Mictrotransit design: fixed-line transit, on-demand mobility, or both?

Alexandre Jacquillat[a], Julia Yan[b], Arthur Delarue[c] and Shriya Karam[a,*]

[a]Sloan School of Management, MIT, Cambridge, MA, USA
[b]Sauder School of Business, University of British Columbia, Vancouver, BC, Canada
[c]Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA
[*]Corresponding Author

Extended abstract submitted for presentation at the 12th Triennial Symposium on Transportation Analysis conference (TRISTAN XII)
June 22-27, 2025, Okinawa, Japan

February 15, 2025

Keywords: Transit systems, on-demand mobility, continuous approximation, integer optimization

1 INTRODUCTION

Current urban mobility systems are struggling to meet evolving customer needs and preferences towards sustainable alternatives. Post Covid-19, transit ridership and revenue remain stagnant, unable to recover to pre-pandemic levels given recent changes in passenger behavior. Toward more flexible modes, door-to-door ride-sharing can provide convenient transportation options for passengers. For instance, in small cities and towns, where transit has been known to suffer due to low demand, ride-sharing has helped to improve mobility gaps (McKinsey, 2021). Yet, practical challenges arise for high demand regions as agencies face increasing operating costs by needing to increase route coverage and expand their fleets (Bloomberg, 2023).

To address the long-standing struggles of classical urban mobility systems, providers are increasingly leveraging microtransit systems that combine fixed-route and demand-responsive elements with high-capacity vehicles (Deloitte, 2022). For instance, operating models have limited the scope of ride-sharing in large areas to serve as a feeder into a transit system, also known as first- and last- mile on-demand service (Maheo *et al.*, 2019). These on-demand feeder systems have been the dominant form of leveraging short, on-demand trips to augment the reach of transit networks, yet may represent a limited view in hybrid system design (Hu & Schneider, 2024). In order to simultaneously maintain transit network coverage while ensuring high on-demand quality of service, providers are also interested in emerging systems that co-design fixed-route transit and ride-sharing (Bloomberg, 2024).

As providers increasingly leverage multimodal strategies with fixed and flexible operations, there are key questions in *how* to design and operate hybrid systems. In this paper, we seek to provide strategic-level insights in designing multimodal systems through analytical models and a data-driven optimization approach. We first find that lower demand favors on-demand modes, but that counterintuitively, larger systems (higher demand and fleet overall) favor on-demand due to the flexibility provided by fleet allocation. Second, we find that while individual locations tend to favor either transit or on-demand operations exclusively, combining multiple modes at the network-level can lead to significant benefits. To build on our theoretical findings, our optimization model using historical passenger requests for Tri-cities, Washington shows the benefits of hybrid operations under spatio-temporal dynamics. Ultimately, our results show that the "right" operating model depends on demand and network characteristics, and can even involve a combination of fixed-route and flexible mobility solutions.

2 ANALYTICAL MODEL

We consider simplified settings to co-design transit and on-demand systems, first seeking the thresholds that favor flexible on-demand modes and fixed-route transit, and second, exploring the benefits of combining different strategies. We minimize total passenger inconvenience using continuous approximation methods, which include detour and waiting times:

Detour approximation. To approximate detour time, or latency, for on-demand trips, we rely on the traveling repairman problem (TRP). Blanchard *et al.* (2024) provide constant-factor estimates of the TRP and find that the optimal latency scales with $\mathcal{O}(n\sqrt{n})$, where n is the number of total vertices in a network. This approximation contrasts with the constant-factor estimates of the traveling salesman problem (TSP), where total travel time for a tour scales with $\mathcal{O}(\sqrt{n})$. Unlike the TSP which exhibits economies of scale through its concave approximation, the convex approximation of the TRP indicates diseconomies of scale. For microtransit systems, each additional passenger creates increasingly costly detours and deteriorates service quality at higher demand levels.

Waiting time approximation. To approximate waiting times for transit and on-demand vehicles, we collapse the temporal dimension of vehicle dispatch. In reality, vehicles are dispatched in real-time based on passenger requests, which we model in our optimization approach in Section 4. In our analytical models, we assume random passenger arrivals over a uniform distribution and equally spaced vehicles, each with a spacing of H. We therefore approximate the average waiting time as $H/2$, or half of the vehicle headway (Osuna & Newell, 1972).

We apply these approximations to a service area partitioned into geographical zones to serve passenger demand across a time horizon T. The operator must allocate vehicles across zones and edges from a fixed fleet F, and allocate passengers to either the transit, on-demand, or first- and last- mile on-demand modes. We minimize total passenger inconvenience which includes on-demand detours, waiting times, and walking time (per passenger walking cost denoted by w).

3 THEORETICAL RESULTS

On-demand for first- and last- mile. To expand the reach of existing transit systems, agencies will typically augment a transit network with on-demand vehicles to serve passengers via first- and last- mile trips. We simulate this operating model with fixed transit frequencies, where our central question seeks the network characteristics that favor transit or first- and last-mile on-demand. We focus on a two-zone network with one directed edge with passenger demand D traveling along the edge between the zones.

Proposition 1 characterizes the structure of the optimal solution. We first find that low fleet sizes favor a transit only solution, as the operator needs a minimum number of vehicles to operate any first- and last- mile on-demand service at all. As the fleet size increases, and when demand is sufficiently high (relative to D^*), we see the benefits of hybrid operations (i.e., a mixture of first- and last- mile on-demand and transit). On the other hand, when demand is low relative to D^*, the operator can offer first- and last- mile on-demand service for all passengers. In this regime, the operator can take advantage of shorter on-demand detours and waiting times associated with low demand and high fleet size conditions.

Proposition 1: If $F \leq T/w$, the transit only solution is optimal. If $F > T/w$ and $\sqrt{D} \geq D^*$, the hybrid solution is optimal. Else, the first- and last- mile on-demand only solution is optimal.

While Proposition 1 shows that lower demand and larger fleet sizes favor first- and last- mile on-demand operations, we observe a counterintuitive scaling effect in Proposition 2. Larger-scale systems, where demand and fleet are both large and scaled by a factor $\alpha \geq 1$, imply that first- and last- mile on-demand is optimal. With both higher demand and fleet sizes, the operator can effectively reduce vehicle occupancies and avoid transit walking costs that scale with demand only. From a practical standpoint, the flexibility provided by fleet allocation can alleviate the negative impacts of on-demand detours and delays.

Proposition 2. Let D^* and F^* be values such that first- and last- mile on-demand is optimal. Then first- and last- mile on-demand is optimal with αD^* and αF^* for all $\alpha \geq 1$.

On-demand for all miles. Instead of optimizing first- and last- mile on-demand operations only, we next consider different uses of on-demand vehicles altogether. For instance, can vehicles instead be used to serve origin-destination on-demand trips, rather than serving first- and last-mile on-demand trips only? To address this question, we consider a model that allocates vehicles and passengers to either fixed-route transit or origin-destination on-demand trips.

Propositions 3 and 4 show that while at individual locations, "local" hybrid operations (i.e., a mixture of transit and on-demand) are not beneficial, "global-hybrid" operations (at least one transit or on-demand edge) can have arbitrarily large benefits. Proposition 3 suggests that at individual, localized edges, it is sufficient to invest in either transit or on-demand operations exclusively. However, network-level operations do not feature this "one-size-fits-all." As shown in Proposition 4, there exist instances where the cost difference between the "global-hybrid" solution and the transit and on-demand only solutions becomes unbounded. That is, as the number of edges n in a network approaches ∞, it becomes increasingly worse to operate transit or on-demand exclusively. Altogether, these results show the benefits of hybrid operations for high-dimensional networks, therefore motivating the need for the co-existence of modes through strategic system design.

Proposition 3. For each edge in a network of n edges, transit only or on-demand only is optimal.

Proposition 4. Let q_{OD}, q_T, and q_{GH} be the the optimal on-demand only, transit only, and the global-hybrid costs. There exists γ such that $\min(q_{OD}, q_T) - q_{GH} \geq \gamma n$ for all $n \geq 1$.

4 DATA-DRIVEN ASSESSMENT

To validate and extend the theoretical results, we develop an optimization model to design hybrid transportation systems utilizing historical passenger requests from Tri-cities, Washington. The optimization model jointly considers both fleet and passenger operations. For fleet operations, the operator must allocate vehicles from a shared operating budget B, comprising two key decisions: vehicle allocation across zones for on-demand trips and frequency setting for fixed transit lines. For passenger operations, passengers must be assigned to modes, either transit only, on-demand only, or first and/or last mile on-demand. In contrast with the analytical model, the optimization model includes temporal coordination and customer-station assignment.

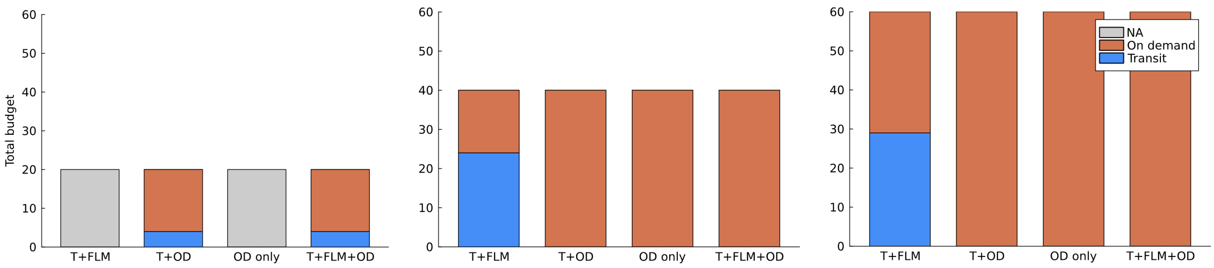
Table 1 summarizes the performance (objective, vehicle load, passenger mode split) and average levels of service (wait, in-vehicle, and walk times) as the operating budget size varies. We compare these metrics across the following models: transit and first and/or last mile on-demand only ("T+FLM"), transit and full on-demand only ("T+OD"), full on-demand trips only ("OD only"), and the hybrid model with all modes ("T+FLM+OD"). For the transit and first and/or last mile on-demand only ("T+FLM"), it is important to note that we allow for "local" on-demand trips that occur within zones only, but not "full" on-demand trips across zones.

Table 1 shows two operating regimes based on the overall budget, which demonstrate the need for a "no-one-size-fits-all" design approach. First, we find that for operating conditions with constrained budgets ($B = 20$), there is a 14.58% cost improvement by integrating first- and last- mile on-demand with fixed-route transit and full on-demand trips. Primarily, hybrid operations allow replacing long walking trips to transit stations with on-demand trips that serve these connections, or even origin-destination on-demand trips. Second, under flexible regimes, we observe cost savings ranging from 62.58% ($B = 40$) to 55.37% ($B = 60$) by operating an on-demand only system. As shown in Figure 1, the operator can leverage the entire budget for on-demand service, rather than "split" vehicles across modes as in the transit and first- and last-mile on-demand only model ("T+FLM"). The practical implication is that the "right" operating model can involve a combination of fixed-route and flexible solutions based on the overall budget, motivating the need for a "no-one-size-fits-all" design approach.

Table 1 – *Model performance and service levels with varying budgets for a vehicle capacity of 4.*

		Performance			Transit			On-demand		
Budget	Model	Obj. (%)	Avg. load (%)	Mode split (%)	Wait	In-vehicle	Walk	Wait	In-vehicle	Walk
20	T+FLM	Inf.	–	–	–	–	–	–	–	–
	T+OD	0.00%	44.25%	83/17/0	6.38	14.63	53.38	3.39	15.00	0.00
	OD only	Inf.	–	–	–	–	–	–	–	–
	T+FLM+OD	-14.58%	42.21%	83/7/10	8.59	27.97	35.21	3.28	12.43	0.00
40	T+FLM	0.00%	40.55%	38/5/57	5.12	16.15	27.40	3.00	8.63	0.00
	T+OD	-62.58%	27.55%	100/0/0	–	–	–	0.04	12.67	0.00
	OD only	-62.58%	27.55%	100/0/0	–	–	–	0.04	12.67	0.00
	T+FLM+OD	-62.58%	27.55%	100/0/0	–	–	–	0.04	12.67	0.00
60	T+FLM	0.00%	30.30%	38/1/61	4.77	14.61	21.07	2.32	7.95	0.00
	T+OD	-55.37%	25.00%	100/0/0	–	–	–	0.00	12.65	0.00
	OD only	-55.37%	25.00%	100/0/0	–	–	–	0.00	12.65	0.00
	T+FLM+OD	-55.37%	25.00%	100/0/0	–	–	–	0.00	12.65	0.00

"Obj.": normalized cost relative to the optimal cost of the highest cost model. "Inf.": infeasible model.
"Avg. load": average load across on-demand vehicles, relative to the vehicle capacity.
"Mode split": percentage of demand assigned to (full) on-demand, (full) transit, and first-/last-mile on-demand.
All average transit and on-demand cost metrics are reported in minutes per passenger. For waiting times, solutions reported are the realized times, rather than the solutions that achieves the lowest cost waiting time.



(a) $B = 20$ (b) $B = 40$ (c) $B = 60$

Figure 1 – *Transit and on-demand allocation across models for varying budgets B.*

References

Blanchard, Moïse, Jacquillat, Alexandre, & Jaillet, Patrick. 2024. Probabilistic bounds on the k-traveling salesman problem and the traveling repairman problem. *Mathematics of Operations Research*, **49**(2), 1169–1191.

Bloomberg. 2023. Accessed Feb. 04, 2025.

Bloomberg. 2024. Accessed Feb. 04, 2025.

Deloitte. 2022. Accessed Oct. 14, 2024.

Hu, Lingqian, & Schneider, Robert J. 2024. Panacea or Band-Aid? First-/Last-Mile Connection in Public Transit Systems. *Journal of the American Planning Association*, 1–8.

Maheo, Arthur, Kilby, Philip, & Van Hentenryck, Pascal. 2019. Benders decomposition for the design of a hub and shuttle public transit system. *Transportation Science*, **53**(1), 77–88.

McKinsey. 2021. Accessed Feb. 04, 2025.

Osuna, Edgar Elias, & Newell, Gordon F. 1972. Control strategies for an idealized public transportation system. *Transportation Science*, **6**(1), 52–72.