The Impact of Public Transport Reform: an Assessment of Deregulation Policies

Mouwen, A.M.T.

2016

Chapter 1

Introduction

The importance of public transport (PT) in sustaining a healthy and accessible environment is increasingly recognized (European Commission, 2013). Especially in cities, PT can contribute to solving congestion and reducing CO_2 and other pollutants. However, PT services in general are only viable with the aid of large amounts of public subsidies. Due to the recurrent fiscal crises and budget deficits in many Western countries and cities, PT services are becoming harder to secure. Under these circumstances, the call for subsidy savings and the introduction of competition by means of deregulation, witnessed in many countries, is not surprising. Of course these policies are also driven by the specific institutional context, habits and legislation in countries, but the dominant argument and common factor in deregulation policies is the need to cut budgets.

In this PhD thesis the topics addressed concern the effects of introduction of competition into the PT sector by imposing deregulation. These public policies may be essential in achieving efficiency and equity objectives. The approach of this PhD thesis is to perform theoretical and empirical analyses of the economic structure and the institutional context of PT. By doing so the aim is to clarify the complex mechanisms underlying PT provision, demand and finance in a (de)regulated environment dominated by competitive tendering (CT), and to assess the impact of CT on passengers, PT authorities, and PT firms.

The scope of this thesis is restricted to bus, tram, metro and regional train services. In the remainder of this work the object of this study is referred to as 'public transport' (PT). Heavy rail, and technical and infrastructural topics are explicitly excluded With regard to the actors involved, the primary perspective of this thesis is the authorities as we mainly evaluate regulatory policies. However, the perspective of the passenger and the PT firm is also frequently examined. PT authorities (PTAs), in their role as passengers' representatives, have to be acquainted with passengers' demands and preferences. The firm acts as an intermediate level between the authority and the passenger: PT services in a regulated regime are procured by authorities, provided by operators and consumed by customers/passengers. Thus, firms are an indispensable link, and authorities wishing to impose regulatory policies have to take the underlying economic and financial structures of firms into account.

1.1 Research questions

A literature review was performed regarding three main topics: the economic fundamentals of PT, passenger satisfaction with PT, and financing and funding of PT.

On the topic of economic fundamentals the review revealed that the circumstances under which PT is supplied, namely providing excess capacity under network conditions, poses some serious challenges for regulatory policies. This observation formed the basis of the first block of research questions on regulation and economic structure: why do public authorities intervene in markets? What are the arguments for (de)regulation of an industry, in what way do these arguments apply to the PT industry, and in what manner do strategic contractual decisions made by authorities and firms affect the efficiency of PT service procurement?

On the topic of satisfaction, in the prevalent literature rather little attention is paid to the quantitative relationship between introduction of competition and satisfaction with PT. Therefore the second block of research questions deals with: how is the concept of satisfaction and service attribute importance to be defined, what are the main drivers for PT passenger satisfaction, in what way do socio-demographic and trip characteristics affect satisfaction and to what extent does competitive tendering affects overall and attribute specific customer satisfaction with PT?

On the topic of efficiency the literature review revealed that a great number of studies is performed on the efficiency effects of deregulation policies, however in these studies the efficiency effects of competition are hard to identify properly due to lack of unbiased data. Additionally, the review on this topic showed that often not all cost factors of regulatory change are accounted for as these costs (especially the transaction costs) are not known to the objective observer. Contracts serve as the formal stipulation of arrangements between parties, and the literature showed that contract renewal can be a powerful incentive driver. Concerning efficiency topics, the third block of research questions to be answered in this thesis is: what is the impact of PT contract renewal on operational costs, subsidies and PT ridership?

1.2 Outline

To answer the first block of research questions, Chapter 2 of this thesis provides an introduction to the economic fundamentals of PT firms and deals with topics such as network conditions and production technology. In Chapter 2 also the arguments for regulation and the concepts of competitive tendering (CT), contracts and transaction costs are described in greater detail.

The quantitative core of this thesis is laid out in Chapters 4, 5 and 6. Methodological issues are not dealt with in a separate chapter, but are included in these three chapters. The empirical tests of the models are based on data of the Netherlands. Chapter 3 therefore serves as an introduction on PT in the Netherlands and describes the main trends of PT demand, supply, finance, governance and market structure. Also the level of transaction costs is estimated in this chapter. The research questions in block 2 are answered in Chapters 4 and 5. In Chapter 4 the drivers of customer satisfaction are explored by building a number of econometric models of overall satisfaction and transaction-specific (or attribute) satisfaction, taking account of customer and environmental characteristics such as age, negative social safety experiences and level of urbanization. Hypotheses on these issues are formulated and empirically tested.

After an introduction to the definition and measurement of the dimensions of satisfaction, in Chapter 5 empirical models on the relationship between CT and satisfaction are constructed and tested. In Chapter 6 the research questions in block 3 are explored. This chapter deals with the effect of contract renewal on costs, subsidies and ridership under a competitive tendering regime. Panel data for the period 2001–2013 on the level of concession areas are collected and employed. A number of econometric models are constructed to assess the aforementioned relationships. This thesis is finalized in Chapter 7 with a summary statement of the main findings, conclusions, and recommendations for further research and policy.

Chapter 2

Regulation in the public transport industry

This chapter discusses the economic and political arguments of (de)regulation and considers the most likely market outcome in the case that no regulation is imposed upon the sector. Arguments are found in the production, costs and market structure of PT provision. Procurement of PT services by way of competitive tendering (CT) is a common used regulatory instrument. Therefore, we first describe the costs and production characteristics. Next, we discuss considerations for regulation and competitive tendering including transaction costs.

2.1 Costs and production characteristics

This section provides basic information concerning the cost and production characteristics of PT. We restrict the description to those characteristics that are relevant in assessing the potential effects of regulation, i.e. demand fluctuations, network conditions, production technology and subsidies.

2.1.1 Fluctuations in demand for PT

Demand for PT services largely follows societal activity patterns and therefore is spatially and temporally differentiated. If firms fully comply with this differentiated demand, the PT network, routes and services have to be designed to accommodate maximum (peak hour and peak direction) demand, instead of average demand which would be economically more feasible.[1] In that situation, costs and production structure are based on excess capacity; i.e. in peak hours, capacity determines production and costs. This characteristic is manifest particularly in the vehicle fleet size and the number of drivers. As production capacity is defined based on peak performance, in off-peak hours many buses stand idle.[2] Numerous studies have come to the conclusion that the actual fleet size exceeds the optimal size necessary to produce the observed output (Viton, 1981; De Borger, 1984; Obeng, 1984; Karlaftis and McCarthy, 2002). These studies show large short-term economies of capital utilization.

2.1.2 Network conditions

PT firms perform their services within a network. This feature is of great importance in understanding the firm's cost structure, as it affects costs in two ways: (i) a relatively high proportion of fixed costs in total cost and (ii) the largely fixed network size and layout. For rail services, it is evident that fixed rail infrastructure costs are an important determinant, but bus services are also provided under conditions of large fixed costs, for instance pertaining to terminals, depots, maintenance and repair shops and information systems. The costs of these specific fixed facilities can be defined as sunk costs for the firm as few alternative uses are

1 The decision to comply with peak demand may not be voluntary but may stem from requirements that PTAs impose on the firms, but even when firms only facilitate average demand, the back-haul problem causes production under excess capacity.

2 On top of peak capacity for peak demand, firms have to account for 10–25% extra reserve capacity (break down, maintenance, etc.).

possible and/or they render lower revenues when sold.[3] Production under these conditions easily leads to economies of scale as capacity can be increased at less than marginal costs (Berechman and Giuliano, 1985; Farsi et al., 2007). As with capital (vehicles), certainly in the short term, the firm can hardly adjust the network size and structure as this would mean changing the production structure completely and would lead to falling demand, passengers being accustomed to the network. Given the network and excess capacity characteristics, PT is undertaken under economies of scale and density (Berechman, 1993; Filippini and Prioni, 2003).

Another characteristic of services performed within a network is that these services are perceived by passengers as unique. Alternative routes are seen as imperfect substitutes, for instance being slower and having longer waiting times, more interchanges, etc. Thus, a PT network can be interpreted as a vector of unique – spatially differentiated – products. Spatial differentiation diminishes as the network density increases, thereby increasing the chance of routes being substitutes. Therefore, in urban settings, passengers will regard routes more as substitutes than in rural settings.[4]

2.1.3 Fixed input production technology

PT firms have limited ability to swap capital for labour as the number of vehicles, drivers and maintenance personnel is proportional, i.e. labour and capital are not easily mutually substitutable; the elasticity of input factor substitution is low.[5]

2.1.4 Subsidies

There has been little discussion thus far on the effect of subsidies on PT productivity and costs. It appears that subsidies negatively affect these variables. De Borger et al. (2002) provide an overview of empirical results and add the important suggestion that causality runs from subsidies to (operational) costs and not vice versa. The main body of studies on this relationship dates from before 2006; however, Sakai and Takahashi (2013) recently showed for the local bus market in Japan that PT efficiency decreases as the subsidy ratio increases.

---

3 In combination with the excess capacity characteristic of supply, vehicle costs may also considered as largely sunk, especially in case of an abundant supply of used vehicles in the market, implying that the price of used vehicles is quite low, as is for instance the case in the Netherlands (Van Woelderen, 2009). In this case, the fixed costs of buying vehicles become sunk.

4 Market entrance for newcomers in a deregulated situation (see Section 2.2) implies entrance to an urban market is more difficult than to a rural market as in the former it will be hard for a new entrant to find a market niche, especially because demand patterns in cities invoke high frequency routes over dedicated infrastructure in high volumes. For a new entrant, having to establish product differentiation can be a strong barrier to entry.

5 For instance, the price elasticity of labour to capital is approximately 0.04 and for capital to labour approximately 0.1 (Berechman, 1993, p. 134; Karlaftis and McCarthy, 2002). This feature also implies that a standard Cobb–Douglas production function is not appropriate for analysing the costs of PT.

2.1.5 *Optimization strategies*

The implication of the characteristics described above is that strategies for increasing efficiency available to operators are limited. Theoretically, given the high share of labour costs in total costs (for buses more than 50%, according to Koolen and Stoelinga, 2005), substantial cost efficiencies can be achieved by reducing the direct labour unit price or amount, or rescheduling driver shifts. Other potential optimizations are redesigning routes, maintenance and fuel conservation activities, optimizing vehicle capacity (bringing small or articulated buses into operation) and reductions in indirect personnel (management). However, in practical terms there are several impediments to these cost-saving strategies, such as the strong unionization of the sector, the requirements set by authorities and production characteristics, as described above. For instance, given excess capacity, the fixed proportion of labour to capital and strong unions, strategies aimed at making direct cuts to labour costs are hardly viable.[6] Route redesign and optimizing vehicle capacity are often restricted by the requirements of the authorities and maintenance and fuel conservation activities are restricted by the fixed vehicle fleet size. Therefore, firms have little room to optimize production and the first and most important cost reduction strategy firms apply in practice is cutting indirect personnel costs.

2.2 Market structure and potential deregulation effects

Section 2.1 discussed the characteristics of PT provision. In this section, we discuss PT market structure in a situation in which there are no barriers to entry and exit and firms can set prices without restrictions.[7] What would be the market outcome in that situation? We use the cost and production characteristics set out in the previous sections to answer this question theoretically. In the debate on why and how to regulate or privatize the sector, this is an important issue.

As noted in section 2.1, the PT sector produces under conditions of excess capacity, several types of economies (scale, density) due to network conditions and therewith a relatively large associated proportion of fixed (sunk) costs for infrastructure and production facilities. Economic theory in these circumstances predicts that without regulation, a monopolistic or oligopolistic market structure would be the outcome as incumbent firms can effectively hinder new entrants (Kahn, 1988). There has been a long-lasting political debate on whether this outcome is preferable or not (Baumol, 1982; Sharkey, 1982; Shleifer, 2010). Free market advocates state that monopoly would not lead to resource misallocation as the self-regulating mechanism of 'market contestability' becomes effective (Baumol et al., 1982; Tirole, 1988).[8] If contestability conditions prevail in the PT market, theory predicts that following deregulation

6 In many countries these strategies are not plausible as authorities restrict the competitiveness of the transit labour market by law, prescribing the transfer of all direct personnel from the incumbent to a new operator under the same labour conditions.

7 This implies a theoretical situation of the complete absence of public authority control and interference in the primary strategic parameters of the firm.

8 Misallocation may lead to high prices charged to customers and monopolistic rents.

there will be only one firm supplying PT in a certain market, but this firm will be forced to set prices based on average costs, implying no monopolistic earnings. If this firm were to charge above-average prices, new entrants would successfully penetrate the market by undercutting the incumbent's prices. Thus, according to contestability theory, the threat of entry will force the incumbent not to behave in a monopolistic fashion. However, several authors have shown that in the PT sector the primary conditions for contestability (free market entrance and exit and average cost pricing) probably do not apply (Bannister et al., 1992; Berechman, 1993). The two main impediments are: (i) the lack of possibility for new firms not currently in the market to adopt a hit-and-run strategy as sunk costs prevent such a strategy[9] and (ii) the fact that the incumbent does not act under rigid/fixed price conditions as demand price elasticity is low. As a consequence, in response to the potential entry of a new firm, the incumbent will lower prices.[10]

Given the PT cost and production properties, we conclude that in the absence of any regulation, it is unlikely that transit markets are contestable, implying an incumbent can successfully deter competitors and therefore oligopolistic or monopolistic market outcomes are more likely. Taking a closer look at market segments (urban versus regional markets), we hypothesize that new operator entry is more likely in inter-urban/regional markets than in intra-urban markets. Although urban markets potentially generate more PT demand than regional markets, in urban markets it is more difficult to differentiate products and find a market niche as in an urban market routes are more likely to be substitutes. For a new entrant having to establish product differentiation, this can be a strong barrier to entry. Moreover, urban networks are denser and their operation depends more on the availability of expensive fixed assets, such as central (hub) bus terminals and depots, than is the case in inter-urban/regional markets. To enter an urban market, a new firm has to invest heavily and the incumbent – already owing these fixed facilities – will be capable of effectively deterring new entrants. To summarize, under complete deregulation, intra-urban market segments are most likely to be characterized as monopolistic, whereas more competitive conditions are likely to prevail on inter-urban/regional routes.[11]

2.3 Arguments for regulation

What is the fundamental reason for the public sector to intervene in the PT market? Essentially, government organizations impose regulatory policy as a mechanism to minimize economic inefficiency and to make welfare distribution more equitable. General theory puts forward three main fields of arguments for regulation: economic, political and social (Demsetz, 1968;

9 As we have described, many amenities are sunk. Also, the fact that a potential entrant has to publish its timetable and route network makes it possible for the incumbent to alter its routes and prices accordingly, thereby making hit-and-run for the new entrant very unprofitable.

10 Empirical studies show demand is very price inelastic, especially in the short term (Goodwin, 1992; Litman, 2004).

11 Empirical evidence for the UK and the USA seems reasonably consistent with this conclusion (see De Borger and Kerstens, 2006).

Stigler, 1971; Van Delden and Veraart, 2001).[12] Before we elaborate on these arguments, we first define 'regulation' in the PT sector as a situation in which a public entity imposes restrictions on one or several key decision parameters for PT firms. Following Berechman (1993), we characterize the following parameters as essential for firms: fare level and fare structure, use of input factors, type, composition and quality of services supplied, route network makeup and design, vehicle fleet size and vehicle type, freedom of market entry and exit, investment policy, legal ownership of firms and fixed facilities. As such, in the absence of restrictions on the above parameters, we term such a situation completely non-regulated.[13]

2.3.1 Economic arguments for regulation

The most important theoretical economic argument underlying regulation is market failure. Market failure may be attributed to many factors; however, we concentrate on public goods and natural monopoly as possible causes.[14] Under certain conditions, markets may fail to allocate resources effectively because the marginal costs faced by individuals in consuming the goods and services are not equal to the societal marginal cost as a whole or are not equal to market prices. This leads to inefficiencies, such as over- or underproduction or consumption, or costs for society (Vickers and Yarrow, 1991). The main reasons for market failure are externalities, public goods and (natural) monopolies (see e.g. Button, 1993; Hillier, 1997; Shleifer, 2010). A characteristic of (negative) externalities is that the costs related to them are not (fully) borne by the parties that caused them, but fall onto society. Regulation is then needed to repair this. Externalities become manifest in transportation, for instance as traffic congestion, air pollution and traffic unsafety. In the case that public goods are involved, markets may fail.

A good or service is defined as 'public' if either of two conditions apply: non-paying customers cannot be prevented from consuming it (non-excludable) and for any level of production the cost of providing it to a marginal (additional) individual is zero (non-rivalrous). Left to market forces, these goods will not be provided for. The provision and maintenance of (non-congested) transport infrastructure is often explained as a case of a public good (Samuelson and Nordhaus, 1985, pp. 48–49), but there is much discussion regarding this claim (see e.g. Block, 1993; Skousen, 1997).

If a monopoly is a natural monopoly, there may be reasons for public regulation. Natural monopolies occur when production technology causes long-run average total costs to decline as output expands (economies of scale and density), for instance due to relatively high fixed costs (Baumol et al., 1982). In this situation, competition may not be a viable strategy

12 The political and economic arguments related to market failure are brought together by Dempsey (1989).

13 This does not however imply that restrictions are completely absent, as in every situation authorities have to impose restrictions to prevent externalities and collusion, to exert administrative control and to safeguard the interests of specific groups, such as the elderly and disabled.

14 Imperfect information, destructive competition, incomplete markets and stimulating innovation are also often mentioned as rationales for regulation; however, in general, these are seen as less important (Shleifer, 2010). We will not discuss these topics further.

as firms tend quickly to reduce to one again due to mergers or failures, or production will consume more resources than is efficient (Tirole, 1988). In theory, in this situation, one supplier will be able to produce more efficiently than two or more suppliers. Higher prices will result if two or more suppliers provide the services. The government has to protect these monopolies through regulation, for instance by granting franchised monopolies with price caps.[15] Public transport is often claimed to be a natural monopoly (Mills, 1980; Savage, 1997; De Borger et al., 2002). However, as was the case with the public goods argument, the natural monopoly argument has also attracted criticism. Austrian School adepts, for instance, deny the existence of natural monopolies and argue that the only reasons public authorities want to control monopolistic markets are political (DiLorenzo, 1996; Shleifer, 2010).[16]

2.3.2 Political arguments for regulation

The political argument most commonly put forward for regulation is that public ownership of a firm is the best way to safeguard employment and social rights for the workers (see e.g. Melly and Puhani, 2013). Another is that in urban planning, transit is seen as a prerequisite for healthy urban development. Land use planning and transport planning interact, as offices in central business districts (CBDs), for instance, are preferably built within easy access to public rapid transport, so as to reduce hinder by car in these areas aiming at more sustainable cities. It is feared that PT deregulation will reduce this feeding role and will make integrated land use/transport planning – transit-oriented design (TOD) – more difficult (Herala, 2003; Wegener and Fuerst, 2004). Finally many authors in PT point to general budget deficits as a prime determinant for imposing regulation on the PT sector (Andersen, 1992; Banister et al., 1992; Hensher and Wallis, 2005; Ongkittukul and Geerlings, 2006). These more pragmatic – budget-steered – policy objectives for regulation are often considered more influential than the aforementioned theoretical arguments (Berechman, 1993; Hensher and Wallis, 2005).[17]

2.3.3 Social arguments for regulation

The equity principle is often mentioned as an argument for regulation (Berechman, 1993). From this perspective, some goods and services are valued as essential for individuals' basic welfare. Non-discriminating service supply has to be assured by regulation to protect the weak and the poor in society and to strengthen social cohesion, safety and public health (Gwilliam and Van de Velde, 1990; Preston and Rajé, 2007). Government may consider public transport as essentially a merit good and therefore sustain a socially accepted minimum level of public transport for specific target groups (such as elderly or disabled people and people

15 As is/was the case in many public utility goods such as electricity, telecommunications and rail transport.

16 The Austrian School adepts do not deny the existence of economies of scale and density in utility services and goods, but this is in no way an argument for wanting to exert public controls as competition is always most preferable.

17 In the practice of policy, political and economic arguments interfere and are willingly mixed by advocates and opponents of regulating alike.

living in remote, low density areas). In this context, it is defensible that services are provided at affordable prices to all, or at prices that do not fully cover costs. The ability of the PT industry to subsidize low-use, loss-making routes internally using the earnings from heavy-used profitable routes and to compensate off-peak tariffs with peak hour revenues (cross-subsidization) is often used as a reason for regulation. In the absence of regulation, competition could lead to charging above-marginal prices on profitable routes and not supplying low-used routes (cream-skimming). Indeed in the period after deregulation in the UK outside London, this situation seems to have occurred (Preston and Almutairi, 2013).

2.4 Competitive tendering: the best of both worlds?

We have seen that public transport services are network services and are tied to public welfare. The first characteristic may lead to market failure due to natural monopoly features and the irreversibility of investments, the latter characteristic to service meritorization. Therefore, the classic 'good market' model (the first best solution is self-regulation by means of competition within the market) cannot simply be transposed to these services and a second best solution may be optimal.[18] This second best solution does not however mean that total exception for competition must be made (Cox, 2003) as regulatory failure lies in wait (Dempsey, 1989).

The proponents of regulation and of deregulation agree that, under the condition that the network remains a (natural) monopoly, given the particular PT cost and production (network) characteristics separating the network and the services over the network may be a feasible second best solution. Also the potential beneficial effects of competitive tendering (CT) on costs and subsidy reduction are hardly disputable (Berechman, 1993; Cox, 2003, Gautier and Yvrande-Billon, 2013; Preston and Almutairi, 2013). CT, also called 'competition for the market', 'competitive bidding', 'franchising' or 'contracting out', is a regulatory regime initiated and organized by authorities who set the requirements, grant the right to operate the network and pay a subsidy. By means of CT, firms are incentivized to reveal their true costs, therefore providing authorities with the necessary market information. The theory underpinning CT rests on a long history of economic debate on how to regulate 'natural' monopolies (see section 2.3.1). As early as 1859, Chadwick wrote "the existence of economies of scale in water, electricity, or other 'public utilities' in no way necessitates either monopoly or monopoly pricing, as a competitive bidding system for the services of private utility franchises can eliminate monopoly pricing as long as there is there competition for the field" (Chadwick, 1859, p. 389). If, for whatever reason, public control is deemed necessary to control a market, when properly used, CT may be an instrument by which this can be carried out with the minimum loss of efficiency. A number of authors argue that the transaction cost of tendering

18 For technical and economic reasons, parallel competition between networks is not feasible and parallel competition by more than one service provider within a given network does not correspond to the natural monopoly characteristics of PT service supply. Private suppliers will not easily enter the market as huge investments (which are partly sunk) are involved and the margins are small. These investments may further act as a market exit barrier.

should be accounted for when assessing the effects of competitive tendering. Therefore in the next section we deal with transaction costs.

2.5 Transaction costs

Contracts are a powerful instrument for governing procurement transactions between principals (PTAs) and agents (operators), as they make it possible to allocate risk and introduce incentives for agents. Contract theory (Laffont and Tirole, 1986) and transaction cost economics (TCE) theory (Williamson, 1985) address several issues that have to be taken into account when assessing the cost effects of regulatory reform.[19] One of the most important issues is the occurrence of transaction costs.

The magnitude of transaction costs in the case of the procurement of PT services is not known (Van de Velde and Beck, 2010), but is of interest when assessing the effects of a competitive tendered award. In this section, we aim to define these costs and also to identify potential disadvantages of a competitive tendering regime according to TCE theory. We start with a description of two fundamental concepts in procurement decisions. Then, based on TCE theory, we discuss CT problems in general terms. In Section 2.5.3. we focus on the planning function of PT. In Chapter 3 we perform a 'back-of-the-envelope calculation' of transaction costs of competitive tendered PT procurement in the Netherlands.

2.5.1 Award mechanism, contract, and type of transaction cost

To assess PT-service procurement, it is useful to distinguish between the awarding procedure and the contract type.

We define the awarding procedure as the procedure that results in the right to operate PT for a given period of time in a geographical area or on a route. We distinguish authority-initiated awarding procedures and market-initiated procedures. The two most common forms of authority-led initiatives are an award after competitive tendering and a direct or negotiated award.[20] In market-initiated procedures the award by authorities is rather uncomplicated, as in most cases there are no subsidy transfers and authorities only exert a limited review of, for instance, vehicle safety and pollution regulations in order to provide operators with a licence.

Contracts belong to two broad types, that is, fixed-price (FP) contracts (net or gross cost contracts) and cost-plus (CP) contracts.[21] These types differ in their degree of risk-sharing and incentive power. CP contracts (also called management contracts) are lowpowered, as the PTA collects traffic receipts and fully reimburses operators' production costs and pays an

19 Contract theory analyses how parties to a contract make decisions under uncertain conditions, when there is asymmetric information. Information asymmetry denotes a situation in which the agent has more/better information regarding the transaction than the principal. Contracts can be incentivised in order to promote certain outcomes, but can also contain a level of moral hazard (Laffont and Tirole, 1993). Moral hazard pertains to a situation in which one of the parties does not enter the process with fair intentions and/or provides misleading information on its assets.

20 Competitive tendering in contract theory is very similar to auctions.

21 In practice, many hybrid contract types are used. See examples in Van de Velde et al. (2008).

additional pre-specified amount (the "plus"). Profits are independent of realisations, therefore the operator bears no risk; the risks of over-subsidization are with the authority. FP contracts are high powered. In a fixed-price gross cost contract, the operator bears production risks and the PTA the revenue (commercial) risks. In fixedprice net cost contracts, the operator bears both production and commercial risks. Finally, a fixed-price contract type that is not commonly used is what Gautier and Yvrande-Billon (2013) define as a 'concession contract', that is, a net cost contract but where the operator is additionally in charge of investments in, for instance, dedicated infrastructure.[22]

Although in theory the award procedure and contract type are independent of each other, in practice logical combinations prevail. For instance, CT usually implies fixed-price contracts, as price is the most important award criteria in this procedure. Further, TCE theory argues that, for complex, longrun, unpredictable services, a negotiated awarding procedure with a cost-plus contract is the most economically viable combination (Williamson, 1976, 1985). TCE theory distinguishes between three phases in procurement procedures that bring forth transaction costs: the contact, contract, and execution phase. For procurement of PT services through CT, contact phase transaction costs are for instance costs made for the invitation to tender, pre-selection of suppliers, drafting procedures and award criteria, preparing bid documents, and processing bid competition. Costs in the contract phase apply to the costs of the award and settlement of the transaction. This implies costs of drafting and agreements once trading partners are identified and include costs paid to lawyers, accountants, and notaries and the time spent by personnel of the principal and agent for these activities. The execution phase refers to agreements made to curb the opportunistic behaviour of transaction partners. Costs associated with the execution of the contract include costs of monitoring transactions, performance measurement, judging conformance to the agreement, identifying and solving disagreements, renegotiations and adjustment of the agreement, enforcement and application of sanctions, awarding bonuses and penalties, etc.

2.5.2 Difficulties with competitive tendered awards according to TCE theory

TCE theory identifies some serious difficulties in awarding and executing competitive tendered contracts (Williamson, 1976). According to TCE theory, these problems arise due to contractual disabilities of parties and may lead to considerable transaction costs. In this section we discuss these problems in general. In Chapter 3 we assess the occurrences of these problems in the Netherlands.

The most prominent problem pertains to the complexity of the services to be procured and thereby to the complexity of the contract and the choice of contract type. The seminal theory concerning incomplete contracts was developed by Grossman and Hart (1986), arguing that due to long term uncertainties, contract parties cannot commit themselves to mutually

22 The contract of the City of Lille is an example of such a contract.

beneficial renegotiations later on in their contract relationship. A hold-up problem may occur (see also Schmitz, 2001). Williamson (2002) put forward that, 'as all complex contracts are unavoidably incomplete [. . .] parties will be confronted with the need to adapt to unanticipated disturbances by reason of gaps, errors, and omissions in the original contract' (Williamson, 2002, p. 174). A complex contract may therefore not only lead to transaction costs, but also to strategic behaviour of bidders and thereby to adverse selection. In that case, opportunistic bidders, keen to discover specification deviancies, place a low bid as they gamble on renegotiation or contract adaption.[23] Opportunistic behaviour of bidders will increase with the complexity of the service offered and the chance of unexpected occurrences, especially if they are out of the bidder's control. Therefore, complex incomplete contracts will likely be associated with considerable transaction costs according to TCE theory.

Regarding procurement decisions by principals, a tension may exist between principals' wish to provide as many ex ante incentives as possible by choosing a specific contract type and their wish to avoid ex post transaction costs due to costly renegotiation (Bajari and Tadelis, 2001). For example, high incentive FP contracts reduce costs ex ante, but will probably lead to priceincreasing contract adaptations (especially when the service to be procured is complex and the design of the service is therefore left incomplete). On the other hand, low-incentive CP contracts discourage ex ante cost savings and may lead to opportunistic behaviour of providers, but the (transaction) costs of contract adaption are low as the reimbursement process is uncomplicated and well defined. In longer contracts these issues become more problematic, as contract incompleteness is causally related to contract length, and it is not possible for authorities to make provisions for all conceivable relevant future circumstances. Moreover, it is not economically feasible to take all these possible occurrences into account in the contract, as it is not certain whether they really occur; this would lead to excessive transaction costs in the form of monitoring, inspection, etc.

Contracts for the procurement of PT services are by nature complex, as they aim to capture complex arrangements and inter-dependent issues, such as operations, costs, infrastructure quality, asset specificity and ownership, subsidies, etc. Moreover, many conditions may change during the execution of the contract: traffic volumes may increase, roads may be renovated, or new dedicated PT infrastructure may become available. It is therefore difficult for PTAs to fully specify the service design and network ex ante. Especially traffic conditions and the condition of PT infrastructure are hard to predict but exert a large influence on operations and operational costs, as described in Chapter 2. Concerning award mechanisms, Hensher and Stanley (2008) put forward the argument that in complex situations competition is not always preferable. For complex bus and multi-modal areas, directly awarded performance-based contracts (PBC) perform better than competitive tendered contracts, as the latter implies high

23 Hensher and Stanley (2008) observed such strategic behaviour in PT procurement under a CT regime with a fixed price contract.

transaction costs.[24] Especially when public goals not only include cost reduction objectives, but also patronage growth and qualitative requirements—such as innovations, planning expertise, drivers' behaviour, and vehicle cleanliness (see also Chapter 5 on the effect of CT on satisfaction with these quality attributes)—negotiated PBC contracts may have advantages over other contractual arrangements.[25] These notions resonate with the results of TCE theory, showing that it may be preferable that simple uniform products or services are better procured (bought), whereas complex goods/services can best be produced (made) by the principal, as Tadelis and Williamson (2012) argue.

TCE theory also implies other problems that may be connected to CT, such as strategic asset ownership (which may deter market entrance, and especially matters when contract incompleteness leads to distorted ex ante investments and sunk costs), inadequate (use of) awarding criteria, bidding imparity between the incumbent and competing firms, and collusion (Williamson, 1985; Cox, 2003; Yvrande-Billon, 2006; Hensher and Stanley, 2008).[26] All of these items have to be managed and mitigated by authorities. This places a heavy burden on them and requires professionalism and competence. Authorities—specifically small ones—often lack experience and expertise to do this properly. Moreover, in order to steer properly during execution of the contract, open and free access to operational performance, patronage, and revenue data has to be secured.

2.5.3 Contracts and planning function control

The way contracts deal with the tactical level of operations is of great importance for the risk coverage of PTAs and operators (Van de Velde and Sleuwaegen, 1997; Van de Velde et al., 2008). The tactical level refers to network planning and service design and includes, for example, settings of fare levels and level-of-service. In an authority-led regime—such as a CT regime—the authority decides whether the planning function is positioned within the realm of the operator or the PTA. Decisions made at this level strongly influence the costs and subsidies of operations (see Chapter 2), but may also impact on contractual arrangements and transaction costs. We will elaborate on this in Chapter 3.

---

24 Bajari et al. (2009) empirically showed for private sector building contracts that directly awarded negotiated contracts performed better than auctions. The authors also demonstrated that auctions are the preferred award mechanism for cases in which the goods or services are simple to define, design, and monitor and do not change technically during the execution of the contract.

25 The main disadvantage of a PBC is the risk of budget uncertainty for the authorities (open-ended arrangement). For instance, in the initial Adelaide model the operator was so successful in achieving patronage growth in an open-ended PBC that money ran out and the PTA had to reduce the total payment (Hensher and Wallis, 2005). A PBC becomes stronger when supported by the threat of competition in the case of the non-compliance of the operator chosen after negotiation, as this incentivises the operator to perform in accordance with the stated intentions. Finally, one has to be aware that negotiated PBC may also involve significant transaction costs, such as establishing appropriate benchmarks and monitoring systems to evaluate performance.

26 Yvrande-Billon (2006) demonstrated that collusion was indeed a problem in the French PT award system and that the specific characteristics of the French award mechanism (negotiation after pre-selection), did not result in competition and optimal welfare outcomes.

PTAs may exert input or output control over the planning function. Input control implies that the authority specifies the services to be delivered in great detail and procures them on the basis of detailed terms of requirements (ToR), leaving hardly any room for the innovation of operators and placing few incentives on them. Output control implies that PTAs only prescribe the minimum production levels and standards, thereby leaving room for the operator to develop the network and services and 'search for the customer'. The latter types of contract are suitable for patronage-incentive payments.[27] In input-controlled contracts, the incentives for the operator are mainly directed towards the PTA, whereas in output-controlled contracts incentives are directed towards the passenger.

Probably the most prevalent contract type in Europe is the net-cost contract (Van de Velde et al., 2008). In Figure 2.1, for net-cost contracts both aforementioned methods of control are visualised when competitive tendering is used as the award mechanism.

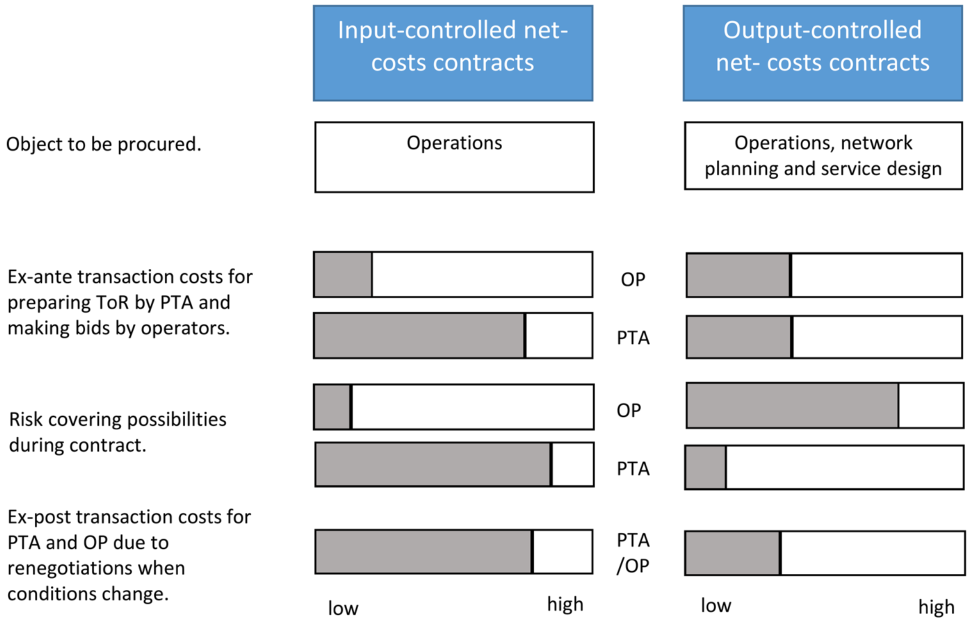


Figure 2.1. *Two models of control in net-cost contracts*

When the PTA is responsible for the planning function, all lines, routes, frequencies, etc. need to be specified. This implies high ex ante transaction costs for preparing the tendering documents. In the case that the PTA restricts itself to descriptions of the service level and intends to steer on output rather than input, these transaction costs may be relatively low. The

27 Given that many factors affecting patronage are outside the control of the operator, the appropriate level of patronage incentive payments in a contract may be fairly modest.

opposite holds for the ex ante transaction costs of the operator; these are lower in the case of input-controlled contracts than in output-controlled contracts, although even in input-controlled contracts considerable bidding costs may be needed to prepare the bid. In input-steered contracts there are many risk coverage possibilities for operators, as the operator is able to adjust his production level in case expectations fail. In net-cost contracts, for instance, the incentive for revenue rests with the operator. In the case that a certain line under-performs on ridership and revenue, in output-controlled contracts the operator is allowed to adjust production levels on this line downwards (limited to a minimum standard set by the PTA) and increase frequency on an over-occupied line to attract additional passengers and revenues. In an input-controlled contract this is not possible without negotiation, as all line frequencies are determined by the PTA. Therefore in output-controlled contracts the risk is easier to govern and renegotiation or contract adaption is less likely/usual than in input-controlled contracts. Therefore it is likely that input-controlled contracts will result in higher transaction costs during execution of the contract than output-controlled contracts. On balance, it is very difficult to say what the total transaction cost outcome of either of these models is, but information from operators and PTAs suggests that output-steered contracts will result in lower transaction costs than input-steered contracts. In Chapter 3 we will provide details of these transaction costs for the situation in the Netherlands.

2.6 Conclusion

The PT sector produces under conditions of excess capacity (due to fluctuating demand patterns) and a relatively large proportion of fixed (sunk) costs for infrastructure and production facilities. These characteristics lead to several types of economies (scale, density and scope). The economic theory in these circumstances predicts that, without regulation, a monopolistic or oligopolistic market structure would be the outcome as incumbent firms can effectively hinder new entrants (Kahn, 1988). Given the specific characteristics of PT markets, we conclude that, under complete deregulation, intra-urban market segments are most likely to be characterized as monopolistic, whereas more competitive conditions are likely to prevail on inter-urban/regional routes.[28.] As public transport services are tied to public welfare (equity principles), a monopolistic market outcome may not be a welfare-optimal outcome, as market failure lies in wait.[29] However, PT service meritorization may also easily endanger the optimal welfare. Neither outcome is optimal. Policy makers and scholars agree on the potential beneficial effects of competitive tendering (CT) on costs and subsidy reduction (Berechman,

28 Although urban markets potentially generate a greater PT demand than regional markets, in urban markets it is more difficult to differentiate products and find a market niche as in urban markets routes are more likely to be substitutes. Moreover, urban networks are denser and their operation depends more on the availability of expensive fixed assets.

29 Besides economic arguments, there are also political and social arguments for regulation. However, in practice, budget-steered policy objectives often act as the prime determinant for imposing regulation on the PT sector (Andersen, 1992; Banister et al., 1992; Berechman, 1993; Hensher and Wallis, 2005; Ongkittukul and Geerlings, 2006).

1993; Cox, 2003; Gautier and Yvrande-Billon, 2013; Preston and Almutairi, 2013). If regulation is deemed necessary to control a market, when properly used, CT may be an instrument by which control can be carried out with the minimum loss of efficiency. When assessing the welfare effects of this regulatory instrument, transaction costs should be taken into account.

We conclude that in the Netherlands there is room to improve the total performance of PT by making contractual choices that are more in line with PTAs' objectives and by taking transaction costs explicitly into account. Competitive tendering with fixed-price contracts may perform poorly regarding transaction costs when PT services are complex and contracts are incomplete. As the trend in the Netherlands is towards longer and more complex contracts, it is likely that PTAs do not take the level of transaction costs and competitive pressure into consideration when making strategic decisions on the contract duration, complexity and type, size of the services to be procured and award mechanism.

Additionally we conclude that the contract type in the Netherlands is in many cases not consistent with the most appropriate award mechanism. The dominant contract type is a net-cost contract coupled with input control by the PTA. Under these circumstances the operators are held responsible for the fare box revenues; however, they are hardly given the freedom to develop the services in the network and 'search' for the passengers. This inconsistent behaviour of authorities may be the consequence of their risk aversion and the dominant image of PT in the Netherlands as a social/merit good, but it is questionable whether this conduct leads to an optimal welfare outcome.

Chapter 3

Main trends in demand, supply, finance and markets in the Netherlands

In the previous chapter, theoretical arguments for regulation in the PT sector were discussed and it was concluded that regulation by imposing competitive tendering (CT) may be a good alternative for a complete free market. In this chapter we focus on the situation in the Netherlands, where competitive tendering was introduced in 2000. We describe the way in which CT in the Netherlands is implemented and provide quantitative information on demand, supply, governance and transaction cost of competitive tendering.[30] This provides a background for subsequent chapters.

3.1 Competitive tendering in the Netherlands

In the Netherlands, as in many European countries, CT became popular in a period of economic crisis and rising subsidies for PT in the mid-70s and 80s, as this instrument was deemed potentially able to reduce costs and subsidies. Therefore, the introduction of this regime, albeit often defended on economic grounds, was more justified in practice by political objectives. Discussions on new governance arrangements for the PT industry in the Netherlands resulted in legislation that became effective in 2000. This legislation provided for phased CT introduction. The intention was that – spread across the nation – by the beginning of 2004/2007, 35% of the turnover in the regional and municipal markets would be tendered and by 1 January 2008 all concessions had to be tendered.

The results of a mid-term evaluation of the actions and effects (Appelman et al., 2004) gave rise to the opinion of the central government that CT should be obligatory, but the original time path was judged to be too optimistic. Moreover, it was decided that three main cities should be released from the obligation to tender (Amsterdam, Rotterdam, The Hague), although administrative independence of the municipal companies was imposed.[31] Outside these cities, CT was compulsory, but the regional and local public transport authorities (PTAs) were left free to develop their own specific governance mechanisms. This applies, for instance, to rolling stock and infrastructure ownership. Except for the three municipal firms, in nearly all regional concessions the vehicles are owned by the operating firm and PT infrastructure ownership largely rests with the authorities.[32] Importantly, regional authorities have the freedom to retain the responsibilities for decisions at the tactical level (planning function) at their own discretion or position them with the operator. This is a crucial decision in market performance terms and functioning as it defines whether the PTA or the operator is in control of designing the route network, fare level and level of service (see also Chapter 2). Although the aim of central government is to position this planning function with the operators, the practice is that nearly all PTAs keep tight control of this, resulting in a situation in which the primary decisions concerning how PT is planned are made by politicians and are barely influenced by PT customers and/or firms (Appelman et al., 2004).

30 We also collected similar information on the situation in the UK. Where applicable, we refer to this.

31 Contracts in these three cities are awarded privately (without competition).

32 An exception is made for vehicle repair and maintenance facilities and depots.

3.2 Trends in demand, supply, finance and governance

In this section we present information on PT performance in the Netherlands. We concentrate on the period after competitive tendering became compulsory (after 2000), but also provide some general background information on the period prior to 2001. We report on modal share, PT performance (both operational and financial) and market structure. As the empirical chapters mainly focus on transit (bus, tram and metro), we omit any consideration of long-haul trains here.

3.2.1 Modal share

Figure 3.1 and Table 3.1 show the modal share for the period 1985–2014.

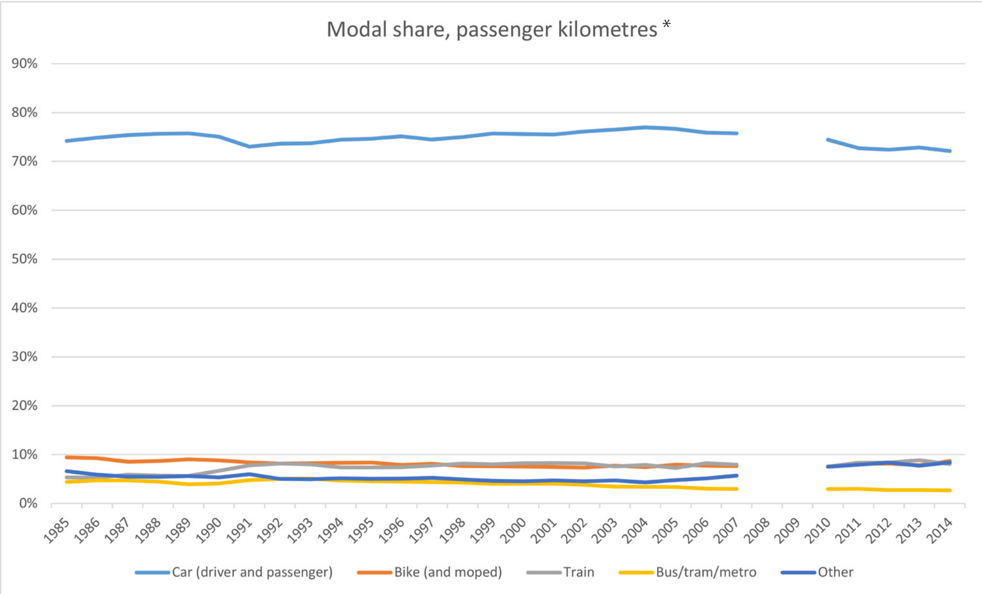


Figure 3.1. *Modal share*

*Source: CBS. * Note: due to survey method alterations, no information is available for 2008/2009.*

What is visible is the rather stable dominant position of the private car (share approximately 75%), although after 2010 a slight fall in car use can be observed (car share in 2014 is 72%). Transport using bicycles and train have stable shares of around 8% each. The share of bus, tram and metro (BTM) transport seems to diminish slightly to around 3% in 2014.[33] However,

33 Similar information for the UK for the same period shows an average car share of 85%, with bus and train each at 6%, and bicycles less than 1%.

the large differences in PT use between areas should be noted. In the large metropolitan areas, for instance, the BTM share is around 6% (see Table 3.1).[34]

3.2.2 BTM performance

Table 3.2 shows supply and demand and Table 3.3 exhibits financial performance indicators.

After 2005, passenger kilometres increased, but given the possible incongruences in data for the period after 2008, this trend is hard to interpret. The supply (number of vehicle kilometres) grows steadily over time, which may be linked to the policy change towards the introduction of competitive tendering described in 3.1.[35] The increasing supply is also manifested in the increasing number of frequent lines relative to infrequent lines. However the number of stops declines, indicating network rationalization. As demand does not keep pace with supply, the occupancy rate declines after 2006.

Table 3.1. *Passenger kilometres per person per day (2014)*

	Netherlands		**Heavily urbanized areas***	
	Number	**%**	**Number**	**%**
Car (driver and passenger)	22.27	72	17.47	63
Bike (and moped)	3.00	10	3.17	11
Train	2.75	9	3.99	14
Bus/tram/metro	0.89	3	1.54	6
Other	1.82	6	1.69	6
Total	30.73	100	27.86	100

*Note: source CBS, *Address-density >2,500 per km*^{2}

The collection of reliable financial data is difficult as definitions change and information is kept confidential. Therefore, we have to be prudent when interpreting Table 3.3. However, what can be deduced is that after deflating the expenditures, the growth in fare box revenues exceeds the growth in total operational expenditures, indicating that the cost recovery ratio improves (see also Figure 3.2).[36]

34 Also notable is that in densely populated areas, the total number of kilometres travelled is lower than in rural areas.

35 For the UK, after 2000, PT demand in London rises, but outside London falls. PT supply shows the same pattern, indicating a stable occupancy rate.

36 In the UK, the net subsidies paid by central and local governments for bus services decreased after 2009.

Table 3.2. *Trends in BTM supply and demand **

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Passenger kilometres (x million)	6,461	6,694	6,364	6,254	6,311	6,224	6,413	6,462	6,355	6,526	6,710
Vehicle kilometres (x million)	483			470	468	470	480	513	529	545	
Number of stops	55,130			51,478	50,554	50,644	49,777	49,551	49,793	50,386	
Number of infrequent lines (< 2x/h)	1,107			1,035	987	985	976	893	913	927	
Number of frequent lines (>= 2x/h)	629			613	630	655	675	736	743	768	
Total number of lines	1,736			1,648	1,617	1,640	1,651	1,629	1,656	1,695	
Occupancy rate (pax kms/veh.kms)	13.3			13.2	13.4	13.2	13.3	12.5	12.0	11.9	

**Source: KNAW/WROOV (2013). Including Student PT tickets. After 2008 less reliable due to a change in the ticketing registration system. After 2009 no uniform data at disposition.*

Table 3.3. *Financial performance (BTM)*

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Nominal fare receipts (million Euro)*	496	509	504	565	577	597	604	637	658	675	737	808	821	
Idem, relative to 2000	-2%		-1%	11%	13%	17%	19%	25%	29%	33%	45%	59%	61%	
Expenditures on operations (current prices, million Euro):														
By central government**		966	1,034	1,110	1,079	1,097								
By metropolitan auth.***							710	729	736	788	851	911	900	843
By provinces***							440	470	546	629	801	820	918	821
Total gov. expenditures		966	1,034	1,110	1,079	1,097	1,150	1,199	1,282	1,417	1,652	1,731	1,818	1,664
Idem corrected for inflation (1999 prices)		942	964	1,001	953	958	987	1,018	1,071	1,155	1,331	1,377	1,413	1,262
Idem, relative to 2000			2%	6%	1%	2%	5%	8%	14%	23%	41%	46%	50%	34%

** Source: KNAW/WROOV (2013). Exclusive of para transit, inclusive of student contract. **Source: Ministerie Verkeer en Waterstaat (2000-2005). All expenditures on PT are funded by central government, but after 2004 no proper registration at central level is possible due to combining infrastructure and operational subsidy transfers from central to decentralized level.***Source: Koopmans et al. (2013).*

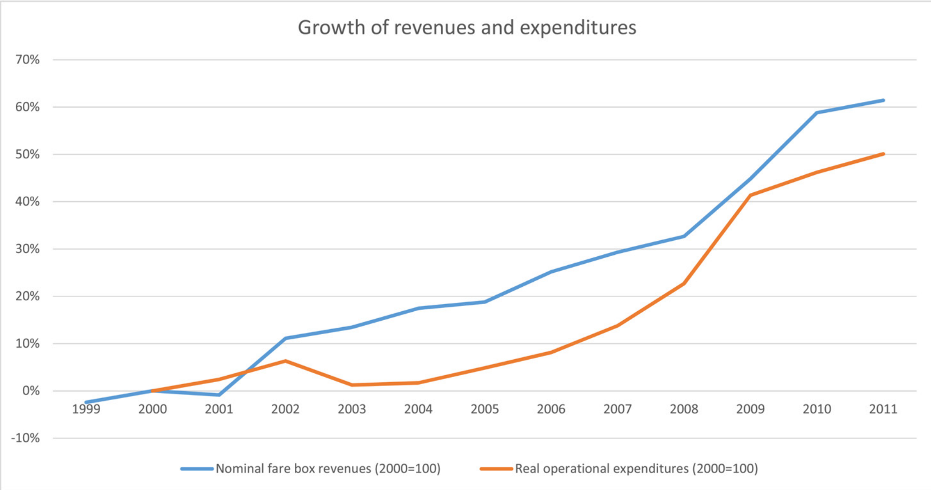


Figure 3.2. *Growth of revenues and expenditures on PT operations*

3.2.3 Market situation

Since 2001, the responsibilities for public transport in the Netherlands have been decentralized. The responsibility now (as of 2015) rests with 14 PTAs: 12 provinces and 2 metropolitan authorities. In the main metropolitan cities (Amsterdam, Rotterdam and The Hague), companies owned by the municipality are responsible for PT operations. The remainder of PT in the Netherlands is operated by (semi) private firms. As of the opening up of the market in 2001, a number of foreign companies have acquired formerly Dutch-owned companies.

Dutch central government policy is aimed at enforcing competition by means of CT. The former state-owned VSN group monopoly had to be broken down, as – according to the Passenger Transport Law 2000 – excessive market power is prohibited. Thus, before competition could take place, the VSN group had to be dismantled. Therefore, in 1998 VSN was split into VSN-1 and VSN-2.[37] VSN-1 was later restructured and renamed Connexxion Holding.[38]

In the pre-tendering phase (before 2001), the market in the Netherlands was dominated by only a few operators who held the initiative rights. The regional market was dominated by Connexxion Holding and the big municipal markets by the municipal companies. In 2001, Connexxion – directly or by means of its subsidiaries – held more than 50% of the market (see

37 VSN-2 was intended to be a temporary holding of which the firms were supposed to be sold externally. A number of firms were indeed sold (to Arriva, UK) and to Connex, France), but Connexxion also acquired ownership over a former VSN-2 firm.

38 Till 2007 the state kept the shares of Connexxion. In 2007 67% was sold (to Transdev), in 2012 the remaining shares were also sold to Transdev Transdev and Veolia merged in 2011, so actually Connexxion and Veolia now operate under the same ownership. Transdev-Veolia owns 75% of Connexxion.

Table 3.4). To assess market power, in Table 3.4 the trend in market share and the Herfinfdahl–Hirschman index (HHI) are shown.[39]

Table 3.4. *Trend in market share*

	2000	2004	2009
Connexxion-Transdev	51.5%	46.6%	42.2%
Arriva	16.5%	16.6%	10.0%
Veolia	1.7%	4.7%	17.2%
Other regional firms	10.0%	10.4%	11.6%
Municipal firms	20.3%	21.6%	19.0%*
Number of firms in the market	13	11	8
HHI for regional markets	0.31	0.26	0.25

Note: vehicle kilometres, source: KNAW/WROOV (2013). Not available after 2009.

* *In 2009 the municipal operator of Utrecht was acquired by Connexxion.*

After competitive tendering was imposed on the sector, the dominant player, Connexxion, lost market share, especially to Arriva and Veolia. Although the number of competing firms decreases due to take-overs of former independent regional firms by Connexxion, the decreasing Herfinfdahl–Hirschman index (HHI) implies that in 2009 market shares are more evenly distributed relative to 2000.[40]

Concerning the effects of competitive tendering, Table 3.5 shows more detail.

We identify a sharp decrease in the percentage of incumbent firms that stay in charge after competitive bidding (83% in 2002, 33% in 2014), indicating that the market is maturing and the comparative first mover advantage (Lieberman and Montgomery, 1988; Cox, 2003) is diminishing.[41] Also noticeable is the increase in the contract duration. The average number of bids per procedure stays more or less the same (around 3); however, in more than 80% of procedures, only the three big contenders (Arriva, Veolia, Connexxion) placed a bid and took a share of the market.

39 See Calkins (1983).

40 The municipal market is not subject to tendering.

41 This advantage pertains to economic efficiency due to proximity to the PTA and knowledge of the network and local conditions.

Table 3.5. *Details of competitive tendering **

	2002	2005	2009	2014
Number of bidding procedures	12	8	9	5
Number of concessions privately awarded (**)	0	0	3	2
Number of competitive tendered bidding procedures	12	8	6	3
Incumbent stays after competitive tendering	10	5	4	1
Only one bid after competitive tendering	3	0	1	0
Average number of bids per tender	2.4	3.5	3.2	2.8
Average contract duration (year)	5.5	5.9	6.8	9.3

** Note: area concessions only. ** In most cases, short (one year) contract extensions as to level mutual concession durations. Source 2002: CROW/KPVV (2015). Source 2005, 2009, 2014: KNV (2015).*

3.2.4 Public governance

In Table 3.6, the trends in public governance are shown. It is clear that public governance after 2000 becomes more concentrated (fewer concession areas and fewer authorities). The decreasing number of concession areas is of interest if placed in the setting of transaction cost economy theories (see Chapter 2). The number of PT-competent authorities decreases due to the initial decentralization of competences from the central to the regional level, as accompanied the introduction of competition in 2000 and the further regional concentration of competences and budgets in 2014.[42]

Table 3.6. *Trends in public governance*

	1999	2000	2004	2009	2014
Number of area concessions	n.a.	74	72	54	39
Number of PT authorities	35	19	19	18	14

Source: CROW/KPVV (2015)

3.3 Transaction costs and procurement problems

3.3.1 Introduction

In Section 2.5 we described potential problems with the awarding of PT services by means of competitive tendering. In this section we assess procurement problems for the Netherlands and will focus on the effects of these problems on transaction costs. Transaction costs must be taken into account when assessing the effects of tendering PT, as these costs eventually

42 This thesis does not assess the effects of decentralization. For further reading on this subject, see Van de Velde and Pruijmboom (2003).

settle as costs for society. In the long run, the ex-ante bid costs of the operators are eventually paid for by the authorities, as they—one way or another—are incorporated into the contract price. Transaction costs fall apart in ex ante and ex post costs, as we discussed in Section 2.5.

Data on individual contractual arrangements and on transaction costs are not publically available, as they contain confidential and competition-sensitive information. We found a number of consultants, PTAs, and operators willing to share information with us under the condition that we present it in an aggregated way. As this information came from a selected sample of respondents and could not be validated, it cannot be used to perform statistical analyses and/or reliably levelled up for the Netherlands. The information does, however, provide us with valuable estimates of the level of transaction costs of both parties involved (PTAs and operators) and indicates the order of magnitude of these additional competitive-tendering costs. The 'back-of-the-envelope calculation' of transaction costs we performed contains PTAs' procurement costs and operators' bidding costs. For both parties we also indicate additional external costs for legal advice and support (lawsuits). The estimates of PTAs' transaction costs are based on information of managers of a medium sized PTA and two consultants. We are of the opinion that these estimates are rather reliable and representative for all Netherlands PTAs, as ex ante transaction costs are likely not strongly related to the size of the concession area, but rather to the formal and long-term (EU-regulated) procurement procedures.[43] The estimates of operators' transaction costs are based on information obtained from a number of operators that together have a market share in Netherlands PT of approximately 45%. Apart from transaction cost estimates, our respondents provided us with information on the occurrences in the Netherlands of the potential CT problems we described in Section 2.5.2.

3.3.2 The level of transaction costs for the period 2001–2015: a rough estimate

As described above, we were able to roughly estimate the ex ante bidding cost and costs made for lawsuits for both authorities and operators, but we have no information on ex post transaction costs. We have, however, the impression that the latter costs may be substantial, as circumstances during contract execution frequently change, resulting in renegotiations and contract adaptations.

We estimated the ex ante transaction costs for the period 2001–2015 for the Netherlands. In this period 107 concessions were granted.[44] We divided the awarded concessions into three groups based on annual turnover: small (less than €30 million), medium (€30 to €70 million), and large (more than €70 million). We estimated that for the period under study there were 20 lawsuits disputing the concession award intention of PTAs. In the case that the award is legally disputed, two operators go to court. Following Table 3.5, the average number of bidders we assume is three and the average contract length is 7 years. Based on these data, in Table 3.7

43 On average, for PTAs the procedure for granting a concession takes 1.5 to 2 years.

44 We disregarded very small concessions and fast ferry concessions.

to 3.9 the estimates of the aggregated ex ante transaction costs for the period 2001–2015 are given.

For the period 2001–2015 for the Netherlands, the total ex ante transaction costs of competitive tendering are roughly estimated to have been 216 million euro to 530 million euro.[45] On a yearly basis this is € 14.4 to 35,5 million. From Table 3.7 it becomes clear that – if the transaction costs are categorized – the bid costs made by operators are dominant related to the procurement costs of the PTAs. We assume that the transaction costs of the operators are included in the bid price, so we may expect that without competitive tendering the PTA expenditures on PT operations may be considerably lower. To assess the latter, we disaggregated the transaction cost estimates in Table 3.7 into estimates of average transaction costs per concession awarded (see Table 3.8).

Table 3.7. *Estimated ex ante transaction costs for the period 2001–2015.**

						Average costs per year	
	Number of concessions awarded by PTAs	**Number of bids per conces-sion**	**Total units**	**Lower limit bid costs per unit (€)**	**Upper limit bid costs per unit (€)**	**Lower limit €)**	**Upper limit (€)**
Bid costs (operators)							
Small	54	3	162	400,000	1,000,000	4,320,000	10,800,000
Medium	32	3	96	600,000	1,700,000	3,840,000	10,880,000
Large	21	3	63	1,000,000	2,500,000	4,200,000	10,500,000
Cost of lawsuits (operators)							
	Number of rewards disputed	Number of lawsuits per award					
	20	2	40	50,000	250,000	133,333	666,667
Subtotal operators						12,493,333	32,846,667
Procurement costs (PTAs)			107	250,000	350,000	1,783,333	2,496,667
Cost of lawsuits (PTAs)			20	100,000	150,000	133,333	200,000
Subtotal PTAs						1,916,667	2,696,667
Total						14,410,000	35,543,333

** Note: exclusive of operator–operator transaction costs. Authors' own calculation.*

45 The 2015 price level.

Hoofdstuk 3

Table 3.8. *Ex ante transaction costs per concession award (millions of euros)*

	Lower limit	Upper limit
Small	1.5	3.5
Medium	2.1	5.6
Large	3.3	8.0

As we will describe in Chapter 6, we conducted a survey among PTAs on costs, subsidies, and efficiency per concession. We use the information gathered in this survey to weight the transaction cost estimates against the total subsidies and subsidy reduction due to the introduction of competitive tendering. Based on the sample, the average annual subsidies for small concessions is €6.2 million, for medium concessions €20.9 million, and for large concessions €54.3 million. We take an average contract period of seven years as a starting point. Table 3.9 shows the calculation of the transaction costs (TC) related to subsidy expenditures and subsidy reduction due to competitive tendering (CT).

Table 3.9. *Transaction costs and subsidy savings per concession (total contract period)*

	Total subsidy (millions of euros)	TC as % of total subsidies		Total subsidy savings due to CT (millions of euros)*	TC as % of CT-caused subsidy reduction	
		Lower	Upper		Lower	Upper
Small	43.4	3	8	8.7	17	40
Medium	146.3	1	4	29.3	7	19
Large	380.1	1	2	76.0	4	10

** In Chapter 6 we will indicate that the subsidy-reduction effects of first time contract renewal amount to at least 20%.*

Transaction costs are mainly fixed and hardly vary with contract volume. Therefore, in small concessions, transaction costs may account for 17% to 40% of the subsidy savings initiated by competitive tendering. As contract volumes increase, the relative share of transaction costs decreases, but in large concessions (with an average annual subsidy of €54.3 million) the share still amounts to 4% to 10%. Although the information provided to us cannot statistically be validated, based on a 'back-of-the-envelope calculation' we conclude there are strong indicators that transaction costs linked to competitive tendered award mechanisms may be considerable compared with the subsidy-saving potential of this mechanism.[46]

46 We expect the transaction costs of CT to increase in the future as the trend is the Netherlands is towards longer-term, higher-volume contracts, as we showed in previous sections.

3.3.3 Other potential problems associated with CT

In Section 2.5.2 we described five potential problems associated with CT: complex contracts, asset ownership, award criteria, bidding parity, and collusion. We assess these potential problems for the Netherlands.

In the Netherlands the PT services to be procured are complex in nature, as they are area-based, contain a great number of inter-dependent issues (such as the complex relationship between operations and infrastructure), and are in the long run liable to unpredictable changes. As services are complex, so are the contracts. Due to this complexity there are indicators for opportunistic bidding and operator-led contract renegotiation, which may lead to operators speculating on flaws and loopholes in the contract (the danger of moral hazard).

Ownership of the infrastructure and other assets is often mentioned as a potential problem for CT (Hensher and Stanley, 2008 and Chapter 2). For the Netherlands, however, this is less of an issue as ownership of most strategic PT infrastructure (terminals, stations, shelters, and traffic and information systems) is in public hands. Depots, repair and maintenance shops, and bus stop signs are owned by operators or by municipalities and rented by operators, but the information provided to us gives no indicators that transfers of these assets from the incumbent to new operators would raise problems.[47]

In the Netherlands the award procedure leaves room for subjectivity and ambiguity, as authorities use a mix of qualitative and quantitative criteria to assess the bids. Problems associated with award criteria and procedures are the main causes of legal disputes. Especially after 2010, it is a rule rather than an exception that the intention to grant a concession is lawfully disputed by the non-awarded bidders arguing that the bid assessment was not conducted properly. It is plausible that the bidding costs (see Table 3.7) are a main trigger for these lawsuits. As the costs of starting a lawsuit are relatively low compared to the bid costs, and the potential reward (the profit from the contract) is high, it is not surprising that operators make such attempts in court.

Following our informants, bidding imparity is, no problem in the Netherlands. PTAs encourage as much as possible a level playing field by supplying all the necessary information on the current operations, state of the infrastructure, fare box revenues, qualities, and numbers of personnel, and so on to all potential bidders.[48]

Finally, in other countries, such as France, problems have been reported regarding colluding bidders.[49] There is no evidence of such a behaviour in the Netherlands, although the possibility cannot be excluded.

47 The transfer from the incumbent to a new operator causes transaction costs, as parties have to assess the state of these assets, negotiate, and stipulate the agreement in a contract. The associated costs are not known to us.

48 Direct personnel (e.g. drivers) are lawfully protected, as regulation requires direct staff to be hired by the contract winner on the same employment terms.

49 Yvrande-Billon (2006) indicates collusion was prominent in the French PT procurement mechanism.

3.4 Conclusion

In this chapter we indicated the trends in the demand, supply and governance of the bus, tram and metro segment (BTM) in PT in the Netherlands. We conclude that, over time, the share of BTM has slightly decreased and the supply and the cost recovery ratio have increased. As regards the market situation and governance, after 2000 (the implementation of CT), we identify a clear concentration tendency on the authority as well the firm side: the number of authorities, area concessions and firms in the market has decreased. After competitive tendering was imposed on the sector, the dominant player, Connexxion, lost market share, especially to Arriva and Veolia. Descriptive statistics on competitive tendering in the Netherlands are shown in more detail in Table 3.5.

For the period 2001–2015 the annual average ex ante transaction costs of competitive tendering are roughly estimated to have been 14.4 million euros to 35.5 million euros.[50] We assume that the transaction costs of the operators are included in the bid price, so we may expect that without competitive tendering the PTA expenditures on PT operations may be considerably lower. In small concessions the transaction costs may account for 17% to 40% of the subsidy savings initiated by competitive tendering. In large concessions (with an average annual subsidy of €54.3 million) the share of transaction costs in the subsidy savings is 4% to 10%. Although the information on transaction costs provided to us cannot be validated statistically, based on our estimates, we conclude that there are strong indicators that the transaction costs linked to competitively tendered award mechanisms may be considerable compared with the subsidy-saving potential of this mechanism.[51]

In Section 2.5.2 we described five potential problems associated with CT: complex contracts, asset ownership, award criteria, bidding parity and collusion. In this chapter we assessed these potential problems. In the Netherlands the PT services to be procured are complex in nature, as they are area-based, contain a great number of inter-dependent issues (such as the complex relationship between operations and infrastructure) and in the long run are liable to unpredictable changes. As the services are complex, so are the contracts. Due to this complexity there are indicators for opportunistic bidding and operator-led contract renegotiation, which may lead to operators speculating on flaws and loopholes in the contract (the danger of moral hazard). Ownership of the infrastructure and other assets is not a problem, as most of the strategic PT infrastructure (terminals, stations, shelters and traffic and information systems) is in public hands (see also Hensher and Stanley, 2008 and Chapter 2). The award procedure leaves room for subjectivity and ambiguity, as authorities use a mix of qualitative and quantitative criteria to assess the bids. The problems associated with award criteria and procedures are the main causes of legal disputes. It is plausible that the

50 The 2015 price level.

51 We expect the transaction costs of CT to increase in the future as the trend in the Netherlands is towards longer-term, higher-volume contracts.

bidding costs (see Table 3.7) are the trigger for these lawsuits, as the costs of starting a lawsuit are relatively low compared with the bid costs and the potential reward (the earnings from the contract) is high. Finally, bidding imparity is not an issue in the Netherlands. PTAs provide a level playing field by supplying all the necessary information on the current operations, state of the infrastructure, fare box revenues, qualities, number of personnel, and so on to all potential bidders.[52]

52 Direct personnel (e.g. drivers) are lawfully protected, as regulation requires direct staff to be hired by the contract winner on the same employment terms.

Chapter

4

Drivers of Customer Satisfaction with Public Transport Services

This chapter is based on Mouwen, A., Drivers of customer satisfaction with public transport, Transportation Research Part A, 78, 2015:1-20

4.1 Introduction

In previous chapters PT is seen from an authority and firm perspective. In this and the following chapter the focus is shifted to the perspective of the customers; the PT passengers. In this chapter we study the drivers of customer satisfaction as understanding of these drivers is essential for the assessment of the effects of competitive tendering (CT) on satisfaction we perform in Chapter 5.

Especially in urban settings, public transport (PT) plays an important role. The share of PT usage in the modal split in the EU-19 overall is only approximately 17%,[53] but in urban areas, the share of PT is considerably higher: up to 50% (European Platform on Mobility Management, 2014). Congestion in the EU in most cases is located in and around urban areas and costs nearly 100 billion euros annually, or 1% of the EU's GDP (European Commission, 2013). Moreover, urban mobility accounts for 40% of all CO_2 emissions of road transport and up to 70% of other pollutants from transport (European Commission, 2014). Public transport can contribute to solving these problems. A recent study (Replogle and Fulton, 2014), for instance, calculates a 40% reduction in urban passenger transport emissions by 2050 if expanding the use of public transportation, walking, and cycling in cities.

In transportation studies, the attractiveness of PT is primarily considered by concentrating on the technical aspects of the PT services (for a synthesis see Currie and Wallis, 2008); the view of the customer is often neglected. In economics and marketing, however, the management of consumer services is a topic that is widely studied (see for instance Bougie et al., 2003; Curasi and Kennedy, 2002; Zeithaml et al., 1990). Few authors bridge the gap between measuring the technical quality of PT service provision and measuring passengers' perception of PT quality. Eboli and Mazulla (2011) are among the few to take both these perspectives into consideration. A challenge for a customer-based PT policy is that the PT service perception of non-PT users may be biased. These users may perceive PT as mainly homogeneous and not tailored to their individual needs compared with, for example, bicycles and cars. People who are acquainted with PT are better able to perceive and evaluate the quality and heterogeneity of the actual PT supply.

A specific aspect of PT services in most Western countries is that they are provided by operators acting under a contract with a public transport authority (PTA). In many cases, these services are offered by private or semi-public contractors operating in almost monopolistic conditions. Under such circumstances, PT operators are not bound by competitive market requirements in meeting passengers' needs. Operators may be inclined to focus on the needs of the PTA instead of the needs of the passengers. However, it is in the interest of customers, authorities, and operators that PT systems are more focused on passengers' needs. PTAs can enhance customers' orientation of operators by imposing on them (or negotiating with

53 In the EU-15, in 2011, passenger cars had a share measured in passenger kilometres of 83%, buses and coaches 8.4%, train 7.2%, and tram and metro 1.4% (European Commission, 2013).

them) quality-based incentive payment schemes (see e.g. Hensher and Hougton, 2004). The question for PTAs is how their steering mechanisms may be adjusted to bring PT services more in line with different user preferences and at the same time warrant the best possible value for (subsidy) money.

This chapter aims to contribute to the knowledge of how customer and trip characteristics moderate the satisfaction of PT users in the light of the steering mechanisms and governance of PTAs acting in a tendering environment. These topics are considered from the perspective of the PTA acting in the role of the representative of the customer/passenger. Under the conditions of putting a PT contract out to tender, the authority sets the rules of the game; therefore, PTAs have to acquire knowledge concerning which incentives that they impose on the operators will enhance the customer satisfaction and which will not.

In this chapter, we model overall satisfaction as a function of user satisfaction with 15 attributes of PT services. We also model the moderating effect of customer characteristics on attribute-level satisfaction. In this way, we are able to study the aspects that matter to PT customers and demonstrate that the factors that are important to customers differ significantly across socio-demographic and trip characteristics. We use a data set containing the satisfaction scores and background characteristics of PT in the Netherlands for the period 2010–2011.

This chapter contributes to the existing literature on PT quality assessment in three ways. First, it highlights the complex interaction between satisfaction and negative social safety experiences (NSSEs). Second, our study takes both the level of satisfaction and the composition of satisfaction into consideration. Third, due to the rich data set used (180,000 observations), we are able to attain robust results for a variety of PT modes (bus, tram, metro, regional train) and several urban contexts.

We structure this chapter as follows. In Section 4.2, we provide a theoretical framework and elaborate on the central concepts that build satisfaction. Based on this theory, we postulate a number of generic hypotheses. In Section 4.3, we derive three models that describe the relationships between satisfaction with PT and characteristics of the customer and of the PT trip. Section 4.4 describes the results obtained from applying the models to a sample of Dutch PT users for the period 2010–2011 and tests the hypotheses. We conclude in Section 4.5 with a summary of the main findings, implications for policy, and recommendations for further research.

4.2 Conceptual model and hypotheses

In this section, based on a literature review, we describe the concepts that influence customers' satisfaction and behaviour in PT and visualize the relations between these concepts. As the customer is the main research topic of this chapter, we restrict the description and visualization of the conceptual model to the customer/passengers side. Based on the theoretical discussion, in this section we also formulate some generic hypotheses concerning the relationship between customer and trip characteristics and the level and composition of satisfaction.

4.2.1 Consumer setting: expectations, perceptions and behaviour

From the perspective of customers, a service relation is a value-based relation, i.e. the relationship is a function of the costs and benefits that accrue from the relationship. Customers' assessments of a service depend on the balance between sacrifices and benefits, both monetary and non-monetary. Sacrifices and benefits are moderated by customers' tastes and characteristics. The needs of individual customers (passengers) are significant determinants of the level of satisfaction the customers' experience (Bryant and Cha 1996) and of repurchase intentions (Mittal and Kamakura 2001). These needs differ among individual customers; they are heterogeneous. Many authors in transportation research point to this heterogeneity in passengers' perceptions of the different aspects of the service provided (Andreassen, 1995; Baltes, 2003; Burton et al., 2003; Tyrinopoulos and Antoniou, 2008; Eboli and Mazulla, 2009; Brons and Rietveld, 2009; Dell'Olio et al., 2010; Diana, 2012).

When focusing on the relationship between satisfaction and behaviour, the above-mentioned studies strongly suggest that satisfaction is a determinant of behaviour, but many also suggest that 'there must be more'. For example, it has been observed that not all satisfied customers return for a repurchase and not all dissatisfied customers disappear (Jones and Sasser, 1995; Cronin, 2003). Numerous concepts that are related to or influence behaviour have emerged and "it still remains unclear if we really understand all fundamental issues relative to the conceptualization and measurement of service quality" (Cronin, 2003, p.332). In the course of time, many concepts that aim at better predicting customer behaviour have been added to the basic model. We consider the inclusion of the concepts of attitudes, involvement, emotions and loyalty to be among the most valuable contributions in improving the theoretical framework. In the context of this chapter, there is no room to elaborate on these concepts; the interested reader is referred to Anable (2005), Lai and Chen (2011), Sánchez-Pérez, et al. (2007a), Pantouvakis and Lymperopoulos (2008), Curasi and Kennedy (2002) and Lee et al. (2001).

Amongst researchers and practitioners, also in the field of transportation studies, agreement is widespread that overall satisfaction results from disconfirmation between expectations and perceptions (Oliver, 1980: Fornell 1992) and that perceived service performance and satisfaction drive purchase intentions and behaviour (LaBarbera and Mazursky, 1983; Joewono and Kubota, 2007; Nathanail, 2008; Sánchez-Pérez, et al., 2007a). An interesting aspect concerning the satisfaction/behaviour relationship is non-linearity; negative experiences (losses) leading to dissatisfaction are perceived to be of greater weight than gains of an equal amount (Andreassen, 1995). The impact of poor performance may carry a greater consequence than the benefits of performance excellence (Cronin, 2003). This understanding is especially useful when analysing the effect of negative social safety experiences (NSSEs). Concerning NSSEs, we postulate that people who have had an NSSE will be more dissatisfied with safety at stops and on-board the vehicle compared to people who have not had an NSSE. Also, we would expect that having an NSSE moderates the weight of the attributes 'safety at stops' and 'safety on board' in a positive way; i.e. people who have had an NSSE – relative to people who have not – put more weight on these attributes. We thus formulate the following hypothesis:

> H1: Negative social safety experiences negatively moderate the level of satisfaction of the attributes safety on board and safety at stops (H1a) and positively moderate the importance of these attributes (H1b).

4.2.2 Transaction-Specific Satisfaction and Overall Satisfaction

A distinction must be made between customers' satisfaction with respect to specific transactions or service encounters and customers' global or overall evaluation of a service (Gustafsson and Johnson, 2004). As this is also a valuable distinction in the context of PT services, we elaborate on this. Transaction-specific satisfaction (TSS) is the result of a cognitive judgment of transactional service encounters (Andreassen, 1995; Cronin, 2003; Dell'Olio et al., 2010; Lai and Chen, 2011).[54] In contrast, overall satisfaction is an affective/emotional response to a perceived discrepancy between expectations and perceptions. Overall satisfaction is a more holistic affective construct after a service delivery experience, whereas transaction-specific satisfaction refers to transaction-specific (attribute-based) cognitive evaluation of service encounters. Consequently, satisfaction with individual attributes pertains to an ex post evaluation of services and acts as an antecedent for overall satisfaction (Dell'Olio et al., 2010, 2011).[55] This is also the way we treat the relationship between transaction-specific satisfaction and overall satisfaction in the PT sector.

Several empirical studies imply that consumers' perceptions of service attributes (e.g. waits) are more likely to be the basis for consumers' attribute experience and behaviour than objectively measured attribute service levels (Cameron et al., 2003; Zakay, 1989). In addition, several studies provide evidence that passengers are poor estimators of actual attribute values (e.g. waiting time duration). In our study, we focus on perceived rather than actual attribute service performance.[56]

Because TSS, consisting of attribute-based evaluations of the service, is an antecedent of overall satisfaction, we assume, as many authors do (see, amongst others, Anderson et al., 2008; Eboli and Mazulla, 2009; Gustafsson and Johnson, 2004; Brons and Rietveld, 2009) that

54 Perceived service performance (see Figure 1) may be equated to transaction-specific satisfaction.

55 For this reason, concerning the design of a consumer satisfaction survey (CSS), it is strongly advised to place the question on overall satisfaction *after* the question on satisfaction with attributes. Sánchez-Pérez, et al. (2007a, 2007b) show in the context of PT satisfaction research a significant difference in passengers' OS evaluations before and after passengers reflect on the importance of service attributes.

56 An additional practical argument is that reliable nationwide data on actual PT-performance are not available.

the evaluations of the individual attributes reflect the weight of these attributes in overall satisfaction. The weights of the service attributes correspond to the relative importance customers attach to the individual attributes and define the composition of satisfaction. We postulate the following two hypotheses:

> H2: Customer and trip characteristics moderate the level of overall satisfaction (H2a) as well as the level of service attribute satisfaction (H2b);
> H3: Customer and trip characteristics moderate the composition of satisfaction (importance of service attributes).

4.2.3 Customer characteristics: need for segmentation

For policy and managerial reasons it is efficient to focus resources or stimuli on specific groups of (potential) customers and herewith focus on those who are most likely to change behaviour. Essentially, segmentation is simply the act of defining meaningful sub-groups of individuals. It is about reducing the number of entities being dealt with into a manageable number of groups that are mutually exclusive and share the same well-defined characteristics (Anable, 2005). The goal of segmentation is to make it possible to target specific marketing measures or policies to specific groups and to predict the response to these stimuli.

A customer's assessment of value depends on sacrifice (i.e. the monetary and non-monetary costs associated with use of the service) and the customer's frame of reference (Zeithaml et al., 1990). Thus, differences in customers' assessment of service value are likely to be due to differences in monetary and non-monetary costs, customers' attitudes towards the service, prior experiences, situational circumstances, and socio-demographic and behavioural characteristics of the customers (Anderson et al., 2008; Dell'Olio et al., 2010; De Oña et al., 2012; Tyrinopoulos and Antoniou, 2008; Diana, 2012).

Two fundamentally different means of segmentation are described by Anable (2005), i.e. a priori and post hoc. A priori segmentation relates to selection of groups from a population in advance based on known characteristics such as sex, age, car use, et cetera. In contrast, post hoc segmentation uses statistical analyses to identify segments (see, for instance, De Oña et al., 2012). In the latter approach, respondents are clustered according to their multivariate profiles and thus the segments are determined by the data, not by the researcher. A potential drawback of the latter method is that the resulting segments are 'meaningless' and hard to interpret, so making it difficult to derive managerial or policy implications. In the empirical part of our study, we use the a priori method to segment the population.

The above-mentioned relationships between the concepts are visualized in the conceptual model shown in Figure 4.1. Please note that in the structural model developed in Section 4.3, we consider only the relations that are shown with the solid lines and highlighted text boxes.

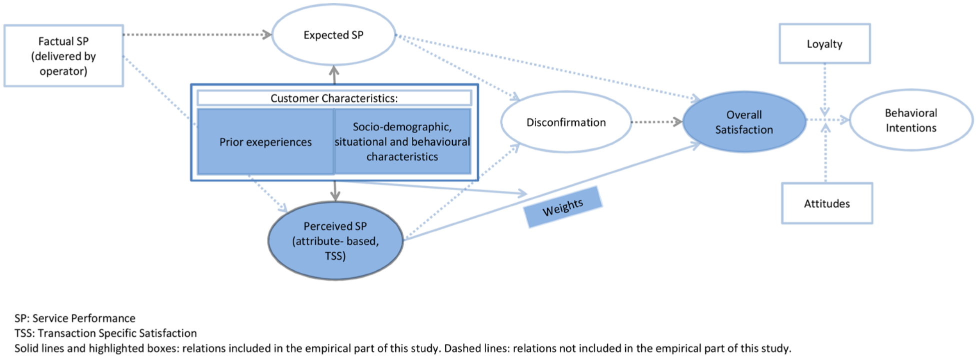


Figure 4.1. *Conceptual model: determinants of service performance (SP) with PT services.*

4.2.4 Summary of relevant literature

In Table 4.1, we summarize the main characteristics of twelve quantitative studies that are closely related to our study. This is useful in positioning our study in relation to similar studies. A comprehensive summary of these studies is reported in Appendix 4A.

Most studies consider either the level of satisfaction (S) or the composition of satisfaction (importance of attributes, I) as their research topic. Our study examines both parameters. Concerning the sample size, it is apparent that this is dependent on the survey method used. When stated preference (SP) techniques are used, the number of respondents is relatively small. A bigger sample is needed for revealed preference (RP) surveys. The sample size for RP surveys varies between 388 and 4,123 respondents. There is one outlier: the study by Brons and Rietveld (2009). Our study pools observations for 2 years and contains 180,000 observations. The dominant survey method used is revealed preference. Only 2 studies use stated preference methods. Unlike most of the other studies, our study encompasses all PT modes. It is conspicuous that few studies refer to satisfaction with PT by train. If service attributes are pooled or clustered, a priori methods (factor analysis and principal component analysis) are the dominant methods used. We do not cluster attributes. Few studies consider negative social safety experiences (NSSEs) as a determinant of service satisfaction; our study constitutes an exception. Finally, Table 4.1 shows that half of the studies described here, as well as our study, take the interaction between transaction-specific satisfaction and customer characteristics into consideration; the other half only reports direct effects.

Hoofdstuk 4

Table 4.1. *Summary of previous studies on satisfaction with PT*

Study	Importance (I) of and/or satisfaction (S) with service attributes	Sample size	PT mode	Survey method	Method used to cluster service attributes	NSSE in scope?	Interaction between TSS and customer characteristics
Andreassen (1995)	S	600	Tram, bus, train	RP	Expert opinion	No	Yes, only with urban context and freq. of use
Brons and Rietveld (2009)	Both	350,000	Train	RP	PCA	No	No
De Oña et al. (2012)	I	858	Bus	RP	Classification tree (post hoc)	No	No
Dell'Olio et al. (2010)	I	768	Bus	RP	No clustering	No	Yes
Dell'Olio et al (2011)	I	36	Bus	Focus groups and SP	No clustering	No	Yes
Diana (2012)	S	4,123	Bus, car (multimodal)	RP	No clustering	No	No
Friman et al. (2001a)	S	997	Bus, tram	RP	PCA	Yes	Yes, only with NCI
Friman et al. (2001b)	S	95	Bus, tram	SP	FA	Yes	Yes, only with NCI
Lai and Chen (2011)	n.a.	763	Metro	RP	FA	No	No
Pantouvakis and Lymperopoulos (2008)	Both	388	Ferry	RP	FA	No	No
Sánchez-Pérez et al. (2007a)	I	1,000	Bus	RP	No clustering	No	No
Tyrinopoulos and Antoniou (2008)	I	1,372	Bus, trolley, tram, metro	RP	FA	No	Yes, only with gender
This study	Both	180,000	Bus, regional train, tram, metro	RP	No clustering	Yes	Yes

FA: factor analysis (a priori method)

PCA: principal component analysis (a priori method)

RP: revealed preference

SP: stated preference

4.3 Model estimation

The relationships between attribute-based satisfaction, overall satisfaction and customer characteristics are studied by means of constructing three linear models. In the first model, overall satisfaction is the dependent variable and attribute-based satisfaction combined with customer characteristics are the explanatory variables, leading to importance scores or weights of the attributes for several groups of customers. These weights make up the composition of satisfaction. The second and third models are used to study the ways in which the level of satisfaction is directly subject to customer characteristics.

Model parameter estimations may be based on several techniques; for instance, normalized pair-wise estimation, principal components regression or multiple regression (see Gustafsson and Johnson, 2004). We use multiple regression to study the interaction between overall satisfaction and a number of customer characteristics. The dependent variables of the models are the measured satisfaction scores, scaled 1 to 10. We estimate the coefficients applying ordinary least squares (OLS).

4.3.1 Service attributes and customer characteristics

The specific transactions a passenger encounters are associated with specific PT service attributes, e.g. walking to a bus stop, waiting for the bus, or entering the vehicle. In our models, we distinguish 15 PT service attributes. We group the service attributes in core attributes, which display 'what' is delivered, and peripheral attributes, which display 'how' the service is delivered (see, similarly, Iacobucci et al., 1994; Anderson et al., 2008; Lee et al., 2001). Within peripheral attributes, we distinguish interactional and physical attributes (Table 4.2).

The satisfaction scores of these variables can take a range from 1 (very dissatisfied) to 10 (very satisfied). As can be seen in Appendix 4.B, the correlations between the values of satisfaction for different attributes are significant at $p < .01$ and as expected, all correlations are positive. The majority of the correlations lie around 0.2, indicating a small positive correlation between the evaluations of attribute satisfaction. The level of satisfaction with the attributes *personnel behaviour* and *driver's behaviour*, between *on-time performance* and *travel speed* and between *safety on board* and *safety at stops* are mildly correlated. We choose for presenting the model results on the level of individual attributes, and not clustering them into factors using a factor analysis as the latter would diminish the explanatory power of the model. Moreover, factors determined by using statistical methods, are hard to interpret and operationalize in real world.

The division of the data into age groups is based on a commonly used rule of thumb in PT research in the Netherlands where the legal driving age is 18 years old (below that age, all PT users are captive) and the retirement age for most individuals is around 65 years. The group limits of 27 and 40 years are chosen to represent respondents starting in the labour market and experienced employees respectively. The division into urban classes is based on quartiles, dividing the 70 areas we distinguish in the database into four classes, each class containing an equal number of areas.

Table 4.2. *Definition of service attributes*

	NAME	CLARIFICATION
	Core attributes	
1	On-time performance	Accuracy of the realized departure times in relation to the schedule.
2	Travel speed	Appreciation of travel speed and time.
3	Service frequency	Number of departure opportunities per hour.
4	Prices of the tickets	Price of various types of ticket and season cards.
	Peripheral interactional attributes	
5	Personnel behaviour	Behaviour of the several types of personnel (e.g. drivers, station guards) when dealing with passengers.
6	Driver's behaviour	Driving performance of the driver.
	Peripheral physical attributes	
7	On-board information on delays	On-board information provision (static, dynamic, vocal) on delays.
8	Ticket-selling network	Ease of obtaining a ticket from on- and off-board selling points.
9	Information provision at stops	Information available for passengers at terminals and stops (static, dynamic).
10	Safety at stops	Safety at terminals and stops as perceived by passengers when waiting.
11	Vehicle tidiness	Level of cleanliness of the vehicle in general.
12	Ease of boarding and alighting	Ease of boarding and alighting from the vehicle.
13	Seating capacity	Chance of getting a seat.
14	On-board noise	Level of noise in the vehicle.
15	Safety on board	Sense of safety during the trip.

4.3.2 Interaction between composition of satisfaction and characteristics of customers

There are two main ways of determining attribute importance. The first is to ask customers directly how important service attributes are to them, for instance using scale ratings. The alternative – the way we used – derives the importance of attributes indirectly via regression estimations.

In general terms, the function between overall satisfaction (OS), transaction-specific satisfaction (TSS) and customer characteristics can be expressed as:

$$OS_i = f(TSS_i, C_i) \tag{4.1}$$

where OS_i is the overall satisfaction of individual *i*, TSS_i is a vector of attribute satisfactions of individual *i* and C_i is a vector of characteristics of individual *i*.

More specifically, if customer *i* experiences satisfaction with *N* different service attributes and has *M* different characteristics, the model can be written as in Eq. 4.2, where u_i is a random disturbance term:

$$OS_i = f(TSS_{1i}, TSS_{2i} \dots\dots\dots\dots, TSS_{Ni}, C_{1i}, C_{2i} \dots\dots\dots\dots, C_{Mi}) + u_i \tag{4.2}$$

We will assume a linear relationship:

$$OS_i = a_0 + \sum_{j=1}^{N} b_j \, TSS_{ji} + \sum_{g=1}^{M} c_g \, C_{gi} + u_i \tag{4.3}$$

We will report b_j of model 4.3 as a standard estimate (no interaction). As we are primarily interested in the interactional effects between TSS_{ji} and C_{gi} for individual *i*, we add to Eq. 4.3 the interactions between 1 to *N* attribute satisfactions and 1 to *M* characteristics:

$$OS_i = a_0 + \sum_{j=1}^{N} b_j \, TSS_{ji} + \sum_{g=1}^{M} c_g \, C_{gi} + \sum_{j=1}^{N} \sum_{g=1}^{M} d_{jg} \, TSS_{ji} C_{gi} + u_i \tag{4.4}$$

The value of the coefficient d_{jg} informs on the effect of a characteristic of the customer on the importance of each service attribute for the customer; b_j denotes the importance of the service attributes of the reference group; c_g denotes the contribution of the characteristics to the constant term of the equation.

Interaction is restricted to a first-level interaction: we only model the interaction of attribute satisfaction with each of the eight individual characteristics. In this way, the effects of the characteristics are not mutually dependent.

4.3.3 *Interaction between level of satisfaction and characteristics of customers*

The second type of model – see Equations (4.5) to (4.7) – relates the level of overall satisfaction (OS) and transaction-specific satisfaction (TSS) to customer characteristics. Thus:

$$OS_i = f(C_i) \qquad\qquad TSS_i = f(C_i) \tag{4.5}$$

In linear form:

$$OS_i = g_0 + \sum_{g=1}^{M} r_g \, C_{gi} + v_i \tag{4.6}$$

$$TSS_i = h_0 + \sum_{g=1}^{M} s_g \, C_{gi} + w_i \tag{4.7}$$

where OS_i denotes overall satisfaction of individual *i*; TSS_i denotes transaction-specific satisfaction and, r_g and s_g are the coefficients to be estimated. We develop separate models for each service attribute. As we have 15 different service attributes we use 15 separate models to estimate the coefficients that represent the interaction between the level of satisfaction and customer characteristics.

4.4 Dataset and empirical results

4.4.1 Context

Dutch PT acts under highly regulated market conditions. The characteristics of the PT market structure in the Netherlands are such that this market can be described as only sligthly competitive. Due to the system of tendering imposed upon the sector as of the year 2001, PT supply in a given geographically delineated area is the prerogative of a single supplier.[57] We term the conditions in the Netherlands after 2000 as a temporary monopoly (Van de Velde and Pruijmboom, 2003). Under these conditions, providers have hardly any incentives to compete for customers' satisfaction and providers are therefore not inclined to provide more heterogeneous services of their own volition. By means of setting the terms of requirements in a tendering procedure, PTAs may aquire a more heterogeneous service supply.

4.4.2 Dataset

We used a dataset containing the satisfaction scores of urban and regional Dutch PT users (bus, tram, metro and regional train) for the years 2010 and 2011. A random nationwide sample of approximately 90,000 passengers yearly was taken by means of a written questionnaire handed out to the passengers in the vehicles. Data collection was stratified for region, workday/weekend and peak/off peak. Satisfaction scores (scaled 1–10) were collected for the satisfaction with the total trip (overall satisfaction) and for the satisfaction with several service attributes (see Table 4.3). In addition, background characteristics of the respondents are asked for, such as gender, age, trip frequency, car use, and trip motive. As the location where the passengers are questioned was also known, we were able to add a location tag to the dataset (see Table 4.4).

4.4.3 Descriptive statistics

In Table 4.3, basic descriptive statistics concerning attribute satisfaction are given and Table 4.4 shows the profile of respondents. As can be seen from Table 4.3, the average level of overall satisfaction in the period studied is 7.28. Positive outliers are satisfaction with the attributes *seating capacity*, *ease of boarding and alighting* and *safety on board*. Negative outliers are satisfaction with *on board information on delays* and *prices of the tickets*. As may be expected, satisfaction evaluation is not normally distributed.

57 See Mouwen and Rietveld (2013) for the relationship between tendering and satisfaction with PT.

Table 4.3. *Basic descriptive statistics on satisfaction*

		N	Mean	Std. Deviation	Skewness	Kurtosis
	Core attributes					
1	On-time performance	164,131	7.12	2.231	-.897	.432
2	Travel speed	163,414	7.34	1.847	-1.014	1.470
3	Service frequency	160,129	6.75	2.190	-.751	.243
4	Prices of the tickets	136,063	4.96	2.504	.055	-.737
	Interactional attributes					
5	Personnel behaviour	150,444	7.30	1.816	-.894	1.310
6	Driver's behaviour	161,771	7.18	1.708	-.886	1.538
	Physical attributes					
7	On-board information on delays	140,238	5.00	2.619	-.045	-.996
8	Ticket-selling network	140,222	7.19	2.368	-.817	.121
9	Information provision on stops	158,721	7.14	2.134	-.890	.614
10	Safety at stops	155,606	7.69	1.611	-.819	1.623
11	Vehicle tidiness	168,081	6.88	1.827	-.657	.696
12	Ease of boarding and alighting	166,967	8.45	1.645	-1.464	3.028
13	Seating capacity	168,343	8.45	2.256	-1.605	1.967
14	On-board noise	165,804	6.25	1.972	-.505	.123
15	Safety on board	154,959	8.03	1.481	-.964	2.265
	Overall satisfaction	165,594	7.28	1.384	-.862	2.613

4.4.4 Results: standard estimates, hypotheses and interaction effects

Introduction

In this section we describe the findings based on estimations of the models (4.3), (4.4), (4.6) and (4.7). In the left-hand panel of Table 4.5 estimates of the attribute weights when no interaction effects are included are presented (model (4.3)). The right-hand panel shows the effects of interactions on the attribute weights in model (4.4). For example, the figure 19% refers to d_{2_5} in model (4.4), showing the interaction between service attribute 2 (travel speed) and characteristic 5 (NSSEs). With respect to the interaction effects, we focus on the main findings. We restrict the description of the interaction effects to a selection of characteristics (i.e. age, PT mode, grade of urbanization and NSSEs). Full results are available on request.

Table 4.4. *Profile of survey respondents*

Gender	Percent	Car availability	Percent	NSSE	Percent
Panel A					
Men	43.4	yes	34.7	no	93.5
Age[a]	Percent	PT-mode	Percent	Urban class (# inh/sq. km build on area)	Percent
Panel B					
<18	17.0	bus	52.2	low urbanized (15–25)	18.3
18–27	44.9	tram	7.7	medium (25–35)	24.9
28–40	12.8	metro	7.6	high (33–44)	18.6
41–64	20.5	regional train	31.6	very high (44–70)	38.2
>65	4.7	ferry	0.9		
Number of PT trips past week[b]	Percent	Trip motive	Percent		
Panel C					
<=1	23.1	live	33.7		
2-3	24.6	work	15.6		
>3	52.3	study	21.3		
		other	29.4		

[a] *Mean age 29.6 year (sd. 15.9)* [b] *Mean trip frequency 3.5 trips per week (sd. 1.9).*

We defined the reference group as follows: younger than 18 years of age, questioned on a trip in a rural environment, travelling by bus and not having had NSSEs. The interaction results of target groups are all to be interpreted relative to the reference group. To gain a better understanding of the findings, we transformed the interaction coefficients into percentages indicating the relative weight or importance score for each attribute/segment combination.

We describe the results on two levels: (1) by the attributes most influenced (potentially) by PTAs; (2) by customer characteristic. The first level of analysis is prompted by the perspective we take in this chapter. We focus on the topic of satisfaction seen in the light of the steering mechanisms of PTAs in a tendering environment. Some attributes are largely in the control of PTAs, whereas others are not. PTAs in a tendering environment can exert the greatest influence on attributes associated with *service frequency* and vehicle attributes such as *on-board noise* and *on-board information on delays*. These attributes can be prescribed in the terms of reference (TOR) in a tendering procedure, so the PTA is the decision-making body in this respect. At the other end of the spectrum are service attributes in which many parties (operators, road authorities, other road users such as cars and bicycles, central government, et cetera) are involved. PTAs have only little control over these attributes. To this category we assign the attributes: *ticket selling network*, *prices of the tickets, on-time performance and travel speed*. The remainder of the service attributes we assume to be in between; to achieve goals in relation to these aspects, the PTA is very dependent on the actions and performance of the operator.

Standard estimates

The results without taking interaction into account (left-hand panel of Table 4.5) show that the attributes *travel speed, on-time performance* and *service frequency* (core attributes) are the most important determinants of overall satisfaction. This outcome is also found in many other studies (see Pantouvakis and Lymperopoulos, 2008 for Greek ferries; Anderson et al., 2008 for US airlines; Brons and Rietveld, 2009 for Dutch long-distance rail).[58] What is remarkable, however, is that the interactional attributes *personnel behaviour* and *driver's behaviour* and the physical attribute *vehicle tidiness* are also considered very important. Concerning the latter, Lai and Chen (2011) found cleanliness having a significant influnce on passengers behavioural intentions. Bordagaray et al. (2013) found driver's kindness to be especially important for individuals between 35 and 55 years old, and for frequent PT users. Tyrinopoulos and Antoniou (2008) found that vehicle cleanliness is of great importance for PT users; driver behaviour however is in their study not rated among the most important attributes.

58 Pantouvakis and Lymperopoulos (2008) and Anderson et al. (2008) relate their finding to the time customers spend in a technical environment. Time spent in a technical environment might be connected to the importance customers attach to physical attributes. The relatively high weights passengers attach to physical attributes for trips using ferries, airlines or trains may be explained by the relatively long time spent in the vehicle.

Table 4.5. *Attribute weights for a selection of characteristics without (left hand panel), and with interaction (models 4.3 and 4.4)*[a].

		Standard estimate, no interaction		Interaction effects								
				Reference group	Age		Mode		NSSE	Urban density		
		b	%	<18, bus, no NSSE, low urban density (%)	28-40 vs. <18 (%)	>65 vs. <18 (%)	metro vs. bus (%)	regional train vs. bus (%)	Had versus had not an NSSE (%)	Medium vs. low (%)	High vs. low (%)	Very high vs. low (%)
	Core attributes											
1	On-time performance	0.10**	11**	4**	4	8*	6	7*	6*	7*	10*	5
2	Travel speed	0.14**	15**	15**	21*	17	15	18*	19*	13	13	16
3	Service frequency	0.10**	11**	12**	13	16*	8	11	13	12	13	14
4	Prices of the tickets	0.03**	4**	3*	3	7*	7*	2	3	2	2	0*
	Interaction attributes											
5	Personnel behaviour	0.08**	9**	11**	14*	14	6	11	11	11	12	9
6	Driver's behaviour	0.08**	8**	14**	12	16	7*	11	12	12	9*	14
	Physical attributes											
7	On-board information on delays	0.04**	4**	4**	4	2	8*	7*	2	3	3	3
8	Ticket-selling network	0.03**	4**	3*	1	-3*	2	0*	3	2	5	5
9	Information provision on stops	0.04**	4**	4**	7*	7	4	4	1*	4	6	4
10	Safety at stops	0.01**	1**	3	-4*	-2	-1	4	5	7*	9*	10*
11	Vehicle tidiness	0.08**	8**	9**	7	4	12	7	9	6	6	9
12	Ease of boarding and alighting	0.04**	5**	7**	6	5	8	9	3*	6	3*	5
13	Seating capacity	0.04**	5**	4**	4	2	5*	6	5	3	4	4
14	On-board noise	0.04**	5**	3	2	0	5	4	6*	6*	4	4
15	Safety on board	0.06**	6**	4	5	6	6	0	1*	5	2	-1*

*** $p < .01$, *$p < .05$, other: $p \geq .05$ (two-tailed)*

[a] *Significance for the standard estimates and reference group defined relative to zero. Significance for the interaction effects defined relative to the reference group.*

Table 4.6. *Effects of a selection of characteristics on the level of overall satisfaction, and attribute satisfaction (model 4.6 and 4.7).*

		Age		**Mode**		**NSSE**		**Urban density**	
		28–40 vs. <18	>65 vs. <18	Metro vs. bus	Regional train vs. bus	NSSE vs. none	Medium vs. low	High vs. low	Very high vs. low
	Overall satisfaction	0.094*	0.763*	-0.292*	-0.427*	-0.416*	-0.071*	0.021	0.037
	Core attributes								
1	On-time performance	0.303*	0.893*	-0.171*	-0.224*	-0.477*	-0.334*	-0.324*	-0.035
2	Travel speed	0.164*	0.795*	0.046	-0.377*	-0.421*	-0.038	0.025	0.116*
3	Service frequency	0.189*	1.071*	-0.177*	-0.462*	-0.375*	0.104*	0.239*	0.558*
4	Prices of the tickets	0.555*	2.071*	-0.612*	-0.475*	-0.474*	0.000	0.079	-0.106*
	Interaction attributes								
5	Personnel behaviour	0.135*	0.802*	-0.8*	-0.45*	-0.496*	-0.216*	-0.195*	-0.364*
6	Driver's behaviour	0.089*	0.615*	-0.165*	-0.13*	-0.394*	-0.163*	-0.037	-0.207*
	Physical attributes								
7	On-board information on delays	-0.113*	0.668*	1.21*	1.719*	-0.371*	0.135*	0.186*	0.359*
8	Ticket-selling network	-0.224*	0.196*	-0.287*	-0.048*	-0.634*	0.063*	-0.046*	-0.272*
9	Information provision on stops	0.078*	0.605*	-0.079	-0.058	-0.42*	-0.03	0.023	0.043
10	Safety at stops	0.022	0.441*	-0.508*	-0.264*	-0.759*	-0.082*	-0.081*	-0.13*
11	Vehicle tidiness	0.14*	0.754*	-0.785*	-0.576*	-0.469*	-0.033	0.014	-0.101*
12	Ease of boarding and alighting	-0.181*	-0.254*	0.217*	0.09*	-0.455*	0.097*	0.079*	0.022
13	Seating capacity	0.157*	0.597*	-0.517*	-0.607*	-0.45*	-0.068*	0.023	-0.261*
14	On-board noise	0.282*	0.767*	-0.022	-0.375*	-0.297*	-0.272*	0.051	-0.009
15	Safety on board	0.074*	0.437*	-0.563*	-0.239*	-0.735*	-0.043	-0.013	-0.088*

$*p < .05$, other: $p \geq .05$ (two-tailed)

The results in Table 4.6 represent the outcomes of models (4.6) and (4.7) and indicate how the level of satisfaction is moderated by customer characteristics. The table shows that many user and trip characteristics have a significant effect on overall satisfaction and on attribute satisfaction.

Testing the hypotheses

Using the results in Tables 4.5 and 4.6, we test the hypotheses postulated in section 4.2.

H.1: Negative social safety experiences negatively moderate the level of satisfaction of the attributes safety on-board and safety at stops (H.1a) and positively moderate the importance of these attributes (H.1b).

Hypothesis 1a is supported by our data. Table 4.6 shows that having an NSSE has a significant negative impact on the satisfaction with the attributes *safety on-board* and *safety at stops*. However, hypothesis 1b is not supported by our data. Table 4.5 shows that having an NSSE significantly moderates the relative weight passengers attach to the attribute *safety on board*, but the impact of the change is not in line with our hypothesis (weight decreases compared to the reference group). Moreover, no significant influence on the relative weight of the attribute *safety at stops* can be observed.

H2: Customer and trip characteristics moderate the level of overall satisfaction (H.2a) as well as service attribute satisfaction (H.2b).

H.2a: Table 4.6 shows that all customer and trip characteristics – except in highly and very highly urbanized areas – are significantly associated with the level of overall satisfaction.

H.2b: Table 4.6 also shows that the majority of customer and trip characteristics are significantly associated with the level of satisfaction with the attributes; 98 of 120 interactional coefficients are significant. We may conclude that hypothesis 2 is supported by our data.

H.3: Customer and trip characteristics moderate the composition of satisfaction (importance of service attributes).

The right-hand panel of Table 4.5 shows that this hypothesis is not supported by our data for most customer and trip characteristics. The differences in the weights of the service attributes relative to the reference group can only in a few cases significantly be attributed to differences in customer characteristics.

Influence of age on composition and level of satisfaction

Concentrating on the three attributes within the control of PTAs, it is conspicuous that elderly people attach significantly more weight to service frequency, implying that a policy aimed at increasing service frequency in general mainly leads to a shift in attribute importance for elderly people (Table 4.5). As service frequency is linked to waiting time, this shift in importance may be correlated to the declining physical condition of elderly people. Table 4.5 also shows that elderly PT users place more emphasis on the attributes *price*, *on-time performance* and *service frequency*, and less emphasis on *ticket-selling network* than young PT users (reference group). Compared to youngsters under 18, medium age passengers (28-65) attach considerable more importance to personnel, and drivers behaviour. Probably this is an international sociological phenomenon, caused by increasing infirmity, and is also found in Bordagaray et al. (2013). It is remarkable, Dell' Olio et al. (2010) find a contrary result. Our results may imply that it is not so much that age moderates importance directly, but rather that the influence is to a certain extent exerted indirectly via differences in travel patterns between the age groups. In the Dutch situation, elderly people tend to travel more in off-peak periods when seat availability is not a problem and they also travel more on seasonal tickets than young people do. Therefore, they may attach less weight to the probability of getting a seat and to the quality of the ticket-selling network.

Concerning the level of satisfaction, it is very likely that policies aimed at the introduction of new buses and raising service frequency will lead to more satisfied elderly people (Table 4.6). Table 4.6 also shows that the level of satisfaction (overall and attribute-level) rises with age. This holds for all but one attribute. In particular, the satisfaction with the attributes *price* and *trip frequency* of elderly PT users relative to young PT users is remarkable. The high satisfaction of elderly people with the attribute ***price*** may be related to the reduced fares for elderly people. The negative satisfaction score of elderly people compared to youngsters in relation to the attribute *ease of boarding and alighting* is probably related to the declining physical condition of elderly people. Our findings are in line with those of other empirical studies (Mittal and Kamakura, 2001; Anderson et al., 2008; Bryant and Cha, 1996).[59]

Influence of mode choice within PT on composition and level of satisfaction

Mode choice within PT in most circumstances is not a free choice because no alternatives are offered. However, analysing differences in users' satisfaction with metro and regional trains compared to bus services, and the importance accorded the different modes of transport,

59 The sample reveals that 50% of the respondents older than 65 years of age have a car at their disposal, which could be used in place of the PT trip. For PT users younger than 18 years of age this is only 16%. Therefore, elderly people are less constrained (captive) in terms of using PT than youngsters. However, we have not found any indicator in our data that car availability moderates the level or the composition of satisfaction. This may be due to the lack of clarity in the questionnaire regarding car availability. Unfortunately, we are therefore unable to test the assumption that elderly users of PT are more satisfied than youngsters because of the more voluntary nature of their PT use.

Hoofdstuk 4

offers useful information for the governance of PT systems. PT mode choice has significant impacts on the relative importance of the attribute *on-board information on delays*. It may therefore be expected that a PTA policy of replacing bus lines with metro lines will lead to a radical change in attribute importance (Table 4.5). Moreover such a policy would probably lead to an increase in the level of satisfaction of passengers in relation to the attributes *on-board information on delays* and less satisfaction with *on-time performance* and *service frequency* (Table 4.6).

We observe that PT users travelling by metro place more emphasis on *on-board information* and *price* and less emphasis on *driver's behaviour* than users of buses (Table 4.5). Concerning regional trains, PT users who travel by regional trains place more emphasis on *on-board information, on-time performance* and *speed* and less emphasis on the *ticket-selling network* than users of buses (reference group). For both modes, the high importance consumers attach to ***on**-board information* is remarkable. This high weight may be connected to the lack of visual orientating points outside the vehicle due to the high speed and/or underground routes. Consumers need on-board information systems to orientate themselves on the metro more than they do on a surface mode such as buses.

In accordance with the findings of Bordagaray et al. (2013) that PT users are perceiving heterogeneity in actual supply and quality of bus lines, in our study we found a strong effect of PT mode use on satisfaction: users of rail services (especially users of metros and regional trains) are *overall* less satisfied compared to users of bus services (Table 6). The attributes that negatively differentiate these users compared to those of buses are in particular *cleanliness*, *personnel behaviour, price* and *reliability*. Metro and regional train passengers are relatively very satisfied with *on-board information* and *ease of boarding and alighting the vehicle*. What is striking is that passengers even evaluate *on-time performance* and *service frequency* in metro services as more dissatisfying compared to bus services, whereas in general metros and regional trains outperform buses with respect to these service attributes. These results may indicate that passengers' satisfaction is moderated by their expectations; they expect metros to run on schedule and frequently and compare these expectations with perceived performance. These findings differ from the outcomes of several transportation studies indicating the existence of a 'rail bonus' (see, for instance, Axhausen et al., 2001). It is probable that these findings are also related to specific circumstances in the Netherlands during the study period. The period 2010–2011 was characterized by the decentralization of responsibility for regional trains from central government to regional PTAs. Regional PTAs were obliged to put these lines out to tender. The first round of bidding for regional train services did not go smoothly. In a number of cases, the new operators encountered difficulties in acquiring the specified rail vehicles and had many hurdles to overcome before services ran at the desired quality. These start-up problems may have had an effect on passengers' service evaluations.

Influence of negative critical incidents on composition and level of satisfaction

Negative critical incidents may play an important role in service evaluation (Friman et al., 2001a, 2001b; Voorhees et al., 2009). Negative experiences have more impact on service evaluation (satisfaction) than positive experiences. Negative experiences, for instance, elicit negative word-of-mouth, whereas positive experiences tend not to be shared with others (Bougie et al., 2003).

Our analysis includes NSSEs, i.e. encounters with, for instance, theft, robbery, harassment, aggression and bodily harm in the PT setting (at stops and on board the vehicle).[60] Our results indicate that introducing new vehicles or increasing the frequency of service provision will probably not have a major impact on the relative importance or the satisfaction of passengers who have had NSSEs. The moderating effect these incidents have on the attribute weights is depicted in Table 4.5. At first sight, it is a counter-intuitive outcome in that passengers who have had an incident with social safety weight *safety during the trip* significantly lower than passengers who have not encountered such an incident, but this may be an indicator that these incidents mainly occur at stops and not in the vehicle. We note that PT users who have had an NSSE place more emphasis on *speed* and *on-time performance* (core attributes) and less emphasis on *information provision at stops* and *ease of boarding and alighting* than PT users who have not experienced a social safety incident.

Table 4.6 shows that NSSEs have a significant negative effect on *overall satisfaction* and on attribute-level satisfaction (the regression coefficients for all 15 attributes are statistically significant). Friman et al. (2001a) found a similar strong effect of NSSE on satisfaction. These are remarkable results: passengers who have experienced one or more incidents related to safety not only rate satisfaction as relatively low for the attributes that are most likely to be related directly to such incidents (i.e. *safety during the trip, safety an stops*), but also attributes that seem to have no relationship with the incident itself.[61] It might well be that negative incidents impact heavily on attitudes towards PT in general. People who had a bad experience blame the whole PT system for not having 'protected' them. Thus, their focus in the objective evaluation of attribute performance is blurred.

Influence of urban environment on composition and level of satisfaction

The final segmentation variable analysed concerns level of urbanization, based on the location in which the consumer survey was administered. This urbanization measure is not the residential location or destination of the passenger, but represents the spatial environment of the trip.

60 Grade of urbanization and NSSEs are positively correlated (see Table 4.C2 in Appendix 4.C); the chance of being involved in an unpleasant social safety situation increases as the grade of urbanization increases.

61 Friman et al. (2001a) found in a sample of Göteborg (Sweden) PT users that the frequency of negative experiences exerts an impact on overall satisfaction via attribute-specific satisfaction.

Table 4.5 indicates the importance of the attribute *safety at stops*, which is strongly moderated by urban density. In urban environments, the chance of being confronted with an NSSE at stops is considerably higher than in rural environments (see Table 4.C2 in Appendix 4.C). So the importance that PT customers attach to the attribute *social safety* might be highly correlated to negative experiences with social safety. Another notable result is that PT users in highly urbanized environments place more emphasis on the attribute *safety at stops* and less emphasis on *safety on board* and *price of the tickets* than PT users in less urbanized areas.

Table 4.6 shows that for the segment 'very highly urbanized versus low urbanized' the satisfaction related to eight of 11 significant service attributes is lower in very highly urbanized environments compared to low urbanized environments. The three attributes for which the opposite holds true have a positive magnitude that over-compensates the negative impact. This concerns the attributes *service frequency, travel speed* and *on-board information on delays*. Also, the relatively low satisfaction for the attribute *personnel behaviour* in very highly urbanized areas compared to low urbanized areas is striking. Furthermore, we note two clear significant trends: (1) the level of satisfaction with the attributes *ticket-selling network* and *safety at stops* clearly decreases as the level of urbanization increases, whereas (2) the opposite holds for the attributes *on-board information on delays* and *service frequency*. These effects may indicate that people in smaller cities are more satisfied with PT than people in metropolitan areas. This may be caused by differences in actual service performance in dense urban areas compared to low urbanized areas.[62]

When evaluating the impact a PTA policy of introducing new vehicles and/or raising service frequency may have on differences in the perceptions of importance and satisfaction in relation to urbanization, we may conclude in general that the relative importance of attributes are scarcely likely to be influenced by such a policy. However, it is probable that increasing the frequency of service and the introduction of new vehicles will have a major positive effect on the level of satisfaction of passengers using PT in a highly urbanized area.

4.5 Conclusion and recommendations

4.5.1 Conclusion

We found that heterogeneity in the perception of PT (both the level and the composition of satisfaction) significantly depends on the customer characteristics, situational conditions (e.g. urban setting), and experiences such as NSSEs. Overall, the attributes *on-time performance, travel speed*, and *service frequency* are seen by PT users as the most important, followed by *personnel/driver behaviour* and *vehicle tidiness*. A generic policy aimed at achieving these attributes may yield favourable results with respect to satisfaction.

62 In Table 4.C1 in Appendix 4.C we show that the PT service level in low urbanized areas is indeed a lot worse compared to highly urbanized areas.

Further, we demonstrated the influence of specific customer characteristics on satisfaction and attribute weights and placed the results in the context of steering mechanisms of PTAs in a tendering environment. We took as a starting point the fact that PTAs in a tender can exert the greatest influence on the attributes associated with *service frequency* and vehicle attributes such as *on-board noise* and *on-board information on delays*. A policy aimed at increasing the service frequency and putting new vehicles into operation is likely to lead to more satisfied older people (>65), passengers travelling by regional train, and people living in dense urban areas. Concerning the attribute weights, specifically older people (>65), people travelling by metro or regional train, people who have experienced an NSEE, and people in medium-densely populated areas attach relatively great importance to these attributes.

Our theoretical framework assumes that actual customer behaviour is not only triggered by service supply characteristics, but also by concepts, such as attitudes, involvement, emotions, loyalty, and satisfaction. We concentrated the empirical part of this study on satisfaction only, thereby abstracting from the other above-mentioned constructs that drive behaviour. Within the concept of satisfaction, we indicated that a distinction has to be made between the level and the composition of satisfaction, and we consider the combination of these two as one of the determinants of behaviour. Our empirical results, however, do not support the hypothesis that the composition of satisfaction between user groups differs significantly.

From a practical viewpoint, the methodology provides a relevant contribution to the previous studies, as it stresses the impact of NSSEs, urban settings, and PT mode use. The method is easy to handle for PTAs and operators alike.

4.5.2 Policy recommendations

The performance and quality of PT systems should correspond as much as possible to the diverse demands of (potential) PT customers. It is therefore important for authorities and operators to acquire knowledge concerning the preferences and evaluations of customers. Authorities and operators may increase their efficiency by focusing their resources and strategies on retaining their existing customers rather than trying to attract new customers (Dell'Ollio et al., 2011). In this respect, it may be wise for them to look for the most cost-efficient measures and aim specific instruments and policies at specific target groups. For instance, we found that trip frequency and speed impact heavily on satisfaction. When considering several possible policies, PTAs should take into account the problem that a policy aimed at increasing the service frequency to raise satisfaction is possibly not the most efficient, as it may be very costly. A more cost-efficient measure may be to increase the travel speed, as such a measure latches onto both satisfaction and operational efficiency. Therefore, we recommend that PTAs should perform supplementary calculations of the costs associated with the measures. In the case that the investment costs of measures are known, the results of this study make it possible to develop a cost/effect matrix, an instrument that indicates the expected increase in satisfaction from investing a monetary unit in a certain measure, and to identify which target group will presumably profit most in terms of a satisfaction increase.

We further recommend that PTAs should take notice of the following two observations. First, PT services are considered by passengers as intermediary services and thus the value they attach to these services is a derivate of the value of the activity at the final destination. This implies that PTAs should design their policy measures in accordance with knowledge concerning the trip motives of customers. Second, the majority of PT users are 'captive customers', constrained to use the PT service because they have no private alternatives, such as cars, at their disposal. Dissatisfied captive passengers have no transport alternatives and no way of expressing their dissatisfaction with these services other than complaining. Both observations imply that it may well be the case that PT patronage is rather inelastic with regard to satisfaction. In addition, PTAs intending to use satisfaction evaluations as part of incentive payment schemes towards providers should complement these schemes with objective measurements of performance, such as reliability of operations and/or actual numbers of passengers transported.

4.5.3 Discussion and recommendations for further research

The database we used for this study has a number of limitations. First of all, like many other studies (see Appendix 4.A), this study is based on evaluations of existing PT users only. This implies a lack of generalizability of the outcomes for non PT users. More satisfaction research should be undertaken on non-captive users and non-PT users and specifically on the relationship between car availability and PT service satisfaction as the ultimate objective of PTAs is to stimulate the shift to more sustainable modes.

In this study we took the perspective of the steering possibilities of PTAs in a tendering environment. We have touched upon a number of drivers for satisfaction with PT, but it is probable that other drivers of interest for PTAs wishing to place incentives on operators' performance remain to be revealed. In particular, we recommend more research on (1) actual duration of travel time elements and (2) thresholds, keeping in mind that the drivers of satisfaction may be either more general or case specific.

Pantouvakis and Lymperopoulos (2008) show that differences in the importance of PT service attributes may be connected to the time customers spend in the different environments of the PT system (technical environment versus interpersonal environment). The more time spent in a certain surrounding, the greater the chance of being affected by this setting. In-depth analyses of travel time proportions (in-vehicle time, walking, waiting) related to satisfaction may be helpful.

Experience-based satisfaction/behavioural thresholds are most probably present in continuous services such as PT (see Bolton and Drew, 1991; Mittal and Kamakura, 2001; Van Hagen, 2011). These thresholds imply (dis)satisfaction arises beyond a certain tolerance level because a minimum service level is expected. These thresholds are probably not universal but are moderated by customer characteristics and so will vary between user segments. More knowledge of these thresholds is essential for PTAs to enable them to develop an incentive-based bonus/penalty arrangement regarding PT services to be delivered by operators.

Our study supports the findings of many other studies also showing that in the PT sector segmentation matters; however, we shed little light on why it matters. We recommend further research in the PT sector focusing on the distinction between general and situational circumstances that exert an influence on satisfaction. General circumstances can be defined as those which are universally valid, as opposed to situational (case-specific) circumstances that influence satisfaction only under specific conditions and/or in an indirect way. Age, for instance, is an exponent of a general determinant. International literature (Bryant and Cha, 1996; Anderson et al., 2008) strongly suggests that elderly people react differently to the same stimuli, whether concerning tranportation or having a haircut, than youngsters with respect to satisfaction evaluation.

Urbanization is an example of a situational explanatory variable. We find in our sample that urbanization is an important determinant of satisfaction and attribute-level importance, especially influential in terms of the relative importance of the attributes *on-time performance* and *safety at stops*. It is most likely that actual service performance plays a major explanatory role, but it may also be the case that attitudes towards PT differ between city-dwellers and those from villages. More research efforts should be directed towards these possible determinants of satisfaction with PT.

Appendix 4.A. Review of relevant literature

Study	Importance and/ or Satisfaction of service attributes	Origin of sample	Sample size	PT-mode	Study object	Survey method
Andreassen (1995)	S	Norway	600	Tram, Bus, Train	PT-users	RP
Brons and Rietveld (2009	Both	Netherlands	350,000	Train	PT-users	RP
De Oña et al. (2012)	I	Spain	858	Bus	PT-users	RP
Dell'Olio et al. (2010)	I	Spain	768	Bus	PT-users	RP
Dell'Olio et al (2011)	I	Spain	36	Bus	PT-users and potential PT-users	Focus groups and SP
Diana (2012)	S	Italy	4,123	Bus, car (multimodal)	PT-users	RP
Friman et al. (2001a)	S	Sweden	997	Bus, Tram	PT-users	RP
Friman et al. (2001b)	S	Sweden	95	Bus, tram	PT-users	SP
Lai and Chen (2011)	n.a.	Taiwan	763	Metro	PT-users	RP
Pantouvakis and Lymperopoulos (2008	Both	Greece	388	Ferry	PT-users	RP

Segmentation variables used	Method used to cluster service attributes	Functional form of model	NCI in scope?	Interaction between TSS and customer characteristics
Freq. of use Urban context	Expert opinion	LISREL (max. likelihood)	No	Yes, only with urban context and freq. of use
Freq. of use Age Car avail.	PCA	Lineair regression	No	No
Gender Age Occupancy Travel behaviour	Classification Tree (post hoc)	CART, (classification regression tree)	No	No
Freq. of use Gender Age Income Car avail.	No clustering	Ordered probit	No	Yes
Freq. of use Gender Age Income Car avail.	No clustering	Multinominal logit	No	Yes
Freq. of use Urban context Attitude	No clustering	Multi variate correspondence analysis	No	No
None	PCA	LISREL (max. likelihood)	Yes	Yes, only with NCI
None	FA	LISREL (max. likelihood)	Yes	Yes, only with NCI
Involvement	FA	SEM (max. likelihood)	No	No
Freq. of use (proxy for loyalty)	FA	Linear regression	No	No

(continued on next page)

Study	Importance and/ or Satisfaction of service attributes	Origin of sample	Sample size	PT-mode	Study object	Survey method
Sánchez-Pérez et al. (2007a)	I	Spain	1,000	Bus	PT-users	RP
Tyrinopoulos and Antoniou (2008)	I	Greece	1,372	Bus, Trolley, Tram, Metro	PT-users	RP
This study	Both	Netherlands	180,000	Bus, Regional train, Tram, Metro	PT-users	RP

FA: Factor analysis (a priori method).

PCA: Principal component Analysis (a priori method).

RP: Revealed preference.

SP: Stated preference.

Segmentation variables used	Method used to cluster service attributes	Functional form of model	NCI in scope?	Interaction between TSS and customer characteristics
Freq. of use Gender Age Income Education Occupancy	No clustering	Ordered logit	No	No
Gender	FA	Ordered logit	No	Yes (only with gender)
Freq. of use Gender Age Trip motive Car avail PT mode NCI Urban context	No clustering	Linear regression	Yes	Yes

Appendix 4.B. Correlations between service attribute satisfaction

		Seating capacity	Vehicle tidiness	Personnel behaviour	Driver's behaviour	On-board noise	Ease of boarding and alighting	Information provision on stops
Seating capacity	Pearson Correlation	1.00	.292**	.300**	.270**	.174**	.294**	.200**
	Sig. (2-tailed)		0.00	0.00	0.00	0.00	0.00	0.00
	N	168,343	167,385	149,749	161,063	165,076	166,212	158,008
Vehicle tidiness	Pearson Correlation	.292**	1.00	.431**	.382**	.364**	.256**	.255**
	Sig. (2-tailed)	0.00		0.00	0.00	0.00	0.00	0.00
	N	167,385	168,081	149,624	160,958	165,016	166,032	157,864
Personnel behaviour	Pearson Correlation	.300**	.431**	1.00	.474**	.286**	.282**	.293**
	Sig. (2-tailed)	0.00	0.00		0.00	0.00	0.00	0.00
	N	149,749	149,624	150,444	145,067	147,728	148,615	142,563
Driver's behaviour	Pearson Correlation	.270**	.382**	.474**	1.00	.380**	.301**	.275**
	Sig. (2-tailed)	0.00	0.00	0.00		0.00	0.00	0.00
	N	161,063	160,958	145,067	161,771	159,223	159,954	152,429
On-board noise	Pearson Correlation	.174**	.364**	.286**	.380**	1.00	.226**	.223**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00		0.00	0.00
	N	165,076	165,016	147,728	159,223	165,804	163,939	156,008
Ease of boarding and alighting	Pearson Correlation	.294**	.256**	.282**	.301**	.226**	1.00	.334**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00		0.00
	N	166,212	166,032	148,615	159,954	163,939	166,967	157,068
Information provision on stops	Pearson Correlation	.200**	.255**	.293**	.275**	.223**	.334**	1.00
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	
	N	158,008	157,864	142,563	152,429	156,008	157,068	158,721
On-board information on delays	Pearson Correlation	.085**	.213**	.240**	.255**	.230**	.128**	.373**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	139,578	139,504	128,557	135,314	137,940	138,738	134,621

On-board information on delays	Ticket-selling network	Prices of the tickets	On-time performance	Travel speed	Service frequency	Safety on board	Safety at stops	Overall satisfaction
.085**	.212**	.175**	.254**	.245**	.178**	.238**	.181**	.361**
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
139,578	139,549	135,337	163,363	162,657	159,382	153,955	154,560	164,816
.213**	.219**	.251**	.274**	.289**	.234**	.314**	.238**	.456**
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
139,504	139,409	135,215	163,212	162,530	159,249	153,881	154,471	164,688
.240**	.289**	.272**	.321**	.328**	.250**	.326**	.245**	.486**
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
128,557	127,128	123,014	146,432	146,013	143,342	137,813	138,128	147,530
.255**	.242**	.222**	.327**	.365**	.252**	.331**	.244**	.478**
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
135,314	134,699	130,667	157,364	157,029	153,769	148,401	148,802	158,846
.230**	.169**	.251**	.226**	.282**	.237**	.214**	.165**	.381**
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
137,940	137,674	133,487	161,173	160,572	157,294	152,036	152,549	162,608
.128**	.275**	.125**	.266**	.289**	.217**	.326**	.275**	.364**
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
138,738	138,638	134,371	162,244	161,585	158,285	152,951	153,512	163,636
.373**	.303**	.232**	.355**	.319**	.302**	.251**	.233**	.423**
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
134,621	133,779	129,573	154,997	154,070	151,423	145,603	146,120	155,791
1.00	.241**	.257**	.324**	.256**	.274**	.138**	.122**	.356**
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
140,238	120,620	116,104	137,371	137,058	135,109	129,664	129,599	137,966

(continued on next page)

		Seating capacity	Vehicle tidiness	Person-nel be-haviour	Driver's beha-viour	On-board noise	Ease of boarding and alighting	Informa-tion provision on stops
Ticket-selling network	Pearson Correlation	.212**	.219**	.289**	.242**	.169**	.275**	.303**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	139,549	139,409	127,128	134,699	137,674	138,638	133,779
Prices of the tickets	Pearson Correlation	.175**	.251**	.272**	.222**	.251**	.125**	.232**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	135,337	135,215	123,014	130,667	133,487	134,371	129,573
On-time performance	Pearson Correlation	.254**	.274**	.321**	.327**	.226**	.266**	.355**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	163,363	163,212	146,432	157,364	161,173	162,244	154,997
Travel speed	Pearson Correlation	.245**	.289**	.328**	.365**	.282**	.289**	.319**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	162,657	162,530	146,013	157,029	160,572	161,585	154,070
Service frequency	Pearson Correlation	.178**	.234**	.250**	.252**	.237**	.217**	.302**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	159,382	159,249	143,342	153,769	157,294	158,285	151,423
Safety on board	Pearson Correlation	.238**	.314**	.326**	.331**	.214**	.326**	.251**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	153,955	153,881	137,813	148,401	152,036	152,951	145,603
Safety at stops	Pearson Correlation	.181**	.238**	.245**	.244**	.165**	.275**	.233**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	154,560	154,471	138,128	148,802	152,549	153,512	146,120
Overall satisfaction	Pearson Correlation	.361**	.456**	.486**	.478**	.381**	.364**	.423**
	Sig. (2-tailed)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	N	164,816	164,688	147,530	158,846	162,608	163,636	155,791

*** Correlation significant at $p < .01$ level (2-tailed).*

On-board information on delays	Ticket-selling network	Prices of the tickets	On-time performance	Travel speed	Service frequency	Safety on board	Safety at stops	Overall satisfaction
.241**	1.00	.306**	.256**	.272**	.233**	.262**	.234**	.366**
0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00
120,620	140,222	125,209	136,947	136,338	134,050	128,321	128,649	137,701
.257**	.306**	1.00	.268**	.264**	.255**	.189**	.174**	.368**
0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00
116,104	125,209	136,063	133,146	132,397	130,235	124,393	124,634	133,565
.324**	.256**	.268**	1.00	.459**	.334**	.268**	.204**	.520**
0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00
137,371	136,947	133,146	164,131	159,487	156,318	150,633	151,171	161,206
.256**	.272**	.264**	.459**	1.00	.417**	.305**	.233**	.560**
0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00
137,058	136,338	132,397	159,487	163,414	156,153	150,310	150,726	160,874
.274**	.233**	.255**	.334**	.417**	1.00	.225**	.214**	.488**
0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00
135,109	134,050	130,235	156,318	156,153	160,129	147,440	147,915	157,636
.138**	.262**	.189**	.268**	.305**	.225**	1.00	.554**	.398**
0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00
129,664	128,321	124,393	150,633	150,310	147,440	154,959	149,307	151,962
.122**	.234**	.174**	.204**	.233**	.214**	.554**	1.00	.312**
0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00
129,599	128,649	124,634	151,171	150,726	147,915	149,307	155,606	152,463
.356**	.366**	.368**	.520**	.560**	.488**	.398**	.312**	1.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
137,966	137,701	133,565	161,206	160,874	157,636	151,962	152,463	165,594

Hoofdstuk 4

Appendix 4.C. PT proximity and social-safety incidents

Table 4.C1. *Influence of urban density on PT proximity.* **Source: CBS Statline**

Density	Population per hectare built-on area	Distance to nearest train station (km)
Low	19.9	10.9
Medium	28.2	5.9
High	36.7	5.5
Very high	53.7	4.1

Note: correlation between population and distance: -0.8461.

Table 4.C2. *Number of times victim of a social-safety incident with public transport*

Density	1 - 4 times	5 times+
Low	18%	17%
Medium	23%	17%
High	18%	21%
Very high	42%	46%
Number	3,875	66

Chapter 5

Does competitive tendering improve customer satisfaction?

This chapter is based on Mouwen and Rietveld, Does competitive tendering improve customer satisfaction with public transport? A case study for the Netherlands, Transportation Research Part A, 51, 2013: 29-45

5.1 Introduction

Reform in the public transport sector is taking place in many countries. One of the aims is to change public transport gradually from production-oriented towards customer-oriented. Service contracts are in most cases the method used to set bilateral conditions between private operators and public authorities. Contracts serve as an instrument to induce private operators in naturally non-competitive markets to act in line with social targets. Corresponding with the aims of the reform, in public transport contracts a shift towards incentive contracts based on quality requirements can be observed (see, e.g., Hensher and Houghton, 2004; Marcucci and Gatta, 2007). With a good definition of service quality and a good measuring method, authorities are attempting to impose strong incentives on operators.

Reform in Dutch public transport takes the form of competitive tendering of concessions. Following the international trend, over the course of time in the Netherlands operators and authorities have tried also to become more and more customer oriented. Inclusion of quality aspects in contracts has become common practice.[63]

The relationship between tendering and efficiency is widely studied (Hensher et al., 2003; Hensher and Houghton, 2004; Van de Velde and Pruijmboom, 2003; Walter, 2009), and is also the subject of Chapter 6. Few studies, however, explicitly focus on the relationship between tendering and satisfaction. The latter is the subject of this chapter that builds on the insight into customer satisfaction drivers of Chapter 4. The analytical results presented here are based on the situation in the Netherlands between 2001 and 2010.

Figure 5.1 provides the broader context of this chapter: travel behaviour of public transport passengers is influenced by their satisfaction with the quality of public transport services. As we saw in Chapter 4, the level of satisfaction depends on a large number of regional and individual factors, and on the institutional settings within which the service supplier is functioning. We pay particular attention to competitive tendering as a possible driving force for service quality enhancements and study the relationship between tendering and satisfaction. The relationship between tendering and the objective or factual performance of public transport in the Netherlands and other influencing factors is only briefly touched upon. In this chapter we pay no attention to changes in travel behaviour due to tendering. These aspects are dealt with in Chapter 6.

63 In many contracts incentives to improve on the quality of travel information, reliability, cleanliness of the vehicles and social security are incorporated.

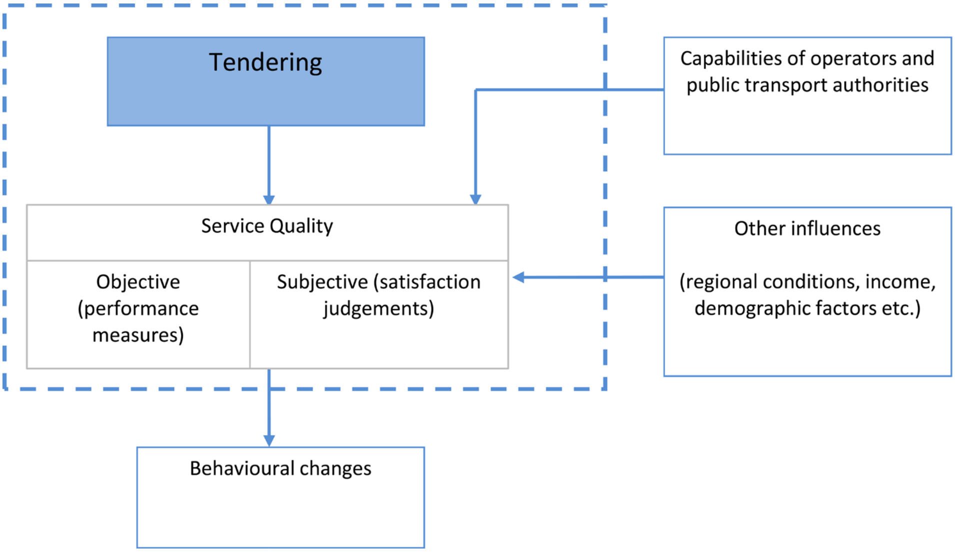


Figure 5.1. *Tendering and passenger satisfaction*

After a short literature review in Section 5.2, the regulative setting in the Netherlands is described in Section 5.3. In Sections 5.4 and 5.5 the research questions and methods are presented. In Section 5.6 the focus is put on quantifying the relationship between tendering and the satisfaction of the total trip. In Section 5.7 these outcomes are studied in more detail by looking at the contribution of the underlying service attributes. Section 8 deals with the central question of this study, i.e. whether tendering is the cause of the observed changes in satisfaction. The chapter is finalized with conclusions.

5.2 Literature on service quality and customer satisfaction

The focus of this chapter is on customer satisfaction and the tendering of public transport concessions. This section provides a review of the literature on the construct of satisfaction with service quality and on the measurement of satisfaction.

5.2.1 Defining service quality

The origin of the definitions of service satisfaction lies in the field of service marketing. Service marketing is a relatively new field of research that combines components from the economic sciences, as well as from psychology and sociology. At the end of the 1980s a debate ensued concerning the definition and dimensions of the concept 'satisfaction'. In that period Zeithaml et al. (1990) developed the SERVQUAL model for measuring service quality. The model could be used as a diagnosis for the shortcomings of service deliverance. The central thesis of the

SERVQUAL model is that service quality can be defined as the difference (gap) between expectations, and perceptions and therefore marketing efforts should be mainly focused on closing this gap. SERVQUAL is still widely used, but its central thesis has faced criticism. The main exponents of this criticism are Cronin and Taylor (1992) and Buttle (1996). According to them, and supported by empirical studies, the central thesis of Zeithaml et al. (that the gap between expectation and perception determines service quality judgments of customers) does not hold. In practice, clients (passengers) are not able to separate expectations and perceptions, which means that, in their judgment, the difference between expectations and perceptions has already been taken into account. Cronin and Taylor (1992) found that humans in their evaluation process do not make an explicit assessment between ex-ante expectations and ex-post perceptions and that these concepts do not originate independently. Expectation and perception are constructs that cannot be measured as independent concepts. Oliver (1980) had earlier defined satisfaction as the non-linear assessment of customers of expectations and experience, in the course of which they include subjective factors such as affection and previous experience. Cronin and Taylor (1992) proposed Oliver's definition of satisfaction as a more adequate construct for customer judgment of service quality than the SERVQUAL definition. The criticism of Cronin and Taylor is widely shared and most authors now agree on taking satisfaction as a measure of customers' quality valuation.

In our opinion expectations and perceptions are in the eyes and minds of customers one construct that precedes satisfaction. Therefore the outcomes of models that are based on the use of both constructs as independent variables should be treated with caution.

5.2.2 Dimensions of service quality

Another theme in the literature concerns the relevant dimensions of service quality. Zeithaml et al. (1990) distinguished five dimensions of service quality upon which the consumers base their expectations, and hence their satisfaction: (1) reliability of service delivery (service outcome); (2) assurance (ability to create trust); (3) empathy; (4) responsiveness; and (5) tangibles (physical environment (setting) of the service). The first dimension is related to the satisfaction of the service delivery itself (outcome). The other four dimensions determine the satisfaction of the service process.

The weights of the dimensions differ according to specific situations and conditions, but several studies have shown that reliability of service delivery is the most important dimension (Strandvik and Liljander, 1994; Iacobucci et al., 1994). Iacobucci et al. also concluded that within reliability of service delivery the perception of the core business is the most important dimension, and that this perception acts as a threshold, a minimal condition for the origination of satisfaction (dissatisfaction). Customers base their overall satisfaction mainly on non-core peripheral attributes (satisfiers). So, if in service delivery these basic conditions are not met, every effort to invest in enhancing other peripheral attributes, will not lead to an increase in satisfaction. In public transport, Van Hagen (2011) showed that, in the heavy rail sector the attributes safety and reliability of the service[64] act as a threshold, and that these attributes may be defined as the core attributes of train service delivery.

Empirical studies also show that demographic characteristics (such as race, gender, marital status, age, and income), experience with the service and environmental factors (weather condition, crowdedness, et cetera) lead to significant differences in service perception and hence satisfaction (Bishop-Gagliano and Hathcote 1994; Anderson et al., 2008). The work of the latter is of particular interest for our study. Anderson et al. showed for the airline sector in the USA that the importance and satisfaction with the core and the peripheral airline service attributes is moderated by customer (demographic) characteristics. They also state that, in models of customer satisfaction, both the service concept and customer characteristics should be incorporated.

5.2.3 Measuring service quality and satisfaction with public transport

Measuring service quality has two dimensions: (1) the objective dimension where service quality can be objectified in performance indicators, such as speed, reliability, and frequency; and (2) the subjective dimension of service quality that only can be measured by means of customer judgments. In this subsection literature on satisfaction surveys in public transport is discussed.

Eboli and Mazzulla (2011) have developed a method in which both subjective and objective measures of transit quality are combined in a single output measure. They state that taking into consideration passenger satisfaction alone can lead to biases, especially when passengers are heterogeneous. On the other hand a specific objective transit performance indicator alone could not be appropriate for evaluating a transit service aspect since the valuation of the passengers is not taken into account. Tyrinopoulos and Antoniou (2008) propose a methodology based on using factor analysis and ordered logit modelling to assess the quality implications of the variability of users' perceived satisfaction across operators. They distinguish several market segments based on demographic variables, several types of operators (rural and metropolitan) and disentangle total service into several service attributes. The output of the importance survey was used as input for the factor analysis. The satisfaction scores were used for ordered logit modelling. The authors found cleanliness and reliability as overall important attributes. Like Tyrinopoulos and Antoniou, Hensher et al. (2003) developed a method that can be used to evaluate the performance among different operators. They start by identifying 13 potentially important service attributes, and then establish a way to measure the relative importance of these attributes. The importance of the attributes is derived by Hensher et al. from a combination of Revealed Preference scores concerning the current trip and the outcomes of Stated Preference choice experiments.

64 In public transport research this is often narrowed to the reliability of the time table.

The above-mentioned authors tested their models empirically in a static situation. In contrast, Friman (2004) examined whether quality improvements have an effect on satisfaction with public transport services and frequency of perceived negative critical incidents. Friman assessed 18 quality improvements by 13 Swedish operators by means of a satisfaction survey. The most important finding of the study is that the satisfaction that passengers experience is only influenced by the quality improvements to a very limited extent. Furthermore, the effect tended to be opposite, in that respondents reported less satisfaction after the implementation of the service improvements. This is an intriguing result that stimulates us to carry out a more or less similar study in a different context, i.e. in the context of regulative reform in the Netherlands.

We study tendering as a possible determinant of satisfaction judgments in public transport. In the literature we came upon a number of possible decisive factors for satisfaction such as the demographic characteristics of the respondents (e.g. age, gender, anxiety, et cetera) and environmental factors (e.g. degree of urbanization, quality of the infrastructure). In this chapter we do not explore these determinants in depth as this was done in Chapter 4, but focus on tendering. Note that in the present chapter we do not investigate the relationship between objective measurable service performance and perceived overall satisfaction since objective data on performance are not collected systematically for all concession areas (with the exception of frequency). Therefore it is not easy to say whether tendering has actually led to improvements in objective service performance.[65]

5.3 Regulatory reform in the Netherlands

In the year 2000 a new transport law came into action that changed the regulatory setting of public transport in the Netherlands drastically. Before that year, the regulative environment can be described as a public-owned monopoly (see also Van de Velde, 1999; Berechman, 1993). In that period the authorities imposed absolute power over the operators and prescribed in detail the services to deliver to them. In rural public bus transport a strict administrative/normative schedule ruled, and, for instance, replacement of the fleet was regulated based on economic parameters such as technical and economic depreciation and not on quality and/or passenger objectives. All deficits were fully covered by the government. In that pre-tendering regime, neither authorities nor operators were explicitly focused on passenger needs. Steering parameters for the authorities were production-/supply-based (scheduled hours). Neither actor considered stimulating improvements aimed at attributes such as on-time performance, travel speed, or service frequency.

With the Transport Law 2000, the Dutch central government imposed the obligation upon regional transport authorities for the competitive tendering of their public transport. According to government, tendering would lead to more efficient and innovation oriented companies,

65 In Chapter 6 we will quantify this relationship in more depth.

and to a better quality of service for the passengers. Several evaluations (Berenschot, 2004; Mu-consult, 2003) showed, however, that the authorities and operators in the first tenders continued with their old habits and mainly steered on non-quality-based, supply parameters. After a period of habituation, authorities overcame the reluctance for more innovative performance-based measures. In later tenders -especially by way of implementing Bonus-Penalty arrangements- a focus on reliability is observed (Rekenkamer, 2009).

During the whole study period, in the tendered regions, compared with the non-tendered regions, supply (service frequencies) increased. Table 5.1 shows that the annual production growth in the tendered regions is higher than in the non-tendered regions.

Table 5.1. *Annual growth of public transport production.**

	concessions tendered	**concessions not tendered**
Relative change in public transport service supply		
2001-2002	n.a	n.a
2002-2003	-2%	-6%
2003-2004	5%	-3%
2004-2005	1%	-3%
2005-2006	10%	2%
2006-2007	25%	3%
2007-2008	33%	1%
2008-2009	7%	2%
2009-2010	n.a	n.a

**It is hard to obtain consistent data for the whole period 2001-2010. The information for the period 2002-2005 is in vehicle hours (source: own calculations based on timetable data), for the period 2006-2009 in vehicle kilometres (source: KPVV, 2011, edited by the authors).*

The afore-mentioned evaluations also show that the authorities tend to use the opportunities of tendering to impose network adaptations aimed at increasing travel speed (e.g. straightening of lines). Actual competition on the Dutch public transport market is modest. During the period 2001-2010, on average 3 bidders per concession area contested for the right to operate (KNV, 2015).[66] A tendency can be observed for the number of competitors in the bidding phase to decrease. The market is divided among three big international conglomerates (Veolia Transport, Keolis and Arriva) [67] and a relatively small Dutch offspring of Netherlands Railways (Q-Buzz). The winner of the tendering procedure is awarded the concession by the authority and, after a period of preparation, will start operations in the awarded region. The year preceding this point in time is treated by us as the ex-ante situation, the year following the

66 This is an open source administered by the joint Public Transport-operators.

67 The former operator Connexxion has been acquired in 2011 by Veolia, but still acts under its own name. Keolis operates in the Netherlands under the name of her subsidiary Syntus.

start of implementation as the ex-post situation. Concessions tend to begin in the month of December. Customer satisfaction surveys are carried out in the month of November.

The forming of concession areas is in practice a dynamic process - neighbouring authorities make arrangements with each other for integrating their areas - so the boundaries of concession areas are not at all constant over the years. On account of these arrangements between authorities, the scale of the concession areas in the Netherlands has increased over time. In the year 2001 the Netherlands was divided into 74 areal concessions. In 2010 this number was reduced to 48 areal and 18 line concessions.

5.4 Data considerations.

As mentioned in Section 5.3, with the regulatory reform of public transport, the Dutch government aimed at increasing efficiency (reduction of costs) and at improvements that would benefit the passengers. We test in this chapter whether the introduction of tendering on the Dutch public transport market leads to changes in passenger satisfaction.

Since the starting point of the reform in 2000, the Dutch government (local, regional and national) supported unified collection of data about public transport passenger satisfaction. Some 90,000 passengers are annually interviewed regarding their perceived satisfaction on a wide range of service attributes. The results of the satisfaction surveys are widely used by both authorities and operators for evaluation, marketing purposes, benchmarks and Bonus-Penalty arrangements.[68] The data collection is based on a stratified sample of public transport trips by bus (both regional and local), tram, metro, and regional train in the Netherlands. Heavy long distance rail users are not interviewed. The data are collected in the November of each year. This is about 10 months after the implementation of a new concession in the pertaining concession areas. The sample per research region is stratified for workday/weekend day, and peak/off peak. For our survey, the unweighted raw data is used.

The survey method used before and after 2004 differs. Before 2004 the survey was commissioned as an oral questionnaire that was administered at stops and terminal points. From 2004 onwards the method changed to a written questionnaire handed out in the vehicles. We use a year dummy to check whether this affects the outcomes.

The main part of the questionnaire consists of questions related to the perceived satisfaction with some 15 quality attributes. Passengers were also asked to give their satisfaction judgment of the total trip. Scores (or marks) are on an interval scale: 1 to 10, where 1 is bad and 10 is excellent. This scaling every Dutch resident knows, since it is used in education at all levels. In addition some background characteristics of the respondents are asked and recorded, such as gender, age, trip frequency, captivity to public transport use, et cetera.

The data set contains regional stratified satisfaction scores on line and area level. Since

68 The survey is commissioned by KPVV, and we are very obliged to KPVV for making the data set 2001-2010 available for this research.

the topic of our study is the relationship between tendering and passenger satisfaction, we enriched the data set with background information on the concessions and the contracts (pre- and post-tendering) originating from KNV, 2015.

The raw data is aggregated in survey regions. The survey regions are not equivalent to concession areas in every case. When necessary we converted the survey region into concession areas. A concession area is defined as a spatial demarcated area containing public transport lines that are operated by one operator. In case a new concession is granted by the authority either the incumbent continues his operations or a new operator takes over. Thus for our study the distinction between tendered and non-tendered ***lines*** within a concession area is not relevant as we concentrate our analysis on areas.

5.5 Research method

A model is formulated to study the relationship between tendering and satisfaction. The model takes into account yearly changes and regional variation in satisfaction scores. The variable ***year*** represents the overall trend applying in the whole country during the year of the data intake. Examples of factors affecting the trend are: the general attitude of people towards social safety; the state of the economy (consumer confidence); or changes in the way the survey data is collected. The variable 'area' corrects for differences between transport areas, for instance, age composition, car ownership and urban density. Note that the area variable also incorporates the overall quality of an operator and of the public transport authority. Further, a ´tendering´ variable is introduced to check whether a new concession possibly leading to a new operator leads to a change in satisfaction. All independent variables are dummy variables.

The model is used with the total trip satisfaction as the dependent variable, as well as for individual service attributes.

If the spatial units (concession areas) were constant all the time, the model can be formulated as:

$$Q_{r,t,i} = b_0 + \sum_{t'} \alpha_{t'} year_{t',i} + \sum_{r'} \beta_{r'} area_{r',i} + \sum_{r',t'} \gamma_{r',t'} \, tendering_{r',t',i} + \varepsilon_{r,t,i} \quad (5.1)$$

where $Q_{r,t,i}$ is the satisfaction of the total trip or service attribute in region *r* and year *t* of individual *i*; b_0, the constant term; α_t , β_r and $\gamma_{r,t}$, are the coefficients of the dummy variables. These are specific constants for, respectively, the influence of the year, the region, and when the services in year *t* were preceded by a tendering procedure; $year_{t',i}$= 1 when *i* is interviewed in t'; =0 else; $area_{r',i}$ = 1 when *i* is interviewed in r'; =0 else;: $tendering_{r',t',i}$ = 1 when tendering took place immediately before year *t'*; 0 = else and $\varepsilon_{r,t,i}$ is the error term. The $\gamma_{r',t'}$ indicates the contribution of tendering in region *r'* in year *t'* to the satisfactions of the passengers affected.

Hoofdstuk 5

The problem is that this formulation is based on the assumption that the area dummy ***r*** applies to the same region during the whole 10-year period. However, since the spatial demarcation changes regularly this is not a valid approach: ***r*** may change meaning several times during the 10 years considered. Therefore, we reformulate the above model for pairs of two years and we take special measures to harmonize the area definitions of the first and the second year in each year pair. As a rule we converted the observations of the ex-ante situation to the area definition of the ex-post situation. In the case of simply merging a number of smaller areas into a bigger area, this can easily be done by linking observations in the first year to the spatial areas in the second year. However, when borders between areas are shifted, their observations also have to be shifted. Since the observations are not only linked to the areas, but also to the public transport lines in the areas, we were able to perform this task. In Appendix 5.A this recode routine is described and visualized.

Thus, after harmonizing the regional codes ***r***, year pairs contain observations based on the regional classification of the second year. We arrive at 9 sets of year pairs (2001/2002, 20002/2003, …, 2009/2010). For each year pair we compare the development of passenger satisfaction in regions where tendering took place with the development of satisfaction in regions where tendering did not take place. Thus we estimate Eq. (5.1) not in one step over the whole 10 year period, but for 9 separate sub-periods and ***t*** has only two possible values each time Eq. (5.1) is estimated (2001 versus 2002 in the first sub-period, etc.).

Estimation of the model took place using the Ordinary Least Square method. To avoid perfect correlation a reference region has to be chosen. In every year pair the reference case is a medium-dense populated area in the southern part of the Netherlands ('de Kempen').[69]

The model building and analysis is based on the comparison per year-pair of regions tendered versus regions non-tendered (in that specific year-pair). So we compare in the next sections the effects on satisfaction of tendered regions relative to non-tendered regions.

5.6 Tendering and total trip satisfaction

In Table 5.2 some aggregated descriptive statistics concerning the satisfaction of the total trip are summed up. The average satisfaction marks in Table 5.2 concern the total of all regions, irrespective of whether they are tendered or not. During these 10 years, 61 concession regions have been tendered for the first time, and 11 regions have been tendered for the second time. The table shows relatively low shares for tendered regions, reflecting that the largest concession areas (the four largest cities) have not been tendered yet. A gradual increase in

69 The 'Kempen' region was chosen as reference region because it represents an 'average' region in the Netherlands consisting of a medium dense city surrounded by rural area. The econometric approach with fixed effects for the concession regions implies that the estimation results for the tendering effects are guaranteed to be not affected by the choice of a specific reference region. The only thing that is affected are the region specific constants βr in (1) but these are not of interest in the present study, reason why they have not been reported in the estimation results. This is a standard result in econometric models with nominal variables represented by dummy variables.

passenger satisfaction of the total trip during this decade is identified.[70] In particular during the first period there were substantial increases.

The aim of our analysis is to find out to what extent the improvement of customer satisfaction can be attributed to the competitive tender procedures applied. In Table 5.3 the main findings of the regression models for the tendering variable are shown. The coefficients are shown for the period 2001-2010 as an aggregated mean of the nine year pairs. In the period 2001-2010, 72 regions are indicated as tendered. The analysis distinguishes between, on the one hand, the total number of tendered regions, and on the other the regions tendered for the first time (61) and those tendered for the second time (11). The reported values of the coefficients can be interpreted as the deviation of the satisfaction scores of the total trip in the regions where tendering took place, compared with regions where tendering did not take place, and controlled for disturbances on the satisfaction scores that may be caused by yearly or regional influences. The findings indicate that, after controlling for year influences and regional-specific conditions, in 42 out of 72 tendered regions, total satisfaction increased after tendering, whereas in the other 30 tendered regions total satisfaction decreased (compared with non-tendered regions).

Table 5.2. *Descriptive statistics total trip satisfaction*

	Number of observations	Average satisfaction total Trip	Std. deviation	No. of regions tendered for the first time	No. of regions tendered for the second time	No. of observations in tendered regions as % of total
2001	68,333	6.84	1.098	0	0	n.a.
2002	70,976	6.70	1.142	8	0	12%
2003	68,222	6.90	1.066	6	0	8%
2004	87,690	7.14	1.460	9	0	10%
2005	82,356	7.14	1.479	14	0	18%
2006	83,524	7.10	1.482	9	2	14%
2007	83,744	7.12	1.498	1	1	4%
2008	83,783	7.25	1.405	6	6	16%
2009	86,821	7.32	1.368	6	2	7%
2010	83,652	7.25	1.399	2	0	2%

70 The data collection method changed in 2004 from orally administered questionnaires to hand-out (written) questionnaires.

Table 5.3 *Descriptive statistics tendering, 2001-2010; total trip satisfaction*

2001_2010	**All tendered regions *(N=72)***	**Regions tendered for the first time *(N=61)***	**Regions tendered for the second time *(N=11)***
No. of regions with positive significant effect of tendering (α ≤ .05).	22	20	2
No. of regions with positive non-significant effect of tendering (α >.05).	20	15	5
No. of regions with negative significant effect of tendering (α ≤ .05).	12	9	3
No. of regions with negative non-significant effect of tendering (α >.05).	18	17	1
Average impact on satisfaction in regions with a positive significant tendering impact (α ≤ .05).	0.320	0.327	0.242
Average impact on satisfaction in regions with a non-significant tendering impact (α ≥ .05).	0.074	0.065	0.103
Average impact on satisfaction in regions with a negative significant tendering impact (α ≤ .05).	-0.268	-0.277	-0.242
Average impact on satisfaction in regions with a negative non-significant tendering impact (α ≥.05).	-0.049	-0.052	-0.001
Average impact on satisfaction in regions with a significant tendering impact (α ≤ .05).	0.112	0.140	-0.048
Average impact on satisfaction in tendered regions.	0.061	0.068	0.025

The results also show that the first round of tendering in a region has a more positive effect on satisfaction than the second round of tendering. This can be seen in the values of the changes of the coefficients. In those regions that underwent a second round of tendering, the average satisfaction increased by 0.025 points (relative to the non-tendered regions). In contrast: in regions where tendering took place only once, an increase of the overall satisfaction of 0.068 points is found. If only the significant cases are taken into consideration, these outcomes hold, and become more pronounced. In a next section we will give in-depth analysis of this finding on second round tendering. As we will see this is not only a matter of sequence (first versus second), but also of timing (early versus late).

To better understand the mechanisms behind satisfaction changes, next, in Section 5.7, an analysis of the service attributes that underlie total service satisfaction is carried out.

5.7 Tendering and satisfaction with service attributes

In the previous section the relationship between satisfaction with the total trip in tendered regions compared with non-tendered regions was dealt with. In this section this relationship is deepened by looking at the service attributes that underlie total trip satisfaction. We are interested in which service attributes contribute the most to the change in satisfaction that was observed in Section 5.6. This is of interest for policy reasons, since only this disaggregated level of satisfaction can be linked to actual measures and actions by authorities and/or operators.

5.7.1 Service attributes

The survey commissioned contains information on passengers satisfaction judgments of 15 service attributes. The selection of the attributes used in the surveys is performed by the commissioner KPVV and is based on an extensive literature survey. The chosen attributes are directly linked to the several stages of public transport service performance. Seen through the eyes of the passengers the attributes can be interpreted as separate transactions that form part of the total service chain. The attributes chosen are common in public transportation research (see among others Stradling et al., 2007; Dell'Olio et al., 2011; De Oña et al., 2012; Diana, 2012; Tyrinopoulos and Antoniou, 2008).

The attributes used for our study are shown in Table 5.4. In Table 5.5 the average satisfaction values of the individual attributes are shown as an average for the whole period 2001-2010. These values are not yet corrected for the weights of the attributes, as will be done in Section 5.7.2. Remarkable are the large variations in average satisfaction between the attributes and in particular the poor satisfaction judgments concerning on-board information on delays and price of the tickets. The highest satisfaction score is obtained for seating capacity, which means that crowded vehicles must be rather exceptional.

5.7.2 Weights of the attributes

Passengers find certain attributes of greater importance than other attributes. Customers' importance judgments can be collected by means of surveys (e.g. Eboli and Mazzulla, 2011; Tyrinopoulos and Antoniou, 2008), but also by means of Stated Preference experiments (Hensher et al., 2003) or in-depth interviews (Beirão and Sarsfield-Cabral, 2007). A number of authors deployed surveys to collect importance judgments of transit passengers and used these data to calibrate predictive models (Tyrinopoulos and Antoniou, 2008; Iseki and Taylor, 2008; Hensher, et al., 2003).

The surveys we used for our study were only aimed at collecting satisfaction judgments. No direct information is available on importance judgments. We developed a simple procedure for estimating weights of the attributes based on the assumption that total trip satisfaction is a weighted average of the satisfaction scores of the 15 attributes. The coefficients of the attribute variables represent the relative weight of these attributes as contributors to the satisfaction of the total trip.

Table 5.4. *Service attributes and their meaning*

Name	Clarification
General aspects of the transit system	
On-time performance	Accuracy of the realized departure times in relation to the schedule.
Travel speed	Appreciation of travel speed and time.
Service frequency	Number of transit vehicles per hour.
Personnel behaviour	Behaviour of the several types of personnel (e.g. drivers, station guards) when dealing with passengers.
Ticket-selling network	Ease of obtaining a ticket from on- and off board selling points.
Prices of the tickets	Price of various types of tickets and season cards.
Terminals and stops	
Information provision on stops	Information available for passengers on terminals and stops (static, dynamic, personnel).
Safety at stops	Safety on terminals and stops as perceived by passengers when waiting.
Vehicles	
Vehicle tidiness	Level of cleanliness of the vehicle in general.
Driver's behaviour	Driving performance of the driver.
On-board information on delays	On-board information provision (static, dynamic, vocal) on delays.
Ease of boarding and alighting	Ease of boarding and alighting the vehicle.
Seating capacity	Chance of getting a seat.
On-board noise	Level of noise in the vehicle.
Safety on board	Sense of safety during this trip.

Table 5.5 *Descriptive statistics: service attributes, 2001-2010*

2001-2010	Number of observations	Average satisfaction	Std. deviation
On-time performance	880,809	6.84	2.17
Travel speed	881,429	7.08	1.84
Service frequency	872,868	6.56	2.17
Personnel behaviour	847,174	7.19	1.83
Ticket-selling network	763,702	7.78	2.43
Prices of the tickets	764,131	5.62	2.97
Information provision on stops	853,315	7.00	2.06
Safety at stops	859,411	7.50	1.61
Vehicle tidiness	896,470	6.68	1.82
Driver's behaviour	871,346	7.00	1.70
On-board information on delays	756,654	4.87	2.72
Ease of boarding and alighting	895,831	7.99	1.74
Seating capacity	900,686	8.09	2.24
On-board noise	887,926	6.24	1.89
Safety on board	858,728	7.80	1.51
Overall satisfaction	883,009	7.11	1.38

To determine the weights, a linear regression model is designed with the overall satisfaction of the total trip as the dependent variable and the satisfaction with the individual service attributes as explanatory variables. Ordinary Least Squares is used for calculating the coefficients.

The model reads for individual *i*:

$$Q_{t,i} = b_0 + b_1 X_{1_i} + b_2 X_{2_i} + \cdots + b_n X_{n_i} + \varepsilon_{t,i} \quad (5.2)$$

where $Q_{t,i}$ is overall trip satisfaction in year *t* for individual *i*; b_0 is the constant term; b_1 b_2 , b_n are the weights of the service attribute variables; X_{1_i} , X_{2_i} are the satisfaction scores of service attributes 1 through 15 (see Table 5.4) and $\varepsilon_{t,i}$ is the error term.

Deviating from the data analysis described in Section 5.6 that is based on year pairs, for determining the weights of the attributes, the observations for all 10 years (more than 900,000 cases) are pooled. The outcomes are presented in Table 5.6.

Table 5.6. *Weights of the service attributes*

	Unstandardized coefficients*	
	B	**Std. Error**
(Constant)	.789	.009
On-time performance	.085	.001
Travel speed	.145	.001
Service frequency	.110	.001
Personnel behaviour	.072	.001
Ticket-selling network	.033	.001
Prices of the tickets	.010	.001
Information provision on stops	.052	.001
Safety at stops	.014	.001
Vehicle tidiness	.070	.001
Driver's behaviour	.075	.001
On-board information on delays	.028	.001
Ease of boarding and alighting	.047	.001
Seating capacity	.050	.001
On-board noise	.047	.001
Safety on board	.061	.001

** All coefficients significant (α≤ .001).*

It can be concluded from Table 5.6 that the most important service attributes are *travel speed, service frequency, and on-time performance*. Passengers highly value these attributes. These are – not surprisingly – the attributes that are related to the primary function of a public transit system: namely, to supply frequent, fast and on-time public transport. These outcomes for the Dutch situation are in accordance with survey results for other countries (Hensher et al. 2003; Tyrinopoulos and Antoniou, 2008; Eboli and Muzzalla, 2010; Beirão and Sarsfield-Cabral, 2007; Berechman, 1993).

5.7.3 Tendering and satisfaction with weighted service attributes

In this section first the relationship between the satisfaction with each of the 15 attributes in the tendered regions (as opposed to the non-tendered regions) is assessed. Secondly the findings on the relative weights of the service attributes are combined with the satisfaction scores in the tendered regions. The weights of the attributes can be interpreted as the importance customers attach to the different trip attributes. This provides a crucial basis for our analysis because it is important to know whether competitive tendering affects important attributes or less important attributes.

Satisfaction with service attributes in tendered regions

For each year pair, a model is formulated for determining the statistical relationship between tendering and each of the 15 service attributes. Each of the models is identical to Model 5.1 described in Section , but the dependent variable $Q_{r,t',i}$ is now the satisfaction with each of the underlying service attributes in region *r* and year *t* or year *t+1* for individual *i*. Based on the data for the 72 concession regions, for each of the 9 year pairs the coefficients of the 'tendering' variable are estimated for each of the 15 attributes. The main results for the period 2001-2010 are shown in Table 5.7.[71] As a reference, in the last row of the table the results for the satisfaction with the total trip – already given in Table 5.3 – are entered. The correlation matrix is shown in Appendix 5.B.

To clarify the results: in 57 out of 72 tendered regions in the period 2001-2010, a positive effect of tendering on the attribute *vehicle tidiness* can be observed. In 51 out of these 57 regions the positive effect is also significant. The average satisfaction with vehicle tidiness in the tendered regions – in relation to the non-tendered regions – increased by 0.41 points (on a scale of 1 to 10), whereas the overall trip satisfaction increased by 0.061 points (see also Table 5.3).

In general, it can be concluded that 12 out of the 15 attributes contribute in a positive way to the tendering effect of overall trip satisfaction. The items that contribute most to the change in total trip satisfaction in the tendered regions are *vehicle tidiness, on-board noise, ease of boarding/alighting from the vehicle,* and *service frequency* (in that order).

The values of the attributes that are linked to *information* and to *on-time performance* in the tendered regions are negative, meaning that the satisfaction with these attributes gets worse compared with non-tendered regions. The common factor in these attributes is that they all refer to reliability. Probably reliability suffers as regions get tendered.[72]

Weighted satisfaction with service attributes in tendered regions

In this section, the weights of the attributes are combined with the values of the average satisfaction with the attributes in the tendered regions. The resulting weighted coefficients represent passengers current satisfaction with the service level, conditioned for the relative importance to the passenger of service attributes.

71 The output is limited to the most important outcomes. Other outcomes are available on request.

72 Another possible explanation is postulated by Friman (2004). She supposes that the alleged blessings of improving service quality (in our case by means of tendering), are a priori communicated to the passengers, leading them to raise their expectations. These passengers are disappointed when the actual service delivery does not match these expectations. The difficulty with this explanation is that it is not clear why it holds for the reliability attributes, and not for the other attributes.

Table 5.7. *Tendering and satisfaction with service attributes, 2001-2010*

Dependent variable	No. of regions with positive effect of tendering.	No. of regions with positive significant effect of tendering (α ≤ .05).	No. of regions with negative effect of tendering.	No. of regions with negative significant effect of tendering (α ≤ .05).	Average impact on satisfaction (all cases)	Average impact on satisfaction, significant cases (α ≤ .05).
On-time performance	32	18	40	24	-0.058	-0.102
Travel speed	39	17	33	15	0.018	0.026
Service frequency	48	33	24	12	0.207	0.331
Personnel behaviour	48	31	24	9	0.122	0.213
Ticket-selling network	43	18	29	11	0.075	0.16
Prices of the tickets	39	22	33	19	0.055	0.104
Information provision on stops	33	18	39	20	-0.016	-0.027
Safety at stops	45	19	27	7	0.056	0.152
Vehicle tidiness	57	51	15	7	0.41	0.515
Driver's behaviour	42	18	30	14	0.059	0.112
On-board information on delays	29	17	43	27	-0.149	-0.209
Ease of boarding and alighting	59	41	13	8	0.22	0.303
Seating capacity	41	22	31	17	0.059	0.087
On-board noise	51	39	21	11	0.223	0.319
Safety on board	46	23	26	8	0.066	0.162
Overall satisfaction	42	22	30	12	0.061	0.112

In Table 5.8, for each attribute, the weights of the attributes are combined with the coefficients of the tender variable for the significant cases as well as for all tendered cases.[73]

Weighted satisfaction refers to the passengers' satisfaction in the tendered regions related to the non-tendered regions and corrected for the relative importance of the attributes. If the rank order of the weighted satisfaction scores is compared with the non-weighted scores, the rank order changes as the importance of the attributes is accounted for. So weighting – also taking the importance judgments into account – does make sense.

Concerning tendering, the weighted results show a clear and potentially policy-relevant outcome. Compared with non-tendered regions, in regions where tendering has a significant positive effect on weighted satisfaction, as well as in regions with a significant negative effect of tendering, the attributes that contribute most are *service frequency, on-time performance, travel speed*, and *vehicle tidiness*. The importance of *vehicle tidiness* is unexpected, but authors such as Eboli and Mazzulla (2010) and Tyrinopoulos and Antoniou (2008) also report the importance of that attribute.[74] When the net effect is considered, the last column of Table 5.8 indicates that 3 out of the 4 most-important attributes in the tendered regions may be linked to the vehicle itself (tidiness, on-board noise, and ease of boarding and alighting).

We conclude that -if the importance of the attributes is taken into account the rise (or fall) in average satisfaction in the tendered regions is mainly determined by the rise (or fall) in satisfaction with the attributes service *frequency, on-time performance, travel speed*, and *vehicle tidiness*. Other attributes only contribute in a limited way. Moreover, an important part of the net effect of the increase in weighted satisfaction may well be connected to launching new vehicles as part of the tender. This will be discussed in more detail in the next section.

73 The weights differ from those of Table 5.6 because the constant term is not taken into account here.
74 In Hensher et al., (2003) however, vehicle tidiness is of no importance.

Table 5.8. *Weighted contribution of service attributes in tendered regions (average 2001-2010)*

Explanatory variables	Weights of the service atrributes	Average coefficients of tendering variable: Tendered regions with sign. pos. effect of tendering	Average coefficients of tendering variable: Tendered regions with sign. neg. effect of tendering	Average coefficients of tendering variable: All tendered regions *(N=72)*	Average coefficients of tendering variable, weighted by importance of service attribute: Tendered regions with sign. pos. effect of tendering	Average coefficients of tendering variable, weighted by importance of service attribute: Tendered regions with sign. neg. effect of tendering	Average coefficients of tendering variable, weighted by importance of service attribute: All tendered regions *(N=72)*
On-time performance	0.094	0.384	-0.467	-0.058	0.036	-0.044	-0.005
Travel speed	0.162	0.362	-0.354	0.018	0.059	-0.057	0.003
Service frequency	0.123	0.598	-0.402	0.207	0.073	-0.049	0.025
Personnel behaviour	0.080	0.359	-0.293	0.122	0.029	-0.023	0.010
Ticket-selling network	0.036	0.460	-0.329	0.075	0.017	-0.012	0.003
Prices of the tickets	0.011	0.698	-0.583	0.055	0.007	-0.006	0.001
Information provision on stops	0.058	0.375	-0.388	-0.016	0.022	-0.022	-0.001
Safety at stops	0.016	0.306	-0.267	0.056	0.005	-0.004	0.001
Vehicle tidiness	0.078	0.643	-0.418	0.410	0.050	-0.032	0.032
Driver's behaviour	0.083	0.405	-0.264	0.059	0.034	-0.022	0.005
On-board information on delays	0.031	0.548	-0.686	-0.149	0.017	-0.022	-0.005
Ease of boarding and alighting	0.052	0.417	-0.284	0.220	0.022	-0.015	0.012
Seating capacity	0.056	0.448	-0.380	0.059	0.025	-0.021	0.003
On-board noise	0.053	0.520	-0.395	0.223	0.027	-0.021	0.012
Safety on board	0.068	0.311	-0.266	0.066	0.021	-0.018	0.004
Overall satisfaction	n.a.	0.320	-0.268	0.061	n.a	n.a	n.a

5.8 In depth analysis of tendering benefits

It is tempting to assign the above-mentioned outcomes to the implementation of tendering itself, but caution is necessary. In this section, we link some general observations on the effects of tendering in the Netherlands to the changes in satisfaction we observed. We derived the additional information needed for this analysis from general public sources and assess these items more qualitatively.

Evaluations of tendering in the Netherlands (Berenschot, 2004; Mu-consult, 2003) indicate three important effects of tendering: (1) new vehicles were introduced; (2) focus on supply-oriented steering by the authorities; (3) replacement of the incumbent operator.

The authors would like to add to this the observation that over the course of time, operators and authorities became more and more experienced in using the instrument of tendering, so time, as a proxy for learning, may also have an impact on satisfaction.

5.8.1 New vehicles

An important observation is that in nearly all tendered cases in the Netherlands, new vehicles were required by the authorities. This led to the situation that the old buses - even if they were not yet fully depreciated- were replaced by new low floor buses equipped with dynamic on-board information systems and comfortable seats.

In line with Zeithaml et al. (1990), it is possible that the vehicles are perceived by passengers as an important tangible environmental dimension of service deliverance, which therefore influences the passenger satisfaction. This may be an important explanation for the observed rise in overall customer satisfaction after the introduction of tendering. Moreover, new vehicles may well have an impact on (the satisfaction with) the following service attributes:

Seating capacity	+ or -
Vehicle tidiness	+
On-board noise	+
Ease of boarding/alighting	+
On-board information	+
On-board safety	+ or -

We tested this assumption by separating the tendered cases into four categories of tendered regions depending on the proportion of new vehicles that were introduced. We distinguish the following categories of tendered regions: (1) no new vehicles; (2) information on vehicles not available; (3) partly new vehicles, or the intake of new vehicles is spread over more years; and (4) the complete fleet is renewed as of the start of operations. Table 5.9 shows, per category, the average coefficients for the tendering variable of Model 5.1.

Hoofdstuk 5

Table 5.9. *Effect of new vehicles on satisfaction*

Attribute	Average coefficient of tendering variable				Delta	
	Tendered regions, no new vehicles *(N=16)*	Tendered regions, new vehicles unknown *(N=14)*	Tendered regions, partly new vehicles (N=9)	Tendered regions, complete new fleet as of beginning operations *(N=33)*	Completely new vs. no new vehicles	Partly new vs. no new vehicles
On-time performance	-0.112	-0.066	0.014	-0.048	0.064	0.126
Travel speed	-0.070	0.016	0.104	0.038	0.108	0.174
Service frequency	0.405	0.094	0.270	0.141	-0.264	-0.135
Personnel behaviour	0.038	0.161	0.140	0.142	0.105	0.102
Ticket-selling network	-0.001	0.165	-0.018	0.100	0.101	-0.016
Prices of the tickets	-0.086	0.147	-0.066	0.117	0.202	0.020
Information provision on stops	-0.154	0.038	0.075	0.004	0.158	0.229
Safety at stops	-0.034	0.099	0.140	0.058	0.092	0.174
Vehicle tidiness	0.266	0.217	0.368	0.574	0.308	0.102
Driver's behaviour	0.002	0.122	0.137	0.040	0.038	0.136
On-board information on delays	-0.288	-0.041	-0.216	-0.109	0.179	0.073
Ease of boarding and alighting	0.151	0.243	0.187	0.252	0.102	0.036
Seating capacity	-0.036	0.043	0.114	0.097	0.134	0.150
On-board noise	0.104	0.143	0.302	0.292	0.188	0.198
Safety on board	-0.037	0.081	0.208	0.071	0.109	0.245
Overall satisfaction	0.014	0.030	0.105	0.085	0.071	0.090

We may conclude that new vehicles highly impact on overall satisfaction, as well as on satisfaction with many attributes. The highest impact on satisfaction of introducing new vehicles concerns the satisfaction of *vehicle tidiness*, *on-board noise*, and *seating capacity*. These outcomes are in line with our expectations. It is however striking that the positive effect on satisfaction of introducing new vehicles is not restricted only to vehicle-linked attributes; in addition the satisfaction with non-vehicle linked attributes such as *information provision on stops*, and *personnel behaviour* increase with new vehicles. It seems that new vehicles contribute to a positive general perception of public transport use. Again – consistent with the literature – this is a sign that in satisfaction judgments, subjective and environmental factors play a significant role. The introduction of new vehicles seems to have a negative impact on the satisfaction of the attribute *service frequency*. Detailed analyses showed that this is however probably a coincidence, since by accident in a number of tendered cases where no new vehicles were introduced, service frequency rose sharply, leading to a significant rise in satisfaction with this attribute.

5.8.2 New operator and experience with tendering

In the period under study, in 41 of the 72 tendered regions the incumbent won the tender (59%) and stayed in control. In 31 tendered regions operations shifted to a new operator (41%)[75]. One might expect that the change of operator, as result of the tendering procedure, has an effect on satisfaction judgment. The argumentation is that a new operator is more willing, and is more challenged, by the authority to change its performances and services than an incumbent.

It is also likely – as we showed in Section – that it matters for satisfaction whether a concession region is tendered for the first time or for the second time. The expectation is that the increasing experience with the tendering instrument of both operators and authorities may impact on satisfaction. As time goes by, both actors may, for instance increase their knowledge on the needs, valuations, and satisfaction of their passengers by means of using the results of the yearly satisfaction surveys performed by KPVV.

We decided to incorporate these possible explanations per attribute in a regression model that has the tendering effect (the output of Model 5.1) as dependent variable. We incorporated the variable experience in the model by way of a time trend and by considering whether a concession is tendered for the first or the second time. So the predictors of the model are: (1) time (year of tendering); (2) regions tendered twice versus once; and (3) new operator versus incumbent (see Eq. 5.3).

75 The average duration of a contract is 8 years. Each year on average 10% of all regions enter the process of competitive tendering.

$$B_{t,i} = b_0 + b_1\, t_i + b_2 tendered\ twice\ vs.\ once_i + b_3 new\ operator\ vs.\ inc._i + \varepsilon_i, \quad (5.3)$$

where t = 1,……10 for the years 2001,…..2010.

The output of Model 5.3 is shown in Table 5.10. In the discussion of the findings, we focus on the four attributes we showed to contribute most to the weighted satisfaction in tendered regions compared with not-tendered regions, i.e. the attributes *service frequency*, *on-time performance*, *travel speed* and *vehicle tidiness* (see Table 5.8).

Table 5.10. *Determinants of tendering effects*

	Regression coefficients model 5.3.		
Attributes	**Trend**	**Twice tendered vs. once tendered**	**New operator vs. incumbent**
On-time performance	-0.032	0,161	-0.128
Travel speed	-0.013	0.045	-0,030
Service frequency	-0,010	0.059	0.097
Personnel behaviour	-0.027	-0,050	0.062
Ticket-selling network	-0,037*	-0.051	-0.034
Prices of the tickets	-0,060*	-0,378*	-0.091
Information provision on stops	-0.022	0,110	-0,166*
Safety at stops	-0,040*	0.033	-0.001
Vehicle tidiness	0.001	-0.265	0.082
Driver's behaviour	-0.031	0.049	0,010
On-board information on delays	0.014	0.043	-0,316*
Ease of boarding and alighting	-0.012	-0.127	-0.056
Seating capacity	-0.014	-0,050	-0.083
On-board noise	-0,020	-0.128	0.017
Safety on board	-0,035*	0.011	-0.029
Overall satisfaction	-0.015	0,000	-0,020

** significant at the .05 level (2-tailed)*

Although the decrease is only significant for four attributes, Table 5.10 shows that the coefficient for the trend variable has dominantly a negative sign, meaning that in two consecutive years the effect of tendering on satisfaction with most attributes is smaller for late versus early tenders. The attributes *prices of the tickets*, *on-board safety*, and safety *at stops* show the sharpest decrease.[76] A positive (though not significant) trend is found for the effect of tendering on the

76 Note that these negative figures do not necessarily mean that tendering had an adverse effect on satisfaction of these service attributes, but that the - possible positive - effect of tendering is decreasing in time.

satisfaction of the attributes *vehicle tidiness*, and *on-board information on delays*. These two attributes show an increase in satisfaction between two consecutive years owing to tendering that is probably connected to the introduction of new vehicles. This mainly negative effect of time on tendering benefits as valued by passengers, is somewhat unexpected, since one would expect that learning would have a positive effect for both tendering authorities and operators. A possible explanation is, that in the course of time, the attention in tendering has shifted from improving passenger satisfaction to efficiency improvements. It is also striking that the effect on overall satisfaction of tendering for the second time is no longer clear. Thus, the effect picked up in Table 5.3, seems not so much a matter of tendering for the first or second time, but reflects lower benefits for passenger satisfaction for late tenders compared to early tenders.

If a new operator takes over from the incumbent, Table 5.10 shows that for 10 out of 15 attributes the satisfaction judgments of passengers are negatively influenced by the change in operator as a result of tendering, and also that the overall satisfaction is lower, although the difference is not significant. The satisfaction with five attributes is positively influenced by a change of operator. Although not significant, the positive satisfaction change due to tendering of a change of operator of the attributes *vehicle tidiness*, and *on board noise*, may well be linked to the introduction of new vehicles rather than to the new operator (see Table 5.10). The increase of satisfaction of service frequency if a new operator takes over from the incumbent is in line with the actual developments in tendered regions (see Table 5.1).

5.8.3 Supply-oriented steering

The evaluations of the Dutch situation in 2004 (Berenschot, 2004; MuConsult, 2003) showed that the dominant steering factor for Dutch authorities in tendering procedures is supply of public transport services (vehicle-hours and kilometres). After the first years of tendering, the authorities also introduced more quality based steering parameters[77], but supply-based steering remained dominant. We showed that the supply of public transport increased considerably as a result of tendering (see Table 5.1). If we define *service frequency* as a proxy for supply, we may conclude that the change in satisfaction concerning this attribute may well be connected to the change of operator; a change that would not have taken place without tendering.

To summarize: concerning the four attributes that contribute most to the weighted effect of satisfaction in tendered regions relative to not-tendered regions (*service frequency*, *on-time performance*, *travel speed*, and *vehicle tidiness* (see Table 5.8). we may conclude that the shift in satisfaction with vehicle tidiness is mainly linked to the introduction of new vehicles. This

77 In the case of quality-based steering, the authorities narrowed the concept of service quality mainly to reliability of the service. In only one of the Dutch cases we found that vehicle cleanliness was specified as an important dimension of policies to increase quality.

introduction is, as was shown, accelerated by the process of tendering. However, new vehicles are not exclusive for tendering. As was already pointed out in Section , in the pre-tendering years, vehicles were also replaced periodically, but we may say that the increase in satisfaction in tendered regions we observed concerning this attribute is – indirectly – the effect of tendering itself.

As concerns the contribution of tendering to the positive change in satisfaction with *travel speed* in tendered regions, the same holds: the introduction of new vehicles impacts positively on the satisfaction with this attribute. The change in satisfaction with *service frequency* in tendered regions relative to not-tendered regions, may well be connected to the change in operator, so therefore to tendering.

Finally, the satisfaction with the attribute *on-time performance*, seems to take hardly any advantage from tendering. It even seems to be negatively related to tendering. Apparently the shift in attention in the tendered regions to the attributes speed and frequency, may have had adverse effects with (the satisfaction of) *on-time performance*.

5.9 Conclusion

Over the period 2001-2010 an analysis of tendered regions versus non-tendered regions in the Netherlands was conducted. In this period 72 regions were tendered, 34 out of them showed a significant change of overall passenger satisfaction compared with non-tendered regions. The average impact on satisfaction in these tendered regions amounts to 0.112 points (on a 10-point scale). This positive effect in the tendered regions is solely caused by regions that were tendered for the first time. We observed 11 regions that were tendered for the second time. The outcomes for the second round of tendering revealed for the tendered regions with a significant impact a decline in average satisfaction of 0.048 points. A more detailed time analysis showed that this is not only a matter of sequence (first versus second), but also of timing (early versus late). It is possible that this finding is connected to the shift in emphasis of the tendering authorities in the second round from quality objectives towards efficiency objectives.

Although there is a positive effect on satisfaction in the tendered regions, we found over the period 2001-2010, also in the non-tendered regions, a trend of an increase in satisfaction. A 'wind of change' emerged from the introduction of obliged tendering of public transport in the Netherlands as of the year 2001. Authorities and operators felt the pressure to increase quality. It is obvious that – also in the non-tendered regions – this atmosphere of urgency led to a more customer oriented approach of both operators and authorities. This may explain the observed rise in satisfaction in the non-tendered regions.[78]

Furthermore, an analysis on the weighted satisfaction judgments of 15 underlying service attributes, revealed that *service frequency*, *on-time performance*, *travel speed*, and *vehicle tidiness*

78 In a mid-term evaluation of the Dutch experiences this effect of 'threat' has been observed (MuConsult, 2003).

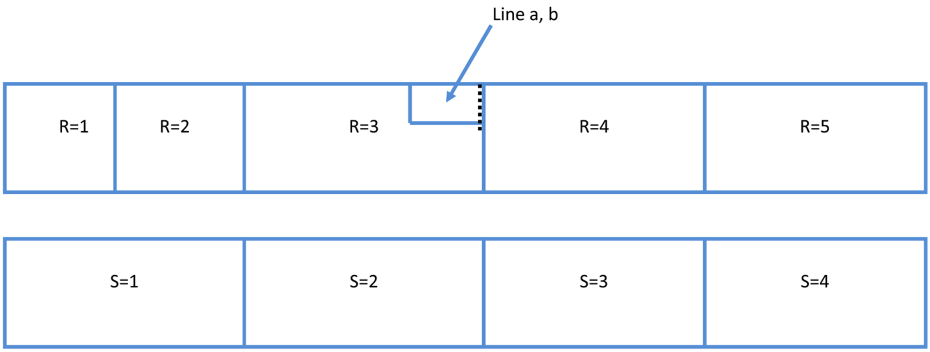
contributed the most to the effect on satisfaction mentioned before in the tendered regions. The latter is also of interest for policy reasons, since it indicates that –besides the traditionally known attributes speed, frequency, and on-time performance- vehicle tidiness also plays an important role in total service satisfaction.

Concerning the question raised in this chapter whether tendering affects satisfaction, we found that new vehicles impact highly on satisfaction and also that -owing to tendering- the introduction of new vehicles was accelerated. We also found that a change of operator due to tendering, in general, negatively impacts on the satisfaction judgments of passengers.

Appendix 5.A. Recoding routine used to cope with changes in demarcations

In the case of a simple merger of two areas R1 and R2 in period *t*=1 into a larger concession area S in period *t*=2 we recoded the satisfaction scores with subscripts R1 and R2 in *t*=1 into scores with subscript S in *t*=2 so that they are immediately comparable with the satisfaction scores available for area s in period *t*=1.

In the case of a border correction between two areas we have to go deeper in the underlying data. Suppose that a part of area R4 is shifted to area S2 between periods *t*=1 and *t*=2. Then we check the public transport lines in R4 that are added to S2. We have satisfaction measurements on all specific lines. Thus we can recode subscripts such that the satisfaction of the travellers in the share of area R4 that becomes part of area S2 in *t*=2 is recoded as S2 in *t*=1. In the figure beneath this procedure is visualized.



Original code t=1		Code harmonized to situation in t=2
R1	→	S1
R2	→	S1
R3	→	S2
R4 (line a, b)	→	S2
R4 (other lines)	→	S3
R5	→	S4

Appendix 5.B. Correlation of tendering variable coefficients (Pearson's *R*)

	On-time performance	Travel speed	Service frequency	Personnel behaviour	Ticket-selling network	Prices of the tickets	Information provision on stops
On-time performance	1	0.599	.315**	.603**	0.171	.344**	.661**
Travel speed	0.599	1	.289*	.513**	0.165	0.192	.574**
Service frequency	0.315	0.289	1	.260*	-0.047	0.223	0.219
Personnel behaviour	0.603	0.513	0.26	1	.500**	.305**	.571**
Ticket-selling network	0.171	0.165	-0.047	0.5	1	.460**	.297*
Prices of the tickets	0.344	0.192	0.223	0.305	0.46	1	.346**
Information provision on stops	0.661	0.574	0.219	0.571	0.297	0.346	1
Safety at stops	0.373	0.396	0.161	0.542	0.487	0.372	0.56
Vehicle tidiness	0.213	0.389	0.134	0.396	0.369	0.488	0.258
Drivers behaviour	0.644	0.612	0.306	0.838	0.397	0.364	0.61
On-board information on delays	0.526	0.298	0.333	0.099	-0.119	0.364	0.367
Ease of boarding and alighting	0.275	0.422	0.077	0.609	0.592	0.406	0.469
Seating capacity	0.225	0.299	-0.069	0.582	0.413	0.045	0.507
On board noise	0.204	0.39	0.355	0.284	0.18	0.391	0.242
On-board safety	0.507	0.524	0.151	0.675	0.523	0.38	0.642
Total trip	.675**	.777**	.421**	.765**	.422**	.407**	.682**

**Correlation is significant at the 0.05 level (2-tailed)*

***Correlation is significant at the 0.01 level (2-tailed)..*

Safety at stops	Vehicle tidiness	Drivers behaviour	On-board information on delays	Ease of boarding and alighting	Seating capacity	On board noise	On-board safety	Total trip
.373**	0.213	.644**	.526**	.275*	0.225	0.204	.507**	.675**
.396**	.389**	.612**	.298*	.422**	.299*	.390**	.524**	.777**
0.161	0.134	.306**	.333**	0.077	-0.069	.355**	0.151	.421**
.542**	.396**	.838**	0.099	.609**	.582**	.284*	.675**	.765**
.487**	.369**	.397**	-0.119	.592**	.413**	0.18	.523**	.422**
.372**	.488**	.364**	.364**	.406**	0.045	.391**	.380**	.407**
.560**	.258*	.610**	.367**	.469**	.507**	.242*	.642**	.682**
1	.258*	.562**	0.194	.497**	.466**	.299*	.810**	.563**
0.258	1	.396**	0.153	.529**	0.224	.736**	.472**	.577**
0.562	0.396	1	0.184	.575**	.502**	.456**	.743**	.766**
0.194	0.153	0.184	1	-0.121	-.295*	0.211	0.17	.274*
0.497	0.529	0.575	-0.121	1	.661**	.438**	.656**	.639**
0.466	0.224	0.502	-0.295	0.661	1	0.148	.605**	.536**
0.299	0.736	0.456	0.211	0.438	0.148	1	.448**	.552**
0.81	0.472	0.743	0.17	0.656	0.605	0.448	1	.715**
.563**	.577**	.766**	.274*	.639**	.536**	.552**	.715**	1

Chapter 6

The effect of contract renewal and competitive tendering on public transport costs, subsidies and ridership.

This chapter is based on Mouwen and van Ommeren, the effect of contract renewal and competitive tendering on public transport costs, subsidies and ridership, Transportation Research Part A, forthcoming.

6.1 Introduction

In the previous chapter we evaluated the customer side. In this chapter we change the perspective to costs, subsidies and efficiency especially relevant for PT firm and authorities.

After 1990 all over the western world reform of the public transport (PT) industry has taken place. In Europe this reform has intensified due to the EU directive 1191/69/EU that put forward competitive tendering for procurement of exclusive PT services as the preferred way.[79] The new regulative framework for PT aims to enable an efficient and effective transfer of subsidies from the public transport authority (PTA) to operators. A reduction in PT ridership (mainly caused by increasing popularity of the car), combined with universal service obligation, led in the 60s and 70s to increasing operational deficits all over the western world, as fare box revenues increasingly failed to cover operational costs. Public budget constraints forced many Western PTAs to implement regulatory reforms. The new regulatory framework generally aimed to introduce incentives for operators to increase efficiency, cut down subsidies and increase social welfare.[80] Triggered by these regulatory changes, a large number of studies have examined the determinants of operational efficiency, and specifically the effect of regulative change on firms' performance. In line with this literature we aim to assess the impact of contract renewal and competitive tendering (CT), on operational costs, subsidy, and PT ridership. We employ panel data for the period 2001-2013 on the level of concession areas in the Netherlands and take the most relevant contract attributes into account.[81]

Our focus on contract renewal is a natural one, as contracts serve as a formal stipulation of arrangements between operators and authorities and govern risk-sharing between PTAs and operators. The role of contract renewal as an incentive driver has long been recognized (Laffont and Tirole, 1993; Dalen et al., 2006; Gautier and Yvrande-Billon, 2013). Contract renewal in a market characterized by CT is pivotal as it allows several operators to bid for a new contract.

One of the main econometric issues we identified in Chapter 2 is that network characteristics of the concession area impact heavily on PT efficiency, and it is therefore essential to control for (exogenous) relevant network characteristics. Most studies control for variables such as network length, average speed, number of stops and lines, however other network characteristics that may influence firms' productivity cannot be assumed away, thereby potentially biasing the analysis. Our approach avoids this issue by using panel data with concession area fixed effects, thereby controlling for all time-invariant unobserved area circumstances that may influence production efficiency of the firm.[82] To our knowledge we are the first to do this.

---

79 The directive is formally finalized with the 2007 Public Service Obligations' Regulation' (later modified by 1893/91/EU). See Van de Velde and Beck, 2010.

80 For reviews of these phenomena, see various volumes of the THREDBO-series (TREDBO, 2015).

81 As individual contracts are not publicly available, in our econometric analysis we do not control for all attributes in the contracts. However we control for a range of relevant variables including vehicle kilometres and new vehicles that are usually the main attributes.

82 This method does not shed light on the influence of individual time-invariant exogenous circumstances on efficiency. We examine this with a separate analysis on the effect of network characteristics on operational costs.

In section 6.2 we review relevant literature on regulatory change in relation to PT costs and efficiency. Sections 6.3 and 6.4 presents models of PT costs, subsidies, and PT ridership. Section 6.5 describes the institutional context of PT in the Netherland and presents the data we used. In section 6.6 estimation results are given. Section 6.7 contains conclusions and recommendations.

6.2 Literature review on regulative change and efficiency in PT

Reform of the PT industry in the western world has led to a large number of studies on PT efficiency and effectiveness. These studies especially focus on how, and to what extent, public sector interventions affect efficiency of, and budget transfers to, PT firms. In this section, we review literature on issues related to competitive tendering, contract type, firm ownership and network characteristics. We also describe the ongoing debate on the most appropriate measure of PT output in the economic analysis.[83]

After the directive 1191/69/EU, CT has become a popular instrument to organize PT in the EU. Its primary aim is reduction of public subsidies. Subsidy reductions come fairly evenly from reductions in factor prices (especially labor and fuel), reductions in the use of labor and land, and adaptions to the production process (Preston, 2002). These efficiency gains are, via sharper contract biddings, transferred from operators to PTAs. Typically, the first round of tendering shows substantial cost reductions up to 50% when PT services were previously provided by public firms under public monopolies, but subsequent re-tendering delivers minimal subsidy reductions (Hensher and Wallis, 2005). Probably the greatest inefficiencies in PT provision are removed as result of the first contract renewal. Further cost reductions are thought to be minimal because the system has matured: authorities and bidders become more experienced (leading to less bid errors), PTAs ask for more demanding contract specifications in subsequent rounds of tendering (such as new low-floor vehicles) and bidders take a longer-term perspective and aim at higher profit margins (Hensher and Wallis, 2005). German urban PT companies operating in areas where CT is implemented reveal a significantly higher average efficiency than other companies (Scheffler et al., 2013). Karlaftis (2010) and Boitani et al (2013) show with panel data of large European cities in nine different countries that firms selected after CT display approximately a 15 to 20% higher total factor productivity than firms selected under different contract awarding regimes. To summarize: CT effectively increases firm efficiency, and decreases subsidy transfers by PTAs.

Contract type is a powerful instrument for PTAs to govern transactions with the operator, as contracts make it possible to introduce specific incentives (Margari et al., 2007; Roy and

83 This chapter would ideally test the influence of all above mentioned factors on operational costs and ridership empirically. Due to data limitations, the empirical analysis however mainly focuses on competitive tendering, network characteristics and PT output measures.

Yvrande-Billon, 2007; Karlaftis and Tsamboulas, 2012; Gautier and Yvrande-Billon, 2013).[84] An important distinction is between high powered fixed price contracts such as gross and net costs contracts and low powered costs-plus, or management, contracts.[85] Under identical network conditions operators regulated by fixed-price contracts are more efficient than operators regulated with costs-plus contracts (Dalen and Gomez-Lobo, 2003; Piacenza, 2006). This implies that the latter contract arrangements especially common in France, are not the most efficient way to reach efficiency goals (Gautier and Yvrande-Billon, 2013). There is empirical evidence that firms under gross costs contracts (where the PTA receives all fare box revenues and therefore bares the commercial risks) are more efficient than firms under net costs contracts (where the operator receives all fare box revenues and bares the commercial risks), as gross costs contracts provides more incentives for production efficiency than net costs contracts. Firms regulated by gross costs contracts solely aim to reduce costs, for example by optimizing the number of drivers and vehicles, whereas firms under net costs contracts also aim to increase revenues (Margari et al., 2007; Gautier and Yvrande-Billon, 2013).[86] We conclude that high-powered incentive contracts (especially gross costs contracts) seem to perform best on efficiency.

According to economic theory, firm ownership matters because public companies tend to be less efficient than private companies, as their deficits are covered by authorities, and may be forced by politicians to hire an inefficient number of workers to boost local employment (Boycko et al., 1996). However, Berechman (1993) and Scheffler et al. (2013), controlling for competition, do not find unambiguous results. Berechman (1993) claims that ownership type is not the determining factor in transit firms' productivity, but rather the size of the transit system and network, as well as the degree of market competition. The majority of frontier studies (especially the parametric frontier studies) show a positive association between private ownership and efficiency, but the results are mixed (see De Borger et al., 2002).[87] Gautier and Yvrande-Billon (2013) and Boitani et al. (2013) find that private operators in France as well as in large European cities are less inefficient than firms governed under mixed public-private ownership. Karlaftis (2010) reports for 15 European cities that public sector operators have 31% higher operating costs than private firms. Karlaftis and Tsamboulas (2012) also show private firms to outperform public firms. We conclude that empirical evidence on efficiency between

84 These studies are based on the (implicit) assumption that the type of contract is exogenous, so independent of the performance and network characteristics. This implies that the type of contract is often determined by political motives, rather than economic ones (Gagnepain and Ivaldi, 2002).

85 Yardstick regulation is another example of a high powered incentive scheme. It uses benchmarking to reduce the problem of asymmetric information between operator and regulator (Dalen and Gomez-Lobo, 2003).

86 Using different methods, Karlaftis (2010) and Karlaftis and Tsamboulas (2012) are among the few that found opposite results.

87 De Borger et al. (2002) note that most of these studies did not control for level of ownership, degree of competition and the regulatory environment. They argue that for strongly regulated markets like urban transit, ownership itself is of little relevance, but level of competition is.

public and private ownership in the PT sector for the most part favors private ownership, but is not conclusive.

Network conditions are important, so in cost studies economies of scale and density are relevant, but show a wide diversity of outcomes. Robust results however are that firms in small networks produce inefficiently, and scale elasticities decrease as production increases (Croissant et al., 2013). Most studies covered by the meta-study of De Borger et al. (2002) provide evidence for U-shaped marginal costs functions and for economies of density (i.e. returns to traffic density given the network). This holds in the short run (through improved utilization of existing capital stock), as well as in the long run when fleet size can be adjusted (Farsi et al., 2007). Gautier and Yvrande-Billon (2013) show for the French urban PT sector significant economies of scale (0.5 up to 0.7), economies of network length (-0.06 up to -0.12), and economies of network speed (-0.25 up to -0.33). Dalen and Gómez-Lobo (2003), based on panel data for Norwegian bus companies, find strong suggestions of the existence of economies of scale for unregulated long-distance intercity services, which may be connected to higher speed and better utilization of drivers in intercity services. Piacenza (2006) stresses the importance of external network conditions for firms' efficiency, as these conditions often dominate the effects of regulative measures and contractual arrangements. More favorable traffic conditions with higher average speed for PT reduces costs with about 13% up to 36%. Thus, network conditions are a crucial determinant of the efficiency of PT. These conditions are given for the operator and usually correlated to other determinants and type of regulation that are included in the economic analysis. This implies that one has to control for network conditions in order to get a proper assessment of regulatory change on efficiency.

Related to efficiency, few studies report on economies of scope across PT modes. Economies of scope are defined as a situation of change in the unit costs of production in a multi-output production setting, due to the possible use of shared facilities such as management, depots and terminals.[88] Berechman (1993) summarizes the results of (mainly) US studies showing slight economies of scope across modes. Moderate scope economies across modes (around two per cent), and fixed costs reduction due to joint production were also found by Di Giacomo and Ottoz (2010). Economies of scope of 25% were found in a study on Swiss urban PT. These economies decrease with increasing outputs (Farsi et al., 2007).

A fierce debate has been going on concerning the preferred output measurement of PT in efficiency studies (Berechman, 1993; Karlaftis, 2004). Output may be measured in terms of supply, such as vehicle hours or vehicle kilometers, or in terms of demand, such as passenger kilometers or passenger trips.[89] Supply- oriented measures are most popular, but don't reflect

88 Economies of scope imply that the average production costs of joint production are lower than the production costs when each of the products is produced separately.

89 The choice of output measure matters a lot. Berechman & Giuliano (1985) estimate a costs function using a demand related output measure, as well a supply measure, and show increasing returns to scale for the former, and the opposite for the latter.

the economic motive of PT-supply (De Borger et al., 2002; Brons et al., 2005). On the other hand, demand-oriented measures ignore that input factors such as fuel and labor do not systematically vary with demand. In addition, demand for PT can only be influenced by PT operators to a limited degree (Scheffler et al., 2013). De Borger et al. (2002) conclude that supply oriented parameters are preferred, even given that the firms' objectives using this measure are not taken into consideration.[90] We will use both measures in our application.

6.3 Estimation methodology

The idea behind implementing competition and regulatory change is based on the assumption that competition favors both efficiency and service quality, thereby leading to costs and subsidy reduction, and to an increase in ridership as demand will react positively to increased service quality. We will test for this by estimating models using operational costs, subsidies, and passenger kilometres as dependent variables. We also test whether economies of density and scope exist.

The general specification of the model to be estimated can be written as:

$$\mathbf{Y_{i,t} = Y\,(X_{i,t}, Z_{i,t}, C_{i,t}, F_s),} \qquad (6.1)$$

where: $Y_{i,t}$ denotes either operational costs, subsidies or passenger kilometres; X represents a vector of outputs; Z denotes a vector of variables including fixed network, and area inputs; C denotes a vector of contractual issues; F denotes a vector of firm specific issues; subscripts $i, t,$ and s indicate respectively concession area, year, and firm.

One may use several functional forms. A standard log-linear costs model implies a restrictive production function (Berechman and Giuliano, 1985), whereas more general specifications such as the translog form place few a priori restrictions on the underlying production function (Farsi et al., 2007; Roy and Billon, 2007; Scheffler et al., 2013, and Croissant et al., 2013). To use the translog function, rather than the log-linear function, is particularly important in a cross-section setting where differences in the dependent variable (costs, subsidies) are substantial. In our application, where we use panel data and focus on yearly changes in the dependent variable between concession areas, these changes are small. Therefore the translog functional

90 The firms' objective should be to transport passengers, not to drive buses around. However the regulation practice in large parts of Europe is such that the PTA sets the desired level of production. In these cases it is uncertain whether the firms' goals actually are set in terms of enhancing ridership. This especially applies to gross costs contracts, but may also be relevant in net costs contracts in situations where the firm has little room to maneuver in increasing fare box revenues, as is the case in the Netherlands.

form provides essentially identical estimates of the coefficients of interest as a log-linear form. We proceed using the log-lineair form, and estimate the paramters using OLS.[91]

Estimates based on panel data are preferred to cross-section data (Croissant et al., 2013). We include year and concession area fixed effects, so all unobserved concession characteristics such as urbanization degree, network length and number of stops which hardly change over time, are taken into account. Appendix 6.A shows that changes in these characteristics are indeed very limited in our data.

6.4 Operational costs, subsidy and passenger kilometres model

In this chapter we use the following (log) specification for the operational cost and subsidy model:

$$\ln Y_{i,t} = \beta_0 + \beta_1 CR_{i,t} + \beta_2 CT_{i,t} + \beta_3 NV_{i,t} + \beta_4 NO_{i,t} + \beta_5 \ln VH_{i,t} + \beta_6 MS_s + \delta_t + \eta_i + \varepsilon_{i,t,s} \quad (6.2)$$

where: $Y_{i,t}$ denotes either operational costs, or subsidies in concession area *i* in year *t; $CR_{i,t}$* denotes the number of contract renewals between 2001 and *t; $CT_{i,t}$* denotes whether the contract is competitively tendered or negotiated[92]; $NV_{i,t}$ denotes new vehicles in contract; $NO_{i,t}$ denotes a new operator, a proxy for new quality aspects in the contract other than new vehicles; $VH_{i,t}$ denotes vehicle hours; MS_s denotes multi or single-production of the firm; δ_t denotes a year fixed effect; η_i denotes a concession area fixed effect; $\varepsilon_{i,t}$ denotes a random error term. In our data i = 1,…,38; t=2001,…,2013; s= 1,…,15. We explicitly control for differences in contract characteristics (e.g. new vehicles), for time differences (which captures price effects and changes in technology), and for differences between concession areas (which capture unobserved spatial aspects of concession areas such as network conditions).[93]

91 We have also estimated stochastic costs frontier models for panel data (Kumbhakar and Lovell, 2000; Holmgren, 2013). More precisely, we have estimated a linear model with a disturbance following a Battese-Coelli (1992) parameterization. The inefficiency term is modeled as a truncated-normal random variable multiplied by a specified function of time. Given concession area fixed effects, which control for time-invariant inefficiency, one expects that stochastic frontier models and regression generate very similar results. This is confirmed by our results which show that the variance of the random inefficiency term is only 1.5% of the overall variance. The marginal effects of explaining variables hardly change.

92 As an alternative to a competitive tendered contract, a contract can be negotiated, indicating the situation where the PTA directly identifies the supplier of the service, imposing or negotiating some conditions.

93 Factor prices and overall technology change are not explicitly included in the specification, but captured by year fixed effects and concession area fixed effects. In Section 2 we elaborated on the potential important effect of type of contract and firm ownership on operational costs and efficiency. Our data only contains new net costs contract cases. Due to multicollinearity we were not able to test firm ownership empirically. We therefore did not incorporate contract type and firm ownership in the model specifications.

We also estimate a model using (log) passenger kilometres as dependent variable. The independent variables are identical to the operational costs model described above:

$$\ln \mathbf{Pax\ km}_{i,t} = \gamma_0 + \gamma_1 \mathbf{CR}_{i,t} + \gamma_2 \mathbf{CT}_{i,t} + \gamma_3 \mathbf{NV}_{i,t} + \gamma_4 \mathbf{NO}_{i,t} + \gamma_5 \ln \mathbf{VH}_{i,t} + \gamma_6 \mathbf{MS}_s + \zeta_t + \theta_i + \varepsilon_{i,t,s} \quad (6.3)$$

where $Pax\ km_{i,t}$ denotes the number of passenger kilometres (ridership).

6.5 Context and data

6.5.1 Institutional context

In the 1990's PT in the Netherlands has witnessed a decline of patronage, and increasing public deficits. This induced the Dutch government to change the regulative structure and impose regulative reform. Until 2000 contract were privately awarded to in-house operators. From 2000, competitive tendering was imposed upon regional and local PT concessions in a gradual way.[94] Exceptions were made for the four largest cities areas Amsterdam, Rotterdam, The Hague and Utrecht.[95] Except for Amsterdam, all bus concessions have been competitively tendered at least once. The contracts for metro and tram in The Hague and Rotterdam, and for bus, tram, and metro in Amsterdam are awarded to the 'in-house' operator owned by the municipality. The Utrecht concession was competitive tendered all the same.

A substantial concentration in PT governance and provision has occurred in the last 25 years. During this period the number of PTAs has declined from 56 to 14, the number of concession areas has declined from 74 to 39 and the number of PT-providers has declined from 25 to 10 (KNV, 2015; CROW-KPVV, 2015). Since 2001 foreign companies have acquired all Dutch owned regional bus companies. Except for the municipal firms in Amsterdam, Rotterdam, and The Hague and a firm affiliated to Netherlands Railways, to date there is not a single Dutch bus company left.

The agreements between the PTA and the operator are in most cases put down in net costs contracts in which the operator bears all commercial risk and receives total fare box revenues. As tariffs are governed by the regulators, in practice operators have little room to increase revenues. Contracts can be renewed via competitive tendering (publically awarded contract) or via (re)negotiation with the incumbent operator (privately awarded contract).[96] During the period 2001-2013, our period of study, a contract may change several times (in our data up to three times). Contracts are almost always awarded based on a weighted set of supply oriented output criteria (e.g., number of vehicle hours supplied), and quality criteria (such as new vehicles, interconnectivity, marketing efforts and sustainability). In addition

94 Most train services were exempted.

95 It was argued that public transport in these areas is provided as integrated networks of bus, tram and metro, and that ownership structure of infrastructure and rolling stock is complex and may form barriers to market entrance.

96 This applies to the contracts in the three largest cities and to contracts in the first phase of the new regulation, as legislation allows for exceptions for CT in the first years.

PTAs assess the bids on additional vehicle hours offered. Most PTAs use a best value for fixed subsidy procurement strategy. This implies that PTAs determine a minimum volume of vehicle hours – based on a estimation of market prices – or set a subsidy cap restraint.

6.5.2 Data collection and definitions

We have collected data by sending a survey to all 18 PTAs. Ten out of 18 PTAs reacted positively and collected annual data on operational costs, vehicle hours, type of contract, contract duration and contract renewal for 38 concession areas for the period 2001-2013. Based on additional sources, we have collected data on the operator in charge, firm ownership, deployment of new vehicles, passenger kilometres, number of stops and network length (CROW-KPVV, 2015; KNV, 2015; Zwart, 2012).[97] The sample appears representative for the Netherlands and contains small and large, as well as urban and rural concessions. For some years, some PTAs were not able to provide information, so we have unbalanced panel data. For the main specification, we have 301 complete observations.

We use structural expenditures on PT operations to measure costs.[98] We do not observe these expenditures directly, but these can be derived from the subsidy from the PTA to the operator and the operators' fare box revenues which are both reported. Given the assumption that the operators' profit margins are small, operational costs approximately equal to the sum of subsidies and fare box revenues.[99] Descriptives of the main variables of interest are shown in Tables 6.1 and 6.2. In Table 6.1, we present descriptives on operational costs, vehicle hours and passenger kilometres. In Table 6.2 we focus on contract renewals.

6.5.3 Descriptives

Table 6.1 shows that annual operational costs per concession area is about € 32 million, with a subsidy ratio of about 50%, average annual subsidy per concession area is € 14 million. Annual vehicle hours per concession is about 280,000 and annual ridership is about 116 million passenger kilometres.[100] Table 6.1 shows considerable standard deviations for these variables, due to differences in the size of the concession areas, but annual changes in these variables are small which justifies our log-linear approach.

97 Data on passenger kilometres are hard to get, as data collection on PT demand is not uniformed, definitions change during time, and data collection is very costly. This especially hold for obtaining time series and panel data on PT demand. We collected passenger kilometres for the period 2001-2011 on PTA level from an external source (Zwart, 2012), and adapted the data in order to obtain a complete as possible picture of the developments in PT demand. The adaption protocol is available on request.

98 This measure does not contain infrastructural investments costs, infrastructure maintenance costs, infrastructure levies and incidental operational expenses.

99 An additional argument which justifies our approach to measure costs is that ownership of PT firms usually is foreign, so profits (or losses) do not affect Dutch welfare.

100 Due to definition changes, the data on passenger kilometres are less reliable for some areas which changed from ticketing system during the period studied.

Table 6.1. *Descriptives*

	Number of observations	Mean	Std. Deviation
Operational costs (€ x 1,000)	301	31,625	35,302
Annual change operational costs (€ x 1,000)	265	915	3,026
Vehicle hours	332	283,432	241,503
Annual change vehicle hours	294	4,480	31,763
Passenger km's (x1,000)	363	115,969	103,786
Annual change passenger km's (x 1,000)	330	553	9,909
Subsidies (€ x 1,000)	350	14,319	15,698
Annual change subsidies (€ x 1,000)	311	-63	2,508

Table 6.2 shows that in the period under study, 61 contract were renewed. 69% of these were renewed after a process of competitive tendering of which about half were awarded to a new operator. The majority of contract renewals results in introduction of new vehicles. After renewal 24% of the concession areas are operated by a publicly owned firm and 44% by a multi-product firm that operates both bus and metro/tram/light train. Concerning type of contract the situation in the Netherlands is such that the majority of contracts that is renewed after 2001 is of the net costs-type. Our sample contains only net-cost contract renewals, therefore we cannot empirically test the effect of contract type. Although we do not have specific information on contract type, in the period before 2001 most contracts were gross-costs contracts. So, probably most of the first contract renewals in our sample are from a gross-costs to a net-costs contract, and therefore the result show the combined effect of first contract renewal and change of contract type. Subsequent contract renewals are always from net-costs to net-costs contracts. CT is implemented in the Netherlands in phases. This implies that not all contract are competitively tendered directly after 2001. However, once a contract is competitively tendered, also subsequent contract renewals are performed under competitive tendering. For 9 contracts in our sample we are also able to assess the effect of change from a negotiated (privately awarded) contract to a CT contract.

Table 6.2. *Descriptives on contract renewal*

	Number of observations	%
Renewed once	15	25
Renewed twice	40	66
Renewed three times	6	10
Renewed via competitive tendering	42	69
New operator and competitive tendering	22	36
New vehicles	31	51
No new vehicles	19	31
New vehicle unknown	11	18
Contract awarded to multi-product firm	27	44
Contract awarded to publicly owned firm	15	24

6.6 Results

6.6.1 Operational costs, subsidies, contract renewal, and tendering

In Table 6.3 the main results are provided for several specifications of the operational costs model (specifications 1 thru 4) and for a specification using subsidies as dependent variable (specifications 5 and 6). In all specifications, we include number of contract renewals, competitive tendering, log vehicle hours, year and concession area fixed effects.[101] We define a contract to be competitively tendered in case the contract is awarded after a public procedure in contrast to a private awarding procedure. Importantly, by including concession area fixed effects we improve on existing studies as we are able to control for time-invariant area related unobserved variables (e.g. residential density). In Appendix 6.B we show the results of two pooled OLS analyzes in which the area fixed effect are excluded, and show that excluding these fixed effects result in extremely biased coefficients.

101 The year fixed effects, not presented here but available on request, show an annual autonomous 3.8% growth of operational costs. This growth reflects increases in input prices for the operator (e.g., wages, gasoline prices, bus prices). So the identified effects of contract renewal are not due to market developments in input prices.

Table 6.3. *Operational costs and subsidies*

	(1)	(2)	(3)	(4)	(5)	(6)
	Costs	**Costs**	**Costs**	**Costs**	**Subs.**	**Subs.**
Contract renewed minimal once	-0.105 ***	-0.092***	-0.088***	-0.057**	-0.198**	-0.229***
	(0.025)	(0.026)	(0.025)	(0.027)	(0.079)	(0.078)
Contract two or three times renewed	-0.061 **	-0.064***	-0.060***	-0.048**	-0.208***	-0.241***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.060)	(0.059)
Contract three times renewed	-0.037	-0.039		0.017	0.346**	
	(0.046)	(0.048)		(0.049)	(0.150)	
Contract competitively tendered	0.016	-0.006	-0.009	-0.032	0.144	0.172*
	(0.024)	(0.033)	(0.033)	(0.033)	(0.104)	(0.104)
New operator		0.081***	0.079***	0.067***	0.118**	0.135**
		(0.023)	(0.023)	(0.023)	(0.069)	(0.069)
New vehicles unknown		-0.032	-0.038	-0.042	-0.031	0.014
		(0.035)	(0.034)	(0.034)	(0.108)	(0.107)
Partly new vehicles		-0.015	-0.015	-0.037	-0.082	-0.084
		(0.034)	(0.034)	(0.033)	(0.103)	(0.104)
All new vehicles		-0.031	-0.029	-0.015	-0.189*	-0.207**
		(0.033)	(0.033)	(0.032)	(0.102)	(0.102)
Vehicle hours (log)	0.425***	0.446 ***	0.445***	0.379***	0.222	0.234
	(0.052)	(0.054)	(0.054)	(0.056)	(0.160)	(0.161)
Single product firm				-0.085***		
				(0.021)		
Year fixed effects (13)	Yes	Yes	Yes	Yes	Yes	Yes
Concession area fixed effects (38)	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	301	301	301	301	325	325
R^2	0.997 2	0.997 35	0.997 34	0.9975	0.970 1	0.9695

*Note: specifications (1) thru (4): dependent variable is logarithm of operational costs, specification (5) and (6) dependent variable is logarithm of subsidies. ***p < 0.01, ** p< 0.05, *p<0.1. Standard errors between brackets*

The results of the most basic specification (1), imply that contract renewal leads to a substantial reduction in operational costs. When a contract is renewed at least once, costs fall by 10%. Contracts renewed at least twice even lead to an extra costs reduction of 6%, so a total reduction of 16%. A third contract renewal seems to reduce operational costs even further, but this effect is statistically not significant at conventional significance levels, probably due to the small number of observations referring to three renewals (see Table 6.1). These results imply that the effect of contract renewal is diminishing; the first time a contract is renewed yields the greatest operational costs reduction. This finding is in line with previous studies (Hensher and Wallis, 2005; Preston and Almutairi, 2013). We do not find evidence that there is an effect of competitive tendering on operational costs, suggesting that the threat of CT is sufficient in a market where the majority of concessions is competitive tendered. The estimated coefficients of (log) vehicle hours show that considerable economies of density exist within concession areas.[102] The costs elasticity of density is about 0.40, i.e. if production volume on a given networks is increased 10%, costs increase by 4%. Economies of density are often found in empirical studies on network industries such as transportation (see, for instance, Berechman and Guiliano, 1985; De Borger et al., 2002; Farsi et al., 2007 and Gautier and Yvrande-Billon, 2013).

In specification (2), we have added two additional variables to capture changes in the terms of the contracts by including new operator (as proxy for quality aspects in the contract) and new vehicles dummies. It appears that the effects of contract renewal and CT hardly change. In case the CT procedure leads to a new operator, PTAs are willing to pay 8% more compared to the situation the incumbent stays in charge. This effect is unlikely induced by differences in factor input prices between operators, as all bidders (new and incumbents) have to abide legislation, and act under the same conditions.[103] Our explanation for this finding is that, within the boundaries of a subsidy cap set by the PTAs, new operators perform better on quality criteria compared to the incumbent, although their costs are higher than the incumbents'. We do not find evidence that there is an effect of new vehicles on operational costs.[104] One possible explanation is that vehicle costs are only a small part of operating costs (for bus exploitation about 15 to 20%, see Koolen and Stoelinga, 2005). Moreover, operators don't own the fleet, but lease the vehicles and have a return arrangement with the lessee, and therefore the risks are shared. Gautier and Yvrande-Billon (2013) reach similar results, as they show that bus fleet age hardly impact costs. As the number of contracts three times tendered is very

102 Network size remains largely constant in the concession areas (see Table 6.1), so it is more appropriate to interpret this result as economies of density rather than economies of scale. We emphasize that the model specification (1) does not allow for comparisons of economies between concession areas, only within these areas. To assess optimal size of concession areas, in section 6.6.3 we further elaborate on economies of density between concession areas.

103 These result indicate regulatory schemes and operators' efficiency levels are exogenous, and that the operators information level about its technology, and its efforts to reduce costs are greatly unobserved by the PTA. Therefore the theory of regulation under asymmetric information may apply to the PT industry in the Netherlands (see Laffont and Tirole, 1993, and Gagnepain and Ivaldi, 2002).

104 The data for the new vehicle variable originates from additional data sources and are likely of lower quality than those based on our questionnaire, so it is likely that the variable 'new vehicles' has measurement error. In general, if an explanatory variable has (random) measurement error, than the estimated coefficient is biased towards zero, which may explain these results.

small, we have excluded this variable in specification (3). The results hardly change, so we may conclude that this variable does not add to the specification.

Economies of scope across modes may occur when firms supply bus as well as tram/metro (in a certain year). In our sample, the (municipal) firms in Amsterdam, Rotterdam and The Hague supply both bus, and tram/metro. Furthermore, a number of other firms provide bus and light train services, but not necessarily in the same concession area. To examine economies of scope, we distinguish between multiple and single product firms. Multi-product firms supply both bus and rail either within the same concession area or in an adjacent concession area. Otherwise, a firm is defined as a single-product firm. In this way we are able to differentiate between concession areas operated by single-product or multi-product firms.[105] Specification (4) of Table 6.3 shows that controlling for single/multi-product firms reduces the effect of contract renewal somewhat (by about one third). Furthermore, we do not find evidence for positive economies of scope. To the contrary, the results indicate diseconomies of scope; multi-product firms have 8.5% higher costs than single-product firms. Multi-product firms seem to be less cost efficient than single-product firms. This implies that in case PTAs aim to procure multi-modal PT services and restrict providers to multi-product firms, the change is they have to pay more subsidies than if they split procurement by mode, in that way making also bids possible from specialized single-product firms.[106] Our results are in line with Di Giacomo and Ottoz (2010), who also report diseconomies of scope across bus and rail modes, but are not in line with Farsi et al. (2007) who find economies of scope. Finally note that we do *not* control for the size of the firm. If it is true that the size of firms is systematically related to whether a firm is a single or multiproduct firm, then our results will be biased.

One may expect the effect of contract renewal on subsidies is more pronounced than on operational cost as the average subsidy to operational costs ratio is about 50%. As shown in specification (5) of Table 6.3, contracts that are renewed at least once induce a 20% fall in subsidies. An additional contract renewal leads again to a 20% subsidy reduction. Contracts three times renewed shows an opposite – subsidy-increasing – effect, possibly because only two contracts in our sample are three times renewed. As we did in specification (3), in specification (6) we also excluded contracts three times renewed. Again, the results hardly change.

As mentioned above we control for all unobservable time-invariant spatial and network effects by using concession areas dummies in our models. Previous cross-section studies use specific control variables (Farsi et al., 2007: number of stops, Karlaftis, 2010: network length, Boitani et al., 2013: GDP per capita, Gautier and Yvrande-Billon, 2013: network length and speed, Piacenza, 2006: commercial speed, Karlaftis and Tsamboulas, 2012: area surface,

105 For convience we assume firms to be in one group for the whole of the contract duration.In reality the latter will not always be true as mergers and acquisitions did occur especially during the period 2007-2011.

106 However, if not from a cost perspective, from a consumer perspective multi-modal concessions may be preferred, as interchanges between bus and rail can be better organized

population density, Margari et al., 2007; average commercial speed, population density, Sakai and Takahashi, 2013, average route length, Scheffler et al., 2013: population density, network length). These empirical studies indicate that many factors determine operational costs, so that it seems merely impossible to capture all possible individual influences using cross-sectional data, and results therefore will be easily biased.

To demonstrate the importance of including area fixed effects, we have also estimated the model without these fixed effects (a so called 'pooled OLS'), see Appendix 6.B. We find that the results are extremely biased. This also occurs when one controls for network variables such as the number of lines and stops and the degree of urbanization. For example the results in these specifications erroneously indicate that the first round of contract renewal leads to an increase in operational costs. These results are clearly biased because the identification of the parameters of interest is based on cross-sectional variation in the data.[107]

In Section 6.2 we discussed the question of the most appropriate output measure for public transport efficiency. Should PT output be defined in demand, or in supply parameters? In Appendix 6.C – using a slightly smaller set of observations – we report results when using passenger kilometres and vehicle hours as output measure. It appears that the coefficient for passenger kilometres is not significant, indicating the influence of ridership on operational costs is absent and using a supply oriented measure for PT performance seems therefore the most appropriate.[108] As is common in these types of models, the R^2 of all models is (very) high, because variation over time in the dependent variable, given the year fixed effects, is limited, as the area fixed effects capture most cross-sectional variation. Importantly however, the value of the R^2 is unrelated to the question whether coefficients are consistently estimated.

6.6.2 Passenger kilometres, contract renewal, and tendering

We aim to test now the hypothesis that contract renewal and CT may work out favorable on PT quality and provision, and therefore may enhance ridership, using the logarithm of number of passenger kilometres as dependent variable. The results are shown in Table 6.4.

107 As a sensitivity analysis, Specification (3) in Appendix C includes concession area fixed effects and network variables. It shows that adding 10% more stops to a network results in 2% higher costs. Note that the variable grade of urbanization is omitted in this model as this variable is time-invariant.

108 In Section 6.2 we also indicated firm ownership may be of influence on operational costs. We tried to assess the effect of a change of ownership of the firm, however, due to multicollinearity, we were not able to estimate the appropriate function. We may however indicate we control our models for the effect of new ownership.

Table 6.4. *Passenger kilometres*

Contract renewed min. once	0.076 ***
	(0.030)
Contract two or three times renewed	0.043*
	(0.025)
Contract competitively tendered	-0.028
	(0.032)
New operator	-0.040
	(0.026)
Logarithm of vehicle hours	0.173 ***
	(0.060)
Year fixed effects (13)	Yes
Concession area fixed effects (38)	Yes
Number of observations	270
R^2	0.997

*Dependent variable log passenger kilometres. ***p < 0.01,** p< 0.05, *p<0.1 Standard errors between brackets*

Our results indicate a significant positive effect of contract renewal on (log) passenger kilometres. Renewing a contract results in a 7.6% increase in passenger kilometres. A second contract renewal leads to an additional 4.0% increase in passenger kilometres, so contracts that are renewed at least twice induce a 12% increase in passenger kilometres. Unfortunately we do not have accurate information on fare box revenues per year per concession. However, we may assume that the increase in passenger kilometres due to contract renewal implies also an increase in revenues. As we showed that contract renewal works out positively on operational cost, we may conclude that contract renewal favors the cost-benefit ratio. Further, there is no effect of CT and new operator on passenger kilometres. The vehicle hours elasticity of demand is small (equal to 0.17). Hence, a 10% increase in vehicle hours leads only to a 1.7% increase in passenger kilometres. Similar results are found in many other studies (see e.g. Goodwin, 1992; Holmgren, 2007; Currie & Wallis, 2008).

6.6.3 Size of concession areas

Viewed from a policy as well as from an economic perspective, the geographical size of concession areas governed by public transport authorities is of interest (De Borger et al., 2002; Farsi et al., 2007; Croissant et al., 2013). In the Netherlands, after the first round of tendering, this size has sharply increased, which raises the question what is the economic rationale for this concentration drive. As Croissant et al. (2013) frame it: is allotment a useful strategy for PTAs? The idea of geographical allotment is not only relevant in economic theory, but also in governance and institutional theory. The latter assumes that regulation and transaction costs in larger areas may be less than in smaller ones. On the other hand, dividing the network in smaller lots may increase competition, and yield a better market/social outcome.

To analyze concession area size, we estimate a between-area fixed effects model, and analyze economies of scale controlling for time-invariant area specific differences such as urban density, network characteristics (number of lines and stops) and PT mode on operational costs.[109] In this way, we compare concession areas which each other using the – over the years – average values of variables. As the differences between costs and production levels between concession areas are substantial, we now use a translog flexible functional form, with vehicle hours as dependent variable.[110] We allow for production heterogeneity (some firms produces multi-products; i.e. bus and rail) by including a PT mode dummy in the specification.

$$lnOC_i = \alpha + \beta_q(lnQ_i - ln\bar{Q}_i) + \frac{1}{2}\beta_{qq}(lnQ_i - ln\bar{Q}_i)^2 + \beta_d lnD_i + \beta_l lnL_i + \beta_s lnS_i + \beta_m M_i + \epsilon_i \quad (6.4)$$

where *OC* denotes operational costs, *i* denotes a concession area, *m* denotes mode, *Q* denotes production in vehicle hours, $\bar{Q}$ denotes production of sample mean, *D* denotes urban density of the concession areas measured as number of inhabitants per hectare build-up area, *L* denotes number of lines, *S* denotes number of stops, *M* denotes a mode dummy, and ε_i denotes the standard error term.

In Table 6.5, the descriptives per concession area are given for bus and rail separately. It appears that in terms of costs and vehicle hours, mean values of bus and rail are of similar order of magnitude, indicating that controlling for PT mode using a dummy indicator in Eq. (6.4) is not problematic. The results of Eq. (6.4) are shown in Table 6.6.

Table 6.5. *Descriptives*

	BUS			**RAIL**		
	N	**Mean**	**Std. Deviation**	**N**	**Mean**	**Std. Deviation**
Operational costs (€ x 1,000)	24	24,388	19,716	9	39,200	49,666
Vehicle hours	26	327,258	224,902	9	179,704	270,286
Urban density	26	33.6	13.0	12	38.7	14.8
Number of lines	26	30.4	19.9	11	3.7	5.0
Number of stops	26	929.3	625.8	11	107.7	194.4

Note: mean values per area, price level 2000.

109 In the previous section we showed that economies of density within concession areas exist using within fixed effects. This does however not imply that larger concession areas perform better than smaller. Therefor in this section we estimate between-area fixed effects.

110 The translog functional form is a second order approximation to an arbitrary costs function. We use the sample mean as the approximation point.

Table 6.6 *Economies of scale (between-area fixed effects).*

Vehicle hours (log)	1.165 ***
	(0.140)
1/2 vehicle hours $(\log)^2$	-0.009
	(0.059)
Urban density (log)	-0.242
	(0.222)
Bus	0.632 *
	(0.326)
Number of lines (log)	-0.238
	(0.206)
Number of stops (log)	0.025
	(0.159)
Constant	17.838 ***
	(1.078)

*Note: 38 concession areas. ***p < 0.01,** p< 0.05, *p<0.1 . Standard errors between brackets*

These results indicate PT in the Netherlands is produced under a situation of constant economies of scale: the effect of log vehicle hours is 1.165 but not significant different from one where the square of log vehicle hours is essential zero.[111] If any economies exist, slight diseconomies of scale are most probable. This implies that concessions areas – whether big or small – are of optimal size, and operational costs cannot be reduced by changing the size of the concession area by geographical allotment.

6.7 Conclusion and policy implications

In this chapter, we have estimated the effects of regulatory change on operational costs, subsidies and ridership. We contribute to existing literature by using a panel dataset that makes it possible to control for all unobserved time-invariant network and area characteristics. We show that contract renewal substantially reduces operational costs as well as subsidies and increases ridership. We also found strong economies of density within networks, constant returns to geographical scale and diseconomies of scope across PT modes. Based on our results one may conclude that the policy of the Dutch administration, which aims to increase efficiency and ridership in the PT sector by means of competition, is successful. Contract renewal under a CT regime leads to decreasing subsidies, and increasing ridership. Our results are however not conclusive. We find that the immediate effect of competitive tendering is absent, suggesting that the threat of CT is sufficient in a market when the majority

111 We also performed separate analyzes for bus and rail. We found very similar results.

of concessions are competitive tendered. To study the effects of threat of competition in more detail, it is recommended to perform an analyses on the content, procedures, and political pressure exerted of publicly awarded contracts. Our economies of scale results indicate that the geographical size of the current concession areas may not be altered without additional costs. From an operational costs perspective, our study suggests that there is no reason to increase (or decrease) the geographical size of concession areas.

PTAs have an important role setting the network conditions over which the PT services are performed, therefore, besides regulatory policies as discussed in the current chapter, PTAs should develop and implement infrastructural policies. If network conditions become too unfavourable, even under a competitive tendering regime, costs and subsidy effects of contract renewal may run out. Therefore PTAs should also focus on sustaining excellent network conditions, thereby aiming at increasing free flow for the PT system. Finally, our analyses focus on costs, subsidies and ridership only. Therefore we cannot assess total welfare effects of regulatory change. We for instance did not analyze transaction and monitoring costs in detail (although in Chapter 3 an estimation is given), changes in level-of-service, fare increases and external effects. To assess the overall effects of regulatory change in the Netherlands, one needs a comprehensive social welfare analysis (see Preston and Almutairi, 2013; Hensher and Wallis, 2005, and Gagnepain and Ivaldi, 2002).

Appendix 6.A. Network characteristics of concession areas

Year	Mean number of lines	Mean number of stops
2003	23	700
2004	22	693
2005	22	687
2006	22	675
2007	22	665
2008	22	674
2009	24	700

Appendix 6.B. Additional analyzes

	(1)	(2)	(3)
	Costs	Costs	Costs
Contract renewed minimal once	0.416***	0.247**	-0.094***
	(0.114)	(0.114)	(0.033)
Contract two or three times renewed	-0.018	0.057	-0.051*
	(0.082)	(0.114)	(0.030)
Contract three times renewed	-0.162		
	(0.289)		
Contract competitively tendered (1) vs. privately (0) awarded	-0.713***	-0.407***	0.007
	(0.084)	(0.105)	(0.037)
ln # lines		-0.427***	-0.099
		(0.110)	(0.064)
ln # stops		0.089	0.195***
		(0.085)	(0.073)
ln grade of urbanization (inh/ha)		0.000	n.a
		(0.137)	
Logarithm of vehicle hours	0.713***	1.051***	0.446 ***
	(0.026)	(0.078)	(0.078)
Year fixed effects (13)	Yes	Yes	Yes
Concession area fixed effects (38)	No	No	Yes
N	301	175	175
R2	0.844	0.889	0.997

****p < 0.01,** p< 0.05, *p<0.1 Standard errors between brackets.*

Appendix 6.C. Sensitivity analysis on the choice of output measure of PT performance

	Costs
Contract renewed minimal once	-0.091***
	(0.027)
Contract two or three times renewed	-0.055**
	(0.022)
Contract competitively tendered	0.023
	(0.026)
Logarithm of vehicle hours	0.407***
	(0.056)
Logarithm pax km's	-0.044
	(0.084)
Year fixed effects (13)	Yes
Concession area fixed effects (38)	Yes
N	245
R2	0.998

*Note: dependent (log) operational costs. ***p < 0.01, ** p< 0.05, *p < 0.1. Standard errors between brackets*

Chapter 7

Conclusion

7.1 Summary

The importance of public transport (PT) in sustaining a healthy and accessible environment is increasingly recognized. Especially in cities, PT can contribute to solving congestion and reducing CO_2 and other pollutants. However, PT services in general are only viable with the aid of large amounts of public subsidies. This PhD thesis is mainly focused on assessing the effects of introduction of a competitive regulatory regime that was introduced in Netherlands as of the year 2001 on urban and regional PT. This policy change may be essential in achieving efficiency and equity objectives. The approach is based on theoretical and empirical analyses of the economic structure and of the institutional context of PT before and after introduction of competitive tendering (CT). The aim of this study is to clarify the complex mechanisms underlying PT provision, demand and finance in a (de)regulated environment dominated by CT, and to assess the impact of CT on passengers, PT authorities, and PT firms.

In Chapter 2 the costs and production characteristics of the public transport (PT) industry, arguments for the regulation of PT and competitive tendering (CT) are introduced. Essentially, government organizations impose regulatory policy as a mechanism to minimize economic inefficiency and to make welfare distribution more equitable.[112] Why regulate the PT industry? General theory puts forward three main fields of argument for regulation: economic, social and political. The most used economic arguments (the efficiency principle) are that PT holds characteristics of public goods and natural monopolies. However, these claims have also attracted criticism. Social arguments (the equity principle) for regulation refer to PT being essential for individuals' basic welfare, and a non-discriminating service supply has to be assured by regulation to protect the weak and the poor in society and to strengthen social cohesion, safety and public health. Apart from these theoretical arguments, in practice the prime political determinants for imposing regulation on the PT sector are general budget deficits. Section 2.1 demonstrates that public transport services are network services and are tied to public welfare. In regulating PT services, two potential dangers lie in wait. The first danger is under-regulation and market failure due to natural monopoly features and the irreversibility of investments. The second danger is service meritorization (regulatory failure) as PT has characteristics of a public good Therefore, the classic first-best solution (self-regulation by means of competition within the market) cannot simply be transposed onto these services and the second-best solution may be optimal. The proponents of regulation and of deregulation agree that, under the condition that the network remains a (natural) monopoly, separating the network and the services over the network may be a feasible second-best solution. There is also agreement on the argument that, if public control is deemed necessary to control a market, when properly used, competitive tendering (CT) may be an instrument by which control can be carried out with the minimum loss of efficiency. A number of authors argue that the transaction cost

112 We define 'regulation' as a situation in which a public entity imposes restrictions on one or several key decision parameters for PT firms.

of tendering should be accounted for when assessing the effects of competitive tendering. Transaction costs have a bearing, for instance, on the costs incurred by the invitation to tender, pre-selection of suppliers, drafting procedures and award criteria, preparing bid documents, processing the bid competition and monitoring performance.[113]

Chapter 3 serves as an introduction to PT in the Netherlands. The chapter describes the way in which CT in the Netherlands is implemented and provides quantitative information on the demand, supply, governance and transaction cost of competitive tendering (CT). In the Netherlands, as in many European countries, CT became popular in a period of economic crisis and rising general budget deficits in the mid-1970s and 1980s. This regulatory instrument was deemed potentially able to reduce costs and subsidies. Discussions on the new governance arrangements for the PT industry in the Netherlands resulted in legislation that became effective in 2000. This legislation provided for phased CT introduction. Importantly, regional authorities have the freedom to retain the responsibility for decisions at the tactical level (the planning function) at their own discretion or to position them with the operator. This is a crucial decision in market performance terms and functioning as it defines whether the PTA or the operator is in control of designing the route network, fare level and level of service. Although the aim of the central government is to position this planning function with the operators, the practice is that nearly all PTAs keep tight control, resulting in a situation in which the primary decisions concerning PT planning are made by politicians. Section 3.2 discusses the mobility market in the Netherlands since 2000 and shows the relatively small share of bus, tram and metro (BTM) use: on average about 3% and in heavily urbanized areas 6%. The BTM supply (number of vehicle kilometres) has increased steadily over time, which may be linked to the introduction of competitive tendering. As the demand does not keep pace with the supply, the occupancy rate of BTM declined after 2006. After 2002 the growth in the fare box revenues exceeded the growth in the total operational expenditures, implying that the cost recovery ratio improved. The descriptive statistics on CT show a sharp decrease in the number of incumbent firms that stay in charge after competitive bidding (83% in 2002; 33% in 2014), indicating that the market is maturing and the comparative first-mover advantage is diminishing. Over time, the contract duration has increased considerably. The average number of bids per procedure has remained more or less the same (around three); however, in more than 80% of the procedures only the three big contenders (Arriva, Veolia and Connexxion) placed a bid and took a share of the market. Public governance became more concentrated after 2000 (fewer concession areas and fewer authorities). In Section 3.4 the level of transaction costs (TC) is estimated. Data on individual contractual arrangements and on transaction costs are not publically available. We assessed these costs based on a selected sample of respondents. Although this information could not be validated, it provides us with valuable estimates of the level of transaction costs of both parties involved (PTAs and operators). For the period 2001–2015 the average annual ex

113 In Section 3.4 transaction costs of PT services' procurement by way of competitive tendering are estimated.

ante transaction costs of competitive tendering amounted to 14.4 million euros to 35.5 million euros.[114] In small concessions the transaction costs may account for 17% to 40% of the subsidy savings of first-time contract renewal initiated by competitive tendering.[115] As the contract volumes increase, the relative share of transaction costs decreases, but in large concessions (with an average annual subsidy of €54.3 million) the share still amounts to 4% to 10%.[116] We assess the potential problems generally associated with CT in the Netherlands. The PT services to be procured are complex in nature, and so are the contracts. Due to this complexity there are indicators for opportunistic bidding and operator-led contract renegotiation. Ownership of the infrastructure and other assets is often mentioned as a potential problem for CT. For the Netherlands, however according to our informants, this is probably less of an issue as the ownership of the most strategic PT infrastructure (terminals, stations, shelters and traffic and information systems) is in public hands. The award procedures leave room for subjectivity and ambiguity, as authorities use a mix of qualitative and quantitative criteria to assess the bids. Problems associated with award criteria and procedures are the main reason for legal disputes. Our respondents believe that the substantial bidding costs (see Table 3.7) are the main trigger for these lawsuits. Bidding imparity is not much of a problem as PTAs provide as much as possible a level playing field,. Finally, there is no evidence of colluding bidders.

In Chapter 4 the focus is shifted to the customers, that is, the users of public transport. In this chapter the drivers of customer satisfaction with PT are studied. Information on these drivers is relevant to PTAs aiming to enhance customers' orientation of operators. Based on the literature a theoretical model using satisfaction concepts is constructed first. Customers' assessments of a service depend on the balance between sacrifices and benefits, both monetary and non-monetary. Moreover, passengers' perceptions of the different service aspects are heterogeneous. In this thesis a distinction is therefore made between customers' satisfaction with respect to specific transactions or service encounters and customers' global or overall evaluation of a service. Transaction-specific satisfaction (TSS) is the result of a cognitive judgement of transactional service encounters. In contrast, overall satisfaction is a more holistic construct and is an affective/emotional response to a perceived discrepancy between expectations and perceptions after a service delivery experience. The relationships between attribute-based satisfaction, overall satisfaction and customer characteristics are modelled by means of linear models. Overall (OS) and transaction-specific satisfaction (TSS) are modelled as a function of customer characteristics. Accounting for customer characteristics, OS as a function of TSS is also modelled, thus deriving importance scores (or weights) of the service attributes for several customer groups. The models are controlled for interaction between TSS and customer characteristics. A number of hypotheses are tested using data on the

114 The 2015 price level.

115 The costs and subsidy savings due to contract renewal under a CT regime are estimated in Chapter 6.

116 We expect the transaction costs of CT to increase in the future as the trend in the Netherlands is towards longer-term, higher-volume contracts.

satisfaction scores of urban and regional Netherlands PT users (bus, tram, metro and regional train) for the years 2010 and 2011 (N = 90.000 annually). The average level of overall satisfaction is 7.28. Positive outliers are satisfaction with the attributes *seating capacity, ease of boarding and alighting* and *safety on board*. Negative outliers are satisfaction with the *on-board information on delays* and the *prices of the tickets*. The results show that the attributes *travel speed, on-time performance* and *service frequency* (core attributes) are the most important determinants of overall satisfaction. It is also noteworthy that the interactional attributes *personnel behaviour* and *driver's behaviour* and the physical attribute *vehicle tidiness* are considered to be very important. Furthermore, interaction models with respect to the customer segmentation variables age, mode choice, experience with negative social safety experiences (NSSEs) and urban density are estimated. The results demonstrate that elderly people attach significantly more weight to service frequency, implying that a policy aimed at increasing the service frequency in general mainly affects the attribute importance of elderly people. As the service frequency is linked to the waiting time, this may be correlated with the declining physical condition of elderly people. The results also show that elderly PT users place more emphasis on the attributes *price, on-time performance* and *service frequency*. The PT mode choice significantly affects the satisfaction levels and attribute importance, the latter especially with regard to *on-board information on delays*. Replacing bus lines with metro lines will lead to a radical change in attribute importance, to an increase in satisfaction with *on-board information on delays* and to a decrease in satisfaction with *on-time performance* and *service frequency*. Negative critical incidents may play an important role in service. We show in Table 4.6 that NSSEs in PT in the Netherlands have a significant negative effect on the overall and attribute-level satisfaction. These are remarkable results: passengers who have experienced one or more social safety incident not only rate satisfaction as relatively low for the attributes *safety during the trip* and *safety at stops*, but also are less satisfied with attributes that have no relationship with the incident itself. With regard to the impact of urbanization on satisfaction, Table 4.6 reveals that the level of satisfaction with eight out of eleven significant service attributes is lower in highly urbanized areas than in areas of low urbanization. The relatively low level of satisfaction with the attribute *personnel behaviour* in highly urbanized areas is also striking. It is concluded that, especially in highly urbanized areas, it is probable that increasing the frequency of services and introducing new vehicles will have a major positive effect on the level of satisfaction. Our empirical results do not support the hypothesis that the composition of satisfaction between user groups differs significantly.

The relationship between tendering and efficiency is widely studied. Few studies, however, focus on the relationship between tendering and satisfaction. In Chapter 5, building on the results and insight of Chapter 4, that relationship is explicitly investigated. We test whether the introduction of competitive tendering (CT) in the Netherlands affects passenger satisfaction. The model building and analysis are based on the comparison per year-pair of regions tendered versus regions non-tendered (in that specific year-pair). Thus, the effects on satisfaction of tendered regions relative to non-tendered regions are compared controlling for area fixed effects. With regard to the effect of CT on overall trip satisfaction the findings demonstrate for the period 2001-2010 that, after controlling for year and area fixed effects, in 58% of the tendered regions the overall satisfaction increased after tendering, whereas in the other 42% the overall satisfaction decreased (compared with non-tendered regions). The net difference in the overall satisfaction for regions with and without CT is only +0.06 points (on a 10-point scale). The results also show that the first round of CT has a more positive effect on satisfaction than the second round of tendering. We show that this is a matter not only of sequence (first versus second), but also of timing (early versus late). The attributes that contribute most to this change in overall trip satisfaction are *vehicle tidiness, on-board noise, ease of boarding/alighting from the vehicle* and *service frequency*. The attribute values that are linked to *information* and to *on-time performance* in the tendered regions are negative, meaning that the satisfaction with these attributes worsens in comparison with non-tendered regions. The common factor in these attributes is that they all refer to reliability. If the attributes are weighted, these results become more pronounced. It is tempting to assign the above-mentioned outcomes to the implementation of tendering itself, but caution is necessary. In Section 5.8, therefore, the effect of new vehicles and new operators on satisfaction is examined, as these variables are closely related to the introduction of CT. The analysis shows that new vehicles greatly affect the overall satisfaction as well as the satisfaction with many attributes. The greatest impact on satisfaction from the introduction of new vehicles pertains to the satisfaction with the *vehicle tidiness, prices of the tickets, on-board noise, seating capacity, information provision on stops* and *travel speed*. It is striking that the introduction of new vehicles affects both vehicle-linked and non-vehicle-linked attributes. It is therefore plausible that new vehicles contribute to a positive general perception of public transport use. It is concluded that the satisfaction difference between tendered and non-tendered areas is mainly caused by new vehicles' introduction. This introduction is accelerated by the process of CT. If, due to CT, a new operator takes over from the incumbent, the overall satisfaction and attribute satisfaction are negatively influenced. Finally the satisfaction with the attribute *on-time performance* seems to take hardly any advantage from CT; it even seems to be negatively related to tendering. Apparently the shift in attention in the tendered regions to the attributes *speed* and *frequency* have had adverse effects on (the satisfaction with) *on-time performance*.

In Chapter 6 the effects of regulatory change on operational costs, subsidies and ridership are estimated. We employ panel data for the period 2001–2013 on the level of concession areas and take the most relevant contract attributes into account. We showed in Chapter 2 that network characteristics exert a strong impact on PT efficiency, and it is therefore essential to control for (exogenous) relevant network characteristics. Most cross-section studies control for variables such as network length, average speed and number of stops and lines; however, other network characteristics that may influence firms' productivity cannot be assumed away, thereby potentially biasing these analyses. Our approach avoids this issue by using panel data with concession area fixed effects, thereby controlling for all the time-invariant unobserved area circumstances that may influence the production efficiency of the firm. A literature review of the most important factors that affect PT efficiency is performed. We discover four influential determinants: awarding mechanism, contract type, firm ownership and network conditions. Additionally the literature review reveals an ongoing debate on the most appropriate output measure for assessing PT performance. The literature results indicate that CT effectively increases firm efficiency and decreases subsidy transfers. The contract type is seen in many studies as a powerful instrument for PTAs to govern transactions with the operator, as contracts make it possible to introduce specific incentives. The literature review reveals that high-powered incentive contracts (especially gross costs contracts) seem to perform best regarding efficiency. According to economic theory, firm ownership matters because public companies tend to be less efficient than private companies, as their deficits are covered by authorities. The empirical evidence on the efficiency effects of ownership in the PT sector, however, is not conclusive. In PT the network conditions are important. Many studies reach the conclusion that economies of scale and density are relevant; however, the studies present a wide diversity of outcomes. Robust results are that firms in small PT networks produce inefficiently and scale elasticities decrease as production increases. PT output may be measured in terms of supply, such as vehicle hours or vehicle kilometres, or in terms of demand, such as passenger kilometres or passenger trips.[117] Supply-oriented measures are most commonly used in empirical efficiency studies, albeit not reflecting the economic motive of PT supply. On the other hand, demand-oriented measures ignore the point that input factors such as fuel and labour do not vary systematically with the demand. In addition the demand for PT can only be influenced by PT operators to a limited degree. Other authors conclude that supply-oriented parameters are preferable. We use both measures in our empirical application. We develop a number of econometric models to assess the efficiency in PT in the Netherlands and are able to apply the models to a panel data set that we collected for this study. Importantly, by including concession area fixed effects, we improve on the existing studies as we are able to control for time-invariant area-related unobserved variables, such as residential density and network length. In Table 6.3 the main results are provided for several specifications of the operational costs model and for a specification using subsidies as the dependent variable. The results of the most basic specification imply that contract renewal leads to a substantial reduction in operational costs. When a contract is renewed at least once, the costs fall by 10%. Contracts that are renewed at least twice even lead to an extra cost reduction of 6%, resulting in a total reduction of 16%. A third contract renewal seems to reduce the operational costs even further,

117 The choice of output measure matters considerably. Berechman and Giuliano (1985) estimate a cost function using a demand-related output measure, as well as a supply measure, and show increasing returns to scale for the former and the opposite for the latter.

but this effect is not statistically significant at conventional significance levels. These results imply that the effect of contract renewal diminishes over time. We do not find evidence that there is an effect of competitive tendering on operational costs, suggesting that the threat of CT is sufficient in a market in which the majority of concessions is competitively tendered. The estimated coefficients of (log) vehicle hours show that considerable economies of density exist within concession areas.[118] The cost elasticity of density is about 0.40, meaning that if the production volume on a given network increases by 10%, the costs will increase by 4%. In the case in which the CT procedure leads to a new operator, PTAs are willing to pay 8% more than in the situation in which the incumbent stays in charge. This effect is unlikely to be induced by differences in factor input prices between operators, as all bidders (new and incumbents) have to abide by the legislation and act under the same conditions.[119] Our explanation for this finding is that, within the boundaries of a subsidy cap set by the PTAs, new operators perform better on quality criteria than the incumbent, although their costs are higher than those of the incumbent. In Section 6.2 we show that in efficiency studies a supply-oriented measure of PT performance seems to be the most appropriate. Our results indicate a significant positive effect of contract renewal on (log) passenger kilometres. Renewing a contract results in a 7.6% increase in passenger kilometres. As regards the optimal size of concession areas, our results indicate that PT in the Netherlands is produced in a situation of constant economies of scale: the effect of (log) vehicle hours is 1.165 but it is not significantly different from one in which the square of (log) vehicle hours is essentially zero. If any economies exist, slight diseconomies of scale are most probable. This implies that the concessions areas – whether big or small – are the optimal size and the operational costs cannot be reduced by changing the size of the concession area by geographical allotment.

7.2 Conclusions

7.2.1 Conclusions on the research questions

With regard to the research questions that were formulated in Chapter 1, the main conclusions that are drawn can be summarized as follows. On the research questions of block 1, it is concluded that there are solid arguments for public intervention in the PT sector. Most prominent argument is the network feature of PT provision leading to possible market failure. It is also concluded that formal stipulation of arrangements between PTAs and operators in a contract is a powerful instrument in hands of PTAs. Therefore a proper and consistent process

118 The network size remains largely constant in the concession areas (see Table 6.1), so it is more appropriate to interpret this result as economies of density rather than economies of scale.

119 These results indicate that regulatory schemes and operators' efficiency levels are exogenous and that the operators' information level about their technology and their efforts to reduce costs are greatly unobserved by the PTA. Therefore, the theory of regulation under asymmetric information may apply to the PT industry in the Netherlands (see Laffont and Tirole, 1993; Gagnepain and Ivaldi, 2002).

of PT procurement (including a targeted award mechanism and clear contract specifications) is of utmost importance in attaining efficiency and equity objectives. Concerning the research questions in block 2, it is concluded that the main drivers for passenger satisfaction are not –as is suggested in more technical oriented PT studies- restricted to core attributes such as speed, trip frequency and reliability, but also encompass peripheral attributes such a personnel/drivers behaviour and vehicle tidiness. Additionally it is concluded that age, PT mode choice, level of urbanization and negative social safety experiences, exert a great influence on satisfaction with PT. Regarding the effect of CT on satisfaction, a positive effect is demonstrated, however the effect is small (+0.06 points on a 10-points scale). This effect is probably highly correlated to introduction of new vehicles in the concessions. On attribute level, the attributes that contribute most to the change in overall trip satisfaction are *vehicle tidiness, on-board noise, ease of boarding/alighting from the vehicle* and *service frequency*. The attribute values that are linked to *information* and to *on-time performance* in the tendered regions are negative, meaning that the satisfaction with these attributes worsens in comparison with non-tendered regions. With regard to the research questions of block 3, it is concluded that contract renewal leads to a substantial reduction in operational costs. When a contract is renewed at least once, the costs fall by 10%. Contracts that are renewed at least twice even lead to an extra cost reduction of 6%, resulting in a total reduction of 16%. Subsidies fall even further. We did not find evidence that there is an effect of competitive tendering on operational costs, suggesting that the threat of CT is sufficient in a market in which the majority of concessions is competitively tendered. Finally, our results indicate a significant positive effect of contract renewal on (log) passenger kilometres. Renewing a contract results in a 7.6% increase in passenger kilometres.

7.2.2 Policy conclusions

Based on the empirical work it is concluded that the policy of the Dutch administration, which aims to increase satisfaction, efficiency and ridership in the PT sector by means of competition, is successful. We find that the immediate effect of competitive tendering on efficiency and satisfaction is (nearly) absent, suggesting that the threat of CT is sufficient in a market when the majority of concessions is competitively tendered. Our economies of scale results indicate that the geographical size of the current concession areas may not be altered without additional costs. From an operational costs perspective, our study suggests that there is no reason to increase (or decrease) the geographical size of the concession areas.

7.3 Recommendations

7.3.1 Policy recommendations

The performance and quality of PT systems should correspond as much as possible to the diverse demands of (potential) PT customers. It is therefore important for authorities and operators to acquire knowledge on the preferences and evaluations of customers. Authorities and operators may optimize their efforts by focusing their resources and strategies on retaining their existing customers rather than trying to attract new customers. In this respect it may be wise for them to seek the most cost-efficient measures and aim specific instruments and policies at specific target groups. For instance, we found that trip frequency and speed exert a strong impact on satisfaction. When considering several possible policies, PTAs should take into account that a policy aimed at increasing the service frequency to raise satisfaction is possibly not the most efficient method, as it may be very costly. A more cost-efficient measure may be to increase the travel speed, as such a measure latches onto both satisfaction and operational efficiency. Therefore, we recommend that PTAs should perform supplementary calculations of the costs associated with the measures. We further recommend that PTAs intending to use satisfaction evaluations as part of incentive payment schemes towards providers should complement these schemes with objective measurements of performance, such as reliability of operations and/or actual numbers of passengers transported.

From an operational costs perspective, our study suggests that there is no reason to increase (or decrease) the geographical size of the concession areas. When assessing the effects of CT, authorities should take the transaction costs of designing and implementing a competitive tendering regime into account. Finally, if the network conditions become too unfavourable, even under a competitive tendering regime, the costs and subsidy savings of contract renewal may run out. Therefore, PTAs should complement regulatory policies with infrastructural policies aimed at sustaining excellent network conditions for the PT system.

It is concluded in this thesis that the reform of the PT sector initiated by the Dutch government was successful. However there is room for improvement and optimization. Firstly: in non-complex situations such as the regional markets outside of the cities, there is possibly room for less regulation and experiments with free entrance of operators (competition on the road). We recommend authorities to investigate whether these strategies are feasible. Second, related to the costs and subsidy savings of contract renewal under a CT regime, we showed that transaction costs may be substantial. Therefore we recommend PTAs to take the level of transaction costs into consideration when making strategic decisions on the size of the services to be procured, on the award mechanism, and on contractual issues such as duration, complexity and contract type. Further we recommend authorities in evaluating deregulation and decentralization policies, to explicitly take account of the transaction costs effects of these policies. Finally, we noticed some inconsistencies in contractual arrangements made by PTAs and concluded that the combination of net-costs contracts and input controlled governance that is dominant in the Netherlands may be suboptimal with regard to subsidy reduction. If input-control is favoured and subsidy reduction is the prime objective of PTAs, we recommend gross-costs contractual arrangements.

7.3.2 Recommendations for further research

We studied the drivers of satisfaction with PT services and the impact of CT on satisfaction. These studies were performed on satisfaction evaluations of existing PT users only. This implies a lack of generalizability of the outcomes to non-PT users. More satisfaction research should be undertaken on non-captive users and non-PT users and specifically on the relationship between car availability and PT service satisfaction, as the ultimate objective of PTAs is to stimulate the shift to more sustainable modes.

In this study we took the perspective of the steering possibilities of PTAs in a tendering environment. We touched upon a number of drivers of satisfaction with PT, but it is probable that other drivers of interest remain to be revealed. In particular we recommend more research on the actual duration of travel time elements and on thresholds, keeping in mind that the drivers of satisfaction may be either more general or case-specific. Our study shows that in the PT sector segmentation matters; however, we shed little light on why it matters. We recommend further research in the PT sector focusing on the distinction between general and situational circumstances that exert an influence on satisfaction. General circumstances can be defined as those that are universally valid, as opposed to situational (case-specific) circumstances that influence satisfaction only under specific conditions and/or in an indirect way. Age, for instance, is an exponent of a general determinant, and urbanization is an example of a situational explanatory variable. It is most likely that the actual service performance plays a major role in explaining the satisfaction differences between urban and rural areas, but it may also be the case that attitudes towards PT differ between city dwellers and village dwellers. More research efforts should be directed towards these possible determinants of satisfaction with PT.

We empirically studied the relationship between CT and satisfaction, and between CT and ridership. The scope of this thesis was not directed to the relationship between satisfaction and ridership. As PT is an intermediary service, we already remarked in Chapter 4 that PT demand is inelastic, and therefore a strong and direct positive relation between satisfaction and ridership is not plausible. Furthermore, such a relationship is very hard to identify as there are many disturbances. However, as quantitative studies on this subject are scarce, we recommend to study this topic in more detail.

Our analyses of the efficiency effects of CT are focused on costs, subsidies and ridership only. Therefore, we cannot assess the total welfare effects of regulatory change. We did not analyse the transaction and monitoring costs in detail, for instance, nor did we analyse the changes in the level of service, fare increases and external effects due to CT. Consequently, to assess the overall effects of regulatory change in the Netherlands, a comprehensive social welfare analysis is necessary. We concluded that the threat of competition may be sufficient to reach government goals. We recommend studying the effects of the threat of competition in more detail, for instance by performing case studies on the content, procedures and political pressure exerted on publicly awarded contracts under a CT regime. In Chapter 2 it was concluded that contract type may be an important efficiency driver in the hands of PTAs. In the survey data for the Netherlands however, only information on net-costs contracts is provided. Therefore it was not possible to empirically test the relationship between efficiency and contract type. As this topic is relevant, we recommend collecting data for concessions granted under gross-costs and perform further analyses on contract type in relation to costs and subsidies. Another interesting research item in this context concerns analyses on the effect of contract incentives on the one hand and efficiency and satisfaction on the other. In recent contracts in the Netherlands it is a rule rather than an exception that bonus-penalty arrangements are incorporated. What is the effect of these incentives, and what would be the appropriate level of incentives to be most effective? If sufficient information on contract details would become available, it would be possible to incorporate these kind of explanatory variables in the efficiency and satisfaction models we developed, and compare the results with those from other countries (see for instance Hensher and Hougton, 2004). Further study into the theory of incomplete contracts developed by authors such as Grossman and Hart (1986), and Williamson (2002) may be valuable in discovering and better understanding of the contractual relationship between PTAs and operators. In this respect to derive a more accurate estimate of the actual transaction costs made by both actors (including the ex post transaction costs of renegotiation and the costs of asset transfers between operators) than was possible in this thesis, is the final research challenge we would like to recommend.