Public Transport
https://doi.org/10.1007/s12469-023-00348-1

ORIGINAL RESEARCH

Beyond the last mile: different spatial strategies to integrate on-demand services into public transport in a simplified city

Andres Fielbaum[1] · **Sergio Jara-Díaz**[3,4] · **Javier Alonso-Mora**[2]

Accepted: 7 December 2023

Abstract

Integrating on-demand services into public transport networks might be the best way to face the current situation in which these new technologies have increased congestion in most cities. When cooperating with on-demand services rather than competing with them, public transport would not risk losing users, and could attract some passengers from private modes thanks to an increased quality of service. This fact has engendered a growing literature discussing how to design such an integrated system. However, all of that research has imposed that on-demand mobility is to solve the so-called "last-mile problem", serving only as a feeder that connects the exact origins/destinations with the traditional public transit network. As it induces a large number of transfers and it precludes some scale-effects to be triggered, in this paper we challenge that imposition and investigate if this is the best spatial integration strategy. To do so, we study a simplified linear city in a morning peak situation, where we propose seven different line structures, all of them combining a traditional fixed line with on-demand ride-pooling (ODRP): three direct structures, where ODRP can serve full trips, three semi-direct, where a single ODRP vehicle can serve the largest part of a trip, and a base case in which ODRP is restricted to the first and final legs only. Our results show that the base case is optimal only under very specific demand patterns, or when transfer penalties are disregarded. Our analytical approach reveals relevant operational aspects of such integrated systems: namely, that the base case can help increase directness (diminishing detours), and that ODRP can help shorten the routes of the fixed services to decrease operator costs.

✉ Andres Fielbaum
andres.fielbaum@sydney.edu.au

1 School of Civil Engineering, The University of Sydney, Camperdown, Australia

2 Department of Cognitive Robotics, TU Delft, Delft, The Netherlands

3 Department of Civil Engineering, Universidad de Chile, Santiago, Chile

4 Instituto Sistemas Complejos de Ingeniería (ISCI), Universidad de Chile, Santiago, Chile

Published online: 02 September 2024

Keywords Public transport · On-demand mobility · First-mile problem · Ridepooling

JEL Classification L91 · L92 · R40 · R48

1 Introduction

The proliferation of smartphones that can connect users and vehicles in real time is changing the ways in which people move worldwide. On-demand systems have existed for a long time, both non-shared (traditional taxis) and shared (such as “colectivos” in Chile, or shared rickshaws in many developing countries); however, they now operate on much larger scales thanks to the new technologies, with companies such as Uber, Didi, and Beat. So far, this change has mostly affected non-shared modes, in which the same vehicle is utilized consecutively by different users which is now known to increase congestion (Diao et al. 2021, Henao and Marshall 2019, Roy et al. 2020, Tirachini and Gomez-Lobo 2020). Online shared alternatives, where different users can ride the same vehicle at the same time, are also emerging but still at lower scales. These alternatives receive different names in the literature, usually on-demand ride-pooling (ODRP) or ridesharing. However, even these alternatives do not guarantee to reduce congestion, which depends on attracting users from private modes rather than from public transport (Tikoudis et al. 2021; Tirachini et al. 2020). In other words, when on-demand systems **compete** with public transport, there still is a potential increase of vehicle-kilometers-traveled, which would not only worsen congestion; it would cause more pollution and can have other undesirable effects such as increasing urban sprawl and inequalities.

Therefore, a crucial question is how to leverage these new technologies to face congestion problems. A relevant alternative to explore is making ODRP **cooperate** with public transport, that is, to integrate it so that some of the transit lines can be on-demand instead of following fixed routes. As we detail in Sect. 2, this approach has begun to be studied in the last years by different scholars obtaining encouraging results in terms of creating more efficient public transport systems. However, almost all the papers that have studied such an integrated framework have assumed that ODRP is to solve the so-called *last-mile problem*, i.e., as a feeder that connects the origins and destinations with nearby transit stations, where the users board or leave a large-capacity vehicle that is part of the traditional network. Limiting ODRP to serve only the last legs of public transport trips can be troublesome for at least two reasons:

1. It imposes two extra transfers on every user whose origin and destination are not walkable from a large transit station, which could be partially avoidable if there were some vehicles directly connecting origins and destinations. Transfers are undesirable for the users (Garcia-Martinez et al. 2018), and disregarding their relevance can have massive effects when designing a public transit network (Fielbaum et al. 2016, Jara-Díaz et al. 2012). This problem becomes even more important when ODRP is involved, as the total traveling times faced by the users are especially unreliable (Fielbaum and Alonso-Mora 2020), and a slight extra

delay might imply losing the large-capacity vehicle if this is based on timetables rather than on frequencies (such as trains).

2. If on-demand vehicles can be designed to serve a larger number of passengers as well, some sources of scale economics can be triggered (Fielbaum et al. 2023): For users, the utilization of a larger fleet diminishes waiting times (similar to the "Mohring Effect" in public transport), and it is easier to form groups of passengers with nearby routes (which is called the "Better Matching Effect"); for operators, operating at a larger scale enables the use of vehicles with a higher capacity.

Despite these problems, there are some reasons that justify utilizing the flexible vehicles to feed or leave the traditional transit network. First, when scholars have compared fixed versus flexible lines, they have consistently found that ODRP should be preferred in low-demand scenarios when using it as a feeder (Badia and Jenelius 2020; Fielbaum et al. 2022; Li and Quadrifoglio 2010; Papanikolaou and Basbas 2020; Quadrifoglio and Li 2009). Second, the "Better Matching Effect" aforementioned can be easily triggered when all the users travel to (or from) a common transfer station, as it is only their origins that need to be nearby (Fielbaum 2020).

All in all, taking into account the advantages of ODRP there is no clear reason to impose its operation only as a feeder within the context of an integrated public transport network which motivates a very clear research question: under which circumstances would a different integration approach be more efficient? This paper is mostly devoted to filling this research gap. To do this, we do an exhaustive analysis over a simplified city model: we propose several spatial strategies to integrate the flexible system into the network, allowing those vehicles to either serve full trips or to mimic a hub-and-spoke design. We derive explicit expressions for the total costs of each of these strategies, and we identify the best strategy depending on the combination of different exogenous parameters (that represent different conditions over the demand and the city). Extensive numerical simulations are run to identify under which circumstances the feeder-based strategy is indeed the best one, and to study different operational aspects of these integrated systems.

The paper is organized as follows. In Sect. 2 we summarize the relevant literature and show that there is a serious research gap which motivates this paper. Section 3 describes the simplified city model and the different line structures that we consider there. Section 4 explains how we derive the analytical formulation for user and operator costs for all the flexible and fixed lines of the different line structures. Section 5 shows and analyzes the results of the numerical simulations. Section 6 concludes the paper and identifies relevant directions for further research.

2 Related works

Following the emergence and massification of on-demand mobility systems, ODRP has been widely studied in recent years. Research on ODRP has included its potential to reduce the number of vehicles needed to serve a given demand (Fagnant and Kockelman 2018, Qian et al. 2017), its impact on congestion (Ke et al. 2020a, b; Tikoudis et al. 2021; Zhu and Mo 2022) and mode choice (Liu et al. 2019, Mo et al. 2021), as well as several operational challenges like unreliability (Alonso-González et al. 2021; Fielbaum and Alonso-Mora 2020; Kucharski et al. 2021) and pricing (Bian et al. 2020; Fielbaum et al. 2022; Furuhata et al. 2015; Ke et al. 2020a, b; Lu and Quadrifoglio 2019). Reviews of recent trends in on-demand mobility have been done by Narayanan et al. (2020), Wang and Yang (2019), and Zardini et al. (2022).

On the other hand, simplified networks have been thoroughly used in order to study (traditional) public transport design, especially to discuss its spatial aspects: the line structure—how to define the set of lines—and the lines' density, i.e., how separated in space they should be. Line structures are studied in graphs composed by a small number of nodes by Fielbaum et al. (2016, 2020a), Gschwender et al. (2016), Hörcher and Graham (2018), Jara-Díaz et al. (2018), Jara-Díaz and Gschwender (2003), Jara-Díaz and Muñoz-Paulsen (2021), Petruccelli and Racina (2021), and Masing et al. (2022); these studies have proved useful to analyze a number of aspects such as the role of transfers, the presence of scale economies, or the combination of rail and bus-based technologies. Several of those papers are based on the Parametric City Model proposed by Fielbaum et al. (2017). The spatial density of public transport lines has been studied in radial (Badia et al. 2014; Byrne 1975; Tirachini et al. 2010) and rectangular (Chen et al. 2015; Daganzo 2010) cities, as well as in the parametric city model (Fielbaum et al. 2021b), consistently showing that a larger number of users allows for shorter walks; these models in which the space is assumed to be continuously available to allocate public transport lines are sometimes referred to as Continuous-Approximation models (Calabro et al. 2023). As such, these simplified networks are able to reveal some relevant fundamental aspects of public transport design, without the need of modeling at a detailed level. These papers are surveyed, together with other ones dealing with the economics of public transportation, by Hörcher and Tirachini (2021).

Some papers have studied whether on-demand mobility might pose a threat to public transport. Hall et al. (2018) focus on Uber and the U.S. case, finding that the average public transport agency has increased its ridership thanks to Uber, but the effect varies strongly across different cities, a conclusion reinforced by Malalgoda and Lim (2019) who state that, in the US, the overall effect depends on the mode and city. This finding that ride-hailing can be either a complement or a substitute for public transport, depending on city-specific circumstances, also holds for Toronto (Young et al. 2020), as well as several European cities (Cats et al. 2022). Irawan et al. (2020) studied the Indonesian case, considering motorcycle-based ride-hailing, and conclude that it is crucial to enhance the quality of service in public transport to prevent it from losing passengers to this new mode.

Let us now review in more detail the papers that, as this one, study how to combine traditional public transport with on-demand shared mobility. As this is a very broad topic, we limit this review to papers that face design aspects of the problem, and in which the on-demand component is shared. We remark that there exist several other research lines, such as the effects of the associated electrification (Bartłomiejczyk and Kołacz 2020), the development of a corresponding enterprise architecture for public transport companies that offer on-demand services (Würtz and Sandkuhl, 2021), or the interaction of public transport with non-shared on-demand mobility (Salazar et al. 2018). Real-life pilots have also been studied (Mueller et al. 2019, Zuniga-Garcia et al. 2022, Perera et al. 2020, Loyola et al. 2023), usually concluding that ODRP is not used only as a feeder but also for short trips, and that its usage is very sensitive to fleet size and income.

We now focus on the literature studying how to design public transport networks that include traditional lines and on-demand shared vehicles. Two papers do this using simplified networks: Fielbaum (2020) bases his research on the Parametric City Model, but making a zoom within each mode that is considered as continuous; and Calabro et al. (2023), who borrow the continuous approximation model from Chen et al. (2015), but enabling to replace some feeder lines by ODRP when convenient. Crucially, both papers restrict ODRP to act as a feeder, a characteristic that is also present in most of the other papers in this realm. Actually, several papers focus on the specific design of ODRP when it is to be used as a feeder (Bürstlein et al. 2021; Calabrò et al. 2022; Chaturvedi and Srivastava 2022; Chen et al. 2020; Huang et al. 2022; Lau and Susilawati 2021; Ma et al. 2019; Wen et al. 2018).

Still on the assumption of a feeder system, Auad-Perez and Van Hentenryck (2022), Mahéo et al. (2019), and Shen et al. (2018) optimize the operation of ODRP together with the design of the public transport network. It is noteworthy that designing a public transport network by itself is already an NP-hard problem for different specifications (Borndörfer et al. 2007, Fielbaum et al. 2018), let alone if flexible routes are also involved, so restricting the on-demand vehicles to serve the first and last miles only is very helpful from a methodological point of view. Pinto et al. (2020) face an analogous problem without imposing a feeder-trunk system,[1] but with a similar implicit assumption, namely that ODRP is to serve low-demand areas, as it is meant to replace lines whose frequency would be extremely low. Kim and Schonfeld (2014) focus on how to coordinate both subsystems so that the waiting time at the transfer points are kept low. On a different but related note, Périvier et al. (2021) study an approximation algorithm for the theoretical programming problem that emerges when the set of public transport routes includes flexible vehicles.

As such, the assumption of a feeder-trunk scheme is present in almost all of the papers that have studied this integration problem. However, it is not yet known whether and under which conditions this is the best strategy. The main contribution of this paper is to fill this research gap, by comparing feeder-trunk with other integration strategies that prioritize either direct trips or a hub-and-spoke structure.

3 The Model

3.1 The linear city

As in Jara-Díaz and Muñoz-Paulsen (2021), our analysis is based on a linear city composed of three residential or activity zones: a periphery P, a subcenter SC, and a CBD

[1] In a feeder-trunk system the *feeder* lines connect the exact origins and destinations with major transit stations, where users are expected to transfer to take a *trunk* line. Trunk lines connect the major stations using higher-capacity vehicles.

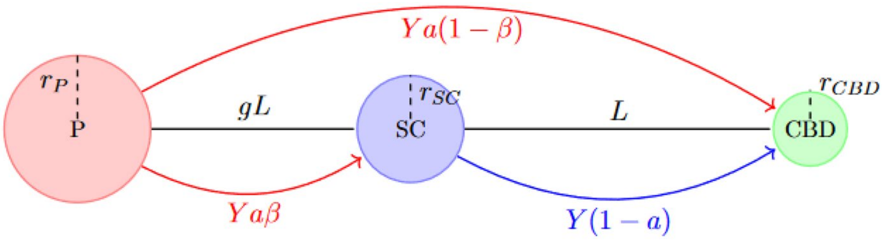


Fig. 1 Linear city representation. Red and blue arrows represent demand flows from the periphery P and from the subcenter SC, respectively. Black solid lines are the connecting streets

(Central Business District). This can be seen as a particular case of the parametric city model by Fielbaum et al. (2017), when this model is limited to just one triad. This linear city is depicted in Fig. 1 with P, SC and the ***CBD*** represented by circles of different radia; it represents a morning peak situation, so that P only generates trips, the CBD only attracts and SC does both. Out of the Y passengers per unit time, a fraction a emerges from P (and $1-a$ from SC), out of which a fraction β goes to SC and $1-\beta$ goes to the CBD. The distances between consecutive zones are gL and L, respectively. As we are interested in what is happening inside the zones, where ODRP provides the almost-door-to-door service, their size is relevant, and we assume that the respective radia fulfill $r_P \geq r_{SC} \geq r_{CBD}$. A glossary containing all the mathematical terms used throughout the paper is provided in Table 4 in the Appendix.

It is worth highlighting how different relevant situations can be captured by this simple model. As discussed above, previous research has suggested that ODRP is more convenient in low-demand scenarios, typically low-density peripheries, which can be modeled via a small a or Y and a large r_P; on the other hand, to analyze the potential of ODRP in compact environments to complement traditional public transport, large values of Y and small circles should be considered. Different intermediate configurations can be achieved by varying the involved parameters.

3.2 Types of lines

In traditional public transport, each line is defined by its route, frequency and vehicle size. In ODRP, there are no fixed routes, but this does not preclude a similar definition. In this context, we will distinguish between the fixed and the flexible lines,[2] where the former are the traditional ones, connecting the bus stops assumed to be placed at the centers of the zones, and the latter act on demand but are restricted to some line-specific origins and destinations. In order to formalize this, let us denote by P_C, SC_C, and CBD_C the central point of the corresponding zones, where the fixed lined stops are assumed to be. Further, when a flexible line departs from the interior of P, we will say that its origin is P_i (where

[2] According to the nomenclature proposed by Vansteenwegen et al. (2022), the flexible lines we study follow a *dynamic offline* type of operation: it is dynamic because users are served on demand, but offline because the vehicle's itinerary is not updated while in operation (it gets full once the users board it, so its route cannot be changed until they are all dropped off).

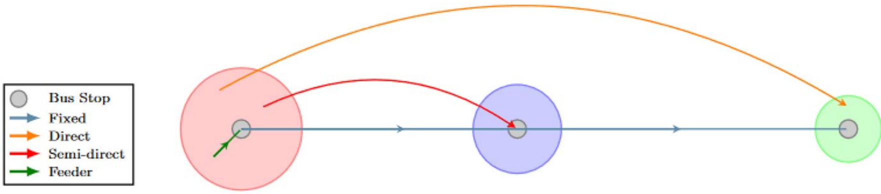


Fig. 2 Types of lines in the linear city: fixed (blue, connecting the three nodes), direct (orange, going from the first to the third node), semi-direct (red, going from the first to the second node), and feeder (green, within the first node) (color figure online)

the letter i stands for "interior"), and we use analogous notations with SC_i and CBD_i. Note that P_i, SC_i and CBD_i do not correspond to specific points in the network, due to the flexible nature of ODRP vehicles. Using this definition, we will consider four different types of lines, represented in Fig. 2.

- *Fixed lines* These correspond to traditional public transport, and are characterized by a route that does not change with the demand, i.e., all stops correspond to one of the P_C, SC_C, and CBD_C, represented in blue in Fig. 2.
- *Feeder lines* These lines connect the interior of a zone with its center. They are called feeders because they are meant to be connected with a different line that moves between different zones. The possible feeder lines are $P_i - P_C$, $SC_i - SC_C$, and $CBD_i - CBD_C$. One feeder line is represented in green in Fig. 2 ($P_i - P_C$).
- *Direct lines* These ones connect directly the origins and destinations of the users that take them (possibly requiring some short walks, as will be explained later), meaning that both the origin and the destination correspond to one of the P_i, SC_i, and CBD_i (not the same). The possible direct lines are $P_i - CBD_i$, $P_i - SC_i$, and $SC_i - CBD_i$. One direct line is represented in orange in Fig. 2 ($P_i - CBD_i$).
- *Semi-direct lines* These ones depart from the exact origins (except some possible short walks) but arrive at the center of a different zone.[3] The possible semi-direct lines are $P_i - CBD_C$, $P_i - SC_C$, and $SC_i - CBD_C$. One semi-direct line is represented in red in Fig. 2 ($P_i - SC_C$).

3.3 Considered line structures

Let us now explain which line structures we are going to consider. In general, a line structure is defined by the routes followed by each of the lines in the system (together with their frequencies and vehicles' size). In this context, they can be described by showing, for each pair of zones, if the corresponding passengers have a direct line (therefore, with zero transfers), a semi-direct line (one transfer) or none of the above (two transfers). This means that we have three alternatives for each of the three pairs of zones, leading to $3^3 = 27$ possible line structures. As studying all of them would be unmanageable, we are

[3] Note that one could define an inverse type of semi-direct line, going from the center of a zone towards an exact destination of a different zone. We do not include those here to maintain a tractable number of line structures to analyze.

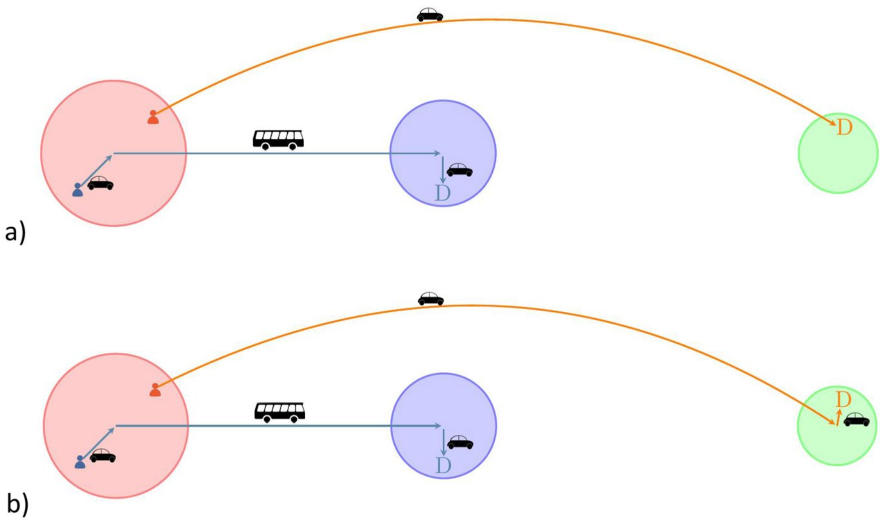


Fig. 3 Examples of Full (top) and Partial (bottom) line structures. FTS has one direct line in orange (its passengers travel without transfers) and no semi-direct ones. PTS has one semi-direct line in orange and no direct ones; all passengers have to make at least one transfer. In both examples, passengers from P to SC do not have a direct nor a semi-direct line, so they need to make two transfers

not considering the cases in which direct and semi-direct lines coexist, prioritizing the "pure" strategies in order to reveal neatly what is the best type of integration depending on the different demand and network conditions.

We begin the description of our line structures with the base case, which is a structure that corresponds to a Feeder-Trunk-Feeder rationale. There is one fixed line connecting the three zones' centers (the trunk line), and three feeder lines: one operating within P (bringing the passengers towards P's center), another within SC (bringing passengers to the center and distributing arriving passengers from the center), and one within CBD (distributing the arriving passengers). Note that in this structure, every traveler has to make two transfers and takes three vehicles: two flexible lines—to move inside the origin and destination zones—and a fixed line to move from the origin zone to the destination zone.

We now define the two classes of line structures that we study as alternatives to the base case, which we call "Full-trip" structure (FTS) and "Partial-trip" structure (PTS). A FTS is characterized by having at least one direct line, and no semi-direct lines, whereas a PTS has at least one semi-direct line and no direct lines. Note that in the case of PTS, users that take the semi-direct line require boarding a feeder line in the destination zone to reach their exact destination.

We now describe FTS and PTS using as examples Fig. 3a and b:

- *FTS* In Fig. 3a we show one direct line of ODRP vehicles, the orange one from P_i to CBD_i. In addition, there are other non-direct lines: the three depicted in

blue (one feeder P_i, one fixed from P_c to SC_c and one feeder SC_i). There is no semi-direct line.
- *PTS* In Fig. 3b, there is one semi-direct line complemented with a feeder as depicted in orange. The blue lines are exactly as in Fig. 3a. This is an intermediate scheme between the base case and FTS.

In a line structure with many direct lines most users will travel directly from their origin to their destination (perhaps including a small walk), without any transfer, i.e., taking just one vehicle. This strategy has the virtue of a strong reduction in the number of transfers compared to the base case, but as they do not share a common intermediate destination (the transfer station), it might be more difficult to gather enough users, leading to longer waiting and walking times, as well as longer detours because the exact destinations might not be close enough.

In a line structure with many semi-direct lines most users will travel from their very origin to the center of the destination zone by ODRP, where they take a second ODRP vehicle, corresponding to the intra-zonal flexible line of their destination zone (similar to Cortés and Jayakrishnan 2002). One transfer is mandatory.

The main idea behind each type of structure is best explained by means of strengths and weaknesses. The basic structure (that dominates the literature) is built to create common intermediate transfer stops for the users, so that it is easier to find passengers that can share a vehicle. For instance, if two users present nearby origins, this can be enough to match them, regardless of their final destination because they can reach such a destination in different vehicles. Moreover, using a fixed line to carry everybody can be very efficient for a large demand, because in that case a high-frequency service is good for the users and large vehicles are good for the operators. However, imposing two transfers on everybody can be quite unpleasant. The two alternative classes of structures are conceived to deal with this trade-off: in simple words, the basic structure values the advantages of fixed lines and passenger collection as more important, FTS diminishes the disadvantages of transfers, and PTS is the most balanced. It is noteworthy that a similar trade-off occurs in traditional public transport, which has been characterized as the *directness* of a line structure and identified as a source of scale economies by Fielbaum et al. (2020a). In our context, the basic structure is the less direct one and FTS are the most direct; further, the more passengers can be transported through a single flexible vehicle, the more direct the structure.

Within the family of Full-trip structures, each structure can be characterized by the pairs of zones that are connected by a flexible line, meaning that the users traveling between them are able to travel without transfers. Let us explain this through an example, in which only the users going from P to CBD travel with one vehicle, while everyone else needs three. In this case, we have:

- One direct line taking users at P_i, traveling to CBD_i, and delivering them there.
- Three feeder lines, one within each zone (we assume that these lines are not taken by the $P-CBD$ users, as they have an alternative route with zero transfers).
- One fixed line going from P to CBD.

Table 1 Description of the three FTS line structures

Name	Pairs of zones served with one flexible vehicle	Flexible lines	Fixed line
FTS-1	$P - CBD$	Four: $P_i - CBD_i, P_i - P_C$, $SC_i - SC_C, CBD_i - CBD_C$	From P to CBD
FTS-2	$P - CBD, SC - CBD$	Four: $P_i - CBD_i, SC_i - CBD_i$, $P_i - P_C, SC_i - SC_C$	From P to SC
FTS-3	$P - CBD, SC - CBD$, $P - SC$	Three: $P_i - CBD_i, SC_i - CBD_i$ $P_i - SC_i$	No

Table 2 Description of the three PTS line structures

Name	Pairs of zones served with one flexible vehicle	Flexible lines	Fixed line
PTS-1	$P - CBD$	Four: $P_i - CBD_C$, $P_i - P_C$, $SC_i - SC_C$, $CBD_i - CBD_C$	From P to CBD
PTS-2	$P - CBD, SC - CBD$	Five: $P_i - CBD_C$, $SC_i - CBD_C$, $P_i - P_C$, $SC_i - SC_C$, $CBD_i - CBD_C$	From P to SC
PTS-3	$P - CBD, SC - CBD$, $P - SC$	Five: $P_i - CBD_C$, $SC_i - CBD_C$, $P_i - SC_c$, $SC_i - SC_C$, $CBD_i - CBD_C$	No

As there are three pairs of zones, and for each of them there are two alternatives (to be connected via a flexible line, or not), there are $2^3 - 1 = 7$ possible FTS (excluding the basic structure with zero direct lines). In order to ease the analysis, here we study three of them, characterized by having one ($P - CBD$), two ($P - CBD$, $SC - CBD$), or three ($P - CBD$, $SC - CBD$, $P - SC$) direct lines as described in Table 1.

The structure FTS-1 includes only one direct line, which connects the pair of zones that would usually have the largest demand ($P - CBD$). FTS-2 includes two direct lines to serve the users going to the most attractive zone (the CBD). In FTS-3, nobody needs to transfer.

Regarding Partial-trip structures, again we can characterize each line through the pairs of zones that are connected via a semi-direct line, meaning that the corresponding users can travel using two vehicles instead of three. Let us take the case in which all the users traveling towards CBD use two vehicles as an example, where we have:

- One semi-direct line that takes users from P_i and delivers them to CBD_C.
- One semi-direct line that takes users from SC_i and delivers them to CBD_C.
- One feeder line within P and another within SC.
- One fixed line from P to SC

Out of the seven possible PTS, we consider three in this paper, described in Table 2.

4 Optimal designs and comparison

4.1 General approach

In this section, we explain how to model the different elements that define user and operator costs. The purpose is to write such costs as a function of some design variables (such as the fleet characteristics), so that an optimal design can be found thereafter.

Crucially, in all of the line structures every user has a unique possible route to go from origin to destination; the only exception are users going from P to CBD in FTS-1 and PTS-1, that we assume optimal for the (semi-)direct alternative rather than a feeder-trunk-feeder alternative that would be available for them but involving two transfers. Moreover, there are no common lines[4] in the system. Therefore, for a given line structure, we can derive the costs of each line independently; the total cost of that structure results from the addition of the costs associated to each of its lines plus the penalties associated to the total number of transfers induced by that structure.

Similar to numerous papers on public transport design (Fielbaum et al. 2016; Daganzo 2010; Calabro et al. 2023; Hörcher and Graham 2018; among many others), our analyses focus on what is the best way to serve a given demand. In other words, the demand is considered as exogenous.[5] Real-life applications require finding an equilibrium between this supply-oriented analysis with demand models and elasticity, a challenge that is beyond the scope of this paper.

We now derive the equations describing user and operator costs for the fixed and flexible lines. In order to find explicit formulae, a number of simplifying assumptions will be required. This is particularly relevant for ODRP, whose specific dynamics depend heavily on the network details, which can be incorporated when running simulations but preclude an analytical approach as we do here. As argued by Calabro et al. (2023) when describing their continuous approximation (CA) approach: "CA models allow to understand the impact of the different decision variables on the

[4] When two fixed-route lines run in parallel in some segments of their route, they become *common lines*. Users traveling within that segment would take the first line arriving at the bus stop, which is why the optimization of those two lines becomes intertwined (Chriqui and Robillard 1975).

[5] From a microeconomic perspective, the resulting optimal design represents the best engineering outcome (combination of resources) for a given demand at given input prices. The economic result is a cost function, i.e. the minimum necessary level of expenses to *produce* (serve) a given demand. As is well known, from such a function associated with the supply side, marginal costs can be obtained. When demand *functions* are introduced, optimal levels of production and optimal prices (welfare or profit maximizing) can be obtained.

performance, in an approximated, concise and computationally efficient way. The results obtained via CA models should be interpreted as high-level trends, which can guide transit planning considerations." In other words, the equations we describe in this section will enable us to understand, from a high-level standpoint, under which circumstances it is better to offer the different line structures described in the previous section.

In the following sections, the variables to be optimized, i.e., those that can be decided by the operator of the system, are highlighted in bold.

4.2 Fixed line optimization

We follow the public transport model proposed by Jara-Díaz and Gschwender (2009), where all the features that define the public transport operation, for both users and operators, are expressed as a function of the frequency $\boldsymbol{f}$ of the fixed line. Recall that the fixed line operates between the centers of the zones, so the exact users' origins and destinations are not relevant for this design. Further, we assume only one fixed line, although it could be possible to consider more than one even for this simplified linear city (see, for instance, Jara-Díaz et al. 2012).

As the characteristics of the line vary depending on the integration strategy, we now define some generic notation. We denote by T the time required to make the whole circuit if there were no stops, i.e., twice the length of the path A (that can either be short when $A = gL$, or long, when $A = (g + 1)L$) divided by the vehicle speed v_c. We denote by Y_1, Y_2, and Y_3 the number of passengers that use the fixed line from-to $P_C \to CBD_C, P_C \to SC_C, SC_C \to CBD_C$, respectively. Note that:

$$Y_1 = Ya(1 - \beta) \tag{1}$$

if there is no flexible line from P to CBD, and $Y_1 = 0$ otherwise. Similarly,

$$Y_2 = Ya\beta \tag{2}$$

if there is no flexible line from P to SC, and $Y_2 = 0$ otherwise. Also,

$$Y_3 = Y(1 - a) \tag{3}$$

If there is no flexible line from SC to CBD, and $Y_3 = 0$ otherwise. We begin with operators' costs C_O which depend on the number of buses B and their capacity $\boldsymbol{K}$, through the expression:

$$C_O = B(c_0 + c_1\boldsymbol{K}) \tag{4}$$

where c_0 and c_1 are exogenous parameters. Note that $\boldsymbol{BK}$ represents the number of seats in the system. The number of buses depends on the cycle time t_c, with

$$\boldsymbol{B} = \boldsymbol{f}t_c \tag{5}$$

The cycle time is given by T plus the time spent at the bus stops. Let us denote by t the time spent by each user to board and to alight (2 t in total), and $Y = Y_1 + Y_2 + Y_3$

the total hourly demand of the fixed line. Therefore, each bus carries Y/f passengers, meaning that:

$$t_c = T + \frac{2tY}{f} \tag{6}$$

Finally, as costs increase with the capacity, this is adjusted to be able to carry all the passengers in the most loaded of the two arcs:

$$\boldsymbol{K} = \frac{max\{Y_1 + Y_2, Y_1 + Y_3\}}{f} \tag{7}$$

We now continue with user costs, which depend on the average waiting time t_w and the in-vehicle time t_v, $C_U = Y(p_w t_w + p_v t_v)$, where p_w and p_v are exogenous parameters representing time values. Average waiting time is calculated as:

$$t_w = \frac{1}{2f} \tag{8}$$

Equation (8) assumes that users arrive uniformly during the interval between two consecutive buses. Average in-vehicle time t_v is the sum of the time in-motion t_m and the time spent at stops t_s. The former can be computed as the time needed to tour the whole path between P and CBD ($T/2$), times the average proportion of such distance traveled by the passengers:

$$t_m = \frac{T}{2} \frac{Y_1(g+1)L + Y_2 gL + Y_3 L}{Y(g+1)L} \tag{9}$$

When a user visits an intermediate stop in her journey (in our case, this only happens for a $P - CBD$ user when stopping at SC), they spend there the whole time that the bus is stopped, where we assume that boarding and alighting occur simultaneously (using different doors). On the other hand, in the final stop, this is reduced, on average, to half that time. Therefore:

$$t_s = \frac{t}{fY}\left[Y_1 \cdot \left(max\{Y_2, Y_3\} + \frac{Y_1 + Y_3}{2}\right) + Y_2 \cdot \frac{Y_2}{2} + Y_3 \cdot \left(\frac{Y_1 + Y_3}{2}\right)\right] \tag{10}$$

Putting everything together, the total value of the resources consumed, as a function of the frequency, is:

$$VRC(f) = \mu f + \frac{G}{f} + \delta \tag{11}$$

With

$$\mu = Tc_0 \tag{12}$$

$$G = 2tYc_1 max\{Y_1 + Y_2, Y_1 + Y_3\} + \frac{p_w Y}{2} + p_v\left[Y_1 \cdot \left(max\{Y_2, Y_3\} + \frac{Y_1 + Y_3}{2}\right) + Y_2 \cdot \frac{Y_2}{2} + Y_3 \cdot \left(\frac{Y_1 + Y_3}{2}\right)\right] \tag{13}$$

$$\delta = 2tYc_0 + c_1 Tmax\{Y_1 + Y_2, Y_2 + Y_3\} + p_v \frac{T}{2} \frac{Y_1(g+1) + Y_2 g + Y_3}{(g+1)} \tag{14}$$

Which implies that the optimal frequency is:

$$\boldsymbol{f}_* = \sqrt{\frac{\boldsymbol{G}}{\boldsymbol{\mu}}} \tag{15}$$

And the costs of the fixed line are:

$$C = 2\sqrt{G \cdot \mu} + \delta \tag{16}$$

4.3 Flexible lines optimization

In order to find closed expressions for the different components of the cost functions for the flexible lines, we adapt the method by Fielbaum (2020), which analyzes the case where all the users travel to the center of a zone (similar to the feeder lines of our model, but with passengers traveling in only one direction). For ODRP, there is no such a thing as frequencies, so the challenge is to write both operators and users costs as a function of other design variables.

As in this model ODRP is part of a public transport network, it is not a strange feature that users have to walk, which has been shown to be very efficient for ODRP systems (Fielbaum 2022; Fielbaum et al. 2021a; Martin et al. 2021; Wang et al. 2022). Specifically, we assume that users gather at a meeting point in the origin, so that nobody walks more than $\boldsymbol{l}$, which is a variable to be optimized. In the destination we do not follow the same rule, as this would force the system to find users whose origins and destinations are nearby (at a walkable distance), which can be too difficult to achieve, reducing the degree of shareability of the system (actually, Fielbaum 2020 shows that such a strategy would preclude sharing at all). Therefore, users are dropped off at their exact destination. Moreover, by including both alternatives (walking at the origin, not walking at the destination), the effect of both options is captured when we compare the line structures and analyze the evolution of total costs when the key parameters change.

The analysis of the cost function distinguishes two different cases.

4.3.1 A line that connects different zones (direct or semi-direct)

Let us begin by noting that, due to the demand structure in which all users travel in the same direction, the vehicle returns empty to its origin zone. We denote by D the distance between the center of the origin and destination zones. The number of hourly users of the line is denoted by y.

As we assume that users are homogeneously distributed in time and space, if they walk no more than $\boldsymbol{l}$, it means that the walking (access) time t_a is given by:

$$t_a = \frac{2\boldsymbol{l}}{3v_a} \tag{17}$$

where $\frac{2l}{3}$ is the average distance between the points within a circle with radius $\boldsymbol{l}$ and its center, and v_a is the walking speed.[6]

To calculate the average waiting time, we follow Fielbaum (2020), by assuming that the system is perfectly coordinated, so that whenever $\boldsymbol{K}$ users have emerged, the vehicle with capacity $\boldsymbol{K}$ will arrive at the gathering point; we will expand on this assumption later. This can be achieved thanks to the regularity assumptions on the users' spatial distribution, which ensures that, given a fixed time lapse, the same number of users will emerge within a circle of radius l. The time lapse required for $\boldsymbol{K}$ users to emerge in a circle with radius $\boldsymbol{l}$ is $\boldsymbol{K}\frac{r_O^2}{y\boldsymbol{l}^2}$, where r_O is the radius of the origin zone. On average, users will wait for half that time (the first user to emerge waits for that full time lapse, but the last one waits zero), hence:

$$t_w = \boldsymbol{K}\frac{r_O^2}{2y\boldsymbol{l}^2} \tag{18}$$

Let us denote by v_c the velocity of the vehicles.[7] We assume that, in order to travel between zones, it is always needed to go from center to center. Therefore, average in-vehicle time at the origin zone is $\frac{r_O}{3v_c}$ and the time to connect the origin and destination zones is $\frac{D}{v_c}$. In the case of the FTS, there is also a detour in the destination's zone to deliver all the users, which can be seen as the solution of a random traveling-salesman-problem in the destination circle visiting $\boldsymbol{K}$ destinations, such that each vehicle covers an expected length of[8] $\gamma r_D\sqrt{\boldsymbol{K}}$ at the destination, with

6 Note that we are disregarding the specific effects at the borders of the circles, in order to allow for an explicit expression of these equations. This means that walking times would be lower than our results, which happens for all the line structures, so it does not have a relevant effect on the comparison between them.

7 The velocity of the vehicles could be different for each type of line, reflecting that they could use different technologies. However, there is no obvious way to do this: for instance, if the fixed line uses traditional buses, it could be slower than the other lines, whereas the contrary would happen if using BRT or Metro. We opt to use the same vehicle's velocity everywhere to avoid biasing our analysis.

8 This is actually a convergence result when the number of points is large (a special case of the Beardwood–Halton–Hammersley Theorem). However, the dependency on the square root of the number of points holds for smaller values.

$\gamma \approx 1.26$ (Applegate et al. 2007). Users spend on average half that time in the vehicle. Putting everything together[9]:

$$t_v = \frac{1}{v_c}\left(\frac{2r_O}{3} + D + \frac{\gamma}{2} r_D \sqrt{\boldsymbol{K}} I_{FTS}\right) \tag{19}$$

where I_{FTS} is a binary function indicating if this is an inter-zonal line in FTS (case in which $I_{FTS} = 1$), or not (so $I_{FTS} = 0$). The factor $\sqrt{\boldsymbol{K}}$ captures that users will face longer detours when the vehicle needs to deliver more passengers, but that such detours increase at a decreasing rate because more users imply that their destinations are situated closer to each other.

Operators' costs depend on $\boldsymbol{K}$ and in the number of small vehicles $\boldsymbol{B}$, following Eq. (4). We now find a relationship between these two variables, leveraging the fact that vehicles and users are perfectly coordinated, i.e., that vehicles do not need to wait for users. Let us define as ρ the average load of the vehicle in time. The following two expressions represent the same amount, namely the pax-hour driven per hour by the flexible line, such that the resulting unit on both sides is in passengers, and the total pax-hours offered by the system would be obtained by multiplying the expressions by the total operational time.

$$\boldsymbol{BK}\rho = y t_v \tag{20}$$

The left-hand side is the total capacity moved by the system multiplied by the average load (supply), and the right side is the number of users multiplied by the average time they spent using the line (demand). Note that for these lines $\rho = \frac{1}{2}$: if vehicles do not deliver the users at their exact destinations ($I_{FTS} = 0$), they spend half of the time full and half of the time empty; otherwise ($I_{FTS} = 1$), an additional path needs to be considered, where users are being dropped one by one, so that on average vehicles are at half capacity.

We now explain how Eq. (20) ensures that the system is coordinated, as mentioned above. In fact, t_v/ρ is the cycle time of the vehicle, so that $t_v/\boldsymbol{B}\rho$ is the headway h between two consecutive vehicles reaching the origin zone, thus Eq. (20) implies that $h = \boldsymbol{K}/y$. In a traditional public transport system, we have the waiting time fulfilling $t_w/h = 1/2$ (Jara-Díaz and Gschwender 2009). In this case, combining Eqs. (18) and (20) we obtain that the same factor gets multiplied by $r_O^2/\boldsymbol{l}^2$, which is the proportion of vehicles useful to a particular passenger. In other words, if we had an inequality instead of Eq. (20), the users' waiting time would be either longer, meaning that the vehicle would not have enough time to arrive when the users are gathered, or shorter, meaning that not enough users would have emerged when the vehicle arrives.

Putting everything together, we can write the total costs of these flexible lines as a function of $\boldsymbol{K}$ and $\boldsymbol{l}$:

[9] We are disregarding the boarding/alighting time for ODRP because these are small vehicles, so that this time has little effect. This is a usual assumption in previous studies as well (Bilali et al. 2019, Fielbaum 2020).

$$VRC(\boldsymbol{l}, \boldsymbol{K}) = \alpha_1 \boldsymbol{l} + \alpha_2 \frac{\boldsymbol{K}}{\boldsymbol{l}^2} + \alpha_3 \frac{1}{\boldsymbol{K}} + \alpha_4 + I_{FTS}\left[\alpha_5 \sqrt{\boldsymbol{K}} + \alpha_6 \frac{1}{\sqrt{\boldsymbol{K}}}\right] \tag{21}$$

With:

$$\alpha_1 = y\frac{2p_a}{3v_a} \tag{22}$$

$$\alpha_2 = \frac{p_w r_O^2}{2} \tag{23}$$

$$\alpha_3 = \frac{yc_0}{\rho v_c}\left(\frac{2r_O}{3} + D\right) \tag{24}$$

$$\alpha_4 = \frac{yc_1}{\rho v_c}\left(\frac{2r_O}{3} + D\right) + y\frac{p_v}{v_c}\left(\frac{r_O}{3} + D\right) \tag{25}$$

$$\alpha_5 = yp_v\frac{\gamma}{2}\frac{r_D}{v_c} + \frac{yc_1\gamma r_D}{2\rho v_c} \tag{26}$$

$$\alpha_6 = \frac{\gamma}{2}\frac{yc_0 r_D}{\rho v_c} \tag{27}$$

Following the standard assumptions of simplified public transport models, we regard $\boldsymbol{K}$ as continuous for this step, so that we can take the derivative of the cost function with respect to $\boldsymbol{K}$. When $I_{TFS} = 0$, the optimal values of the design variables can be easily computed through the first order conditions, and are

$$\boldsymbol{K} = \frac{2^{1/2}\alpha_3^{3/4}}{\alpha_1^{1/2}\alpha_2^{1/4}}, \boldsymbol{l} = \frac{2^{1/2}\alpha_3^{1/4}\alpha_2^{1/4}}{\alpha_1^{1/2}} \tag{28}$$

With a cost of:

$$C_{PTS} = 2\sqrt{2}\left[\alpha_1^{1/2}\alpha_2^{1/4}\alpha_3^{1/4}\right] + \alpha_4 \tag{29}$$

When $I_{TFS} = 1$, due to the presence of the terms that depend on $\sqrt{\boldsymbol{K}}$, the first-order conditions lead to equations of degree larger than 4, hence there is no analytical solution. Therefore, for the analytical investigation, we will use the same expression given by Eq. (28), which is a sub-optimal solution but useful to have some intuition concerning the role played by the different parameters. It is worth noting that the two terms that are disregarded when using this solution (last two terms in Eq. 21) push into opposite directions. This leads to total costs equal to:

$$C_{FTS} = 2\sqrt{2}\left[\alpha_1^{1/2}\alpha_2^{1/4}\alpha_3^{1/4}\right] + \alpha_4 + I_{TFS}\left[\alpha_5\frac{2^{1/4}\alpha_3^{3/8}}{\alpha_1^{1/4}\alpha_2^{1/8}} + \alpha_6\frac{\alpha_1^{1/4}\alpha_2^{1/8}}{2^{1/4}\alpha_3^{3/8}}\right] \quad (30)$$

When we run numerical simulations (Sect. 5), we will use the optimal values for $\boldsymbol{K}$ and $\boldsymbol{l}$.

4.3.2 Feeder lines

As we assume that at the origin users need to walk, but at the destination they are served at their door, the analysis is different when the feeder only takes users towards the center of its zone (as in *P*), when it only distributes users from such a center (as in *CBD*), or when it does both (as it can happen in *SC*). In the first two cases, we denote by y the number of users of the line, while in the third case we need to disentangle $y = y_{in} + y_{out}$, with y_{in} corresponding to the users traveling towards the center and y_{out} being the ones traveling from the center. The radius of the zone is denoted by r.

Case in which users only go to the center:

The access, waiting and in-vehicle times are calculated similar to the analysis above, leading to:

$$t_a = \frac{2\boldsymbol{l}}{3v_a}, t_w = \boldsymbol{K}\frac{r^2}{2y\boldsymbol{l}^2}, t_v = \frac{2r}{3v_c} \quad (31)$$

Moreover, it still holds that $\boldsymbol{BK}\rho = yt_v$, with $\rho = \frac{1}{2}$. Everything together yields:

$$VRC(\boldsymbol{l}, \boldsymbol{K}) = v_1\boldsymbol{l} + v_2\frac{K}{\boldsymbol{l}^2} + v_3\frac{1}{\boldsymbol{K}} + v_4 \quad (32)$$

which implies

$$C = 2\sqrt{2}\left[v_1^{1/2}v_2^{1/4}v_3^{1/4}\right] + v_4 \quad (33)$$

where

$$v_1 = \frac{2yp_a}{3v_a}, v_2 = \frac{p_w r^2}{2}, v_3 = \frac{2c_0 yr}{3\rho v_c}, v_4 = \frac{2yr}{3v_c}\left(p_v + \frac{c_1}{\rho}\right) \quad (34)$$

Case in which users only go from the center:

In this case, all users travel from the center of the corresponding circle and can share the vehicle. Therefore, a vehicle departs just when it gathers K passengers (similar to Li and Quadrifoglio 2010), making:

$$t_w = \frac{\boldsymbol{K}}{2y} \quad (35)$$

There is no walking time, and in-vehicle time is:

$$t_v = \frac{\gamma}{2v_c} r\sqrt{\boldsymbol{K}} \tag{36}$$

It still holds that $\boldsymbol{BK}\rho = yt_v$, with $\rho = \frac{1}{2}$. Everything together yields:

$$VRC(\boldsymbol{K}) = \varepsilon_1 \boldsymbol{K} + \varepsilon_2 \sqrt{\boldsymbol{K}} + \varepsilon_3 \frac{1}{\sqrt{\boldsymbol{K}}} \tag{37}$$

with

$$\varepsilon_1 = \frac{p_w}{2}, \varepsilon_2 = \frac{y\gamma r}{2v_c}\left(p_v + \frac{c_1}{\rho}\right), \varepsilon_3 = \frac{c_0 y\gamma r}{2v_c\rho} \tag{38}$$

The explicit solution for such an equation is extremely long, so we do not write it here[10].

Case in which passengers travel in both directions:

Let us denote by K_{in} and K_{out} the number of users boarding the vehicle in each direction, such that

$$\boldsymbol{K} = max(K_{in}, K_{out}) \tag{39}$$

We assume that all the vehicles serve consecutively one inner and one outer trip, and that thanks to the regularity assumptions the finishing point of one trip coincides exactly with the starting point of the following one, where users have just gathered to board.

The waiting time is calculated as the weighted average between the directions:

$$t_w = \frac{y_{in}}{y_{in} + y_{out}} K_{in} \frac{r^2}{2y_{in}\boldsymbol{l}^2} + \frac{y_{out}}{y_{in} + y_{out}} \frac{K_{out}}{2y_{out}} \tag{40}$$

Average access time only exists for those traveling towards the center:

$$t_a = \frac{y_{in}}{y_{in} + y_{ouu}} \frac{2\boldsymbol{l}}{3v_a} \tag{41}$$

Average in-vehicle time also depends on the direction:

$$t_v = \frac{y_{in}}{y_{in} + y_{out}} \frac{2r}{3v_c} + \frac{y_{out}}{y_{in} + y_{out}} \frac{r\gamma\sqrt{K_{out}}}{2v_c} \tag{42}$$

Finally, with regard to the extra equations to write the passenger-hours-traveled in two different ways (to link the number of vehicles $\boldsymbol{B}$ and their capacity $\boldsymbol{K}$), this now has to happen in both directions, that is

[10] See https://www.wolframalpha.com/input?i=solve+A+x%5E%283%2F2%29+%2B+Bx%2F2+-+C%2F2+%3D+0 (accessed on 25/02/2022).

$$\boldsymbol{B}K_{in}\rho_{in} = y_{in}\frac{2r}{3v_c}, \boldsymbol{B}K_{out}\rho_{out} = y_{out}r\gamma\frac{\sqrt{K_{out}}}{2v_c} \tag{43}$$

The parameter ρ_{in} now represents the average load of people traveling towards the center that are in the vehicle. While the vehicle travels towards the center, the load is 1, and when it is delivering the passengers from the center, the load is zero, leading to:

$$\rho_{in} = \frac{2r/3v_c}{2r/3v_c + r\gamma\sqrt{K_{out}}/v_c}\cdot 1 + \frac{r\gamma\sqrt{K_{out}}/v_c}{2r/3v_c + r\gamma\sqrt{K_{out}}/v_c}\cdot 0 = \frac{2}{2+3\gamma\sqrt{K_{out}}} \tag{44}$$

where we are using the fact that the vehicle alternates one trip in direction "in" and one trip in direction "out". The analysis for ρ_{out} is equivalent, but we note that the average load while distributing the users from the center is $1/2$, which leads to:

$$\rho_{out} = \frac{2r/3v_c}{2r/3v_c + r\gamma\sqrt{K_{out}}/2v_c}\cdot 0 + \frac{r\gamma\sqrt{K_{out}}/v_c}{2r/3v_c + r\gamma\sqrt{K_{out}}/v_c}\cdot\frac{1}{2} = \frac{3\gamma\sqrt{K_{out}}}{4+6\gamma\sqrt{K_{out}}} \tag{45}$$

As we have two ways to write B, straightforward algebra leads to an equality that states that the load per direction is proportional to the demand:

$$\frac{K_{out}}{K_{in}} = \frac{y_{out}}{y_{in}} \tag{46}$$

To put everything together, let us define $\Gamma = max\left(1, \frac{y_{out}}{y_{in}}\right)$ so that $\boldsymbol{K} = \Gamma\cdot K_{in}$. Then[11]:

$$VRC(\boldsymbol{K_{in}}, \boldsymbol{l}) = \phi_1\boldsymbol{l} + \phi_2\frac{\boldsymbol{K_{in}}}{\boldsymbol{l}^2} + \phi_3\boldsymbol{K_{in}} + \phi_4\sqrt{\boldsymbol{K_{in}}} + \phi_5\frac{1}{\sqrt{\boldsymbol{K_{in}}}} + \phi_6\frac{1}{\boldsymbol{K_{in}}} + \phi_7 \tag{47}$$

with

$$\begin{aligned}
&\phi_1 = \frac{2p_a y_{in}}{3v_a}, \phi_2 = \frac{p_w r^2}{2}, \phi_3 = \frac{3p_w y_{out}}{y_{in}}, \phi_4 = \frac{p_v y_{out} r\gamma}{2v_c}\cdot\sqrt{\frac{3y_{out}}{y_{in}}} + \Gamma\cdot\frac{c_1 r\gamma}{v_c}\sqrt{3y_{out}y_{in}},\\
&\phi_5 = \frac{c_0 r\gamma}{v_c}\sqrt{3y_{out}y_{in}}, \phi_6 = \frac{2c_0 y_{in} r}{3v_c}, \phi_7 = \frac{2p_v y_{in} r}{3v_c} + \frac{2c_1 y_{in} r\Gamma}{3v_c}
\end{aligned} \tag{48}$$

These first-order conditions cannot be solved analytically, so when this line exists, it will be optimized numerically.

[11] As now $\boldsymbol{K}$ is written as a function of $\boldsymbol{K_{in}}$, we put the latter in bold in what follows.

Table 3 Parameters used in the simulations

Parameter	a	Y	β	p_v	p_w	p_a	p_T	c_0
Value	0.8	1000 [pax/h]	0.25	2.32 [US$/h]	4.64 [US$/h]	6.96 [US$/h]	0.31 [US$]	4.02 [US$/h]
Parameter	L	g	t	v_c	v_a	r_{CBD}	r_{SC}	r_P
Value	5 [km]	1.5	2.5 [sec]	$25/\sqrt{2}$ [km/h]	$5/\sqrt{2}$ [km/h]	3 [km]	5 [km]	10 [km]

Velocities are divided by $\sqrt{2}$ to account for a Manhattan distance instead of Euclidean. Parameters $a, \beta, p_v, p_w, p_a, c_0, c_1$ are obtained from Fielbaum (2020). The transfer penalty p_T is a moderate value compared to the literature (Jara-Diaz et al. 2022)

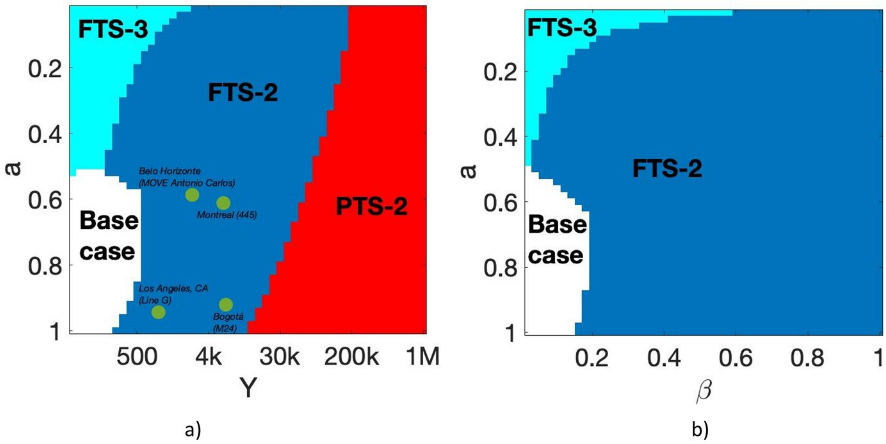


Fig. 4 Line structures with the lowest total costs, for different combinations of the parameters: **a** ***a*** (percentage of users departing from ***P***) and ***Y*** (total number of passengers per hour), and **b** ***a*** and β (degree of polycentricity)

4.4 Cost per structure

Following Tables 1 and 2, we have seven different structures (including the basic one). For each of them, the total cost can be characterized by Eqs. (16), (29), (30), (33), (37) and (47), plus transfer penalties for those users that need to transfer: as argued by Garcia-Martinez et al. (2018), the discomfort of transfers can be captured by the model through a constant penalty p_T, which is multiplied by the total number of transfers in the system. In Appendix A we identify, for each structure, the fixed and flexible lines, the corresponding parameters for each of them, and the resulting costs.

5 Results

5.1 Global results

Let us begin the analysis by identifying which is the best line structure, among the seven defined above, for different combinations of the parameters. To do so, we consider base values for the parameters as defined in Table 3, where the p_T parameter represents 8 equivalent in-vehicle minutes (EIVM), a conservative value indeed as suggested by the specialized literature (see Jara-Diaz et al. 2022, who suggest a universal EIVM between 13 and 18).

We consider three key parameters for this part of the analysis, which characterize the demand pattern: a, β and Y. We depict which is the best structure for a given combination of (a, Y) and (a, β), in Fig. 4 left (Y in logarithmic scale) and right, respectively.[12]

The first conclusion we obtain from Fig. 4, which answers the main question we investigate in this paper, is that **the base case scenario (feeder-trunk-feeder) is only optimal under very specific circumstances**, namely when a is large, i.e., most trips begin at P, and either β approaches 0 (implying that there is almost no demand departing or arriving at the subcenter, so that the load of the buses remains the same during the whole route), or Y approaches 0 (a very small demand, where gathering passengers with different destinations can be crucial). On the other hand, both FTS and PTS structures might be the best ones, depending on the combination of the parameters.

In the left figure we include representative lines from cities where[13] $\beta \approx 0.25$.

Figure 4a reveals that for a small demand, it is better to use FTS-3 or the Base case, depending on the value of a. As Y grows, it becomes better to use FTS-2 and then PTS-2. Remarkably, this does not follow a relevant scale-related result that has been reported in traditional public transport by Fielbaum et al. (2020a): **a larger demand does not imply using more direct line structures when ODRP is also involved** (recall that, in general terms, directness evolves in this scheme from the base case to PTS, and then to FTS, where FTS-3 is the most direct possible structure). The reasons explaining why those results do not apply here, where we also have ODRP, are discussed in the following subsection when we study in more detail the evolution of the system as the demand increases.

[12] It is worth mentioning that, after rounding, the capacity of the small vehicles ranges between 3 and 11 seats.

[13] Estimating these parameters is a complex task that is out of the scope of this paper, so we only use available exogenous information. The internal distribution of the trips was obtained from Jara-Diaz and Olea (2021) for Belo Horizonte, Montreal, and Bogotá, and from Fielbaum et al. (2017) for Los Angeles (CA). The number of users per line was obtained from https://www.transmilenio.gov.co/loader.php?lServicio=Tools2andlTipo=descargasandlFuncion=descargarandidFile=47467 (accessed on 18/06/2022, Bogotá), https://stringfixer.com/tr/Los_Angeles_Metro_Busway (accessed on 18/06/2022, Los Angeles), https://www.itdp.org/2014/05/27/belo-horizonte-launches-second-brt-corridor-as-world-cup-nears/ (accessed on 18/06/2022, Belo Horizonte), and https://globalnews.ca/news/6408983/reserved-bus-lane-slowing-down-papineau-avenue-rush-hour-traffic-stm/ (accessed on 18/06/2022, Montreal). When the hourly information was not available, we assumed the same proportion as in Bogotá between peak morning and total daily trips.

Figure 4b shows that when β is large, i.e. when there are almost no $P - CBD$ trips, it is better to utilize the FTS-2 structure. When a is large, this result reinforces the discussion in the previous paragraph: most of the users are traveling from P to SC, and they are precisely the ones using the traditional system; in other words, the OD pair that exhibits the largest demand is the one with the least direct structure.

On the other hand, when a is small, one could expect that an analogous result would be that passengers going from SC to CBD (the most demanded OD pair in such a case) should also travel using the fixed line, meaning that FTS-1 should be preferred. To understand why this is not the case, it is worth noting that FTS-1 and PTS-1 never appear as the dominant structures. The reason for this is that those structures have a fixed line that goes from P to CBD, but the users of the second leg ($SC - CBD$) are reduced in comparison with the base case, since those going from P to CBD have a direct or semi-direct line, implying that there is an increased idle capacity there (compared to the base case). Following Tables 1 and 2, in FTS-2 and PTS-2 the fixed line cycle is shortened as it does not arrive to the $\boldsymbol{CBD}$, which is why they outperform FTS-1 and PTS-1, leading to lower operational costs (as confirmed in Fig. 6). This suggests a relevant high-level conclusion that can be stated as **if some of the users traveling in the first segments of a fixed-route line can now use ODRP, it might be better to offer ODRP to all the users in such segments, so that the line length can be shortened.**

5.2 Evolution of the costs with the demand

We now show, in Figs. 5, 6, 7, the evolution of the different components of users' and operator's costs as the demand grows. Let us begin with the user costs by depicting average in-vehicle (left), waiting (center), and walking (right) times in Fig. 5.

Regardless of the line structure, in-vehicle time increases with the demand, which happens because a greater demand leads to larger vehicles, which affects all lines:

- In the fixed line, vehicles carry more passengers and hence spend more time waiting for them to board and alight at the bus stops. This is a well-known fact in traditional public transport analysis (Fielbaum et al. 2020).
- In the flexible lines, a larger capacity implies a longer detour when dropping off the passengers at their destinations (Eq. 19).

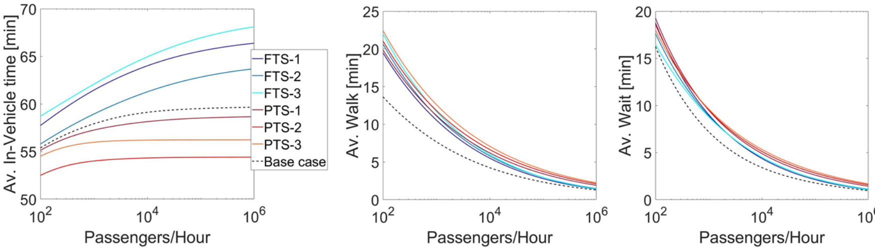


Fig. 5 Evolution of the components of users' average traveling times as the total demand grows

Figure 5 on the left also explains why a greater demand does not imply a more direct structure. The FTS structures present the largest in-vehicle times, with FTS-3 being the largest of all. Moreover, this difference becomes more relevant as the demand grows. This happens because FTS structures are the only ones in which the same vehicle does the following two things: it connects different zones, and it delivers the users to their final destinations. Connecting different zones implies a long traveling time, which leads to bigger vehicles (through α_3 in Eq. 28), namely because otherwise too many vehicles would be required. But large vehicles delivering the users to their final destinations imply a long detour. Therefore, a relevant operational conclusion is that **there is a virtue in splitting the long and short legs of a trip, to tailor the vehicles' sizes and avoid long detours in the short parts**. Note that the increase in in-vehicle time, as the demand grows in ODRP, corresponds to a source of scale diseconomies, identified as the "Flex-Route effect" by Fielbaum et al. (2023).

Waiting and walking times, on the other hand, show a similar pattern. They both decrease with the demand, which is explained because there are more vehicles and it is easier to gather the users together, two sources of scale economies (denoted by Fielbaum et al. 2023 as "Mohring effect" and "Better-matching effect", respectively). All structures present similar values, with the exception of the base case in which both walking and waiting are significantly shorter, which is explained because the two transfers imply that it is very easy to find users that can share a vehicle.

In Fig. 6 we analyze the evolution of the total fleet and the total number of seats, defined as the sum of the capacity of all the vehicles in the system, which characterize the total operators costs (Eq. 4). PTS-3 yields the lowest operator costs, but this is not enough to make it optimal. It is noteworthy that the three structures in which the fixed line route is not shortened, i.e., FTS-1, PTS-1, and the base case, present the highest operator costs. The other structures present similar values.

Finally, in Fig. 7 we show average total costs. It reveals that FTS-1 and PTS-1 are usually the worst structures, whose reasons were already discussed above. The base case is the best one for low demand levels, but it becomes one of the

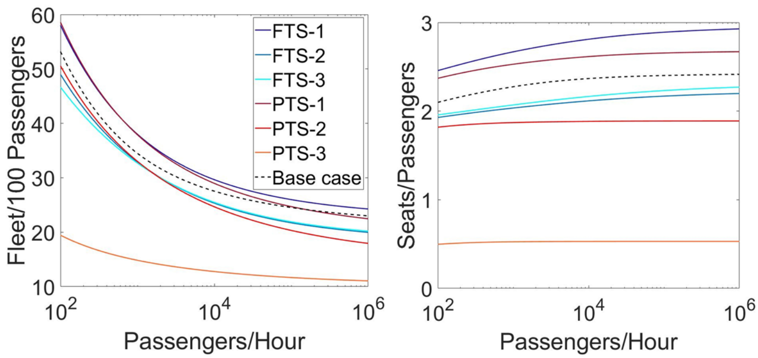


Fig. 6 Evolution of the sources of operator costs as the demand grows

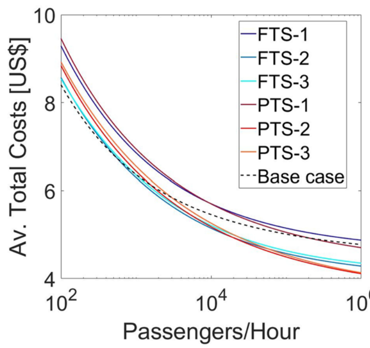


Fig. 7 Evolution of average total costs as the demand increases

worst when the demand is higher. This reinforces the main conclusion of this paper: namely, that other types of spatial integration can be better than assuming ODRP to serve only as a feeder, depending on the circumstances. PTS-3 presents almost the same costs as PTS-2, meaning that it is almost the best structure for many combinations. All curves are decreasing, implying that scale economies prevail, which is expected as this is the case for public transport and for ODRP.

5.3 Role of the transfer penalty

It is a well-known fact that disregarding the discomfort induced by transfers can have a relevant effect on the optimal design of public transit networks (Fielbaum et al. 2016). Moreover, capturing the exact effect of transfers can be challenging from a methodological point of view (Garcia-Martinez et al. 2018). In this subsection, we investigate which would be the dominant line structures if p_T is assumed to be either 0 or twice its original value.

Let us begin with the case $p_T = 0$, depicted in Fig. 8. This figure is equivalent to Fig. 4, only changing p_T. The contrast with Fig. 4 is evident. The base case

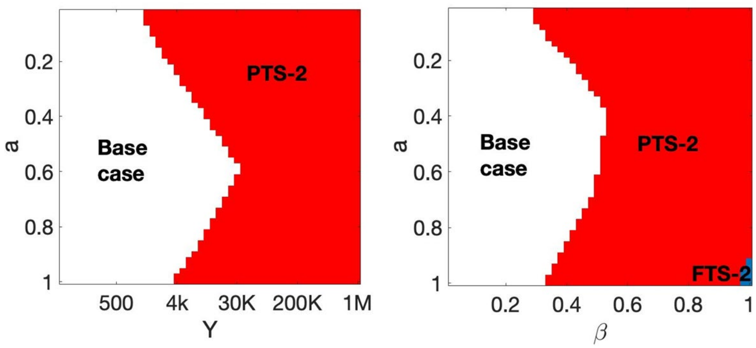


Fig. 8 Representation of the line structure with the lowest total costs, for different combinations of the parameters $\boldsymbol{a}$, $\boldsymbol{Y}$ and β, imposing no transfer penalty

dominates in a much wider zone, and FTS almost disappears. Therefore, we can conclude that **imposing ODRP to serve as a feeder involves an implicit assumption, namely that transfers are not uncomfortable.** On the other extreme, when the value of p_T is doubled (reaching the recommended interval in Jara-Diaz et al. 2022), the only structure that does not involve transfers, FTS-3, is dominant everywhere. Such massive changes reinforce that having an appropriate measurement of the transfer penalty is crucial to have an accurate design.

5.4 Sensitivity analysis

As our model involves several parameters, we analyze the impact of each of them on the different line structures. To do this, we modify each of the exogenous parameters from half to double their original value and compute the average cost for each of the line structures. The modification of the parameters is done one-by-one, and the other parameters are kept fixed. The figures and details can be found in the Appendix B. We summarize here the most relevant conclusions:

- If passengers are traveling between distant points (i.e., larger values of g or L), massive vehicles are more efficient and the base case is favored.
- If the zones P or CBD are smaller (i.e., lower values of r_P or r_{CBD}), it becomes easier to gather them at the origin and to take them to their final destinations, favoring FTS and PTS.
- If vehicles can travel faster (i.e., larger value of v_c), FTS structures are favored because it becomes less relevant to split the long and short parts of a trip.

6 Synthesis and conclusions

There is relevant potential in designing public transit networks that offer jointly traditional fixed-route and on-demand services. Most previous research has addressed this challenge imposing that ODRP should be limited to feed the massive fixed network, hence solving the so-called first-mile and last-mile problems. In this paper, we study whether this assumption is correct, by taking this solution as a benchmark or base case and comparing it with direct-based (FTS) and semi-direct-based (PTS) solutions in a simplified linear city: in FTS, some passengers execute their whole trip using ODRP, while in PTS, some users travel from their origin to a different zone using a single ODRP vehicle, but still face the last-mile through a feeder.

By including a number of simplifications regarding the operation of the public transport system, we derived analytical expressions for all the characteristics defining the costs: traveling times for the users, and fleet attributes for the operators. Such expressions depend on some design variables (namely the size of the fleet and how much should users walk), implying that they can be optimized to minimize total costs. We do so for different combinations of the parameters and obtain the resulting costs for each of the structures we analyze.

Our main conclusion is to observe that the base case situation is optimal only under very specific circumstances, such as low-demand monocentric cities. It can also be optimal if one assumes that transfers are not uncomfortable for the users, as the base case imposes two transfers to every user. Otherwise, the inclusion of flexible direct or semi-direct services can improve the system, as they can reduce the number of transfers and offer competitive operator costs.

We show that when routes are flexible, it is no longer true that line structures become more direct as the demand grows (which does happen in traditional public transport). The reason for this is that the routes' length increases when more users share the vehicle that transport them to their doors. This problem is partially relieved when the longest part of the route is traveled in a large vehicle (which is less costly), and the last leg is done in a smaller vehicle that induces a shorter detour.

The two structures that involve only one direct or semi-direct line (FTS-1 and PTS-1) exhibit the highest costs, regardless of the scenario we study. In both structures, only a small number of the users can board the direct or semi-direct service. In particular, both structures remove a fraction of the fixed line users in its first segment, but not all of them, so that the fixed line cannot be made shorter, but does lose part of its demand and hence scale economies are not leveraged. This fact illustrates the complexity in the relationship between the flexible and the fixed lines.

In all, this paper shows that it is urgent to propose ways to integrate on-demand services into public transit in which ODRP is not constrained to operate as a feeder. Our analysis is mostly limited by the simplifying assumptions we have included in the analysis: a simplified city model, with temporal homogeneity, no explicit street network, and considering that all the design variables are continuous (including vehicles' capacities). Although we believe that our stylized model captures the essence of the problem, the most relevant directions for future research are performing a similar analysis in a more realistic setting, where complex relationships between the network structure and the lines offered might emerge. Particularly, the general problem of designing a traditional public transport network with its frequencies has been thoroughly studied in the last decades, and yet it remains open due to its complexity (Durán-Micco and Vansteenwegen 2022): modifying this problem to include some on-demand lines, and developing heuristics to solve it, might be one of the most relevant research questions in public transport theory in the years to come.

Appendix A: Lines and parameters per structure

1. FTS-1
   Fixed line: $Y_1 = 0, Y_2 = Ya\beta, Y_3 = Y(1-a), A = (g+1)L$
   Feeder line within $\boldsymbol{P}$: $y_{in} = Ya\beta, y_{out} = 0, r = r_P,$
   Feeder line within $\boldsymbol{SC}$: $y_{in} = Y(1-a), y_{out} = Ya\beta, r = r_{sc}$
   Feeder line within $\boldsymbol{CBD}$: $y_{in} = 0, y_{out} = Y(1-a), r = r_{CBD}$
   Direct line $P_i - CBD_i$: $y = Ya(1-\beta), r_O = r_P, r_D = r_{CBD}, D = (g+1)L$
   Number of transfers: $0 \cdot Ya(1-\beta) + 2 \cdot Y(1-a+a\beta)$
   2. FTS-2
   Fixed line: $Y_1 = 0, Y_2 = Ya\beta, Y_3 = 0, A = gL$

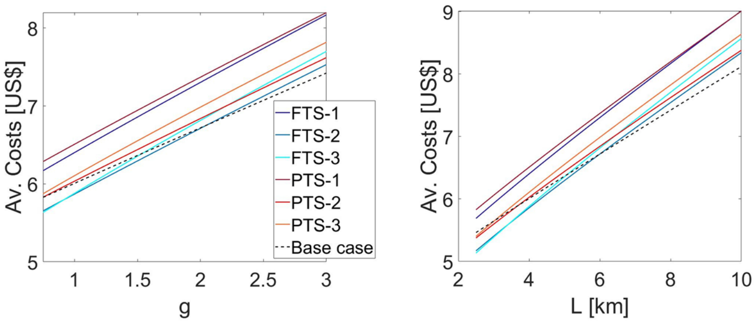


Fig. 9 Sensitivity analysis with respect to the distances between the periphery, subcenter and CBD

Feeder line within $\boldsymbol{P}$: $y_{in} = Ya\beta, y_{out} = 0, r = r_P,$
Feeder line within $\boldsymbol{SC}$: $y_{in} = 0, y_{out} = Ya\beta, r = r_{sc}$
Direct line $P_i - CBD_i$: $y = Ya(1-\beta), r_O = r_P, r_D = r_{CBD}, D = (g+1)L$
Direct line $SC_i - CBD_i$: $y = Y(1-a), r_O = r_{SC}, r_D = r_{CBD}, D = L$
Number of transfers: $0 \cdot Y[a(1-\beta) + 1 - a] + 2 \cdot Ya\beta$
3. FTS-3
Direct line $\boldsymbol{P - SC}$: $y = Ya\beta, r_O = r_P, r_D = r_{SC}, D = gL$
Direct line $\boldsymbol{P - CBD}$: $y = Ya(1-\beta), r_O = r_P, r_D = r_{CBD}, D = (g+1)L$
Direct line $\boldsymbol{SC - CBD}$: $y = Y(1-a), r_O = r_{SC}, r_D = r_{CBD}, D = L$
Number of transfers: 0
4. PTS-1
Fixed line: $Y_1 = 0, Y_2 = Ya\beta, Y_3 = Y(1-a), A = (g+1)L$
Feeder line within $\boldsymbol{P}$: $y_{in} = Ya\beta, y_{out} = 0, r = r_P,$
Feeder line within $\boldsymbol{SC}$: $y_{in} = Y(1-a), y_{out} = Ya\beta, r = r_{sc}$
Feeder line within $\boldsymbol{CBD}$: $y_{in} = 0, y_{out} = Y(1-a) + Ya(1-\beta), r = r_{CBD}$
Semi-direct line $P_i - CBD_C$: $y = Ya(1-\beta), r_O = r_P, r_D = r_{CBD}, D = (g+1)L$
Number of transfers: $1 \cdot Ya(1-\beta) + 2 \cdot Y(1 - a + a\beta)$
5. PTS-2
Fixed line: $Y_1 = 0, Y_2 = Ya\beta, Y_3 = 0, A = gL$
Feeder line within $\boldsymbol{P}$: $y_{in} = Ya\beta, y_{out} = 0, r = r_P,$
Feeder line within $\boldsymbol{SC}$: $y_{in} = 0, y_{out} = Ya\beta, r = r_{sc}$
Feeder line within $\boldsymbol{CBD}$: $y_{in} = 0, y_{out} = Ya(1-\beta) + Y(1-a), r = r_{CBD}$
Semi-direct line $P_i - CBD_C$: $y = Ya(1-\beta), r_O = r_P, r_D = r_{CBD}, D = (g+1)L$
Semi-direct line $SC_i - CBD_C$: $y = Y(1-a), r_O = r_{SC}, r_D = r_{CBD}, D = L$
Number of transfers: $1 \cdot Y[a(1-\beta) + 1 - a] + 2 \cdot Ya\beta$
6. PTS-3
Feeder line within $\boldsymbol{SC}$: $y_{in} = 0, y_{out} = Ya\beta, r = r_{sc}$
Feeder line within $\boldsymbol{CBD}$: $y_{in} = 0, y_{out} = Y(1-a) + Ya(1-\beta), r = r_{CBD}$
Semi-direct line $P_i - SC_C$: $y = Ya\beta, r_O = r_P, r_D = r_{SC}, D = gL$
Semi-direct line $P_i - CBD_C$: $y = Ya(1-\beta), r_O = r_P, r_D = r_{CBD}, D = (g+1)L$
Semi-direct line $SC_i - CBD_C$: $y = Y(1-a), r_O = r_{SC}, r_D = r_{CBD}, D = L$
Number of transfers: $1 \cdot Y$

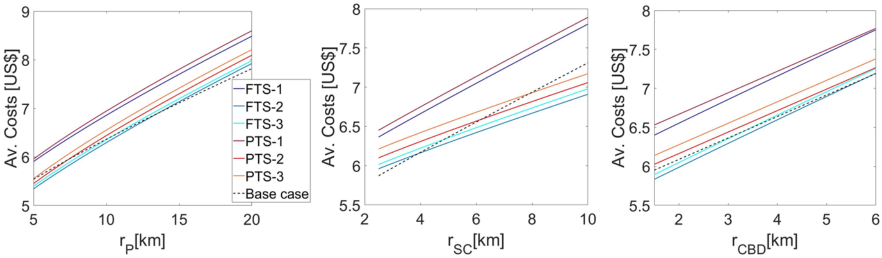


Fig. 10 Sensitivity analysis with respect to the radia of the periphery, subcenter and CBD

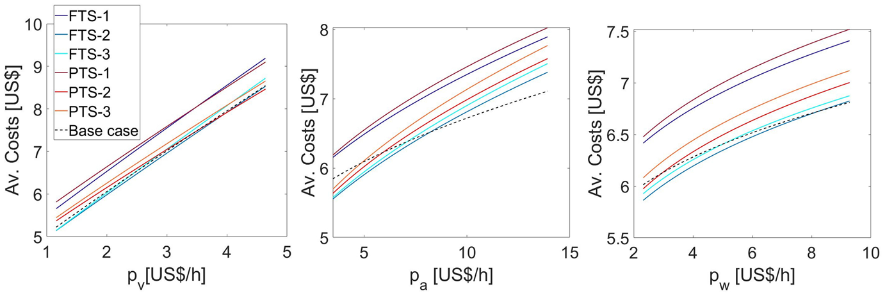


Fig. 11 Sensitivity analysis with respect to cost-related user parameters

7. Base case
Fixed line: $Y_1 = Ya(1-\beta), Y_2 = Ya\beta, Y_3 = Y(1-a), A = (g+1)L$
Flexible line within $\boldsymbol{P}$: $y_{in} = Ya, y_{out} = 0, r = r_P,$
Flexible line within $\boldsymbol{SC}$: $y_{in} = Y(1-a), y_{out} = Ya\beta, r = r_{sc}$
Flexible line within $\boldsymbol{CBD}$: $y_{in} = 0, y_{out} = Ya(1-\beta) + Y(1-a), r = r_{CBD}$
Number of transfers: $2 \cdot Y$

Appendix B: Sensitivity analysis

Let us begin the sensitivity analysis by studying the effect of the parameters g and L, which define the distance between P, SC, and CBD. The results are depicted in Fig. 9, and reveal that average costs increase for all structures, which is an obvious consequence of the need for traveling longer distances. Notably, the comparison among the structures changes by favoring the base case: When the passengers are traveling between distant points, the efficiency of massive vehicles outperforms ODRP.

In Fig. 10 we analyze the effect of the radia of the three circles. As expected, average costs increase due to the longer distances involved. However, the effect on the comparison among structures is different depending on the zone, especially regarding the base case, which becomes more competitive when r_P or r_{CBD} increase,

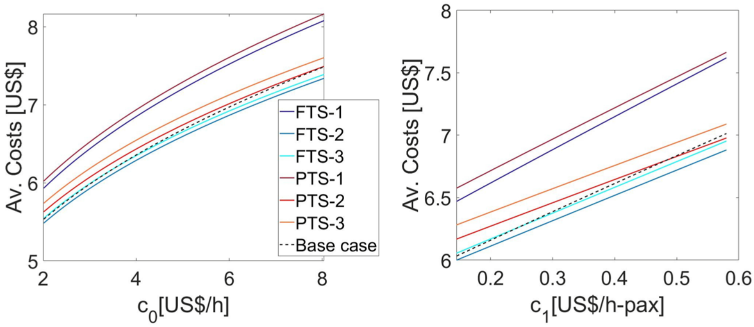


Fig. 12 Sensitivity analysis with respect to operators' cost-related parameters

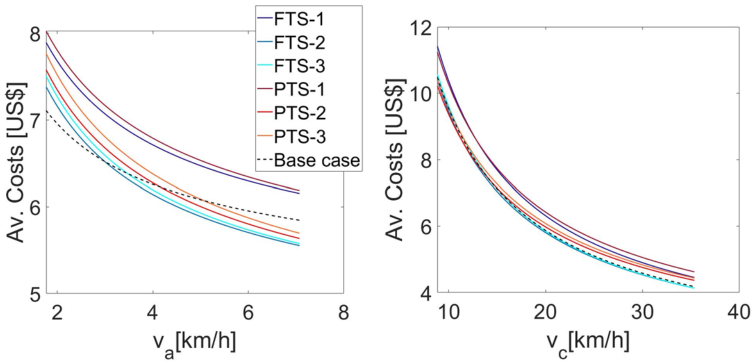
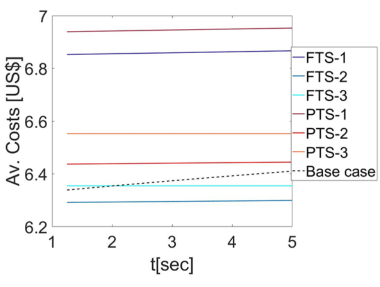


Fig. 13 Sensitivity analysis with respect to walking and vehicle velocities

Fig. 14 Sensitivity analysis with respect to the time needed to board and alight a bus



but the contrary happens with r_{SC}. In the case of *P* and *CBD*, a smaller circle implies that it is easier to serve the users in an (almost) door-to-door fashion. To understand what happens in *SC*, note that FTS-1 and PTS-1 are similarly affected as the base case, and these are the three structures in which the *SC*-feeder line takes passengers in both directions: the unbalance between the two directions becomes more relevant when the radius of the circle increases.

Figure 11 shows the impact of the three users' cost-related parameters. As they define the costs, it is expected that they make total costs increase. Their impact on the comparison between structures follows directly what was analyzed via Fig. 5: as FTS-3 involves the largest in-vehicle times, it is the most affected by p_v; and as the base case presents the lowest waiting and walking times, it gets favored when p_a or p_w increase.

Figure 12 shows the impact of operators' cost-related parameters. As expected, they make total costs increase. The fixed cost per vehicle c_0 does not have a very significant impact on the comparison between structures, whereas a larger c_1 (the fixed cost per seat) makes the base case less competitive, namely because the potential savings of large vehicles become less relevant.

In Fig. 13 we show the impact of the two involved speeds: walking and vehicle. Walking speed presents the exact reverse effect as p_a, which is expected as every time p_a appears in an equation it is divided by v_a. Increasing vehicle speed mostly favors FTS-3 and FTS-2: this happens because the problem that we identified above, namely that these structures do not split the long and the short legs of the trip, becomes less important as now they are toured faster, allowing for smaller vehicles and making detours less relevant.

Finally, in Fig. 14 we depict the effect of t, the time spent by each user boarding or alighting a fixed line. Its impact is generally mild, with the only exception of the base case, because in this structure all of the passengers utilize the fixed line.

Appendix C: Glossary

See Table 4.

Table 4 Glossary of the terms involved in the equations throughout the paper

Symbol	Meaning	Units
City and demand		
Y	Total number of users	pax/hour
a	Percentage of the users departing from *P*	–
β	Out of the users that depart from *P*, the percentage that goes to *SC*	–
L	Distance between *SC* and *CBD*	km
g	Ratio between the distances *P-SC* and *SC-CBD*	–
r_P, r_{SC}, r_{CBD}	Radius of *P, SC,* and *CBD*, respectively	km
Optimization—general		
c_0	Fixed cost per vehicle	US$/hour
c_1	Fixed cost per seat	US$/pax-hour
K	Capacity of a vehicle	pax
t_w, t_v	Average waiting and in-vehicle time of the line	hour
p_w, p_v	Monetary cost of one time unit waiting and in-vehicle	US$/hour
v_c	Average speed of the vehicles when moving	km/h
B	Fleet size	–
Optimization—fixed line		
f	Frequency	1/hour
A	Distance between the first and last node of the line	km
Y_1, Y_2, Y_3	Number of passengers in the line going $P_C \rightarrow CBD_C, P_C \rightarrow SC_C, SC_C \rightarrow CBD_C$, respectively	pax/hour
t_c	Cycle time	hour
T	Time required to make the whole circuit if there were no stops	hour
t_m	Average time spent by users in motion	hour
t_s	Average time spent by users waiting for others to board and alight	hour
Optimization—flexible lines		
t_a	Average walking time	hour

Table 4 (continued)

Symbol	Meaning	Units
l	Maximum walking distance	km
r_O, r_D, r	If the line travels between different zones, r_O and r_D represent the radia of the origin and the destination zone, respectively. Otherwise, r represents the radius of the unique zone	km
y	Number of users of the line	pax/hour
y_{in}, y_{out}	For intra-zonal lines, the number of users traveling towards/from the center of the zone, respectively	pax/hour
D	Distance between the centers of the origin and destination zone	km
γ	Factor that multiplies $\sqrt{K}$ when calculating the average detour at the destination zone	–
ρ	Average load of the vehicle	–
ρ_{in}, ρ_{out}	For intra-zonal lines, average load of the vehicle when going towards/from the center, respectively	–
K_{in}, K_{out}	For intra-zonal lines, maximum load of the vehicle when going towards/from the center, respectively	pax

Acknowledgements Professor Jara-Diaz acknowledges partial funding by Fondecyt, Chile, Grant 1200157, and Conicyt, Chile, Grant AFB230002.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Declarations

Conflict of interest The authors report no conflict of interests.

References

Alonso-González MJ, Cats O, van Oort N, Hoogendoorn-Lanser S, Hoogendoorn S (2021) What are the determinants of the willingness to share rides in pooled on-demand services? Transportation 48(4):1733–1765. https://doi.org/10.1007/s11116-020-10110-2

Applegate DL, Bixby RE, Chvátal V (2007) The traveling salesman problem: a computational study. Princeton University Press, Princeton

Auad-Perez R, Van Hentenryck P (2022) Ridesharing and fleet sizing for on-demand multimodal transit systems. Transp Res Part C Emerg Technol 138:103594. https://doi.org/10.1016/j.trc.2022.103594

Badia H, Jenelius E (2020) Feeder transit services in different development stages of automated buses: comparing fixed routes versus door-to-door trips. Transp Res Proc 47:521–528

Badia H, Estrada M, Robusté F (2014) Competitive transit network design in cities with radial street patterns. Transp Res Part B Methodol 59:161–181. https://doi.org/10.1016/j.trb.2013.11.006

Bartłomiejczyk M, Kołacz R (2020) The reduction of auxiliaries power demand: the challenge for electromobility in public transportation. J Clean Prod 252:119776. https://doi.org/10.1016/j.jclepro.2019.119776

Bian Z, Liu X, Bai Y (2020) Mechanism design for on-demand first-mile ridesharing. Transp Res Part B Methodol 138:77–117. https://doi.org/10.1016/j.trb.2020.03.011

Bilali A, Dandl F, Fastenrath U, Bogenberger K (2019) Impact of service quality factors on ride sharing in urban areas. In: 2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS). https://doi.org/10.1109/MTITS.2019.8883364

Borndörfer R, Grötschel M, Pfetsch ME (2007) A column-generation approach to line planning in public transport. Transp Sci 41(1):123–132. https://doi.org/10.1287/trsc.1060.0161

Bürstlein J, López D, Farooq B (2021) Exploring first-mile on-demand transit solutions for North American suburbia: a case study of Markham, Canada. Transp Res Part A Pol Pract 153:261–283. https://doi.org/10.1016/j.tra.2021.08.018

Byrne BF (1975) Public transportation line positions and headways for minimum user and system cost in a radial case. Transp Res 9(2–3):97–102. https://doi.org/10.1016/0041-1647(75)90044-1

Calabrò G, Le Pira M, Giuffrida N, Inturri G, Ignaccolo M, Correia G (2022) Fixed-route vs. demand-responsive transport feeder services: an exploratory study using an agent-based model. J Adv Transp 2022:8382754. https://doi.org/10.1155/2022/8382754

Calabro G, Araldo A, Oh S, Seshadri R, Inturri G, Ben-Akiva M (2023) Adaptive transit design: optimizing fixed and demand responsive multi-modal transportation via continuous approximation. Transp Res Part A Pol Pract 171:103643

Cats O, Kucharski R, Danda SR, Yap M (2022) Beyond the dichotomy: how ride-hailing competes with and complements public transport. Plos One 17(1):e0262496. https://doi.org/10.1371/journal.pone.0262496
Chaturvedi M, Srivastava S (2022) A multi-modal ride sharing framework for last mile connectivity. In: 14th International Conference on COMmunication Systems NETworkS (COMSNETS), pp 824–829. https://doi.org/10.1109/COMSNETS53615.2022.9668583
Chen H, Gu W, Cassidy MJ, Daganzo CF (2015) Optimal transit service atop ring-radial and grid street networks: a continuum approximation design method and comparisons. Transp Res Part B Methodol 81:755–774. https://doi.org/10.1016/j.trb.2015.06.012
Chen S, Wang H, Meng Q (2020) Solving the first-mile ridesharing problem using autonomous vehicles. Comp-Aided Civ Infrastruct Eng 35(1):45–60
Chriqui C, Robillard P (1975) Common bus lines. Transp Sci 9(2):115–121. https://doi.org/10.1287/trsc.9.2.115
Cortés CE, Jayakrishnan R (2002) Design and operational concepts of high-coverage point-to-point transit system. Transp Res Rec 1783(1):178–187. https://doi.org/10.3141/1783-22
Daganzo CF (2010) Structure of competitive transit networks. Transp Res Part B Methodol 44(4):434–446. https://doi.org/10.1016/j.trb.2009.11.001
Diao M, Kong H, Zhao J (2021) Impacts of transportation network companies on urban mobility. Nat Sustainab 4:494–500
Durán-Micco J, Vansteenwegen P (2022) A survey on the transit network design and frequency setting problem. Public Transp 14(1):155–190. https://doi.org/10.1007/s12469-021-00284-y
Fagnant DJ, Kockelman KM (2018) Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in Austin, Texas. Transportation 45(1):143–158. https://doi.org/10.1007/s11116-016-9729-z
Fielbaum A (2020) Strategic public transport design using autonomous vehicles and other new technologies. Int J Intell Transp Syst Res 18(2):183–191
Fielbaum A (2022) Optimizing a vehicle's route in an on-demand ridesharing system in which users might walk. J Intell Transp Syst 26(4):432–447
Fielbaum A, Alonso-Mora J (2020) Unreliability in ridesharing systems: Measuring changes in users' times due to new requests. Transp Res Part C Emerg Technolog 121:102831
Fielbaum A, Jara-Diaz S, Gschwender A (2016) Optimal public transport networks for a general urban structure. Transp Res Part B Methodol 94:298–313. https://doi.org/10.1016/j.trb.2016.10.003
Fielbaum A, Jara-Diaz S, Gschwender A (2017) A parametric description of cities for the normative analysis of transport systems. Netw Spat Econ 17(2):343–365. https://doi.org/10.1007/s11067-016-9329-7
Fielbaum A, Jara-Díaz S, Gschwender A (2018) Transit line structures in a general parametric city: the role of heuristics. Transp Sci 52(5):1092–1105. https://doi.org/10.1287/trsc.2018.0833
Fielbaum A, Jara-Diaz S, Gschwender A (2020) Beyond the Mohring effect: scale economies induced by transit line structures design. Econ Transp 22:100163. https://doi.org/10.1016/j.ecotra.2020.100163
Fielbaum A, Bai X, Alonso-Mora J (2021a) On-demand ridesharing with optimized pick-up and drop-off walking locations. Transp Res Part C Emerg Technol 126:103061. https://doi.org/10.1016/j.trc.2021.103061
Fielbaum A, Jara-Diaz S, Gschwender A (2021b) Lines spacing and scale economies in the strategic design of transit systems in a parametric city. Res Transp Econ 90:100991. https://doi.org/10.1016/j.retrec.2020.100991
Fielbaum A, Kucharski R, Cats O, Alonso-Mora J (2022) How to split the costs and charge the travellers sharing a ride? Aligning system's optimum with users' equilibrium. Eur J Oper Res 301:956–973. https://doi.org/10.1016/j.ejor.2021.11.041
Fielbaum A, Tirachini A, Alonso-Mora J (2023) Economies and diseconomies of scale in on-demand ridepooling systems. Econ Transp 34:100313. https://doi.org/10.1016/j.ecotra.2023.100313
Furuhata M, Daniel K, Koenig S, Ordóñez F, Dessouky M, Brunet M, Cohen L, Wang X (2015) Online cost-sharing mechanism design for demand-responsive transport systems. IEEE Trans Intell Transp Syst 16(2):692–707. https://doi.org/10.1109/TITS.2014.2336212
Garcia-Martinez A, Cascajo R, Jara-Diaz SR, Chowdhury S, Monzon A (2018) Transfer penalties in multimodal public transport networks. Transp Res Part A Pol Pract 114:52–66. https://doi.org/10.1016/j.tra.2018.01.016

Gschwender A, Jara-Díaz S, Bravo C (2016) Feeder-trunk or direct lines? Economies of density, transfer costs and transit structure in an urban context. Transp Res Part A Pol Pract 88:209–222. https://doi.org/10.1016/j.tra.2016.03.001
Hall JD, Palsson C, Price J (2018) Is Uber a substitute or complement for public transit? J Urban Econ 108:36–50
Henao A, Marshall WE (2019) The impact of ride-hailing on vehicle miles traveled. Transportation 46(6):2173–2194
Hörcher D, Graham DJ (2018) Demand imbalances and multi-period public transport supply. Transp Res Part B Methodol 108:106–126. https://doi.org/10.1016/j.trb.2017.12.009
Hörcher D, Tirachini A (2021) A review of public transport economics. Econ Transp 25:100196. https://doi.org/10.1016/j.ecotra.2021.100196
Huang Y, Kockelman KM, Garikapati V (2022) Shared automated vehicle fleet operations for first-mile last-mile transit connections with dynamic pooling. Comput Environ Urban Syst 92:101730. https://doi.org/10.1016/j.compenvurbsys.2021.101730
Irawan MZ, Belgiawan PF, Joewono TB, Simanjuntak NI (2020) Do motorcycle-based ride-hailing apps threaten bus ridership? A hybrid choice modeling approach with latent variables. Public Transp 12(1):207–231. https://doi.org/10.1007/s12469-019-00217-w
Jara-Díaz SR, Gschwender A (2003) From the single line model to the spatial structure of transit services: corridors or direct? J Transp Econ Pol (JTEP) 37(2):261–277
Jara-Díaz SR, Gschwender A (2009) The effect of financial constraints on the optimal design of public transport services. Transportation 36(1):65–75
Jara-Díaz SR, Muñoz-Paulsen E (2021) Lessons from the strategic design of a bimodal public transport system on a linear city. Res Transp Econ. https://doi.org/10.1016/j.retrec.2021.101167
Jara-Díaz SR, Gschwender A, Ortega M (2012) Is public transport based on transfers optimal? A theoretical investigation. Transp Res Part B Methodol 46(7):808–816. https://doi.org/10.1016/j.trb.2012.02.002
Jara-Díaz S, Gschwender A, Bravo C (2018) Total cost minimizing transit route structures considering trips towards CBD and periphery. Transportation 45(6):1701–1720. https://doi.org/10.1007/s11116-017-9777-z
Jara-Diaz S, Monzon A, Cascajo R, Garcia-Martinez A (2022) An international time equivalency of the pure transfer penalty in urban transit trips: closing the gap. Transp Policy 125:48–55. https://doi.org/10.1016/j.tranpol.2022.04.019
Jara-Diaz S, Olea JP (2021) Estimación de relaciones entre los parámetros de distribución de viajes en el modelo de ciudad paramétrica. 20°Congreso Chileno de Ingeniería de Transporte
Ke J, Yang H, Li X, Wang H, Ye J (2020a) Pricing and equilibrium in on-demand ride-pooling markets. Transp Res Part B Methodol 139:411–431. https://doi.org/10.1016/j.trb.2020.07.001
Ke J, Yang H, Zheng Z (2020b) On ride-pooling and traffic congestion. Transp Res Part B Methodol 142:213–231. https://doi.org/10.1016/j.trb.2020.10.003
Kim ME, Schonfeld P (2014) Integration of conventional and flexible bus services with timed transfers. Transp Res Part B Methodol 68:76–97
Kucharski R, Fielbaum A, Alonso-Mora J, Cats O (2021) If you are late, everyone is late: late passenger arrival and ride-pooling systems' performance. Transportmetrica A Transp Sci 17(4):1077–1100
Lau ST, Susilawati S (2021) Shared autonomous vehicles implementation for the first and last-mile services. Transp Res Interdiscip Perspect 11:100440. https://doi.org/10.1016/j.trip.2021.100440
Li X, Quadrifoglio L (2010) Feeder transit services: choosing between fixed and demand responsive policy. Transp Res Part C Emerg Technol 18(5):770–780
Liu Y, Bansal P, Daziano R, Samaranayake S (2019) A framework to integrate mode choice in the design of mobility-on-demand systems. Transp Res Part C Emerg Technol 105:648–665. https://doi.org/10.1016/j.trc.2018.09.022
Loyola M, Nelson JD, Clifton G, Ho CQ (2023) Factors influencing the patronage of flexible transport in urban and rural areas. A case study in NSW, Australia. Res Transp Econ 99:101295
Lu W, Quadrifoglio L (2019) Fair cost allocation for ridesharing services—modeling, mathematical programming and an algorithm to find the nucleolus. Transp Res Part B Methodol 121:41–55. https://doi.org/10.1016/j.trb.2019.01.001
Ma T-Y, Rasulkhani S, Chow JYJ, Klein S (2019) A dynamic ridesharing dispatch and idle vehicle repositioning strategy with integrated transit transfers. Transp Res Part E Logist Transp Rev 128:417–442. https://doi.org/10.1016/j.tre.2019.07.002

Mahéo A, Kilby P, Van Hentenryck P (2019) Benders decomposition for the design of a hub and shuttle public transit system. Transp Sci 53(1):77–88. https://doi.org/10.1287/trsc.2017.0756
Malalgoda N, Lim SH (2019) Do transportation network companies reduce public transit use in the US? Transp Res Part A Pol Pract 130:351–372
Martin S, Taylor SJ, Yan J (2021) Trading flexibility for adoption: Dynamic versus static walking in ridesharing (SSRN Scholarly Paper ID 3984476). Soc Sci Res Netw. https://doi.org/10.2139/ssrn.3984476
Masing B, Lindner N, Borndörfer R (2022) The price of symmetric line plans in the Parametric City. Transp Res Part B Methodol 166:419–443
Mo B, Cao Z, Zhang H, Shen Y, Zhao J (2021) Competition between shared autonomous vehicles and public transit: a case study in Singapore. Transp Res Part C Emerg Technol 127:103058. https://doi.org/10.1016/j.trc.2021.103058
Mueller A, Beyer S, Kopp G, Deisser O (2019) User-centered development of a public transportation vehicle operated in a demand-responsive environment. In: International Conference on Applied Human Factors and Ergonomics. Springer, Cham, pp 545–555
Narayanan S, Chaniotakis E, Antoniou C (2020) Shared autonomous vehicle services: a comprehensive review. Transp Res Part C Emerg Technol 111:255–293. https://doi.org/10.1016/j.trc.2019.12.008
Papanikolaou A, Basbas S (2020) Analytical models for comparing demand responsive transport with bus services in low demand interurban areas. Transp Lett 13(4):255–262
Perera S, Ho C, Hensher D (2020) Resurgence of demand responsive transit services—insights from BRIDJ trials in inner west of Sydney, Australia. Res Transp Econ 83:100904
Périvier N, Hssaine C, Samaranayake S, Banerjee S (2021) Real-time approximate routing for smart transit systems. Proc ACM Measure Anal Comp Syst 5(2):1–30. https://doi.org/10.1145/3460091
Petruccelli U, Racina A (2021) Feeder-trunk and direct-link schemes for public transit: a model to evaluate the produced accessibility. Public Transp 13(2):301–323. https://doi.org/10.1007/s12469-021-00262-4
Pinto HKRF, Hyland MF, Mahmassani HS, Verbas IÖ (2020) Joint design of multimodal transit networks and shared autonomous mobility fleets. Transp Res Part C Emerg Technol 113:2–20. https://doi.org/10.1016/j.trc.2019.06.010
Qian X, Zhang W, Ukkusuri SV, Yang C (2017) Optimal assignment and incentive design in the taxi group ride problem. Transp Res Part B Methodol 103:208–226. https://doi.org/10.1016/j.trb.2017.03.001
Quadrifoglio L, Li X (2009) A methodology to derive the critical demand density for designing and operating feeder transit services. Transp Res Part B Methodol 43(10):922–935
Roy S, Cooper D, Mucci A, Sana B, Chen M, Castiglione J, Erhardt GD (2020) Why is traffic congestion getting worse? A decomposition of the contributors to growing congestion in San Francisco-Determining the Role of TNCs. Case Stud Transp Pol 8(4):1371–1382. https://doi.org/10.1016/j.cstp.2020.09.008
Salazar M, Rossi F, Schiffer M, Onder CH, Pavone M (2018) On the interaction between autonomous mobility-on-demand and public transportation systems. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE, pp 2262–2269
Shen Y, Zhang H, Zhao J (2018) Integrating shared autonomous vehicle in public transportation system: a supply-side simulation of the first-mile service in Singapore. Transp Res Part A Pol Pract 113:125–136
Tikoudis I, Martinez L, Farrow K, García Bouyssou C, Petrik O, Oueslati W (2021) Ridesharing services and urban transport CO_2 emissions: Simulation-based evidence from 247 cities. Transp Res Part D Transp Environ 97:102923. https://doi.org/10.1016/j.trd.2021.102923
Tirachini A, Gomez-Lobo A (2020) Does ride-hailing increase or decrease vehicle kilometers traveled (VKT)? A simulation approach for Santiago de Chile. Int J Sustain Transp 14(3):187–204
Tirachini A, Hensher DA, Jara-Díaz SR (2010) Comparing operator and users costs of light rail, heavy rail and bus rapid transit over a radial public transport network. Res Transp Econ 29(1):231–242. https://doi.org/10.1016/j.retrec.2010.07.029
Tirachini A, Chaniotakis E, Abouelela M, Antoniou C (2020) The sustainability of shared mobility: can a platform for shared rides reduce motorized traffic in cities? Transp Res Part C Emerg Technol 117:102707. https://doi.org/10.1016/j.trc.2020.102707
Vansteenwegen P, Melis L, Aktaş D, Montenegro BDG, Sartori Vieira F, Sörensen K (2022) A survey on demand-responsive public bus systems. Transp Res Part C Emerg Technol 137:103573. https://doi.org/10.1016/j.trc.2022.103573

Wang Z, Hyland MF, Bahk Y, Sarma NJ (2022) On optimizing shared-ride mobility services with walking legs. https://arxiv.org/abs/2201.12639v1

Wang H, Yang H (2019) Ridesourcing systems: a framework and review. Transp Res Part B Methodol 129:122–155. https://doi.org/10.1016/j.trb.2019.07.009

Wen J, Chen YX, Nassir N, Zhao J (2018) Transit-oriented autonomous vehicle operation with integrated demand-supply interaction. Transp Res Part C Emerg Technol 97:216–234

Würtz MO, Sandkuhl K (2021) Towards a reference architecture for demand-oriented public transportation services. In: 2021 IEEE 25th International Enterprise Distributed Object Computing Workshop (EDOCW), IEEE, pp 122–129

Young M, Allen J, Farber S (2020) Measuring when Uber behaves as a substitute or supplement to transit: an examination of travel-time differences in Toronto. J Transp Geogr 82:102629

Zardini G, Lanzetti N, Pavone M, Frazzoli E (2022) Analysis and control of autonomous mobility-on-demand systems. Ann Rev Control Robot Auton Syst. https://doi.org/10.1146/annurev-control-042920-012811

Zhu P, Mo H (2022) The potential of ride-pooling in VKT reduction and its environmental implications. Transp Res Part D Transp Environ 103:103155. https://doi.org/10.1016/j.trd.2021.103155

Zuniga-Garcia N, Gurumurthy KM, Yahia CN, Kockelman KM, Machemehl RB (2022) Integrating shared mobility services with public transit in areas of low demand. J Public Transp 24:100032