Organisation and performance of public transport: A systematic cross-case comparison of metropolitan areas in Europe, Australia, and Canada

Hirschhorn, Fabio; Veeneman, Wijnand; van de Velde, Didier

DOI
10.1016/j.tra.2019.04.008

Publication date
2019

Document Version
Final published version

Published in
Transportation Research Part A: Policy and Practice

Citation (APA)
Hirschhorn, F., Veeneman, W., & van de Velde, D. (2019). Organisation and performance of public transport: A systematic cross-case comparison of metropolitan areas in Europe, Australia, and Canada. *Transportation Research Part A: Policy and Practice*, *124*, 419-432. https://doi.org/10.1016/j.tra.2019.04.008

Transportation Research Part A 124 (2019) 419–432

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Organisation and performance of public transport: A systematic cross-case comparison of metropolitan areas in Europe, Australia, and Canada

Fabio Hirschhorn*, Wijnand Veeneman, Didier van de Velde

Faculty of Technology, Policy and Management, Delft University of Technology, The Netherlands

ARTICLE INFO

Keywords:
Metropolitan public transport
Public transport governance
Public transport performance
Fuzzy-set qualitative comparative analysis (fsQCA)

ABSTRACT

The paper investigates how the interplay between six organisational elements of public transport systems (conditions) – i.e. integration of planning responsibilities within an authority at the regional/metropolitan level; land-use and transport integration; long-term metropolitan public transport planning; agency over funding; fare integration, and allocation of risks between government and operators - influence two key performance indicators (outcomes) – modal split and cost-recovery. The study focuses on selected metropolitan areas in Europe, Australia, and Canada, and employs Qualitative Comparative Analysis (QCA). QCA can handle multiple explanatory conditions in combination, framing the relationship between conditions and studied outcomes in terms of necessity and sufficiency. The paper reveals three alternative combinations of organisational elements that are sufficient for achieving each outcome, underscoring that modal split and cost-recovery depend on the combined effects of multiple conditions (conjunctural causality), and that different paths can lead to similar results (equifinality). Furthermore, even though both outcomes are linked to higher usage of public transport, findings suggest that each of them might require decision-makers to give attention to different elements. Higher modal split is closely linked to both integration between land-use and transport, and the integration of planning responsibilities within an authority at the regional/metropolitan level. Higher cost-recovery, in turn, requires focus on the way agency over funding and risk allocation strategies shape incentives for savings and/or revenue generation.

1. Introduction

This paper contributes to the literature on the governance of metropolitan public transport (PT)[1] and, in particular, the discussions concerning the connection between organisational form and performance (e.g. Hensher and Stanley, 2008; Preston and Almutairi, 2013; Sørensen and Gudmundsson, 2010). The underlying assumption in this literature is that elements of the organisation of PT, such as the ownership structure of operators (public or private), contractual allocation of risks, fare integration etc. might influence the achievement of performance goals (like sustainability or accessibility). These studies examine the links between the introduction or reform of PT organisational elements and variations in performance indicators (e.g. changes in levels of emissions/

* Corresponding author at: Jaffalaan 5, 2628BX, Delft, The Netherlands.
E-mail address: f.hirschhorn@tudelft.nl (F. Hirschhorn).

[1] Metropolitan public transport refers to all collective modes of land passenger transport services available to the general public within a metropolitan area, and linking it to its direct environment. There is no distinction based on ownership or control: these services can be either publicly or privately operated.

https://doi.org/10.1016/j.tra.2019.04.008
Received 27 May 2018; Received in revised form 2 April 2019; Accepted 16 April 2019
Available online 29 April 2019
0965-8564/

passengers or in number of passengers).

Existing work in this field often takes an incremental view, and tries to isolate the performance outcomes potentially brought by the introduction or reform of one single element of PT organisation. Whilst these analyses provide relevant insights, this paper takes a different approach. Recognising that PT is a complex system made up of interacting and interdependent elements and actors (Macmillen, 2013), and that PT's governance is wicked (Marsden and Reardon, 2017), this study focuses on the combined effects that multiple elements of PT organisation have on performance outcomes.

By building on previous research (Hirschhorn et al., 2018), the paper examines six elements of the organisation of PT - i.e. integration of planning responsibilities within an authority at the regional/metropolitan level; land-use and transport integration; land-use and transport integration; long-term metropolitan PT planning framework; agency (decision power) over funding; fare integration, and allocation of risks between government and operators - as potential explanatory factors (or conditions) of two key performance indicators (outcomes), modal split and cost-recovery. The paper aims to answer *what (combinations of) elements of the organisation of PT are sufficient conditions for higher levels of modal split and cost-recovery*. To this end, two models are developed. The first, for the outcome modal split, compares twenty-two cases: eleven metropolitan areas are examined, each, in two distinct moments in time, 2005 and 2015. The second model, for the outcome cost-recovery, compares fourteen of the first twenty-two cases: seven metropolitan areas are examined, each, in 2005 and 2015.

Methodologically, the paper responds to calls for more qualitative and mixed-method research in PT (Marsden and Reardon, 2017; Schwanen et al., 2011) and employs Qualitative Comparative Analysis (QCA) (Ragin, 2008, 1987). QCA is well-suited for the study of complex policy and social phenomena in case studies (Blackman et al., 2013; Byrne, 2013; Gerrits and Verweij, 2016). The method tackles complexity using set-theory and Boolean algebra techniques, and can identify minimally sufficient combinations of conditions for the outcome(s) of interest. Additionally, QCA can account for conjunctural causation and equifinality, i.e. QCA recognises that most often, it is the interplay of several conditions that leads to an outcome and, furthermore, different combinations of conditions may lead to the same outcome.

Section 2 contextualises this study within existing PT research. Section 3 introduces the methods used. The operationalisation of outcomes and conditions is explained in Section 4, whilst Section 5 presents results of the analysis. Section 6 critically assesses these results, followed by conclusions in Section 7.

2. Public transport performance and the influence of organisational elements

PT performance can be understood in two dimensions. A first broader dimension relates to public values and the goals society expects government to achieve in a policy area (Jørgensen and Bozeman, 2007; Koppenjan et al., 2008). In PT, they can be sustainability, accessibility, minimisation of public subsidies etc. These PT goals translate, broadly speaking, into two often opposing types of objectives: increased ridership (goals that depend on maximising usage) and coverage (goals that depend on amplifying the availability of services despite low usage) (Faivre d'Arcier, 2014; Walker, 2008). The second dimension of performance refers to the translation of those broad goals into quantitative metrics. These metrics are performance indicators, such as emissions/passenger or average distance to PT stops.

A well-established literature investigates how PT performance, measured by performance indicators, can be influenced by different elements of the organisation of PT systems, such as the ownership structure of operators (public or private), contractual allocation of risks, or integration of fares (e.g. Hensher and Stanley, 2008; Preston and Almutairi, 2013; Sørensen and Gudmundsson, 2010). Hirschhorn et al. (2018), for instance, developed an international Delphi survey consulting experts across academia and industry to identify and rate (i) key performance indicators and (ii) main organisational elements in PT able to drive a system's performance. Their survey's results show that experts emphasise ridership-related performance indicators, with PT modal split and cost-recovery amongst the top rated. The first indicator refers to the ratio of PT usage in relation to overall usage of (motorised) transport modes, and is indicative of PT's attractiveness to users. The former indicator refers to the ratio between revenues coming from passenger fares and overall operating costs, and is indicative of PT's financial sustainability. Concerning organisational elements, in turn, results from Hirschhorn et al. (2018) highlight the integration of planning responsibilities within an authority at the metropolitan level; policy integration between land-use and transport; the development of long-term PT plans; availability and stability of funding; ticket and fare integration, and risk allocation strategies. In light of these results, this paper examines how these six elements of PT organisation influence the performance indicators modal split and cost-recovery. A brief review of related literature follows.

Integration has become a key guiding principle for transport policies' institutional and structural development over recent decades (Potter and Skinner, 2000). Transport integration is a multifaceted concept though and can take many forms, e.g. between government levels, across policy areas, within tasks across PT levels of planning and control (strategic, tactical, and operational) (Hull, 2005; May et al., 2006; Preston, 2010; Stead and Geerlings, 2005). One type of integration refers to the consolidation, within one authority, of responsibilities for planning multiple PT modes at the regional/metropolitan level. Kumar and Agarwal (2013), Marsden and May (2006), and Pemberton (2000) advance that PT services can achieve better results when planning is done by an overarching organisation with authority over an area corresponding to major commuter patterns. Such an integrated planning authority, they claim, can avoid conflicting directions from overlapping planning agencies, make policy implementation more coherent, and avoid harmful competition between modes. Authors also call for integration between PT and other policy areas (such as environment or health) (Hull, 2008; Stead and Geerlings, 2005). In this context, the integration between land-use and transport policies, in particular, is seen as core in urban areas (Stanley, 2014), given the strong connection between the built environment and travel patterns (Ewing and Cervero, 2010; Van Wee, 2002; Zegras, 2010). Hickman et al. (2013) and Cervero (2013) point that good integration between urban planning and transport is critical to allow greater PT usage, walking and cycling, whilst Stanley (2014) points that it can also enable greater accessibility and social inclusion. Finally, at the tactical level (van de Velde, 1999), the integration between schedules, ticketing, and fares can help make PT systems simpler for passengers to understand and use, consequently leading to higher ridership levels (Chowdhury et al., 2018; Redman et al., 2013; Sharaby and Shiftan, 2012). These three types of PT integration are accounted for in this study through the conditions *Integrated Planning Authority*, *Land-use and Transport Integration*, and *Fare Integration* (see Section 4.1.2).

The preparation and adoption of long-term metropolitan PT plans and an enabling funding framework are two other factors that can favour better performance. Authors recognise long-term PT plans as critical tools to define clear goals and mission to authorities, setting a long-term vision for PT and how it is supposed to achieve the political goals determined by government. These plans may include issues such as the profit and market share aims, the general description of the services, the area of supply and target groups and also plans for important network elements such as definition of major corridors or interchange hubs (Nielsen et al., 2005; van de Velde, 1999). Long-term plans can promote the stability of transport strategies and of a high quality service, influencing positively transportation patterns and, thus, helping the implementation of an integrated and more attractive PT system (Gwilliam, 2003; May, 2004; Nielsen and Lange, 2007). Similarly, an enabling funding framework ensuring availability and stability of resources is core for PT performance (Bouf and Faivre d'Arcier, 2015; Hess and Lombardi, 2005). Litman (2014) and Faivre d'Arcier (2014) emphasise that earmarked sources of subsidy increase funding autonomy and predictability, as they constitute secure resources to improve the quality of PT supply. Veeneman et al. (2015) highlight how the realisation of public goals in PT will be hindered when funding and decision-making take place in different government levels. These two organisational elements (long-term planning and funding framework) are accounted for in this study through the conditions *Long-term metropolitan PT planning framework* and *Agency (decision power) over funding* (see Section 4.1.2).

Concerning risk allocation strategies between government and operators, studies focus on possible performance outcomes connected to the use of different awarding mechanisms, or of different contractual forms regulating service provision and remuneration (broadly speaking gross-cost and net-cost contracts), as well as the influence of the ownership of operators (public or private). These analyses show mixed results: Mees (2005) concludes that Melbourne's urban rail and tram systems' privatisation and franchising contracts failed to deliver expected increases in ridership and cost-savings. Roy and Yvrande-Billon (2007) find that private bus companies regulated under gross-cost contracts are more efficient than operators under net-cost contracts. Mizutani and Urakami (2003) conclude that public bus companies are less efficient than private companies, whereas no significant differences are found by Filippini and Prioni (2003). Filippini et al. (2015) compare bus lines operated under competitively tendered contracts and performance-based negotiated contracts and their overall results show no considerable cost-efficiency differences. Risk allocation is accounted for in this study through the condition *Risk Allocation to Operators* (see Section 4.1.2).

All literature described above offers important insights to support PT policy decisions, however this work often takes an incremental view and isolates the impacts that the introduction or reform of a single element of PT organisation has on performance. This approach might be the result of the predominance in public transport research of an emphasis on a few disciplines and quantitative methods, eventually overlooking some critical aspects shaping PT governance and decision-making (Marsden and Reardon, 2017; Schwanen et al., 2011). Like other social science and policy problems, understanding the relation between the governance of PT and performance requires a complexity-informed view, acknowledging that outcomes depend on the interaction between multiple system elements and context characteristics (Blackman et al., 2013; Verweij and Gerrits, 2012). This is the case with PT that is made up of interacting and interdependent elements such as infrastructures, technology, social norms, and regulations. PT, thus, cannot be understood as reducible to stable and deterministic relationships between variables (Macmillen, 2013).

Furthermore, PT governance is wicked and decision-making involves networks of multiple actors with diverse and sometimes conflicting perspectives in relation to the goals they expect public policies to achieve (de Bruijn and Dicke, 2006; Koppenjan et al., 2008; Thacher and Rein, 2004). Good PT performance depends on the perspective of the stakeholder considered (Faivre d'Arcier, 2014) and choices between ridership and coverage objectives, or even the alignment between different ridership objectives, entail trade-offs (Walker, 2008). Even though modal split and cost-recovery are both ridership-related indicators, and thus with a tendency to be positively affected by growth in PT usage, they may not always align. Vassallo et. al (2009), for instance, have found that despite a significant increase in PT ridership in Madrid between 1996 and 2004, modal split only increased in trips between the centre and the periphery, whereas it decreased in trips within peripheral areas. Additionally, cost-recovery levels dropped due to increasing subsidies.

Recognising the importance of the interplay between PT's systemic elements and its context, as well as the wicked nature of PT governance, this paper seeks new insights on key mechanisms connecting organisation and performance in PT. To this end, and responding to calls for more mixed-method research in public transport, the paper combines qualitative and quantitative data and employs QCA.

3. Methods

3.1. Qualitative Comparative Analysis

QCA (Ragin, 2008, 1987) is particularly suited for the analysis of complex social phenomena and wicked policy issues (Blackman et al., 2013; Byrne, 2013; Gerrits and Verweij, 2016). QCA can account for conjunctural causation and equifinality: the method recognises that most often, it is the interplay of several conditions that leads to an outcome and that different paths may lead to the same outcome. This is possible due to QCA's set-theoretic nature. Conditions and outcomes are defined as sets in which cases have a

Table 1
Calibration.

Sets	Definition	Calibration thresholds		Scoring
Outcomes	Membership in the set of cases with higher levels of modal split of PT (**SPLIT**)[1]	**1.00**: Above 40%. **0.75**: Between 30.1% and 40%. **0.45**: Between 20.1% and 30%. **0.25**: Between 10.1% and 20%. **0.00**: Up to 10%.		Thresholds defined using EMTA and UITP databases. Cases are grouped across five categories based on their modal split ratio.
	Membership in the set of cases with higher levels of cost-recovery (**CR**)[2]	**1.00**: Above 55%. **0.70**: Between 45.1% and 55%. **0.35**: Between 40.1% and 45%. **0.00**: Up to 40%.		Thresholds defined using EMTA database. Cases are grouped across four categories based on their cost-recovery ratio.
Conditions	Membership in the set of cases with stronger Integrated Planning Authority (**PA**)[3]	Multijurisdictional	**1.00**: An organisation is established with planning authority over PT in an area corresponding to major commuting patterns covering multiple local jurisdictions (or such area is a single jurisdiction). **0.70**: An organisation is established with planning authority over PT in portion of the area corresponding to major commuting patterns across multiple local jurisdictions. **0.00**: No such organisation exists and PT planning takes place only at the local level.	The condition aggregates two attributes and the final score is their minimum (logical AND). Qualitative case studies (Online Material) support case scoring.
		Multimodal	**1:00**: The integrated planning authority plans modes that carry at least 75% of PT demand. **0.70**: Between 50% and 74,9% of PT demand. **0.35**: Between 25% and 49,9% of PT demand. **0.00**: Less than 25% of PT demand.	
	Membership in the set of cases with higher degree of land-use and transport integration (**LUT**)[4]	Institutionalisation of Regional Planning	**1.00**: Regional Planning is institutionalised and regional plans prepared. **0.00**: No institutionalised regional planning or regional plans.	The condition aggregates two attributes and the final score is their minimum (logical AND). Case studies (Online Material) support scoring of the first attribute. EMTA and UITP databases support scoring of the second attribute.
		Mode share walking and biking	**1.00**: Above 42%. **0.70**: Between 27% and 42%. **0.35**: Between 18,1% and 26,9%. **0.00**: Below 18,1%.	
	Membership in the set of cases with metropolitan long-term PT planning framework (**LP**)[5]	**1.00**: A PT strategic planning framework determines the preparation of a long-term PT strategy (10 years horizon at least) at the metropolitan level and also its periodic revision, and/or links the long-term PT strategy to short-term implementation plans. A long-term PT plan is in place and it is not the first time the framework is adopted. **0.70**: Previous requirements apply, however the long-term PT plan is in place for the first time. **0.00**: No PT strategic planning framework. Only *ad hoc* plans.		Qualitative case studies (Online Material) support case scoring.
	Membership in the set of cases with higher degree of agency (power decision) over funding at the regional level (**AF**)[6]	Tax base for PT subsidies	**0.70**: At least 60% of PT subsidies come from local/regional tax base. **0.00**: Less than 60% of PT subsidies come from local/regional tax base.	The condition aggregates two attributes. The final score is their sum. Qualitative case studies (Online Material) support case scoring.
		Earmarked PT funds	**0.30**: At least 20% of PT subsidies are from earmarked sources. **0.00**: Less than 20% of PT subsidies are from earmarked sources.	

(*continued on next page*)

Table 1 (*continued*)

Sets	Definition	Calibration thresholds		Scoring
	Membership in the set of cases with higher degree of fare integration (**FI**)[7]	**1.00**: All kinds of tickets (single, daily, season etc.) offer access to all modes and operators. Price does not vary depending on operators used or transfers. This is valid for modes carrying at least 80% of PT demand. **0.70**: A single trip fare varies depending on operators used or whether transfers are required. Only some multi-journey or multi-operator fares allow access to all modes and operators. **0.35**: Fares vary depending on operators used or whether transfers are required. Multi-journey or multi-operator fares may exist, but cover only part of modes and/or operators. **0.00**: Traveling from point A to point B costs the equivalent of the sum of each trip leg.		Qualitative case studies (Online Material) support case scoring.
	Membership in the set of cases with higher degree of risk allocated to operators (**RO**)[8]	Production risk	**1**: Operators bear production risk, or no contract exists and operator is privately owned. **0**: Operator does not bear production risk, or no contract exists with in-house operator.	The condition aggregates three attributes. The final score is their average. Qualitative case studies (Online Material) support case scoring.
		Commercial risk	**1**: Operator bears commercial risk, or no contract exists with privately owned operator. **0**: Operator does not bear commercial risk, or no contract exists with in-house operator.	
		Long-term risk	**0**: In-house operator. **1**: Not in-house operator.	

Main sources used to support the operationalisation of the outcomes and conditions:
(1) EMTA and UITP; (2) EMTA; (3) (Kumar and Agarwal, 2013; Marsden and May, 2006; Pemberton, 2000); (4) (Cervero, 2013; Hickman et al., 2013; Mäntysalo et al., 2015; Searle, 2016; Suzuki et al., 2013); EMTA and UITP for share of biking and walking; (5) (Gwilliam, 2003; May, 2004; Nielsen et al., 2005; van de Velde, 1999); (6) (Faivre d'Arcier, 2014; Litman, 2014; Veeneman et al., 2015) (7) (Chowdhury et al., 2018; Redman et al., 2013; Sharaby and Shiftan, 2012); (8) (Filippini and Prioni, 2003; Mees, 2005; Roy and Yvrande-Billon, 2007)

Table 2
Raw outcome data and final fuzzy-scores of outcomes and conditions.

			Outcomes		Conditions					
Case	Modal Split (%)	Cost-rec. (%)	SPLIT	CR	PA	LUT	LP	AF	FI	RO
Amsterdam05	16.4	38.2	0.25	0.00	1.00	1.00	1.00	0.00	0.70	0.00
Amsterdam15	23.2	49.8	0.45	0.70	1.00	1.00	1.00	0.00	0.70	0.67
Berlin05	33.3	46.5	0.75	0.70	1.00	0.70	0.70	0.00	1.00	0.75
Berlin15	34.4	55.3	0.75	1.00	1.00	1.00	1.00	0.00	1.00	0.76
Birmingham05	12	–	0.25	–	0.00	0.35	0.00	0.00	0.35	1.00
Birmingham15	14.3	–	0.25	–	0.00	0.00	0.70	0.00	0.35	1.00
Helsinki05	37.8	56.5	0.75	1.00	0.35	0.70	1.00	0.70	1.00	0.33
Helsinki15	39.4	48.3	0.75	0.70	1.00	0.70	1.00	0.70	1.00	0.49
Madrid05	49.5	44.5	1.00	0.35	1.00	0.35	0.00	0.70	0.70	0.47
Madrid15	40.6	51.1	1.00	0.70	1.00	0.70	0.00	0.70	0.70	0.39
Melbourne05	7.3	–	0.00	–	1.00	0.00	0.00	0.70	1.00	0.67
Melbourne15	9.3	–	0.00	–	1.00	0.00	0.00	0.70	1.00	0.67
Montreal05	17.4	53	0.25	0.70	0.00	0.00	1.00	0.70	1.00	0.00
Montreal15	20.6	50.9	0.45	0.70	0.00	0.00	1.00	0.70	1.00	0.00
Oslo05	21.6	56.3	0.45	1.00	0.00	0.35	0.00	0.70	0.70	0.33
Oslo15	34.2	50.5	0.75	0.70	1.00	0.70	1.00	1.00	1.00	0.33
Stockholm05	40.0	36.2	0.75	0.00	1.00	0.70	0.00	0.70	1.00	0.67
Stockholm15	39.1	39.5	0.75	0.00	1.00	0.70	0.70	0.70	1.00	0.83
Turin05	26.7	–	0.45	–	1.00	0.35	0.00	0.00	0.70	0.67
Turin15	27.4	–	0.45	–	1.00	0.35	0.00	0.00	0.70	0.67
Vancouver05	11.8	–	0.25	–	1.00	0.00	0.00	1.00	1.00	0.00
Vancouver15	16.1	–	0.25	–	1.00	0.00	1.00	1.00	1.00	0.00

membership score, and cases, in turn, are operationalised as configurations constituted by these conditions. QCA then works by systematically comparing cases' properties (the outcome and conditions) to identify set-relations between them, and frames this relationship in terms of sufficiency and necessity (Rihoux and Ragin, 2009; Schneider and Wagemann, 2012). A (combination of) condition(s) is sufficient if it can lead to the outcome by itself, although other (combination of) condition(s) can also lead to the same outcome. A (combination of) condition(s) is necessary if it is always present when the outcome occurs, but it does not ensure the outcome will happen. Operationally, QCA can be broken-down into the following steps: calibration; analysis of necessity; construction of the truth table, and Boolean minimisation of the truth table for the analysis of sufficiency.

Calibration is the process by which the researcher, relying on substantive knowledge, converts raw case data into set membership scores defining if a case is either in or out of the set (membership score of 1 or 0). The fuzzy-set variant of QCA (fsQCA) can account for non-perfect superset/subset relation, and fuzzy membership scores can take any value between 0 and 1, where 0 indicates full non-membership, 1 indicates full membership, and 0.5 is the crossover point where there is a maximum ambiguity regarding whether a case is more in or out (Ragin, 2009). With fuzzy set calibration, a researcher can bridge the quantitative assessment of degree of membership between 0 and 1 and qualitative differentiation of cases between full membership, full non-membership and the points where cases are more in a given set than out (Ragin, 2008). As a result, fsQCA is able to model quasi-necessity or quasi-sufficiency probabilistically: a (combination of) condition(s) is sufficient for the outcome if its presence (nearly) always leads to an outcome, whereas a (combination of) condition(s) is necessary if it is (nearly) always present when the outcome is observed (Skaaning, 2011; Thomann and Maggetti, 2017).

After calibration, the researcher should check if any condition is necessary for the outcomes, i.e. if any condition is a superset of the outcome. QCA proceeds, then, with the analysis of sufficiency (the main focus of this paper) by building a truth table. In a truth table, each row displays a logically possible combination of the conditions under analysis, considering both their presence and absence (negation). With six conditions the truth table has 64 rows (i.e., 2ˆ6). Finally, the truth table is minimised using Boolean techniques, and rows are compared to eliminate logically redundant factors: "If two Boolean expressions differ in only one causal condition yet produce the same outcome, then the causal condition that distinguishes the two expressions can be considered irrelevant and can be removed to create a simpler, combined expression." (Ragin, 1987, p. 93). This process reveals all minimally sufficient combinations of conditions for the occurrence of the outcome.

3.2. Case selection

In QCA, it is important to include cases that vary both in relation to hypothesised conditions and outcomes analysed. This, however, cannot be done at the expense of a reasonable degree of comparability to ensure that effects of alternative explanatory conditions are alleviated (Toshkov, 2016). The paper's cases are PT systems in metropolitan areas (i) from Western economies, (ii) from high income countries (as defined by the World Bank's indicator of Gross National Income per capita); (ii) from member-countries of the Organisation for Economic Co-operation and Development (OECD), and (iii) with a population size between 1.5 and

5.5 million inhabitants and density below 1,500 persons per square-km in 2014 (OECD).[2]

Two additional steps complemented case selection. First, cases meeting the criteria above were assessed taking into account research design issues, particularly availability of data and the ease to find local experts for interviews. Potential interviewees were then contacted, and the final list of cases was formed by those metropolitan areas in which interviewees agreed to collaborate: Amsterdam, Berlin, Birmingham, Helsinki, Madrid, Melbourne, Montreal, Oslo, Stockholm, Turin, and Vancouver. The QCA model developed for modal split includes data of these eleven metropolitan areas in 2005 and in 2015, resulting in 22 cases. The QCA model developed for cost-recovery, in turn, looks at data from seven of these metropolitan areas, again in 2005 and 2015, totalling 14 cases.

4. Calibration

4.1. Fuzzy-set calibration

The paper adopts fsQCA as it allows for a more nuanced operationalisation of outcomes and conditions. The main task of the researcher during calibration is to define three qualitative anchors: full membership (1), full non-membership (0), and the crossover point (0.5). These choices should be grounded on substantive and case knowledge. Parameters of underlying raw data (such as prominent gaps) can help as supplementary input for calibration decisions. Other practical guidelines include avoiding the definition of overly skewed sets or scores exactly on the crossover point (Schneider and Wagemann, 2012; Thomann and Maggetti, 2017; Vis, 2009). The calibration of quantitative interval-scale data can follow different techniques. The direct and indirect methods (Ragin, 2008) rely on software support. Another option, described by Schneider and Wagemann (2012) and Verkuilen (2005), is to assign fuzzy-scores based on categorical differences between cases. In this paper, the outcomes are calibrated using the latter technique. The calibration of qualitative information, in turn, is less formalised in QCA literature, but guidelines exist (e.g. de Block and Vis, 2018; Legewie, 2017). For the calibration of the six conditions, the paper develops fuzzy-scores building on the literature presented in Section 2. The next sub-sections and Table 1 explain the calibration of outcomes and conditions. Table 2 presents the final calibrated data. Online Material contains supplementary calibration information and summaries of the cases studied.

4.1.1. Outcomes

Modal Split (SPLIT) and Cost-recovery (CR). The modal split of PT analysed in this paper refers to the number of trips made in PT modes in relation to the total number of motorised trips. Cost-recovery refers to the percentage of operating costs covered by fare revenues. As stated in Section 2, despite being measures of ridership-related objectives, they may not always be aligned. Consequently, it is appropriate to develop a separate model for each outcome. The raw data used for the calibration of outcomes is drawn from the Barometer report from the Association of European Metropolitan Transport Authorities (EMTA) (2017, 2016, 2009) (in the case of modal split, cases not covered by the Barometer have their data coming from the Mobility in Cities Database from the International Association of Public Transport (UITP) (2015, 2006) that employs a compatible calculation method).[3]

Calibrating the ratios of modal split and cost-recovery is complex because no normative theory is available to determine what should be deemed high or low levels of these measures. Assigning fuzzy-scores to them based on categorical differences between cases is a helpful way to distinguish relevant and irrelevant differences in ratios, since the numerical distances between ratios of modal split or cost-recovery do not have the same qualitative meaning. In the case of modal split, EMTA's and UITP's entire database were checked to find a general indication of modal split ratios worldwide. The overall average PT modal split of metropolitan areas in these databases is approximately 30%. When considering the ratios of the twenty-two studied cases, this level is located where a clear gap can be seen, and this is used as reference to determine the crossover point. Other breakpoints are then set using other gaps in the raw data, resulting in a five-value fuzzy-score. In the case of cost-recovery, a similar process is used: overall, average cost-recovery level presented in Barometers' 2006 and 2014 is 46%. This value is positioned in a visible gap in the data of the fourteen studied cases, and thus marks the crossover zone. Other breakpoints are stipulated using gaps in the cases' data to eventually form a four-value fuzzy-score.

4.1.2. Conditions

Integrated Planning Authority (PA). To define the set of cases with stronger planning integration, PA is based on two attributes: first, it considers whether cases have an overarching planning organisation with authority over an area compatible with commuting patterns, and not restricted by local political borders. Second, it evaluates the extent to which such authority is multimodal, i.e. the relevance (based on demand levels) of the modes planned by such authority (regardless if the authority only defines minimum service requirements or specific day-to-day operational activities). The two attributes are aggregated following the weakest link technique (logical AND) (Legewie, 2017).

Land-use and Transport Integration (LUT). To define the set of cases with higher degree of integration between land-use and transport, LUT is based on two attributes: first, it considers whether regional planning is institutionalised and regional plans prepared (regardless of the statutory character of these plans, as this is not a requirement for successful spatial planning (Mäntysalo et al., 2015; Searle, 2016)). A strong planning framework, especially at the regional level, is key to support the development of integrated

[2] World Bank data available at: https://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD. OECD data available at https://stats.oecd.org/Index.aspx?Datasetcode=CITIES (access on April 2018).

[3] The authors contacted the teams responsible for the compilation of each database.

Table 3
Truth table for modal split.

PA	LUT	LP	AF	FI	RO	OUT	Incl.	Cases
1	1	1	1	1	0	1	1.000	Helsinki15, Oslo15
0	1	1	1	1	0	1	1.000	Helsinki05
1	1	0	1	1	0	1	1.000	Madrid15
1	1	0	1	1	1	1	1.000	Stockholm05
1	1	1	1	1	1	1	1.000	Stockholm15
1	1	1	0	1	1	1	0.924	Amsterdam15, Berlin05, Berlin15
1	1	1	0	1	0	0	0.803	Amsterdam05
1	0	0	0	1	1	0	0.706	Turin05, Turin15
0	0	0	1	1	0	0	0.692	Oslo05
0	0	0	0	0	1	0	0.640	Birmingham05
1	0	1	1	1	0	0	0.638	Vancouver15
0	0	1	1	1	0	0	0.588	Montreal05, Montreal15
1	0	0	1	1	0	0	0.524	Madrid05, Vancouver05
1	0	0	1	1	1	0	0.506	Melbourne05, Melbourne15
0	0	1	0	0	1	0	0.385	Birmingham15

Table 4
Truth table for cost-recovery.

PA	LUT	LP	AF	FI	RO	OUT	Incl.	Cases
0	0	1	1	1	0	1	1.000	Montreal05, Montreal15
0	0	0	1	1	0	1	1.000	Oslo05
0	1	1	1	1	0	1	1.000	Helsinki05
1	1	1	0	1	1	1	0.901	Amsterdam15, Berlin05, Berlin15
1	1	1	1	1	0	1	0.900	Helsinki15, Oslo15
1	1	0	1	1	0	0	0.658	Madrid15
1	1	1	1	1	1	0	0.622	Sockholm15
1	1	1	0	1	0	0	0.620	Amsterdam05
1	0	0	1	1	0	0	0.500	Madrid05
1	1	0	1	1	1	0	0.433	Stockholm05

Tables' Note: 'OUT' is the outcome column. It takes a score of 1 if the row passes the frequency and consistency thresholds. 'Incl.' refers to the inclusion level, which is the measure of consistency.

strategies (Hickman et al., 2013; Paulley and Pedler, 2000; Suzuki et al., 2013). Second, LUT considers the extent to which cases display mobility patterns indicative of good integration, manifested by higher levels of walking and biking (Cervero, 2013; Hickman et al., 2013). The attributes are aggregated following the weakest link technique (logical AND) (Legewie, 2017).

Long-term metropolitan PT planning framework (LP). The condition LP assesses whether cases have stipulated a planning framework that mandates the adoption and periodic revision of a strategic long-term plan for PT at the metropolitan level. LP also considers the continuity over time of such planning framework, valuing cases that adopt and maintain an organised long-term planning cycle for a longer period. The condition proposes a three-level fuzzy-score to distinguish cases.

Agency (decision power) over funding (AF). The condition evaluates the availability and degree of agency over funding at the regional level. AF considers two attributes. First, the level of government that is primarily responsible for PT funding, as this can affect the achievement of societal goals (Veeneman et al., 2015). Second, the amount of funds coming from earmarked sources, that indicate funding security and stability (Faivre d'Arcier, 2014; Litman, 2014). The score of both attributes is added up to a maximum score of 1. Earmarked sources of funding are not usual and rarely constitute the bulk of PT subsidies, so the value assigned to this attribute in the proposed fuzzy-score is lower.

Fare Integration (FI). This condition assesses cases' degree of fare integration (it does not consider ticket integration). PT fares can be zone or distance-based, or flat across the region. Regardless of the underlying pricing structure, integrated fares do not vary depending on the number of operators used or whether transfers are required (Sharaby and Shiftan, 2012). Based on this definition, the condition proposes a four-level fuzzy-score.

Risk Allocation to Operators (RO). This condition defines the set of cases in which a higher degree of risk is allocated to operators, considering both short-term and long-term risks. The proxy for short-term risk is the allocation of risks in contracts for PT service delivery: gross-cost contracts allocate production risks to operators and net-cost contracts allocate production and commercial risks to operators. Each of these two types of risks constitutes one attribute. For cases adopting different types of contract, the arrangements used in modes carrying at least 80% of PT demand is considered. The proxy for long-term risk, instead, is the ownership structure of the operator(s). The assumption is that in-house operators are ultimately backed by government and thus less exposed to bankruptcy and possibilities of being taken out of the market. If private and public operators coexist, the nature of prevailing operators (carrying at least 80% of PT demand) determines the attribute's score. The three attributes are scored separately and averaged to compute the final score of RO (Legewie, 2017). Online Material includes information specific to all cases in which demand levels are used to support scoring of RO.

A

Pathways	Consistency	Cases
PA*LUT*AF*FI	1.00	Madrid15, Stockholm05, Helsinki15, Oslo15, Stockholm15
PA*LUT*LP*FI*RO	0.94	Amsterdam15, Berlin05, Berlin15, Stockholm15
LUT*LP*AF*FI*ro	1.00	Helsinki05, Helsinki15, Oslo15
Solution: Consistency: 0.97 and Coverage: 0.58		

a. *Solution Pathways. The uppercase letters indicate that the condition is present. Lower case indicates the negation (absence) of the condition. The * denotes a conjunction (logical AND). The pathway consistency shows how consistent that configuration is with the assertion that the combination of conditions is sufficient for the outcome. The solution consistency indicates the degree to which membership in the solution model is a subset of membership in the outcome. Coverage assesses the degree to which a cause or causal combination "accounts for" instances of an outcome. Thus, solution coverage expresses how much of the outcome is covered by the model (Ragin, 2008).*

b. *Sufficiency relation. Cases above the diagonal line are consistent with the sufficiency assertion. Typical cases (upper right quadrant) display the outcome and are explained by a solution pathway. Deviant cases, instead, can be of two types: at the lower right quadrant, they are members of a solution pathway but do not display the outcome; at the upper left quadrant they display the outcome even though no solution pathway covers them. At the lower left quadrant cases are not in any pathway and do not display the outcome. Points in the chart are jittered as they were too densely clustered.*

B

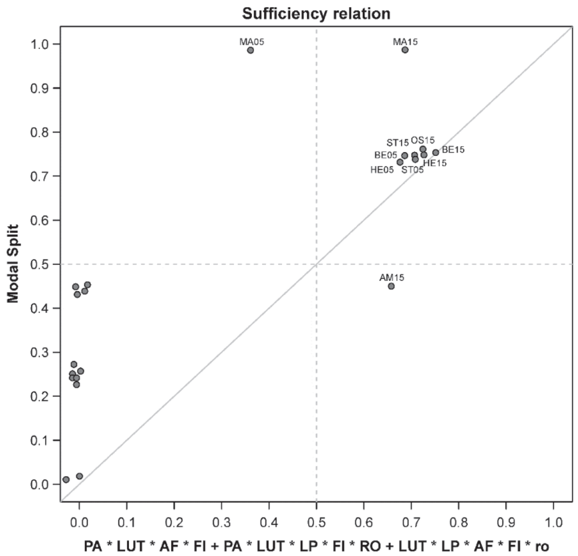

Fig. 1.

A

Pathways	Consistency	Cases
LUT*LP*AF*FI*ro	0.92	Helsinki05, Helsinki15, Oslo15
AF*FI*pa*ro	1.00	Oslo05, Montreal05, Montreal15, Helsinki05
PA*LUT*LP*FI*RO*af	0.90	Amsterdam15, Berlin05, Berlin15
Solution: Consistency: 0.95 and Coverage: 0.73		

a. *As described in Figure 1*

b. *As described in Figure 1*

B

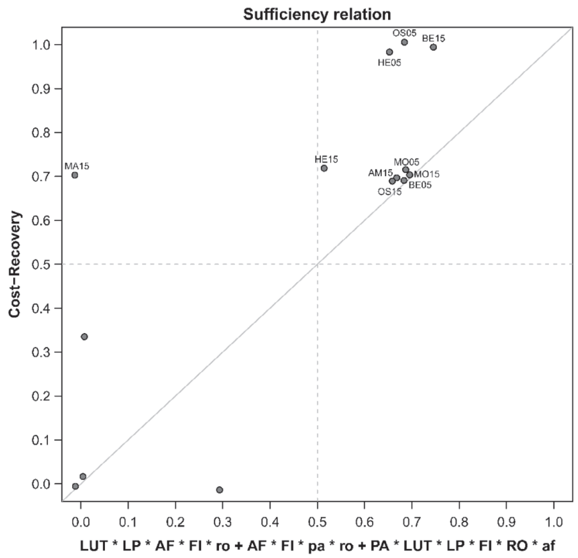

Fig. 2.

5. Results

5.1. Analysis of necessity

A condition is interpreted as necessary when it is a superset of the outcome, and consistency is the parameter of fit indicating the proportion of the outcome that is included in the set of each condition (Duşa, 2018). To claim that a condition is necessary, it must display a consistency level of at least 0.9 (Vis and Dul, 2016). When developing the necessity test, it is also important to inspect if the conditions are relevant or trivial, i.e., a condition might be a superset of the outcome because it occurs virtually in all cases, regardless of a positive or negative outcome (Goertz, 2006; Schneider and Wagemann, 2012). Air, for instance, is a trivial necessary condition for armies to operate (Goertz, 2006, p. 89). A parameter to measure relevance of necessary conditions is RoN (Relevance of Necessity), and Duşa proposes 0.6 as a minimum "decent relevance threshold." (2018, p. 123). The condition FI is the only to pass the consistency threshold, however it has very low RoN value. Therefore, none of the conditions is declared necessary for either modal split or cost-recovery (see Online Material for details).

5.2. Truth table and minimisation

The analysis of sufficiency, main interest in this study, proceeds with the Truth Table Algorithm (Ragin, 2008). In the truth table, each row displays a logically possible combination of the conditions under analysis. Each case is assigned to a row. This is based on the cases' membership in the combination of conditions displayed by each row, calculated with fuzzy multiplication (Ragin, 2008). Rows that do not cover any case are called logical remainders: they represent logically possible configurations that have no empirical manifestation within the cases studied. The researcher, then, must define which configurations of conditions are sufficient for the outcome. Two parameters support this decision: frequency cut-off and consistency cut-off. The first refers to the minimum number of cases a row should cover to be deemed sufficient, conventionally set at one (Rihoux and Ragin, 2009; Schneider and Wagemann, 2012). The second parameter considers the measure of consistency that, in the analysis of sufficiency, indicates the degree to which the combination of conditions displayed in a row is a subset of the outcome. The cut-off is conventionally set at least at 0.75 (Rihoux and Ragin, 2009; Schneider and Wagemann, 2012). Tables 3 and 4 show the truth tables for modal split and cost-recovery built with a frequency cut-off of one and a consistency cut-off of 0.9.

Once the truth table is built, Boolean minimisation follows. This procedure reduces longer expressions into shorter solution formulas revealing combinations of conditions that are minimally sufficient for the outcome. Minimisation works by comparing pairs of truth table rows and eliminating logically redundant factors. According to QCA's Standard Analysis (Ragin, 2008; Ragin and Sonnet, 2005), minimisation can derive three types of solutions that vary according to how they approach the use of logical remainders: complex, parsimonious, and intermediate.[4] The main text presents the intermediate solution, that should be the focus of the substantive discussion (Schneider and Wagemann, 2012). For this solution, minimisation compares all rows passing the frequency and consistency cut-off points and, additionally, part of the logical remainders, the so-called 'easy counterfactuals'. Logical remainders are deemed easy counterfactuals when they include a condition that the researcher, based on existing substantive knowledge, can plausibly expect to be conducive for the outcome. In other words, minimisation of the intermediate solution blocks only difficult counterfactuals. This paper's intermediate solution considers that the presence of all conditions contribute positively to both outcomes, except for RO in relation to which no assumption is made.

Figs. 1 and 2 display two presentations of the solution models for modal split and cost-recovery. First, the solution formulas reveal three alternative minimally sufficient combinations of conditions for each outcome. Second, an XY-plot depicts the superset/subset relationship between outcome and solution. The consistency of the two solution models and of individual pathways are superior to the consistency cut-off set for the truth tables, thus the minimum threshold to declare sufficiency is observed throughout the entire analysis. Moreover, the intermediate solution produces only one model for each outcome, indicating that there are no ambiguities that could weaken results' explanatory power (Baumgartner and Thiem, 2015). Online Material presents supplementary information for all analyses in this Section.

6. Discussion

6.1. Analysis of results

As expected, the condition *fare integration* appears across all pathways, as it passes the consistency threshold for necessity. However, as noted in Section 5.1, it is a trivial condition. The condition *land-use and transport integration*, whilst not necessary, is included in all solution pathways for modal split, showing high empirical relevance for this outcome. This confirms the strong connection between the built environment and modal choice (Ewing and Cervero, 2010; Van Wee, 2002; Zegras, 2010), highlighting the crucial role of this condition for increasing the attractiveness of PT. Amsterdam is the only case in the sample with a high degree of *Land-use and transport integration* but lower level of modal split (an issue further discussed below). An *Integrated planning authority* also shows high empirical relevance for higher modal split, and is part of two solution pathways. In the third pathway, a closer look at the cases shows that only in Helsinki05 no fully integrated planning authority exists. However, even if low, there was some degree of planning integration: YTV, a regional organisation formed by four municipalities (Espoo, Vantaa, Kerava and Kirkkonummi) was responsible for PT services crossing municipal borders (inter-municipal buses and commuter trains) and acted as planning authority for municipal services in Espoo and Vantaa. In the city of Helsinki, the municipally-owned HKL acted as planning authority and operator for the metro, bus and trams. It is plausible to conclude that *integrated planning authority* is also an important enabling condition for higher modal split.

Concerning cost-recovery, *agency over funding* and *risk allocation between government and operators* appear in all solution pathways, being always part of the story explaining this outcome. The way these two conditions are combined seem to shape incentives for cost-

[4] The validity of the different types of solution is the source of debates amongst scholars: see Ragin (2008) or Schneider and Wagemann (2012) for arguments pro the intermediate solution. See Baumgartner (2015) for arguments pro the parsimonious solution. This discussion is beyond the scope of the current study.

savings and/or for revenue generation. In the first two pathways, the conjunction AF*ro is present in cases that display high cost-recovery. The higher *agency over funding* in primarily locally funded PT systems, found in Scandinavian countries for instance, appears to create incentives for planning authorities to minimise PT costs (borne by their shareholders). As a result, contracting transfers operational responsibilities and associated cost risks only. The third pathway, instead, includes the conjunction af*RO that suggests that the lower *agency over funding* in primarily State-funded PT systems, as Amsterdam's for instance, creates less incentives for planning authorities to minimise spending, but to use the allocated budget to maximise production instead (van de Velde et al., 2008). Contracting in this case transfers both planning and operational powers, as well as commercial and production risks. Operators have room to design services and look for ways to cut costs and maximise revenues. Getting the right alignment between *agency over funding* and *risk allocation* might help explain how Amsterdam changed status between 2005 and 2015, and moved to the set of cases with higher levels of cost-recovery (GVB, municipally-owned operator, and the regional planning authority formalised a net-cost contract since 2007). However, just having the right alignment between these two factors is not enough to ensure higher cost-recovery, as the example of Madrid05 demonstrates. These institutional design characteristics involving *agency over funding* and *risk allocation strategy* might result from path dependencies and processes specific to the cases. In Amsterdam and Berlin, for instance, strong municipally-owned operators have historically occupied a prominent role in planning PT. Thus, when a contract was formalised between them and respective planning authorities, the choice of a net-cost arrangement that kept these actors with a significant planning role might have been a reflection of the then existing practices, regardless of (or in addition to) incentives generated by the funding framework.

For some cases, multiple pathways explain the same outcome. Two pathways identified for modal split for instance, cover Oslo15, Helsinki15, and Stockholm15. This raises the question of which of the combinations of conditions is most likely to be operative influencing the occurrence of the outcome. The situation in Oslo is even more interesting because this is the only case that changes status from 'less in' to 'more in' the set of cases with higher modal split. Oslo went through several important changes during the period analysed: a metropolitan planning authority was established (Ruter in 2008), funding agency was enhanced (due both to a commitment by Oslo and Akershus counties to maintain subsidies always at least at their 2007 levels, and to increasing earmarked funds from Oslo's toll ring), fare integration was advanced (important reforms in 2008 and 2011) and a PT planning framework as well as a regional land-use and transport plan were formalised. All these changes can contribute to Oslo's success according to the different solution pathways. However, the regional integrated multimodal planning vision brought by Ruter and the enhanced availability and decision power over funding seem particularly noteworthy: increased and more stable funds enabled the single planning body to implement a series of changes to enhance PT services and attract more users (higher frequency, higher ticket and fare integration, better passenger information systems, new infrastructure etc.). It is plausible to suspect that the pathway including both *integrated planning authority* and *agency over funding* (PA*LUT*AF*FI) is operative. A follow-up within-case analysis could explore this hypothesis.

Other cases, instead, are not fully accounted for by the models, suggesting that factors exogenous to the analysis might be influencing outcomes. Madrid15 has high levels of cost-recovery but is not part of any solution pathway. The same situation occurs with Madrid05 concerning modal split. Amsterdam15, on the other hand, presents a more interesting situation because although it displays a combination of conditions that is sufficient for higher levels of modal split (PA*LUT*LP*FI*RO) it does not achieve the outcome. Amsterdam's modal split grows in the period analysed, however some factor seems to slow down this process and eventually the case does not reach levels of modal split comparable to high performing cases. A possible barrier for PT modal split in Amsterdam is the very high share of bike use. Amsterdam, and The Netherlands in general, have particularly favourable conditions for biking (even topography, dedicated infrastructure, and a strong bike culture based on decades-long supporting policies), and as result, bikes represent a strong competition for PT, especially for shorter trips in congested city areas. However, evidence also indicates that many PT trips take place due to the fact that passengers can use their bikes to access stations (first/last mile of the trip) (Pucher and Buehler, 2008; Rietveld, 2000), suggesting there is a two-way relationship at work between bikes and PT. Madrid and Amsterdam are deviant cases and thus offer leads for follow-up within-case studies that can complement this analysis and unveil new barriers or contributing factors preventing or enabling the occurrence of studied outcomes (Beach, 2018; Schneider and Rohlfing, 2016).

Finally, no combination of conditions is connected to both higher levels of modal split and cost-recovery (no single pathway for both outcomes is the same). This provides another evidence that even though modal split and cost-recovery are measures linked to ridership objectives, they might not be always aligned goals. This does not mean that modal split and cost-recovery are contradictory goals and that a choice between one and the other is necessary. However, achieving each of these goals might ask for a different approach. As just observed, *integration between land-use and transport* and an *integrated planning authority* are core for higher modal split, although not sufficient per se for the outcome. Cost-recovery requires focusing on the way *agency over funding* and *risk allocation strategies* are designed in combination, shaping incentives for cost-savings and/or revenue generation - although again, the combination of these conditions per se is not sufficient.

6.2. Research limitations

The current analysis inevitably has limitations. QCA has been criticised by many authors (Collier, 2014; Lucas and Szatrowski, 2014; Tanner, 2014) (but see also responses in De Meur et al., 2009; Ragin, 2014; Vaisey, 2014). One disadvantage of QCA is that whilst it is able to discern set-relational cross-case patterns, such patterns do not necessarily reflect causation, and the method does not explain the underlying causal processes driving outcomes. Subsequently, the pathways identified by QCA do not clarify whether conditions at play are causal, and can trigger a process, or scope (contextual) conditions, also relevant, but that only constitute a factor that has to be present for a relationship to work (Beach, 2018). Complementing QCA with within-case analyses (e.g. using process tracing) can help address these shortcomings (Beach, 2018; Schneider and Rohlfing, 2016). Finally, QCA's ability to incorporate the time dimension is restricted. It cannot distinguish whether different conditions work in sequence or at the same time for instance. Examining cases in distinct moments in time constitutes one of the possible strategies to add a longitudinal perspective to the analysis (Schneider and Wagemann, 2012; Vis et al., 2013). It allows observing how cases move across different truth table rows or in the XY plot as they change over time.

Furthermore, systematic publication of performance data is still a relatively recent practice amongst authorities. No standard terminology or calculation method exist to define performance metrics. As a result, both data availability and comparability pose difficulties. To mitigate these challenges the paper resorts to databases from renowned institutions that have been promoting benchmarking efforts. Additionally, it is not possible to ensure that every potentially relevant variable has been included in the analysis (as highlighted by deviant cases). This, nevertheless, does not mean the purpose of a study or method breaks down (Radaelli and Wagemann, 2018), and the paper's choice of conditions is backed by expert opinion and academic literature.

7. Conclusion

This paper advances that PT is a complex multifarious system in which technical elements and multiple actors with diverse and conflicting values coexist. As a result, it proposes that the study of PT governance, and particularly the relationship between organisation and performance, should take a complexity-informed approach, recognising that outcomes in PT are the result of the interplay between its systemic elements and context. This claim is aligned with current discussions in PT governance literature that criticise the lack of a systemic comprehension of PT and also of the wicked nature of its governance (Macmillen, 2013; Marsden and Reardon, 2017). To address this gap, the paper examines how the interplay of six PT organisational elements is connected to higher modal split and cost-recovery levels. The study adopts fsQCA, a method well-suited for the analysis of complex social phenomena and wicked policy issues (Blackman et al., 2013; Byrne, 2013; Verweij and Gerrits, 2012).

The analysis identifies three alternative sufficient combinations of conditions connected to each outcome, underscoring PT's expected causal complexity: modal split and cost-recovery depend on the interplay of several conditions, and different paths can lead to the same results. Even though modal split and cost-recovery are related to PT usage maximisation, no single pathway leads to both outcomes. This does not mean these goals are incompatible (cases in this analysis show that it is possible to have both attractive and financially sustainable PT), but that achieving each goal might require policy-makers and transport authorities to focus on different factors: results suggest that *integration between land-use and transport* and an *integrated planning authority* are central for enabling higher modal split, whereas higher cost-recovery is associated to the way *agency over funding* and *risk allocation strategies* shape incentives for savings and/or revenue generation.

These insights can potentially be extended to metropolitan areas beyond those included in the analysis. Generalisations, however, should be circumscribed to similar cases, and primarily analysts or policy-makers interested in other medium-sized high income metropolitan areas in Western economies might benefit from this study's findings. This is because generalisation from case-studies in general, and QCA too, is moderate (George and Bennett, 2005; Gerrits and Verweij, 2016). Differently from *purely* quantitative research methods, that estimate the average effect of independent variables (effects-of-causes approach), case-study designs like QCA search for explanations that are first and foremost linked to the cases being analysed (causes-of-effects approach) (Mahoney and Goertz, 2006).

Methodologically the paper responds to calls for mixed-method research in transport studies (Marsden and Reardon, 2017; Schwanen et al., 2011) by combining the use of qualitative and quantitative data with QCA, a novel method in this field. Concerning the organisation-performance discussion, in turn, the analysis' results produce thicker knowledge on the interplay between diverse elements of PT organisation, providing decision-makers with more leverage to influence strategic outcomes. Findings also open possibilities for future research. Coming work can explore other potential causal relations, such as interdependencies between PT organisational elements, or even the opposite causal direction and the possible influence of good performance on the organisational setup of PT. Furthermore, follow-up studies can help explain underlying causal mechanisms and also expose barriers or enabling factors omitted in this study (Beach, 2018; Schneider and Rohlfing, 2016). These in-depth analyses can focus on other relevant issues that possibly affect the functioning of PT systems, like path dependencies, informal institutions, or the capacity of key actors.

Declaration of interests

The authors have no competing interests to declare.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

The authors are grateful for the interviewees' cooperation. In addition, we thank B. Vis and M. Baumgartner for their generous advice in early stages of this research. Finally, we appreciate the insightful comments provided by the anonymous reviewers.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tra.2019.04.008.

References

Baumgartner, M., 2015. Parsimony and causality. Qual. Quant. 49, 839–856.
Baumgartner, M., Thiem, A., 2015. Model ambiguities in configurational comparative research. Sociol. Methods Res. 1–34.
Beach, D., 2018. Achieving methodological alignment when combining QCA and process tracing in practice. Sociol. Methods Res. 47, 64–99.
Blackman, T., Wistow, J., Byrne, D., 2013. Using qualitative comparative analysis to understand complex policy problems. Evaluation 19, 126–140.
Bouf, D., Faivre d'Arcier, B., 2015. The looming crisis in French public transit. Transp. Policy 42, 34–41.
Byrne, D., 2013. Evaluating complex social interventions in a complex world. Evaluation 19, 217–228.
Cervero, R., 2013. Linking urban transport and land use in developing countries. J. Transp. L. Use 6, 7–24.
Chowdhury, S., Hadas, Y., Gonzalez, V.A., Schot, B., 2018. Public transport users' and policy makers' perceptions of integrated public transport systems. Transp. Policy 61, 75–83.
Collier, D., 2014. Comment: QCA should set aside the algorithms. Sociol. Methodol. 44, 122–126.
de Block, D., Vis, B., 2018. Addressing the challenges related to transforming qualitative into quantitative data in qualitative comparative analysis. J. Mix. Methods Res. 1–33.
de Bruijn, H., Dicke, W., 2006. Strategies for safeguarding public values in liberalized utility sectors. Public Adm. 84, 717–735.
De Meur, G., Rihoux, B., Rihoux, B., Yamasaki, S., 2009. Addressing the Critiques of QCA, in: Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques. Sage Publications, pp. 147–166.
Duşa, A., 2018. QCA with R: A comprehensive resource. Springer International Publishing.
European Metropolitan Transport Authorities, 2017. EMTA Barometer of Public Transport 2015.
European Metropolitan Transport Authorities, 2016. EMTA Barometer of Public Transport 2014.
European Metropolitan Transport Authorities, 2009. EMTA Barometer of Public Transport 2006.
Ewing, R., Cervero, R., 2010. Travel and the built environment: a meta-analysis. J. Am. Plan. Assoc. 76, 265–294.
Faivre d'Arcier, B., 2014. Measuring the performance of urban public transport in relation to public policy objectives. Res. Transp. Econ. 48, 67–76.
Filippini, M., Koller, M., Masiero, G., 2015. Competitive tendering versus performance-based negotiation in Swiss public transport. Transp. Res. Part A Policy Pract. 82, 158–168.
Filippini, M., Prioni, P., 2003. The influence of ownership on the cost of bus service provision in Switzerland - an empirical illustration. Appl. Econ. 35, 683–690.
George, A.L., Bennett, A., 2005. Case Studies and Theory Development in the Social Sciences. MIT Press, Harvard.
Gerrits, L.M., Verweij, S., 2016. Qualitative comparative analysis as a method for evaluating complex cases: an overview of literature and a stepwise guide with empirical application/qualitative comparative analysis. als Verfahren zur Bewertung komplexer Fälle: Eine Literaturübersicht und. Zeitschrift für Eval. 15, 7.
Goertz, G., 2006. Assessing the trivialness, relevance, and relative importance of necessary or sufficient conditions in social science. Stud. Comp. Int. Dev. 41, 88–109.
Gwilliam, K., 2003. Urban transport in developing countries. Transp. Rev. 23, 197–216.
Hensher, D.A., Stanley, J., 2008. Transacting under a performance-based contract: the role of negotiation and competitive tendering. Transp. Res. Part A Policy Pract. 42, 1143–1151.
Hess, D.B., Lombardi, P.A., 2005. Governmental subsidies for public transit. Public Work. Manag. Policy 10, 138–156.
Hickman, R., Hall, P., Banister, D., 2013. Planning more for sustainable mobility. J. Transp. Geogr. 33, 210–219.
Hirschhorn, F., Veeneman, W., van de Velde, D.M., 2018. Inventory and rating of performance indicators and organisational features in metropolitan public transport: a worldwide Delphi survey. Res. Transp. Econ. 69, 144–156.
Hull, A., 2008. Policy integration: What will it take to achieve more sustainable transport solutions in cities? Transp. Policy 15, 94–103.
Hull, A., 2005. Integrated transport planning in the UK: From concept to reality. J. Transp. Geogr. 13, 318–328.
International Association of Public Transport, 2015. Mobility in Cities Database 2015.
International Association of Public Transport, 2006. Mobility in Cities Database 2006.
Jørgensen, T.B., Bozeman, B., 2007. Public values an inventory. Adm. Soc. 39, 354–381.
Koppenjan, J., Charles, M.B., Ryan, N.F., 2008. Editorial: managing competing public values in public infrastructure projects. Public Money Manag. 28, 131–134.
Kumar, A., Agarwal, O.P., 2013. Institutional Labyrinth: Designing a Way Out for Improving Urban Transport Services-Lessons from Current Practice. World Bank, Washington DC.
Legewie, N., 2017. Anchored calibration: From qualitative data to fuzzy sets. Forum Qual. Soc. Res. Sozialforsch 18.
Litman, T., 2014. Local funding options for public transportation. J. Public Transp. 17, 43–74.
Lucas, S.R., Szatrowski, A., 2014. Qualitative comparative analysis in critical perspective. Sociol. Methodol. 44, 1–79.
Macmillen, J., 2013. Mobility as a complex system: key elements and interactions. In: Givoni, M., Banister, D. (Eds.), Moving Towards Low Carbon Mobility. Edward Elgar Publishing, Cheltenham, UK / Northampton, MA, USA, pp. 190–205.
Mahoney, J., Goertz, G., 2006. A tale of two cultures: contrasting quantitative and qualitative research. Polit. Anal. 14, 227–249.
Mäntysalo, R., Kangasoja, J.K., Kanninen, V., 2015. The paradox of strategic spatial planning: a theoretical outline with a view on Finland. Plan. Theory Pract. 16, 169–183.
Marsden, G., May, A.D., 2006. Do institutional arrangements make a difference to transport policy and implementation? Lessons for Britain. Environ. Plan. C Gov. Policy 24, 771–789.
Marsden, G., Reardon, L., 2017. Questions of governance: rethinking the study of transportation policy. Transp. Res. Part A Policy Pract. 101, 238–251.
May, A.D., 2004. Singapore: the development of a world class transport system. Transp. Rev. 24, 79–101.
May, A.D., Kelly, C., Shepherd, S., 2006. The principles of integration in urban transport strategies. Transp. Policy 13, 319–327.
Mees, P., 2005. Privatization of rail and tram services in melbourne: what went wrong? Transp. Rev. 25, 433–449.
Mizutani, F., Urakami, T., 2003. A private-public comparison of bus service operators. Int. J. Transp. Econ. 30, 167–185.
Nielsen, G., Lange, T., 2007. Network Design for Public Transport Success - Theory and Examples, in: Thredbo10 - International Conference Series on Competition and Ownership in Land Passenger Transport.
Nielsen, G., Nelson, J.D., Mulley, C., Tegnér, G., Lind, G., Lange, T., 2005. Public transport – Planning the networks -. HiTrans Best Practice Guide 2.
Paulley, N., Pedler, A., 2000. TRANSLAND: Integration of Transport and Land Use Planning. Final Report for Publication.
Pemberton, S., 2000. Institutional governance, scale and transport policy - lessons from Tyne and Wear. J. Transp. Geogr. 8, 295–308.
Potter, S., Skinner, M.J., 2000. On transport integration: a contribution to better understanding. Futures 32, 275–287.
Preston, J., 2010. What's so funny about peace, love and transport integration? Res. Transp. Econ. 29, 329–338.
Preston, J., Almutairi, T., 2013. Evaluating the long term impacts of transport policy: an initial assessment of bus deregulation. Res. Transp. Econ. 39, 208–214.
Pucher, J., Buehler, R., 2008. Making cycling irresistible: lessons from The Netherlands, Denmark and Germany. Transp. Rev. 28, 495–528.
Radaelli, C.M., Wagemann, C., 2018. What did I leave out? Omitted variables in regression and qualitative comparative analysis. Eur. Polit. Sci. https://doi.org/10.1057/s41304-017-0142-7.
Ragin, C., 2014. Comment: lucas and szatrowski in critical perspective. Sociol. Methodol. 44, 80–158.
Ragin, C., 2009. Qualitative Comparative Analyses Using Fuzzy-Sets (fsQCA). In: Rihoux, B., Ragin, C. (Eds.), Configurational Comparative Methods: Qualitative

Comparative Analysis (QCA) and Related Techniques. Sage Publications Inc, pp. 87–122.
Ragin, C., 2008. Redesigning Social Inquiry: Fuzzy Sets and Beyond. University of Chicago Press, Chicago.
Ragin, C., 1987. The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies. University of California Press, Berkley.
Ragin, C., Sonnet, J., 2005. Between Complexity and Parsimony: Limited Diversity, Counterfactual Cases, and Comparative Analysis. In: Kropp, S., Minkenberg, M. (Eds.), Vergleichen in Der Politikwissenschaft. VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 180–197.
Redman, L., Friman, M., Gärling, T., Hartig, T., 2013. Quality attributes of public transport that attract car users: a research review. Transp. Policy 25, 119–127.
Rietveld, P., 2000. The accessibility of railway stations: the role of the bicycle in The Netherlands. Transp. Res. Part D Transp. Environ. 5, 71–75.
Rihoux, B., Ragin, C.C. (Eds.), 2009. Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques. Sage Publications.
Roy, W., Yvrande-Billon, A., 2007. Ownership, contractual practices and technical efficiency: the case of urban public transport in France. J. Transp. Econ. Policy 41, 257–282.
Schneider, C.Q., Rohlfing, I., 2016. Case studies nested in fuzzy-set QCA on sufficiency: formalizing case selection and causal inference. Sociol. Methods Res. 45, 526–568.
Schneider, C.Q., Wagemann, C., 2012. Set-Theoretic Methods for the Social Sciences A Guide to Qualitative Comparative Analysis. Cambridge University Press.
Schwanen, T., Banister, D., Anable, J., 2011. Scientific research about climate change mitigation in transport: a critical review. Transp. Res. Part A Policy Pract. 45, 993–1006.
Searle, G., 2016. Strategic planning and land use planning conflicts: The role of statutory authority. In: Albrechts, L., Balducci, A., Hillier, J. (Eds.), Situated Practices of Strategic Planning: An International Perspective. Routledge, pp. 317–330.
Sharaby, N., Shiftan, Y., 2012. The impact of fare integration on travel behavior and transit ridership. Transp. Policy 21, 63–70.
Skaaning, S.-E., 2011. Assessing the robustness of crisp-set and fuzzy-set QCA results. Sociol. Methods Res. 40, 391–408.
Sørensen, C.H., Gudmundsson, H., 2010. The impact of governance modes on sustainable transport–the case of bus transport in Greater Manchester. UK. World Rev. Intermodal Transp. Res. 3, 8–25.
Stanley, J., 2014. Land use/transport integration: Starting at the right place. Res. Transp. Econ. 48, 381–388.
Stead, D., Geerlings, H., 2005. Integrating transport, land use planning and environment policy. Views of practitioners from Denmark, England and Germany. Innov. Eur. J. Soc. Sci. Res. 18, 443–453.
Suzuki, H., Cervero, R., Iuchi, K., 2013. Transforming Cities with Transit: Transit and Land-Use Integration for Sustainable Urban Development. World Bank, Washington DC.
Tanner, S., 2014. QCA is of questionable value for policy research. Policy Soc. 33, 287–298.
Thacher, D., Rein, M., 2004. Managing value conflict in public policy. Gov. An Int. J. Policy Adm. Institutions 17, 457–486.
Thomann, E., Maggetti, M., 2017. Designing research with Qualitative Comparative Analysis (QCA): approaches, challenges, and tools. Sociol. Methods Res. 1–31.
Toshkov, D., 2016. Research Design in Political Science, first ed. Palgrave Macmillan.
Vaisey, S., 2014. Comment: QCA works-when used with care. Sociol. Methodol. 44, 108–112.
van de Velde, D.M., 1999. Organisational forms and entrepreneurship in public transport. Part 1: Classifying organisational forms. Transp. Policy 6, 147–157.
van de Velde, D.M., Veeneman, W., Schipholt, L., 2008. Competitive tendering in The Netherlands: Central planning vs. functional specifications. Transp. Res. Part A Policy Pract. 42, 1152–1162.
Van Wee, B., 2002. Land use and transport: research and policy challenges. J. Transp. Geogr. 10, 259–271.
Vassallo, J.M., Pérez De Villar, P., Muñoz-Raskin, R., Serebrisky, T., 2009. Public transport funding policy in Madrid: is there room for improvement? Transp. Rev. 29, 261–278.
Veeneman, W., Augustin, K., Enoch, M., Faivre d'Arcier, B., Malpezzi, S., Wijmenga, N., 2015. Austerity in public transport in Europe: the influence of governance. Res. Transp. Econ. 51, 31–39.
Verkuilen, J., 2005. Assigning membership in a fuzzy set analysis. Sociol. Methods Res. 33, 462–496.
Verweij, S., Gerrits, L., 2012. Understanding and researching complexity with Qualitative Comparative Analysis: evaluating transportation infrastructure projects. Evaluation 19, 40–55.
Vis, B., 2009. Governments and unpopular social policy reform: biting the bullet or steering clear? Eur. J. Polit. Res. 48, 31–57.
Vis, B., Dul, J., 2016. Analyzing relationships of necessity not just in kind but also in degree: complementing fsQCA with NCA. Sociol. Methods Res. 1–28.
Vis, B., Woldendorp, J., Keman, H., 2013. Examining variation in economic performance using fuzzy-sets. Qual. Quant. 47, 1971–1989.
Walker, J., 2008. Purpose-driven public transport: creating a clear conversation about public transport goals. J. Transp. Geogr. 16, 436–442.
Zegras, C., 2010. The built environment and motor vehicle ownership and use: evidence from Santiago de Chile. Urban Stud. 47, 1793–1817.