Regularity (CV)	0.41	0.15	0.02	1.42
Length (km)	18.45	8.58	2.35	57.22
Speed (km/h)	17.83	3.32	1.44	39.29
Daily Ridership (hundreds)	28.52	24.72	0.01	170.64

Notes: This table presents summary statistics for public transit supply characteristics across three levels of aggregation. The sample includes route-day level observations from August 2022 and August 2023 to capture initial and final equilibrium states, excluding transient periods. Panel A shows system-level statistics aggregated across all bundles and routes. Panel B presents bundle-level statistics, where bundles represent groups of routes operated under the same contract. Panel C displays route-level characteristics including service frequency (buses per hour), headway regularity measured by coefficient of variation (CV), route length, average speed, and ridership.



(a) Traffic Flow vs. Inverse Speed

(b) Ridership Market Share

Figure C.4: Equilibrium Outcomes: Traffic Flow, Traffic Speed, and Ridership.

Notes: Panel (a) illustrates the fundamental traffic flow-speed relationship, plotting inverse traffic speed against hourly vehicle flow. The gray points represent individual observations, while the purple line shows the binned averages. Panel (b) displays daily ridership trends across transit firms from August 2022 to August 2023, with each colored area representing a different operator's passenger volume. The stacked area chart shows bus system ridership fluctuating between approximately 1.6-2.2 million daily passengers.

Appendix D Data construction (for online publication)