Does competitive tendering improve customer satisfaction with public transport? A case study for the Netherlands

Arnoud Mouwen*, Piet Rietveld

Department of Spatial Economics, Faculty of Economics and Business Administration, VU University Amsterdam

*Corresponding author. De Boelelaan 1105, room 4A-30, 1081 HV Amsterdam, The Netherlands. T: +31 622808517. F: +31 205986004. E: a.m.t.mouwen@vu.nl

DRAFT

Abstract

During 10 years experience with competitive tendering of regional and local public transport in the Netherlands, national average trip satisfaction of passengers increased from 6.77 to 7.29 (+0.52). This is a remarkable improvement, but a closer look at the data reveals that also in regions without competitive tendering the improvement in satisfaction was substantial. The difference in the improvement for regions with and without tendering is only +0.06. Tendering led in the majority of concession areas to an improvement of average trip satisfaction, but in some 40% of the cases a deterioration was observed. A change of operator in general has a negative impact on satisfaction. We also find that the effect on satisfaction of early tendering is larger than of later tendering. This may well be the consequence of a shift in emphasis of authorities and operators from quality improvement to efficiency improvements.
An analysis concerning the weighted satisfaction judgments of 15 underlying service attributes revealed that *'service frequency, on-time performance, travel speed*, and *vehicle tidiness'* contribute the most to the effect on satisfaction in the tendered regions. We found that new vehicles impact highly on satisfaction with *travel speed* and *vehicle tidiness*. The emphasis in the tenders with increasing *service frequency*, led to an increase in satisfaction but, may have a deteriorating effect on (the satisfaction with) *on-time performance*.

Key words: competitive tendering in public transport, satisfaction, service attributes

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 & Houghton, 2004; Marcucci & Gatta, 2007). With a good definition of service quality and a good measuring method, authorities are attempting to impose strong incentives on operators.

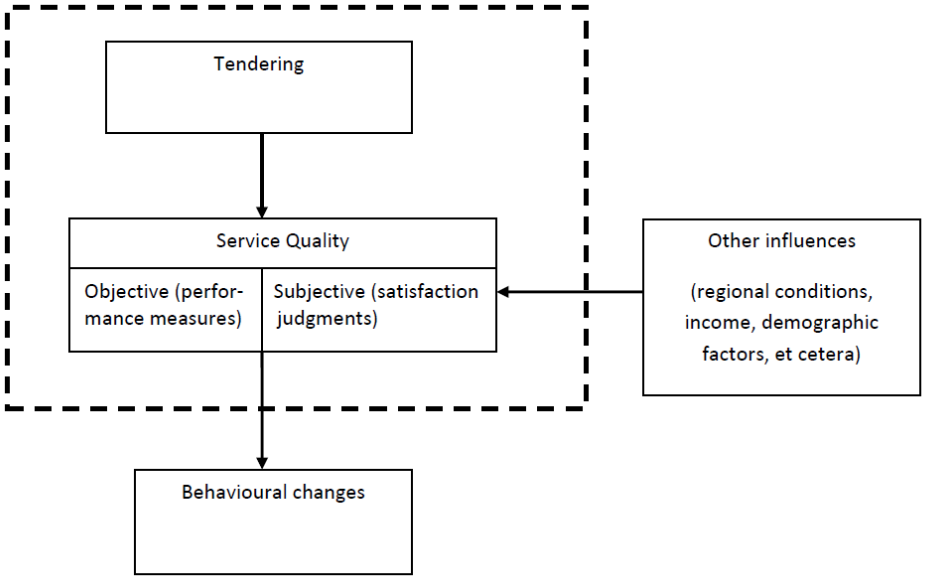
Reform in Dutch public transport takes the form of competitive tendering of concessions. Following the international trend, over the course of time in the Netherlands operators and authorities have tried also to become more and more customer oriented. Inclusion of quality aspects in contracts has become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

The relationship between tendering and efficiency is widely studied (Hensher et al., 2003; Hensher and Houghton, 2004; van der Velde & Pruijmboom, 2003; Walter, 2009). Few studies, however, explicitly focus on the relationship between tendering and satisfaction. The latter is the subject of this paper and the analytical results presented here are derived from the situation in the Netherlands between 2001 and 2010.

Figure 1 provides the broader context of this paper: travel behaviour of public transport passengers is influenced by their satisfaction with the quality of public transport services. 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 performance of public transport in the Netherlands and other influencing factors is only briefly touched upon. In this paper we pay no attention to changes in travel behaviour due to tendering.

DRAFT

Figure 1. Tendering and passenger satisfaction



After a short literature review in Section 2, the regulative setting in the Netherlands is described in Section 3. In Sections 4 and 5 the research questions and methods are presented. In Section 6 the focus is put on quantifying the relationship between tendering and the satisfaction of the total trip. In Section 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 paper is finalized with conclusions.

2 Literature on service quality and customer satisfaction

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

Measuring 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 & Taylor (1992) and Buttle (1996). According to them, and supported by empirical studies, the

DRAFT

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 & 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 & 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.

The authors are the opinion that 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.

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 & Liljander, 1994; Iacobucci, Grayson, & Ostrom, 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, Hagen (2011) showed that, in the heavy rail sector the attributes *safety and reliability of the service*[1] 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, marital status, age, and income), experience with the service and environmental factors (whether condition, crowdedness, et cetera) lead to significant differences in service perception and hence satisfaction (Bishop-Gagliano &

[1] In public transport research this is often narrowed to the reliability of the time table.

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.

Measuring service quality and satisfaction in 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 & 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.

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 paper we do

not explore these determinants in depth, but focus on tendering. Ongoing research of the authors does take these other above-mentioned factors into account.

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 der 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, 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 1 shows that the annual production growth in the tendered regions is higher than in the non-tendered regions.

Table 1. Annual growth of production of public transport(*)

Relative change in public transport service supply		
	concessions tendered	concessions not tendered
2001-2002	n.a	n.a
2002-2003	-2%	-6%
2003-2004	5%	-3%
2004-2005	1%	-3%
2005-2006	10%	2%

DRAFT

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 based on vehicle trip hours and for the period 2006-2009 on vehicle trip kilometres, originating from a different source.

Source 2001-2006: own calculations based on timetable data; 2006-2009: 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. A tendency can be observed for the number of competitors in the bidding phase to decrease. The market is divided among two big international conglomerates (Veolia Transport[2] and Arriva) 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 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 -neighboring 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 83 areal concessions. In 2010 this number was reduced to 48 areal and 18 line concessions.

4 Data considerations.

As mentioned in Section 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 paper 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[3]. The data collection is based on a stratified sample of public transport trips by bus (both

[2] The former operator Connexxion has been acquired in 2011 by Veolia, but still acts under its own name.

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

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 ten 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 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 several external sources.

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.

5 Research method

A model is formulated to study the relationship between tendering and satisfaction. As much as possible, the model is corrected for disturbances on the dependent variable other than tendering. The model corrects for yearly changes and regional variation in satisfaction scores. The variable '*year*' takes into account differences in the outcome imposed by factors that are connected to the year of the data intake. Examples of these factors 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 imposed by general characteristics of areas, for instance, car ownership and urban density. All independent variables are dummy variables. The model is used with the total trip satisfaction as the dependent variable, as well as for specific service attributes.

DRAFT

If the spatial units (concession areas) were constant all the time, the model to be estimated would be:

$$Q_{r,t,i} = b_0 + \sum_{t'} \alpha_{t'} year_{t',i} + \sum_{r'} \beta_{r'} area_{r',i} + \sum_{r',t'} \gamma_{r',t'} tendering_{r',t'} + \varepsilon_{r,t,i} \quad (1)$$

where

- $Q_{r,t,i}$: satisfaction of the total trip or service attribute in region *r* and year *t* of individual *i*;
- b_0 : constant term;
- α_t , β_r and $\gamma_{r,t}$: coefficients of the dummy variables. These are specific constants for, respectively, the influence of the year, the region, and the intervention of tendering leading to new operations in year *t*;
- $year_{t',i}$= 1 when *i* is interviewed in *t;* =0 else;
- $area_{r',i}$ = 1 when *i* is interviewed in *r;* =0 else;
- $tendering_{r',t'}$ = 1 when tendering led to new operations in *r* in year *t;* =0 else;
- $\varepsilon_{r,t,i}$: error term.

In particular $\gamma_{r,t}$ indicates the contribution of tendering in region *r* in year *t* to the satisfaction of the passengers affected.

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 feasible approach. Therefore, we reformulate the above model for pairs of two years. For each year pair we keep the area definition constant. 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.

The actual observations of two years are pooled into year pairs. Year pairs contain the data for both the ex-ante and the ex-post situation. The reformulation is as follows:

$$Q_{r,t',i} = b_0 + \sum_{t'} \alpha_{t'} year_{t',i} + \sum_{r'} \beta_{r'} area_{r',i} + \sum_{r',t+1} \gamma_{r',t+1} tendering_{r,t+1} + \varepsilon_{r,t',i} \quad (2)$$

where

- $Q_{r,t',i}$: satisfaction of the total trip or service attribute in region *r* and year pair *t'* of individual *i*;
- *t'* : *t, t+1*;
- $\alpha_{t'}$, β_r, $\gamma_{r,t+1}$: coefficients of the dummy variables.

Estimation of the model parameters took place using Ordinary Least Squares. Since the data underlying the variables are of categorical scale, to avoid perfect correlation between the variables a reference region in every year pair 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').

DRAFT

Since there are 9 year pairs, the modelling work resulted in 9 regression models. For each of the concession areas the models yield among others coefficients for the tendering effect. In the presentation of the results in the next section, we focus on the tendering effects.

6 Tendering and total trip satisfaction

In Table 2 some aggregated descriptive statistics concerning the satisfaction of the total trip are summed up. During the 10 years, 61 concession regions have been tendered for the first time, and 11 regions have been tendered for the second time. Table 2 shows a gradual increase in passenger satisfaction of the total trip during this decade[4]. In particular during the first period there were substantial increases.

Table 2. Descriptive statistics total trip satisfaction

	N	Average satisfaction total trip	Std. Deviation	No. of regions tendered for the first time	No. of regions tendered for the second time
2001-2002	139,309	6.77	1.123	5	0
2002-2003	139,198	6.80	1.110	6	0
2003-2004	155,911	7.03	1.308	7	0
2004-2005	169,796	7.14	1.469	12	0
2005-2006	165,880	7.12	1.481	7	0
2006-2007	168,013	7.11	1.500	9	3
2007-2008	167,527	7.18	1.454	1	2
2008-2009	170,604	7.29	1.387	9	5
2009-2010	170,473	7.29	1.384	5	1
Total				61	11

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 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. The coefficients for the year dummies and the region dummies, as well as the statistics are available on request.

Table 3. Main findings 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 ($\alpha \leq .05$).	22	20	2
No. of regions with positive non-significant effect of tendering ($\alpha > .05$).	20	15	5

[4] The data collection method changed in 2004 from orally administered questionnaires to a hand-out (written) questionnaires.

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

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).

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 7, an analysis of the service attributes that underlie total service satisfaction is carried out.

DRAFT

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 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.

7.1 Service attributes

The survey commissioned contains information on passengers satisfaction judgments of 15 service attributes. These attributes are shown in Table 4.

Table 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 (eg. drivers, station guards) when dealing with passengers.
Ticket-selling network	Ease of obtaining a ticket from on- and offboard 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	Onboard 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.

In Table 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 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

DRAFT

delays and price of the tickets. The highest satisfaction score is obtained for seating capacity, which means that crowded vehicles must be rather exceptional.

Table 5. Descriptive statistics: service attributes, 2001-2010

2001-2010	N	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
Total trip	883,009	7.11	1.38

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 & Antoniou, 2008; Iseki & 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. Nevertheless, we developed a simple procedure for estimating weights of the attributes.

The routine is based on the assumption that the aggregation of the satisfaction scores of the 15 attributes must add up to the total trip satisfaction. So the set of attributes covers all aspects of the trip. If this assumption is valid, a model can be constructed that predicts the total satisfaction as a function of the given satisfaction judgments of the attributes. The coefficients of the attribute variables represent the relative weight of these attributes as contributors to the satisfaction of the total trip.

DRAFT

To determine the weights, a linear regression model is designed with the 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*:

$$Q_{t,i} = b_0 + b_1 X_{1_i} + b_2 X_{2_i} + \cdots + b_n X_{n_i} + \varepsilon_{t,i} \quad (3)$$

where

- $Q_{t,i}$: total trip satisfaction in year *t* for individual *i*;
- b_0 : constant term;
- b_1, b_2, b_n : coefficients of the service attribute variables;
- X_{1_i}, X_{2_i} : satisfaction scores of service attributes 1 through 15 (see Table 4);
- $\varepsilon_{t,i}$: error term.

Deviating from the data analysis described in Section 6 that is based on year pairs, for determining the weights of the attributes, the data for all 10 years (more than 900,000 cases) are pooled. The outcomes are presented in Table 6.

Table 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

DRAFT

It can be concluded from Table 6 that - if the overall situation in the Netherlands in the period 2001-2010 is taken into consideration- 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 & Antoniou, 2008; Eboli & Muzzalla, 2010; Beirão & Sarsfield-Cabral, 2007; Berechman, 1993).

7.3 Tendering and satisfaction of 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.

Satisfaction with 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 2 described in Section 5, 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 7[5]. As a reference, in the last row of the table the results for the satisfaction with the total trip –already given in Table 3- are entered. The correlation matrix is shown in Appendix 1.

Table 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

[5] The output is limited to the most important outcomes. Other outcomes are available on request.

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
Satisfaction total trip	42	22	30	12		0.061	0.112

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 total trip satisfaction increased by 0.061 points.

In general, it can be concluded that 12 out of the 15 attributes contribute in a positive way to the tendering effect of total 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[6].

Weighted satisfaction with 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.

In Table 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[7]. The top-4 is highlighted.

[6] 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.

[7] the weights differ from those of table 6 because the constant term is not taken into account here

Table 8. Weighted contribution of service attributes in tendered regions (average 2001-2010)

explanatory variables	Weights of the service atrributes	Average coefficients of tendering variable			Average coefficients of tendering variable, weighted by importance of service attribute		
		Tendered regions with sign. pos. effect of tendering	Tendered regions with sign. neg. effect of tendering	All tendered regions (N=72)	Tendered regions with sign. pos. effect of tendering	Tendered regions with sign. neg. effect of tendering	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
Total trip	n.a.	0.320	-0.268	0.061	n.a	n.a	n.a

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 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[8]. When the net effect is considered, the last column of Table 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

[8] In Hensher et al., (2003) however, vehicle tidiness is of no importance.

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.

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. Because we lack reliable information on supply of public transport per concession region, this aspect will be assessed 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.

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) of 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 9 shows, per category, the average coefficients for the tendering variable of Model 2.

DRAFT

Table 9. Effect of new vehicles on satisfaction

	Average coefficient of tendering variable				Delta	
Attribute	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
Total trip	0.014	0.030	0.105	0.085	0.071	0.090

We may conclude that new vehicles highly impact on total 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.

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 and stayed in control. In 31 tendered regions operations shifted to a new operator. 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 6- 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 (B_i: the output of Model 2) 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 Equation 4).

$$B_{t,i} = b_0 + b_1\, t_i + b_2 tendered\ twice\ vs.\ once_i + b_3 new\ operator\ vs.\ inc._i + \varepsilon_i, \quad (4)$$

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

The output of Model 4 is shown in Table 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 8).

Table 10. Determinants of tendering effects

attributes	Regression Coefficients model 4.		
	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
Total trip	-0.015	0,000	-0,020
* significant at the .05 level (2-tailed)			

Although the decrease is only significant for four attributes, Table 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

DRAFT

attributes *prices of the tickets, on-board safety,* and *safety at stops* show the sharpest decrease[9]. A positive (though not significant) trend is found for the effect of tendering on the 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 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 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 9). 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 1).

8.3 Supply-oriented steering

The evaluations of the Dutch situation in 2004 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[10], but supply-based steering remained dominant. We showed that the supply of public transport increased considerably as a result of tendering (Table 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 8), we may conclude that the shift in satisfaction with *vehicle tidiness* is mainly linked to the introduction of new vehicles. This 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 3, 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.

[9] 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.

[10] 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.

DRAFT

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 negative 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.

9 Conclusions

In this paper we studied the relationship between tendering of regional public transport concessions and customer satisfaction.

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 average 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[11]. Possibly the threat of competitive tendering in general, caused a more customer oriented attitude of authorities and operators, even in regions where tendering did not take place.

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

[11] In the last decade, also the satisfaction with train services increased. The percentage of passengers who valued train services with a grade of 7 or higher, increased from 45% in 2001, to 75% in 2010.

Concerning the question raised in this paper 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.

Acknowledgements

We are very much obliged to KPVV for providing us with the satisfaction data for this study. In particular we are indebted to Gerard van Kesteren of KPVV who not only was willing to discuss some fundamental ideas with us but also provided us with a great deal of useful background information on concessions.

References

Anderson, S., Klein-Pearo, L., & Widener, S. (2008). Drivers of Service Satisfaction: Linking Customer Satisfaction to the Service Concept and Customer Characteristics. *Journal of Service Research, 10:* 365-381.

Beirão, G., & Sarsfield-Cabral, J. (2007). Understanding attitudes towards public transport and private car: A qualitative study. *Transport Policy, 14 (6):* 478-489,.

Berechman, J. (1993). *Transit Economics and Regulating Policies.* Amsterdam: North-Holland Publishers.

Berenschot. (2004). *Evaluatie aanbesteding OV-concessies, eindrapport.* Rotterdam: Adviesdienst Verkeer en Vervoer.

Bishop-Gagliano, K., & Hathcote, J. (1994). Customer Expectations and Perception of Service Quality in Retail Apparel Speciality Stores. *Journal of Service Marketing, 8*: 60-69.

Buttle, F. (1996). SERVQUAL; review, critique, research agenda. *European Journal of Marketing, 30* : 8-31.

Cronin, J. J., & Taylor, S. J. (1992). Measuring service quality: a re examination and extension. *Journal of Marketing, 56 (3):* 55-68.

Eboli, L., & Mazzulla, G. (2011). A methodology for evaluating transit service quality based on subjective and objective measures from passengers' point of view. *Transport Policy (18):* 172-181.

Eboli, L., & Muzzulla, G. (2010). How to Capture the Passengers' Point of View on a Transit Service through Rating and Choice Options. *Transport Reviews, 30 (4):* 435-450.

Friman, M. (2004). Implementing Quality Improvements in Public Transport. *Journal of Public Transportation, vol. 7 (4):* 50-65.

Hagen, M. v. (2011). *Waiting experience at train stations.* Delft: Eburon Academic Publishers.

DRAFT

Hensher, D. A., & Houghton, E. (2004). Performance-based quality contracts for the bus sector: delivering social and commercial value for money. *Transportation Research, Part B, 38:* 123-146.

Hensher, D., Stopher, P., & Bullock, P. (2003). Service quality -developing a service quality index in the provision of commercial bus contracts. *Transportation Research, Part A, 37*, 499-517.

Iacobucci, D., Grayson, K., & Ostrom, A. (1994). The calculus of service quality and customer satisfaction: theoretical and empirical differentations and integration. *Advances in Service Marketing and Management 3:* 1-67.

Iseki, H., & Taylor, B. (2008). Style versus Service? An analysis of user perceptions of transit stops and stations in Los Angelos. *Proceedings of the 87th Annual Meeting of the TRB.* Washington.

KPVV. (2011). *Ontwikkeling openbaar vervoer 2000-2009.* Utrecht.

Marcucci, E., & Gatta, V. (2007). *Quality and Public Transport Service Contracts.* working paper no. 586, Pavia.

Mu-consult. (2003). *Decentralisatie en Marktwerking in het Regionaal Spoor-, Stads- en Streekvervoer.* Rotterdam: Adviesdienst Verkeer en Vervoer.

Oliver, R. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of marketing, 17*: 460-469.

Rekenkamer, R. (2009). *Beheer en uitvoering concessies openbaar vervoer, Provincie Zuid-Holland.* Amsterdam: Randstedelijke Rekenkamer.

Strandvik, T., & Liljander, V. (1994). A comparison of episode performance and relationship performance for a discrete service. *Proceedings of Third Marketing Workshop.* Berlin, Germany.

Tyrinopoulos, Y., & Antoniou, C. (2008). Public Transport user satisfaction: Variability and policy implications. *Transport Policy, 15:* 260-272.

Velde, D. v. (1999). Transport Policy -Organisational forms ande enterpreneurship in public transport. Part 1: classifying organisational forms., *Journal of the World Conference on Transport Research Society, 6 (3): 93-98.*

Velde, D. van der , & Pruijmboom, E. (2003). First experience with tendering at the tactical level (service design) in Dutch public transport. *8th Conference on Competition and Ownership in Land Passenger Transport.* Rio de Janeiro (Brazil), September 2003.

Walter, M. (2009). *Efficiency and Competition in Public Transport.* Dresden.

Zeithaml, V., Berry, L., & Parasuraman, A. (1990). *Delivering quality service, balancing customer perception and expectations.* New York: The Free Press.

DRAFT

Appendix 1. Correlation of tendering variable coefficients (Pearson's R, highlighted >.07)

Correlations of tendering variable coefficients (Pearson's R, highligthed >.7)

	On-time performance	Travel speed	Service frequency	Personnel behaviour	Ticket-selling network	Prices of the tickets	on provision on stops	Safety at stops	Vehicle tidiness	Drivers behaviour	information on delays	boarding and alighting	Seating capacity	On board noise	On-board safety		Total trip
On-time performance	1	.599**	.315**	.603**	0,171	.344**	.661**	.373**	0,213	.644**	.526**	.275*	0,225	0,204	.507**		.675**
Travel speed	.599**	1	.289*	.513**	0,165	0,192	.574**	.396**	.389**	.612**	.298*	.422**	.299*	.390**	.524**		.777**
Service frequency	.315**	.289*	1	.260*	-0,047	0,223	0,219	0,161	0,134	.306**	.333**	0,077	-0,069	.355**	0,151		.421**
Personnel behaviour	.603**	.513**	.260*	1	.500**	.305**	.571**	.542**	.396**	.838**	0,099	.609**	.582**	.284*	.675**		.765**
Ticket-selling network	0,171	0,165	-0,047	.500**	1	.460**	.297*	.487**	.369**	.397**	-0,119	.592**	.413**	0,18	.523**		.422**
Prices of the tickets	.344**	0,192	0,223	.305**	.460**	1	.346**	.372**	.488**	.364**	.364**	.406**	0,045	.391**	.380**		.407**
Information provision on stops	.661**	.574**	0,219	.571**	.297*	.346**	1	.560**	.258*	.610**	.367**	.469**	.507**	.242*	.642**		.682**
Safety at stops	.373**	.396**	0,161	.542**	.487**	.372**	.560**	1	.258*	.562**	0,194	.497**	.466**	.299*	.810**		.563**
Vehicle tidiness	0,213	.389**	0,134	.396**	.369**	.488**	.258*	.258*	1	.396**	0,153	.529**	0,224	.736**	.472**		.577**
Drivers behaviour	.644**	.612**	.306**	.838**	.397**	.364**	.610**	.562**	.396**	1	0,184	.575**	.502**	.456**	.743**		.766**
On-board information on delays	.526**	.298*	.333**	0,099	-0,119	.364**	.367**	0,194	0,153	0,184	1	-0,121	-.295*	0,211	0,17		.274*
Ease of boarding and alighting	.275*	.422**	0,077	.609**	.592**	.406**	.469**	.497**	.529**	.575**	-0,121	1	.661**	.438**	.656**		.639**
Seating capacity	0,225	.299*	-0,069	.582**	.413**	0,045	.507**	.466**	0,224	.502**	-.295*	.661**	1	0,148	.605**		.536**
On board noise	0,204	.390**	.355**	.284*	0,18	.391**	.242*	.299*	.736**	.456**	0,211	.438**	0,148	1	.448**		.552**
On-board safety	.507**	.524**	0,151	.675**	.523**	.380**	.642**	.810**	.472**	.743**	0,17	.656**	.605**	.448**	1		.715**
Total trip	.675**	.777**	.421**	.765**	.422**	.407**	.682**	.563**	.577**	.766**	.274*	.639**	.536**	.552**	.715**		1

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

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

DRAFT