Centre for Transport Studies
STOCKHOLM

Satisfaction with crowding in public transport

Isak Rubensson, Maria Börjesson
Department for Transport Science, KTH Royal Institute of Technology

CTS Working Paper 2016:6

Abstract
Customer Satisfaction Surveys conducted among public transport passengers over 15 years in Stockholm show that the satisfaction with crowding has been and is low. Surprisingly, however, crowding does not seem to be important for the passengers' total satisfaction with the public transport service, not even for passengers on the most highly crowded bus services in central Stockholm. A possible reason is that crowding is not perceived as part of the service delivered by the public transport provider. Using less aggregated data than previous studies we also find that the performance of the attributes crowding, reliability and frequency influences satisfaction with these attributes. Moreover, a general result is that when comparing different bus lines and public transport modes, the satisfaction with an attribute is influenced by the performance of this attribute, while the importance level is much more stable across bus lines and public transport modes. There is also a trend increase in the satisfaction with many attributes (but not for crowding), while the importance is more stable over time.
Keywords: Customer Satisfaction, Service Quality, Stated Preference, Public Transport, Preferences, Crowding, Reliability

JEL Codes: R41, R42, R48

These can be found at:
http://www.aeaweb.org/jel/jel_class_system.php#Y

Centre for Transport Studies
SE-100 44 Stockholm
Sweden
www.cts.kth.se

vti

KTH VETENSKAP OCH KONST

1 INTRODUCTION

Stated Choice (SC) (Wardman and Whelan, 2011; Swärdh, 2016) and Revealed Preference (RP) (Tirachini et al., 2016, Hörcher et al., 2016) studies find high valuations of crowding in public transport. Analyzing Customer Satisfaction Survey (CSS) data collected in Stockholm 2008-2016, we find, however, that crowding is the trip attribute which is least correlated with the trip satisfaction of all attributes in the survey, suggesting that crowding is least important for the trip satisfaction. Still, peak hour crowding is a problem in the Stockholm network: during peak hour 80-90 percent of the total passenger travel time in the inner city metro is spent in vehicles with an occupancy rate (number of passengers per seat) above 100 percent. At the central station crowding reaches levels where passengers cannot board the first train that arrives.

Previous research literature exploring the importance of different attributes for the total satisfaction with public transport has, however, left out crowding (Cats et al. 2015; Abenoza et al. 2017). In this study, we therefore explore how satisfied travellers in the Stockholm Public Transport system are with crowding and how important the level of crowding is for trip satisfaction. We measure the importance as the correlation of satisfaction with crowding and the trip satisfaction following the tradition of earlier literature (de Oña and de Oña, 2014). Since crowding levels are increasing over time, we also explore how the satisfaction with crowding and other quality attributes has evolved over the years 2008-2016, and how it differs among public transport modes.

In this study we also explore, using less aggregated data than previous studies, how the performance level of the attributes crowding, reliability and frequency influences satisfaction with these attributes. Specifically, we analyze how the satisfaction with crowding, and its importance, depends on the level of crowding facing the travellers and how that differs over time and across public transport (PT) modes and bus lines, having different levels of crowding but still being within the same city. Previous literature regarding the extent to which increase in the performance level of a public transport attribute leads to an increase in attribute satisfaction has given mixed results. Barabino et al. (2012) find that performance has a large influence whereas Friman and Fellesson (2009), Fujii and Kitamura (2003) and Mackett and Edwards (1998) find evidence of the opposite. Friman and Fellesson (2009) find a weak correlation between the measured actual crowding and the satisfaction with crowding (although the measure of crowding they use is fairly unprecise). However, they find no correlation between measured actual crowding and the trip satisfaction when comparing different cities. From this, we draw the conclusion that the satisfaction with crowding does not influence trip satisfaction in their study, comparing different cities, which is consistent with UK findings and our own.

The huge challenges for the public transport providers in growing urban areas, with increasing congestion and limited resources, call for appropriate and relevant decision support. Here satisfaction studies might play an increasingly

important role. The European Union has commissioned the satisfaction survey Eurobarometer to support decision makers, and customer satisfaction surveys are increasingly used as incentives in tendered PT contracts. PT authorities in several countries have introduced quality-based incentive payments to public transport operators in the past decade (Trafikanalys, 2013; Van de Velde et al., 2008; Hensher and Houghton, 2004) to improve service quality. Since CSS is thus becoming an important decision support for public transport planning and a basis for payments to public transport operators, a better understanding of the travelers' responses to such surveys is important.

To measure the crowding levels and travel time reliability facing the respondents, we use data from automatic passenger count (APC) and automatic vehicle location (AVL) from 2014. From this data, we derive the total travel time that the travelers on different public transport services spend in vehicles with different load factors. In particular, we compare the satisfaction with crowding and its correlation with trip satisfaction for travelers on the most crowded inner city bus line in Stockholm with those found for travelers on the other bus lines. We also compare the travelers' satisfaction with reliability with the actual performance of travel time reliability (also computed from AVL/APC data).

A general result of this study is that when comparing different bus lines and public transport modes, the satisfaction with an attribute is influenced by the performance level, whereas the importance level is much more stable. There is also a trend increase in the satisfaction with many attributes (but not for crowding), but the importance is more stable over time. An exception is the attribute "information of delays", for which both satisfaction and importance have increased over time: possibly because expectations of, and dependence on, delay information have increased. The role of expectations is stressed in earlier literature (Oliver, 1980; Morfoulaki et al., 2007; Chen, 2008).

In particular, we find that higher levels of crowding facing the respondents indeed lead to lower satisfaction with crowding but do not seem to change the impact on (correlation with) trip satisfaction. The latter result is consistent with a separate survey directly asking Stockholm public transport commuters to state the relative importance of different quality attributes: crowding is on average ranked as the least important trip attribute. The importance of crowding is equally low for the travelers on the busiest inner city bus in the peak hour as for travelers in the rest of the system and time-of-day. Our results are also consistent with a CSS conducted in London and in other parts of the UK, Greeno and Joyner (2016) finding that satisfaction with crowding has a limited impact on total satisfaction.

A possible explanation for the low importance of crowding is that travelers don't believe that it can be reduced by the public transport operator because of limited influence on the demand. However, another possible explanation is that crowding is a more affective attribute in contrast to other attributes and thereby evaluated differently. Psychological research distinguishes between attributes that mostly trigger emotional brain processes, affective attributes,

and attributes mostly triggering cogitative brain processes, cognitive attributes (Shiv and Fedorikhin, 1999; Dhar and Wertenbroch, 2000). The brain processes activated when recollecting affective and cognitive attributes are also different (Kahneman, et al., 1997; Olsson et al., 2017) (which is relevant here because respondents in CSS are asked to recollect the satisfaction with attributes on the 'usual' trip[1]). Crowding is by nature an affective attribute, causing unease and stress, while travel time and trip frequency are primarily cognitive attributes. If respondents have cognitive attributes in mind when stating their trip satisfaction, this may explain the low correlation with crowding.

However, when interpreting results from CSS, one should also account for the possibility that it does not trigger logical and rule-based, but more intuitive, brain processes. Brain research and psychology have long made the distinction between two systems of cognitive processes: one characterized by intuition and the other by reasoning (Jacoby and Dallas, 1981; Epstein, 1994; Hammond, 1996; Jacoby, 1996; Kahneman, 2003). These systems are labeled System 1 and System 2 (Stanovich and West, 2000). Kahneman (2003) describes the operations of System 1 as intuitive, fast and effortless, and System 2 as reasoning, effortful, rule-based and logical, less automatic and therefore slower. In CSS, the respondents are asked to grade the satisfaction of different attributes, and satisfaction itself is a feeling and might very well be dealt with by System 1. Choice experiments, however, do not ask about feelings and are presumably more likely to trigger System 2.

Section 2 reviews the theory on consumer satisfaction. Section 3 describes our data and Section 4 includes the descriptive statistics and temporal trends derived from the CSS data. In Section 5 we compare the customer satisfaction with performance by attribute. Section 6 concludes.

2 THEORY

Cost-benefit analysis is a corner stone in transport infrastructure planning in Sweden and in many other countries. Cost-benefit analysis, and in micro-economic theory in general, relies on the assumption that that consumers are rational and have long-term stable preferences that determine their choices. In transport analysis consumers often face a set of discrete alternatives, which can be modeled with a random utility model (RUM) (McFadden, 1973). Moreover, there are typically no markets for goods relevant in transport analysis, for

[1] However, the finding that the importance of crowding is low does not seem to be explained by different brain processes activated when recollecting affective and cognitive attributes. The Stockholm CSS respondents are asked to respond to the survey while they travel, and the questionnaires are collected on-board. On the crowding question, they are asked to respond while having "their usual trip with the particular line they currently travel with" in mind when they respond to the survey. In the UK CSS, resulting in similar results regarding the importance of crowding, the respondents are given the questionnaire and asked to return it later, by mail, or to respond to it online. They are asked questions on crowding on that particular journey where they received the questionnaire.

instance for travel time, travel time reliably, and crowding. The preferences for such non-market goods are therefore typically revealed by revealed preference (RP) or stated preference (SP) data. The non-market valuations work like market prices and can be used in welfare analysis if they are assumed to be long-term stable.

It could be questioned whether there is a long-term and stable preference for crowding, and whether the preference for crowding can be revealed by assuming that respondents are rational utility maximizers. If crowding is an affective attribute, respondents may not be able to make rational and stable decisions regarding choices involving different levels of crowding. However, the large numbers of stated preference studies estimating the value of crowding, and more frequent examples of accounting for crowding in cost-benefit analysis, suggest that many scholars and planners believe that travelers have a long-term and stable preference for crowding. Moreover, the emerging literature with revealed preference studies (Tirachini et al., 2016, Hörcher et al., 2016) give valuations similar to those estimated by the more recent SP studies on crowding, suggesting that respondents choices regarding crowding in RP and SP studies are consistent and determines travelers behavior.

Consumer satisfaction theory was developed in the 1970s by firms to increase the understanding of customers' satisfaction and loyalty to brands and products. Martilla and James (1977) constructed the Importance-Performance (IP) – chart to extract information from quality and satisfaction surveys. The IP chart relates the perceived performance of attributes with how important they are for the customers. The IP chart in Figure 1 has four quadrants, each with a distinctive recommended market strategy for the firm. A successful strategy should be to concentrate the effort to enhance low performance of attributes having the highest importance. It might, however, not be worthwhile improving low performance of an attribute if the attribute has low importance.

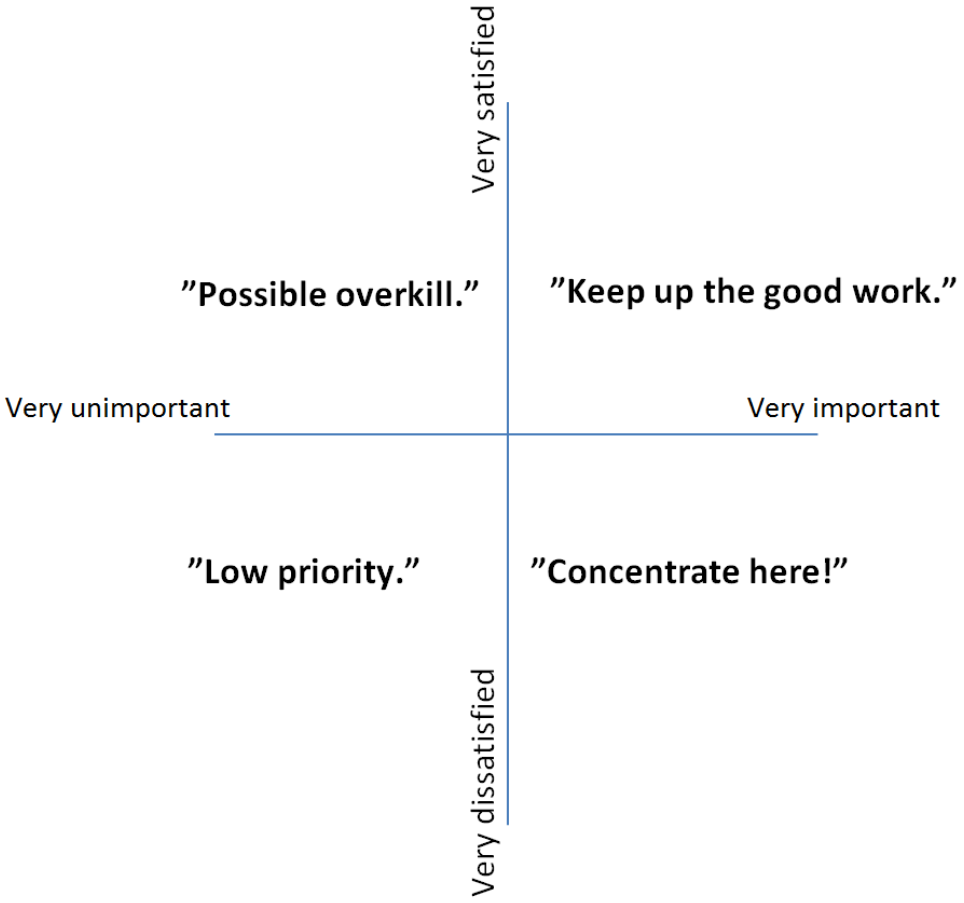


Figure 1 example of Importance-Performance chart

Martilla and James explicitly asked the respondents to state the importance of each attribute. Subsequent literature has two main categories of methods for measuring the importance of each attribute: (a) explicitly asking respondents in the customer satisfaction survey to indicate the importance they attach to each QoSA (Eboli and Mazzulla (2009), Guirao, García-Pastor, and López-Lambas (2016), Yannis and Georgia (2008)); and (b) inferring the importance by modeling the implicit contribution of each attribute to the total satisfaction with the line.(Abenoza, Cats, and Susilo (2017)) To measure the contribution of the satisfaction with a given attribute to the total satisfaction, some form of correlation between them has been employed by researchers, for instance bivariate Pearson correlation, regression analysis, structural equations, and path analysis and neural networks (de Oña and de Oña, 2014). Now, correlation does not mean causality, and there might be a large endogeneity. However, Oliver (1997) advises survey designers to use some form of deduced importance measure for two reasons. First, the questionnaire should be kept short to minimize respondent fatigue, and adding questions of both satisfaction and importance for all survey attributes might increase the number of questions substantially. Second, in surveys where the respondents directly are asked to state the importance of an attribute it is unclear what interpretation of the word important the respondents make. In this paper, we make the same assumption

as in earlier literature, that a high correlation indicates that the satisfaction of the given attribute is important for the trip satisfaction.

Oliver (1997) further points out that the relationship between performance and importance depends on the type of attribute. Similar ideas can be found in Maslow (1943), Kano (1984) and Matzler et al. (2004). The latter divide the attributes into three categories: (1) Attributes having a performance threshold level, which are important primarily when performance drops below the threshold, and have a small impact on the importance when above. (2) Attributes for which importance increases when the attribute is delivered but that cause no dissatisfaction if absent. (3) Attributes for which the importance is independent of performance.

Moreover, as shown theoretically and empirically, in the expectancy-disconfirmation framework the customer satisfaction itself depends on the performance relative to expectations (Oliver, 1980; Morfoulaki et al., 2007; Chen, 2008).

In summary, the earlier literature suggests that satisfaction of an attribute is formed based on the performance of the attribute as well as the consumer's expectation of the performance level. Since the performance of the attribute varies across modes and over time section 4 explores how the satisfaction with a number of public transport attributes varies in these dimensions. This indicates whether the satisfaction with the attributers is influenced by the performance level and the expectation.

Earlier literature also suggest that the satisfaction and importance might interact differently for different attributes. For this reason section 4 also explores the relationship between satisfaction and importance for different attributes, how this relationship has evolved over time and how it differs across modes (over which performance differs). Section 4 also explores whether the methodology of customer satisfaction surveys, not asking respondents to systematically rank or trade attributes or attribute levels against each other, could be the reason for the seemingly low importance of crowding in satisfaction studies.

In section 5 we study in more detail how the performance level of the attributes frequency, reliability and crowding influences the satisfaction and the importance, by distinguishing geographical bus areas and on the most crowded bus line.

3 DATA

We use the Stockholm Customer satisfaction Survey 2008-2016 and automatic location and passenger counts for 2014 from vehicles operating in Stockholm. To study the relationship between the attribute's performance level, satisfaction and importance we also compare the satisfaction data from 2014 with the vehicle gathered location and passenger count data for the same year.

In Section 3.1 we describe the Stockholm CSS, survey method and questionnaire. In 3.2 we describe the location and passenger count data and how it is sampled.

3.1 Stockholm Customer Satisfaction Survey

The Stockholm Customer Satisfaction Survey has been conducted since 2001, with ten survey waves per year.[2] Before 2014, each survey wave included 9 000 interviews, but since 2014 each wave includes 10000 interviews. Paper questionnaires are distributed and collected on board. They include statements concerning the satisfaction with nine quality attributes (see Table 1). The responses are given on a seven-point Likert scale, ranging from '7- agree completely' to '1- disagree completely'. The respondents are instructed to have in mind the usual conditions of the line they are using when stating their satisfaction. Hence, the respondents are tacitly asked to recall the travel conditions from previous trips on the line and state their satisfaction.

After the nine statements, the respondents are asked to grade their trip satisfaction with the line, where one is the lowest grade and seven the highest grade (see Table 1). All statements/questions include the escape-alternative "I have no experience".

Factor	**Wording in survey**	**Response alternatives**
Frequency	I am satisfied with the service frequency	Likert scale 1-7
Reliability	I am satisfied with the travel time reliability	Likert scale 1-7
Information delays	I am satisfied with the provision of information about train delays and timetable changes information on changes and delays	Likert scale 1-7
Cleanness vehicle	The vehicle is clean	Likert scale 1-7
Cleanness platform/stop	The bus stops or platforms are clean	Likert scale 1-7
Crowding	I am rarely troubled by crowding when I travel with this line	Likert scale 1-7
Information delays	Staff are able to answer questions regarding Stockholm Public Transport	Likert scale 1-7
Attitude personnel	Staff are well mannered and service-minded	Likert scale 1-7
Smooth Driving	The driving style is smooth and comfortable	Likert scale 1-7
Trip satisfaction	What is your overall grade of this line?	Grade 1-7

[2] Each wave is conducted during two weeks. One wave per month is conducted January-May and August-December.

Table 1: Statements concerning the satisfaction with nine quality attributes labels, and the question on the trip satisfaction[3]

3.2 Automatic Vehicle Location and Automatic Passenger Counts

To explore the correlation between the actual performance level and the satisfaction with the attributes reliability and crowding, we also use data on travel time reliability of the buses, as well as passenger counts and load factors.

All vehicles in the Stockholm Public Transport system are equipped with Automatic Vehicle Location (AVL). It measures the travel time and reliability for all services on all lines and modes. Approximately 10 percent of all buses and trains are also equipped with Automatic Passenger Counts (APC), measuring number of boarding and alighting passengers and load factors between stops (APC is not used on the Metro). Operators are instructed to operate the APC vehicles in proportion to observed number of passengers by line and service, and at least twice a month for each service.

We use observations for buses where both automatic passenger count (APC) data and automatic vehicle location (AVL) data has been collected. The bus network in Stockholm is divided in 16 tender areas which are the geographical blocks used when preparing calls for tender and tender agreements. Section 5 compares the satisfaction with crowding and reliability with the APC/AVL data on crowding and delays across the 16 tender areas.

4 SATISFACTION AND PERFORMANCE OVER TIME AND ACROSS MODES

Figure 2 shows the share of respondents satisfied (giving the statement for the relevant attribute a grade above four) by quality attribute over the course of the 16 years in which the CSS has been conducted. The trip satisfaction shows a steady trend increase. In fact, the satisfaction with most of the quality attributes displays a trend increase. In particular, the trend of satisfaction with *reliability* follows the trip satisfaction trend closely. The satisfaction with *crowding*, on the other hand, displays a trend decline (although the satisfaction with crowding has increased since 2013). The figure shows the average satisfaction over the seasons in each year and thereby hides seasonal variation. However, the two attributes related to cleanliness (on the stations/stops and in the vehicles) has a clear seasonal variation with low scores in the spring survey waves and high scores in fall survey waves. A possible explanation is that snow, ice and slush make the buses harder to keep clean in the early spring season.

[3] A Likert scale express the level of agreement or disagreement on a symmetrical scale between disagree completely (1) and agree completely (7). The trip satisfaction is not measure on a Likert scale but the respondents are asked to give the service and overall grade, where 1 is the lowest grade and 7 is the highest.

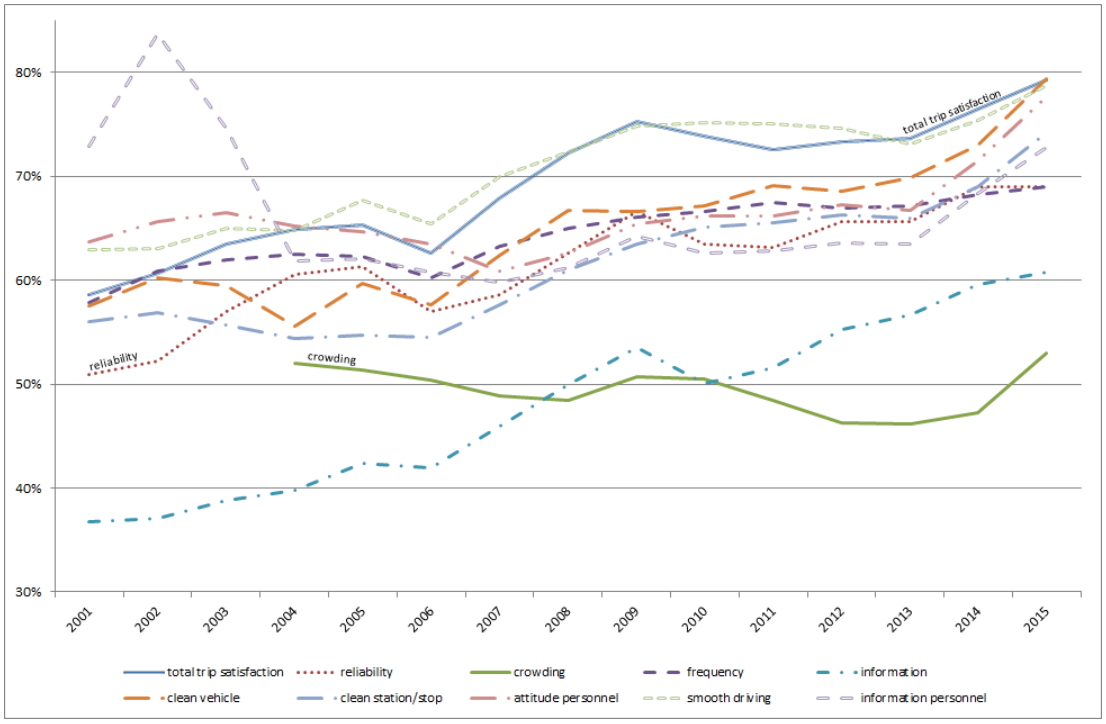
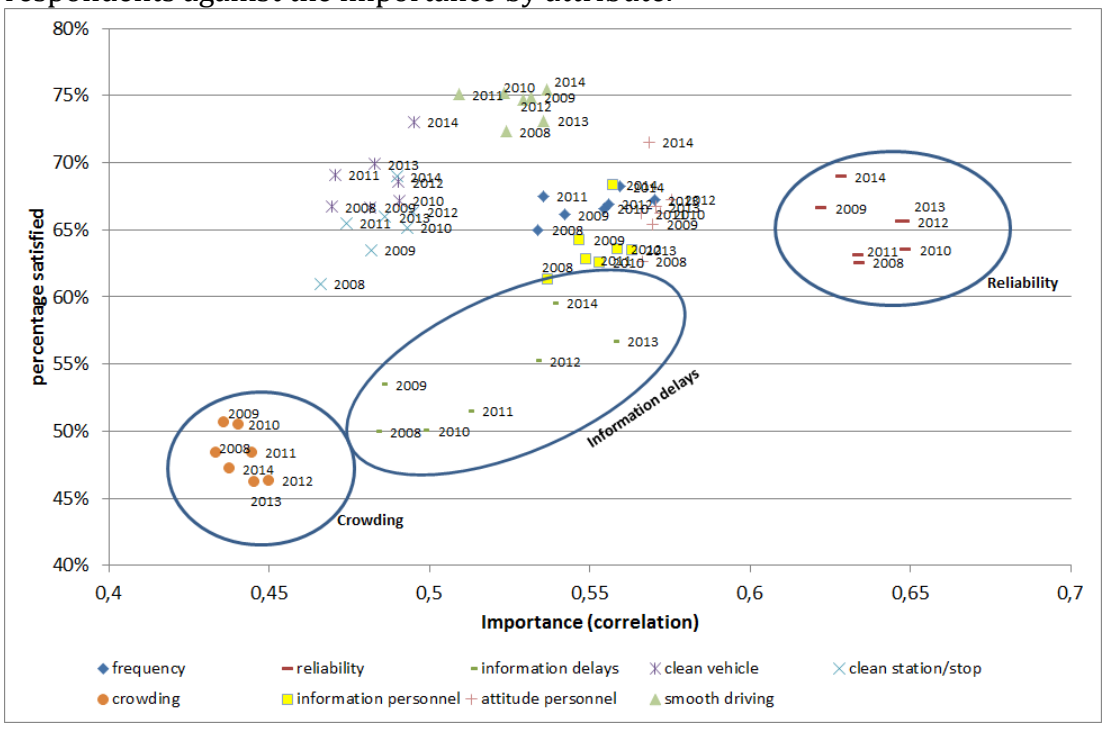


Figure 2 Share of respondents satisfied by quality attribute in each survey conducted since 2001.

Now we assume that the importance of the satisfaction with a given attribute can be measured as the correlation with trip satisfaction. Following Martilla and James (1977) we then make IP charts, plotting the share of satisfied[4] respondents against the importance by attribute.



[4] Share of respondents answering 5, 6 and 7 on the 7-graded Likert scale when stating their level of satisfaction.

Figure 3 Importance-Performance chart using data from 2008-2014. Crowding has the lowest importance and the lowest share of satisfied respondents. Reliability is consistently most important.

The IP chart in Figure 3 shows a similar pattern for most quality attributes: the observations are grouped by attribute and survey wave. The trend increase in satisfaction shown in Figure 2 can be spotted, but it is hard to see any trend regarding the importance of the attributes. However, the attribute *Information delays* shows a trend increase in the share of satisfied respondents and importance. Traffic information has also become increasingly available over the past 15 years due to technical developments. Hence, this is an example of an attribute for which satisfaction indeed increases with performance, but also the importance. This could be due to changed expectations of the information service from the operator or higher dependence on delay information.

In the lower left corner, crowding stands out as being least important and the quality attribute with which passengers are the least satisfied. Travel time reliability is the attribute consistently being most important.

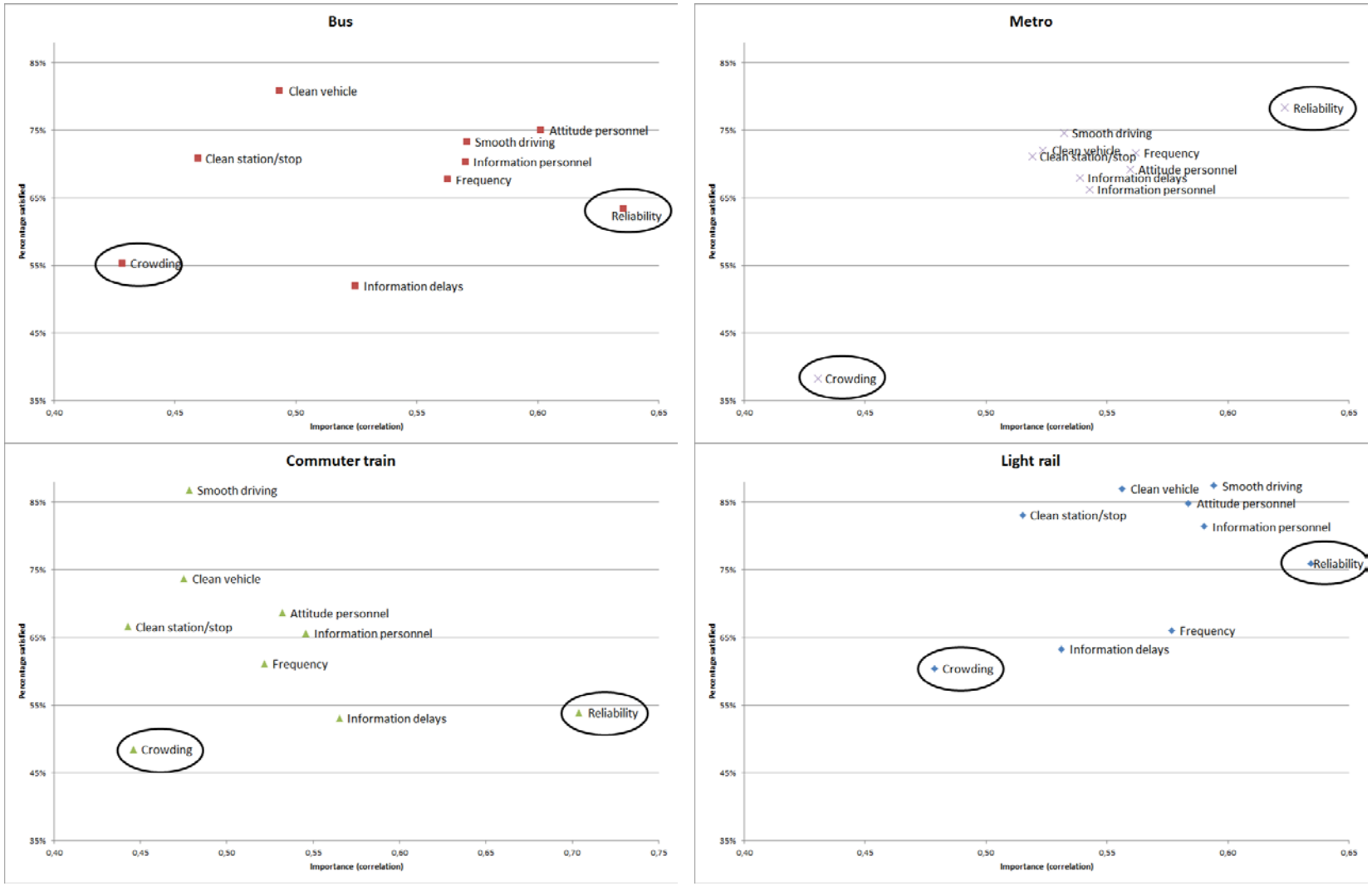

Bus
Clean vehicle
Attitude personnel
Smooth driving
Clean station/stop
Information personnel
Frequency
Reliability
Crowding
Information delays
Percentage satisfied
Importance (correlation)
Metro
Reliability
Smooth driving
Clean vehicle
Clean station/stop
Frequency
Attitude personnel
Information delays
Information personnel
Crowding
Commuter train
Smooth driving
Clean vehicle
Attitude personnel
Clean station/stop
Information personnel
Frequency
Information delays
Reliability
Crowding
Light rail
Clean vehicle
Smooth driving
Attitude personnel
Clean station/stop
Information personnel
Reliability
Frequency
Information delays
Crowding

Figure 4 IP charts for the 4 different transport modes that exist in the Stockholm County Public Transport system. Values are taken from CSS 2014

Figure 4 displays IP charts by public transport mode. The differences across modes regarding the satisfaction with different attributes are consistent with the performance of the modes in the Stockholm public transport system. The metro is the most crowded mode, but has high frequency and reliability. Bus travelers have most interactions with drivers and other staff, and are more dependent on information regarding bus routes. Commuter trains have a long history in Stockholm of poor reliability and information of delays. It has the least reliable travel times because it shares tracks with national and inter-regional trains. Commuter train travelers also have the least flexibility in their choice of route through the public transport system, implying that they suffer more from delays and lack of information.

Light rail travelers have on average higher satisfaction with all attributes and with trip satisfaction. The performance of the quality attributes is not in general higher for light rail but the travelers are richer, because the light rail connect more affluent parts of Stockholm.

Interestingly, the importance of the attributes crowding and reliability is remarkably stable across modes: for all modes, crowding is the least important attribute and reliability is the most important. The importance of the other attributes varies across modes. Bus and light rail passengers place relatively high importance on *frequency* and the attributes related to the performance of the staff (*smooth driving*, *attitude personnel*, *information delays*). This is natural since they also interact more with the staff. Metro passengers in general place a high importance on most attributes (all except crowding). Travelers on commuter train place a relatively lower importance on all attributes but crowding.

In a survey conducted in Stockholm 15 years ago (Olsson et al., 2001) commuters were asked to rank quality attributes in order of importance. The resulting average rankings are shown in Figure 5. It shows that the public transport users ranked the importance of the quality attributes very similarly to what we derived from the SCC data: frequency and reliability are both at the top and crowding at the bottom. This supports the evidence from the satisfaction studies, and that correlation seems to measure importance.

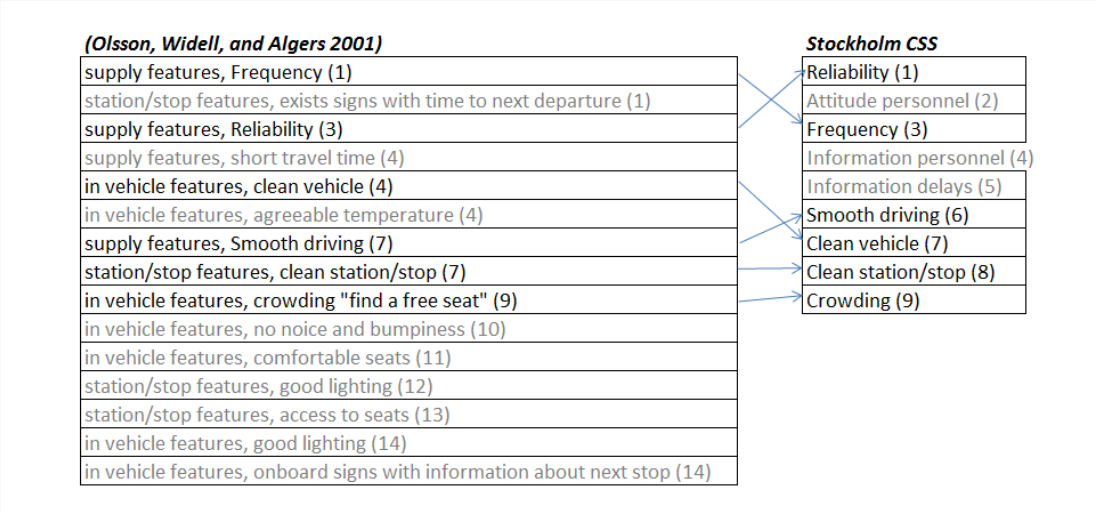


Figure 5 Comparison between rankings of quality attributes made by Stockholm PT-users in 2000 (Olsson, Widell, and Algers 2001) and in 2014 (Stockholm CSS)Change spelling to "noise" in chart

To summarize, a comparison of satisfaction with different attributes across modes is consistent with how the performance of the attributes varies by mode. Moreover, we see a trend increase in the satisfaction with most of the attributes. However, the importance of the attributes is much more stable across modes and over time than the satisfaction. The attribute *information delays*, is an exception in that both satisfaction and importance have increased over time (with performance). This section also demonstrates that crowding is consistently, over time and across modes, ranked as the least important *and* worst performing quality attribute.

5 DOES SATISFACTION DEPEND ON PERFORMANCE?

In this section we analyse in more depth how performance of crowding and delays influences the satisfaction and importance of these attributes. We begin section 5.1 and 5.2 by exploring how satisfaction with crowding and reliability depend on the performance. We compare the average satisfaction with actual performance from public transport vehicles by bus-tender area[5] using data conducted in 2014 only. The actual performance is measured by automatic passenger counts (APC) and/or automatic vehicle location (AVL).

In section 5.3 we compare the satisfaction and importance of all attributes between the travelers on the busiest and most frequent bus line in Stockholm (the inner city bus No. 4) and the travelers on all other buses, still using data from 2014 only.

[5] Administrative geographic units used when defining the extent of a bus tender agreement. An agreement can include one or more bus tender areas.

5.1 Crowding level vs satisfaction with crowding

The AVL data records the actual travel time by line segment (between two stops). If there are delays, this time might be different from the travel time according to the timetable. The APC data records the number of passengers in the vehicle on each line segment. Total travel time on a line segment is computed as the vehicle travel time times the number of passengers. The occupancy rate is computed as the number of passengers per seat.

The level of crowding by tender area is computed as the share of the total travel time that travelers spend in vehicles where occupancy is above 1. Figure 6 plots the computed level of crowding against the share of respondents satisfied with crowding, by bus tender area. The figure also includes the results of a linear regression model, where the slope is -1.2 with t-stat=-3.9. Hence, we thus conclude that higher crowding levels reduce the satisfaction with crowding.

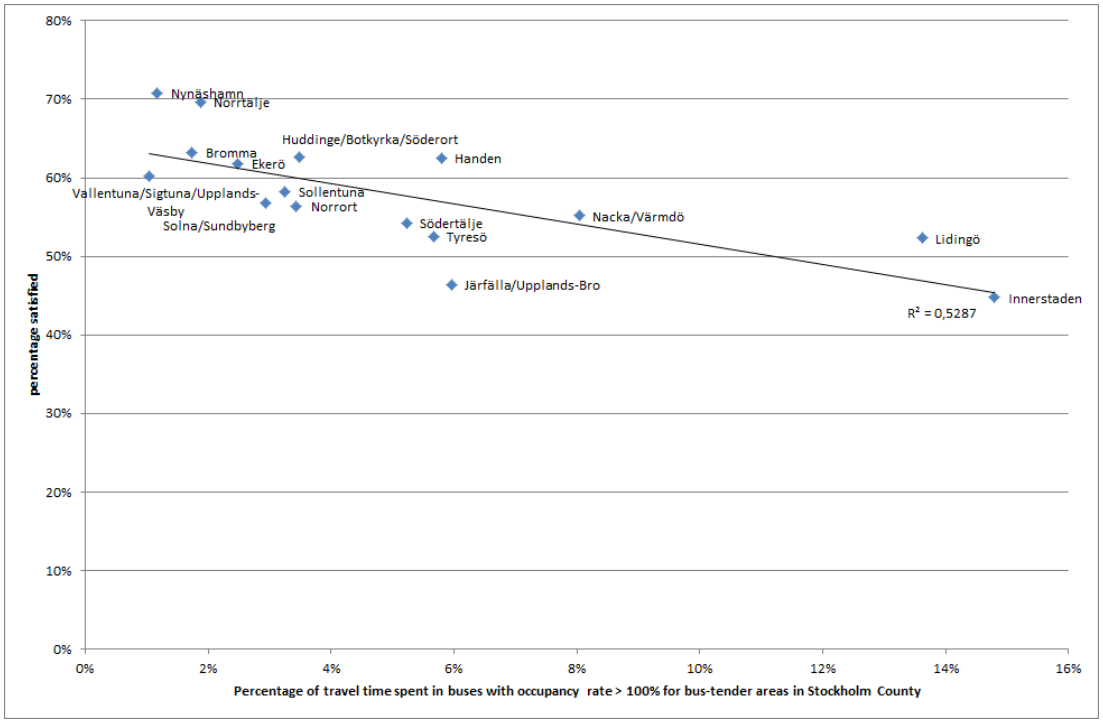


Figure 6 Share of satisfied with crowding vs measured percentage of travel time spent in buses with occupancy rate > 100% for bus-tender areas in Stockholm County.

5.2 Reliability vs satisfaction with reliability

AVL and APC data are used to assess the total delay time experienced by the travelers by bus tender area. The AVL data includes information on the delay of each bus departing from the bus stops within the bus route. All stops are included for all buses. However, since the total delay can only be computed at the end of the trip, when the passenger alights, the total delay time cannot be calculated exactly from only the number of passengers per service and line segment.

We therefore approximate the total delay time experienced by the passengers on bus service j, T_j^d, as

$$T_j^d \cong \overline{x_{\iota\in J}} \max_{i\in j} \delta_i,$$

where x_i is the number of travelers on line segment i and service j, δ_i is the delay of the service for the line segment i. Hence, we assume that all passengers experience the delay of the line segment with the largest delay. This might overestimate the total delay time, but since delay and passenger volumes are correlated, this measure should construct should at least indicate how delay varies across regions.

Figure 7 plots the ratio of the sum over all lines and services of the total delay time and the total travel time of all lines and services in the tender area, $\sum_j T_j^d / \sum_j T_j$, against the share of respondents satisfied with reliability by tender area. The figure shows that satisfaction is influenced by the actual reliability. A linear regression results in a slope of 0.5 with t-statistics -1.9.

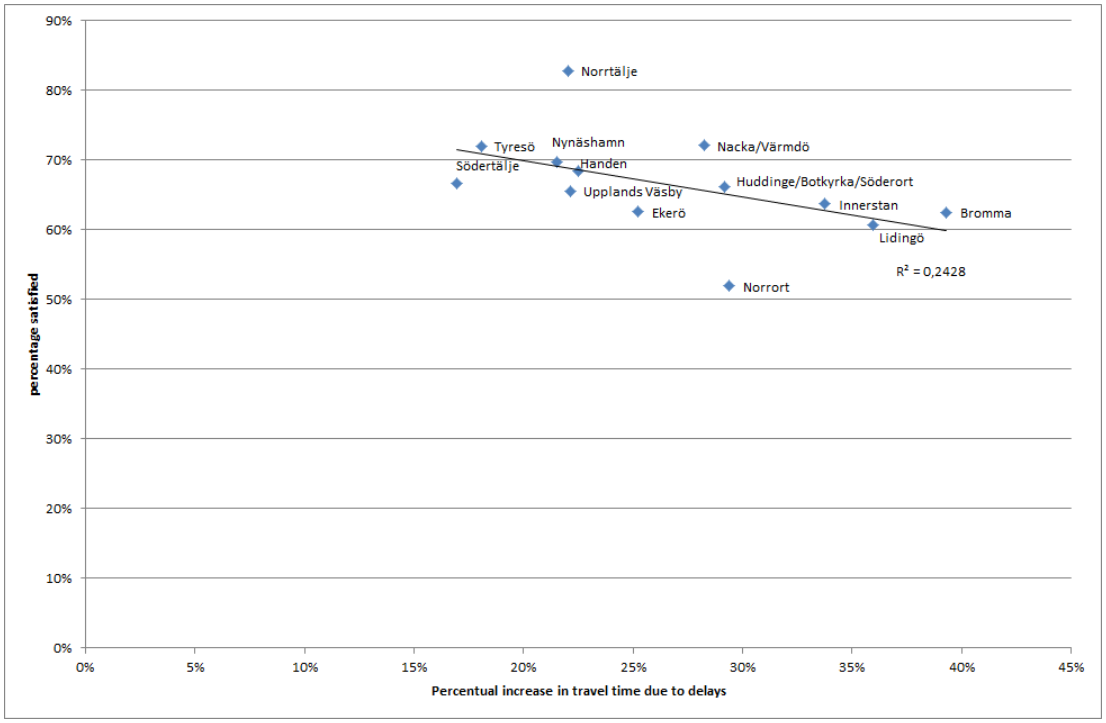


Figure 7 Share of satisfied with reliability vs approximated delays from AVL/APC data for bus-tender areas in Stockholm County[6].

[6] Two data points, Solna/Sundbyberg and Sollentuna, have been excluded from this chart. They make up a part of a region that had a particularly problematic transition to a new winning operator who in early spring of 2013 had considerable difficulties in operating the area reliably. Large delays ensued during January to March of 2013. Looking at the data, the areas show very low satisfaction for reliability which did not abate as fast as the actual problem was resolved. These two data points for satisfaction with reliability during 2014 (1-1,5 year after the delays

caused by the mismanaged transition were gone) show clear outlier tendencies and were omitted.

5.3 Crowding satisfaction crowded line with all bus lines

The inner city bus line No. 4 is the most crowded bus line in Stockholm. It cuts through the city center of Stockholm and has a high frequency (about every 4-6 minutes in rush hour). We have chosen this line because it is the single most crowded line and has a total patronage high enough to enable stable satisfaction data. Table 2 compares occupancy rate for bus line No. 4 during peak period (weekdays 6 am to 9 am) with the average occupancy rate across all bus lines in the county during all time periods (weekdays and weekends).

Occupancy rate	All bus services, average across weekends and weekdays	Bus No. 4 during weekdays peak period
<0.5	61.1	22.5
0.5-0.75	22.7	23.8
0.75-1	10.1	18.9
1-1.25	4.0	14.8
1.25-1.5	1.4	8.7
1.5-1.75	0.5	6.1
1.75-2	0.2	2.5
>2	0.2	2.7

Table 2 distribution of travel time(%) in different brackets of crowding

In the peak hours, travelers on bus No. 4 spend on average 34 percent of their total travel time in vehicles with an occupancy rate over 100%. The corresponding share for all bus lines, services and times-of-day in the county is 6 percent. The crowding is thus substantially higher for travelers on bus No. 4 in the peak than for the average bus travelers. The frequency is also substantially higher for the former travelers.

As a comparison, Table 3 shows, for all attributes, the share of satisfied respondents among those recruited on bus No. 4 in the peak and among those recruited on any line, service and time-of-day. Respondents on bus No. 4 in the peak are more satisfied with the frequency and less satisfied with crowding than bus travelers in general. Hence, the performance level facing the respondents does influence the satisfaction with these two attributes. However, when testing the statistical significance of the difference in the importance of these two attributes between the two groups, we find no difference. This is consistent with the differences across modes in Section 4: the satisfaction increases with performance level but the importance does not.

Note that all the attributes (except reliability and crowding in the survey) have very similar levels of satisfaction between bus No. 4 as the other bus services, which is expected since there are no large differences in performance in these attributes.

Attribute	All bus services, average across weekends and weekdays	Bus No. 4 during week days peak period
Frequency	**67**	**81**
Reliability	64	65
Information delays	50	52
Cleanness vehicle	75	79
Cleanness platform/stop	69	70
Crowding	**55**	**34**
Information	70	73
Staff	74	75
Driving style	73	70

Table 3 Share of satisfied respondents (%) per each attribute for all buses and for bus line 4, frequency and crowding highlighted.

The finding that importance of crowding is low, even when the satisfaction with crowding is low, seems to hold also when looking at other urban areas with similar or higher levels of crowding.[7] Greeno and Joyner (2016) find a low importance of crowding when analyzing National Rail passenger CSS data from the UK: multivariate analyses show that dissatisfaction with crowding explains only 6 percent of the total trip dissatisfaction. The National Rail passenger CSS is conducted on all regional and national rail services.

The expectancy-disconfirmation framework has shown that satisfaction is a product of both performance and expectations. We've seen in this section that, as shown in the previous section, the importance of crowding and reliability is more stable across modes than satisfaction with these attributes are, indicating that the performance level influences the satisfaction but not the importance.

6 DISCUSSION OF FINDINGS

Using more disaggregated data than Friman Fellesson (2009), we find a stronger correlation than they find between performance level and satisfaction with the attributes under study: crowding, reliability and frequency. We also see that satisfaction with different attributes varies across public transport modes in the way that is expected from the performance level and that satisfaction with all attributes except congestion has increased over time for all attributes. However, the importance of the attributes is in general not dependent on performance level. An exception is information on delays, for which the importance (and satisfaction) has increased over time, as the performance level has increased. A likely reason is that the importance has changes as the expectations have increased for this attribute.

[7] Comparing average seat occupancy (crude measure of crowding) all UK Rail operators have higher occupancy than the Stockholm Light rail and Commuter rail systems. The Stockholm metro places third among UK Rail operators with London Overground and First Transpennine Express having higher occupancy rates. (Office of Rail Regulation, 2012; AB Storstockholms Lokaltrafik, 2015).

The puzzling result of this study is the consistency of low importance of crowding. This is not consistent with the high valuation of crowding found in many valuation studies, or the public debate where crowding in the public transport system is seen as a major and increasing problem, justifying mega investments.

The low importance of crowding does not seem to be caused by low crowding levels: the importance of crowding is stable across modes and bus lines with varying crowding levels, for which the satisfaction with crowding differs significantly. Moreover, the importance of crowding is also low in the parts of the London Rail system where travellers are least satisfied with congestion (Greeno and Joyner, 2016).

Stradling et al. (2007) showed that unwanted arousal from affective attributes (feeling unsafe, crowding etcetera) has a prominent role in deterring public transport trips. It is possible that the low correlation between satisfaction with crowding and trip satisfaction (used to measure importance) is due to crowding being an affective attribute, and thereby processed differently in the brain. Another potential explanation of the low importance of crowding can be that the travelers do not perceive the crowding as part of the service provided by the operator, because they don't believe that the operation can or should influence the demand. Folkes (1984) suggests that customers do consider who they believe are responsible for a possible product failure in their total satisfaction with the service. Naturally, an underlying necessary assumption in satisfaction studies is that the consumers have the impression that the service provider can influence the performance of the attribute, and that the attribute is part of the service they are asked to evaluate. All attributes but crowding included in the survey used in this paper are clearly part of the service in the sense that the operator is responsible for their performance.

7 POLICY IMPLICATIONS

Crowding in public transport is increasing in many metropolitan areas, and capacity extensions addressing the increasing crowding often demand huge costs (in particular in European build-up cities). To assess the economic rationale for such extensions is key for an efficient use of public resources and an efficient design of the transport systems: An improved understanding of traveler's satisfaction and valuation of crowding are therefore strongly policy relevant.

Cost-benefit exercises use valuations of crowding derived from stated choice and possibly (a small number of) revealed preference data. This practice relies on the assumption that the valuations of crowding reflect stable preferences based on rationality that therefore impact the travels' behavior. However, the weak impact on satisfaction we find in the present study, possibly due to its affective nature, questions this assumption. A key message to policy makers is therefore to interpret the valuation of crowding with caution until there are more valuations based on RP (and therefore behavior) available.

In tandem with the increasing use of crowding in cost-benefit exercises, customer satisfaction studies are increasingly used in the planning and design process of the public transport system. Customer satisfaction studies are increasingly used to monitor and steer public transport production in Europe; whether they are studies constructed in a single public transport system (as the Stockholm CSS reported in this article) or nationwide (as the Transport focus studies from UK) or as a multilateral benchmarking project in the form of BEST (best2005.net) between Berlin, Copenhagen, Geneva, Helsinki, Oslo, Stockholm and Vienna. The objectives of these surveys include improving the understanding of consumers perceptions and views of the public transport systems, and monitoring the performance of the systems. They are also used to craft strategies to achieve political objectives for the public transport system (related to increased market share for public transport, revenues and customer satisfaction).

Since both valuations (CBA) and CS surveys are used in the planning it helps policy makers to be more consistent and to design more consistent policies, if they understand how the valuations used in cost-benefit exercises and the results of the CS surveys are related. This paper demonstrates that they measure different things.

One of the main applications of the satisfaction studies is to monitor the service level in procured public transport, and to give the operators incentives to maintain or increase the level of service. Here customer satisfaction surveys are an attractive tool for conducting regular check-ups to gauge tender companies' performance as well as influence of exogenous factors such as weather, strikes, break-down in infrastructure and other events outside the tendering companies' responsibility. Results from the customer satisfaction surveys are often linked to bonuses and penalties in the contract with tendering companies'. In the work to alleviate crowding, this attribute is often a good candidate to be included in a customer satisfaction survey. However, based on our study we advise against the use of satisfaction with crowding to assess the tender companies' performance until the interpretation of the survey results are better understood.

Finally, an attractive public transport system is a corner stone for a vital and green city. In the effort of maintaining or building such system the customer satisfaction survey could be a useful tool, provided that the policy makers understand its strengths and weaknesses, which has been the main objective of the paper.

REFERENCES

AB Storstockholms Lokaltrafik. 2015. 'Fakta om SL och länet 2015'. http://sl.se/globalassets/rapporter-etc/fakta_om_sl_och_lanet-2015.pdf.

Abenoza, Roberto F., Oded Cats, and Yusak O. Susilo. 2017. 'Travel Satisfaction with Public Transport: Determinants, User Classes, Regional Disparities and Their

Evolution'. *Transportation Research Part A: Policy and Practice* 95 (January): 64–84. doi:10.1016/j.tra.2016.11.011.

Barabino, Benedetto, Eusebio Deiana, and Proto Tilocca. 2012. 'Measuring Service Quality in Urban Bus Transport: A Modified SERVQUAL Approach'. *International Journal of Quality and Service Sciences* 4 (3): 238–52. doi:10.1108/17566691211269567.

Cats, Oded, Roberto F. Abenoza, Chengxi Liu, and Yusak O. Susilo. 2015. 'Evolution of Satisfaction with Public Transport and Its Determinants in Sweden'. *Transportation Research Record: Journal of the Transportation Research Board* 2538 (January): 86–95. doi:10.3141/2538-10.

Chen, Chih-Chien, and Zvi Schwartz. 2008. 'Timing Matters: Travelers' Advanced-Booking Expectations and Decisions'. *Journal of Travel Research* 47 (1): 35–42. doi:10.1177/0047287507312413.

Dhar, Ravi, and Klaus Wertenbroch. 2000. 'Consumer Choice Between Hedonic and Utilitarian Goods'. *Journal of Marketing Research* 37 (1): 60–71. doi:10.1509/jmkr.37.1.60.18718.

Eboli, Laura, and Gabriella Mazzulla. 2009. 'A New Customer Satisfaction Index for Evaluating Transit Service Quality'. *Journal of Public Transportation* 12 (3). doi:http://dx.doi.org/10.5038/2375-0901.12.3.2.

Epstein, Seymour. 1994. 'Integration of the Cognitive and the Psychodynamic Unconscious'. *American Psychologist* 49 (8): 709–24. doi:10.1037/0003-066X.49.8.709.

Fellesson, Markus, and Margareta Friman. 2012. 'Perceived Satisfaction with Public Transport Service in Nine European Cities'. *Journal of the Transportation Research Forum* 47 (3). doi:10.5399/osu/jtrf.47.3.2126.

Folkes, Valerie S. 1984. 'Consumer Reactions to Product Failure: An Attributional Approach'. *Journal of Consumer Research* 10 (4): 398–409. doi:10.1086/208978.

Friman, Margareta, and Markus Fellesson. 2009. 'Service Supply and Customer Satisfaction in Public Transportation: The Quality Paradox'. *Journal of Public Transportation* 12 (4). doi:http://dx.doi.org/10.5038/2375-0901.12.4.4.

Fujii, Satoshi, and Ryuichi Kitamura. 2003. 'What Does a One-Month Free Bus Ticket Do to Habitual Drivers? An Experimental Analysis of Habit and Attitude Change'. *Transportation* 30 (1): 81–95. doi:10.1023/A:1021234607980.

Greeno, David, and Rebecca Joyner. 2016. 'National Rail Passenger Survey'. Spring 2016. London: Transportfocus.

Guirao, Begoña, Antonio García-Pastor, and María Eugenia López-Lambas. 2016. 'The Importance of Service Quality Attributes in Public Transportation: Narrowing the Gap between Scientific Research and Practitioners' Needs'. *Transport Policy* 49 (July): 68–77. doi:10.1016/j.tranpol.2016.04.003.

Hammond, Kenneth R. 2000. *Human Judgment and Social Policy: Irreducible Uncertainty, Inevitable Error, Unavoidable Injustice*. Oxford University Press on Demand. https://www.google.com/books?hl=sv&lr=&id=VxRnDAAAQBAJ&oi=fnd&pg=PR9&dq=Human+judgement+and+social+policy:++Irreducible+uncertainty,+inevitable+error,+unavoidable+injustice+&ots=xuRnDFCp7L&sig=OQN1oRYdM8XzZcQswmFkaa57fT4.

Hensher, David A., and Erne Houghton. 2004. 'Performance-Based Quality Contracts for the Bus Sector: Delivering Social and Commercial Value for Money'. *Transportation Research Part B: Methodological* 38 (2): 123–146.

Hörcher, Daniel, Daniel J. Graham, and Richard J. Anderson. 2016. 'Link Between Crowding Pricing and Seat Supply in Public Transport'. In . https://trid.trb.org/view.aspx?id=1392972.

Jacoby, Larry L. 1996. 'Dissociating Automatic and Consciously Controlled Effects of Study/Test Compatibility'. *Journal of Memory and Language* 35 (1): 32–52. doi:10.1006/jmla.1996.0002.

Jacoby, Larry L., and Mark Dallas. 1981. 'On the Relationship between Autobiographical Memory and Perceptual Learning'. *Journal of Experimental Psychology: General* 110 (3): 306–40. doi:10.1037/0096-3445.110.3.306.

Kahneman, Daniel. 2003. 'Maps of Bounded Rationality: Psychology for Behavioral Economics'. *The American Economic Review* 93 (5): 1449–1475.

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. 'Back to Bentham? Explorations of Experienced Utility'. *The Quarterly Journal of Economics* 112 (2): 375–405.

Kano, Noriaki, Nobuhiku Seraku, Fumio Takahashi, and Shinichi Tsuji. 1984. 'Attractive Quality and Must-Be Quality'. *JOURNAL OF THE JAPANESE SOCIETY FOR QUALITY CONTROL*. http://en.journals.sid.ir/ViewPaper.aspx?ID=283197.

Mackett, Roger L., and Marion Edwards. 1998. 'The Impact of New Urban Public Transport Systems: Will the Expectations Be Met?' *Transportation Research Part A: Policy and Practice* 32 (4): 231–45. doi:10.1016/S0965-8564(97)00041-4.

Martilla, John A., and John C. James. 1977. 'Importance-Performance Analysis'. *Journal of Marketing* 41 (1): 77. doi:10.2307/1250495.

Maslow, Abraham Harold. 1943. 'A Theory of Human Motivation.' *Psychological Review* 50 (4): 370.

Matzler, Kurt, Elmar Sauerwein, and Kenneth Heischmidt. 2003. 'Importance-Performance Analysis Revisited: The Role of the Factor Structure of Customer Satisfaction'. *The Service Industries Journal* 23 (2): 112–29. doi:10.1080/02642060412331300912.

McFadden, Daniel, and others. 1973. 'Conditional Logit Analysis of Qualitative Choice Behavior'. https://elsa.berkeley.edu/reprints/mcfadden/zarembka.pdf.

Morfoulaki, Maria, Yannis Tyrinopoulos, and Georgia Aifadopoulou. 2010. 'Estimation of Satisfied Customers in Public Transport Systems: A New Methodological Approach'. In *Journal of the Transportation Research Forum*. Vol. 46. http://journals.oregondigital.org/trforum/article/view/981.

Office of Rail Regulation. 2012. 'Costs and Revenues of Franchised Passenger Train Operators in the UK'. http://orr.gov.uk/__data/assets/pdf_file/0019/4933/toc-benchmarking-report-2012.pdf.

Oliver, Richard L. 1980. 'A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions'. *Journal of Marketing Research*, 460–469.

———. 2010. *Satisfaction: A Behavioral Perspective on the Consumer*. M.E. Sharpe.

Olsson, Camilla, Jenny Widell, and Staffan Algers. 2001. 'Comfort factors and their influence on bus and rail demand. Traveller values, models and forecasts for local work trips. March 2001'. VR 2001/06. Vinnova Rapport. Stockholm. https://www.google.se/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwi5o5GR0I3MAhVHYZoKHVVVA7oQFggcMAA&url=http%3A%2F%2Fwww.vinnova.se%2Fsv%2FAktuellt--publicerat%2FPublikationer%2FProdukter%2FKomfortens-betydelse-for-spar--och-busstrafik-%2F&usg=AFQjCNFcDA3gZt3jeomQ6llaR7lAHg2d-g&sig2=8eK6LWB5L84YJJxYxafmSQ.

Olsson, Lars E., Tommy Gärling, Dick Ettema, Margareta Friman, and Michael Ståhl. 2017. 'Current Mood vs. Recalled Impacts of Current Moods after Exposures to Sequences of Uncertain Monetary Outcomes'. *Frontiers in Psychology* 8 (January). doi:10.3389/fpsyg.2017.00066.

Oña, Juan de, and Rocio de Oña. 2014. 'Quality of Service in Public Transport Based on Customer Satisfaction Surveys: A Review and Assessment of Methodological Approaches'. *Transportation Science* 49 (3): 605–22. doi:10.1287/trsc.2014.0544.

Shiv, Baba, and Alexander Fedorikhin. 1999. 'Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making'. *Journal of Consumer Research* 26 (3): 278–92. doi:10.1086/209563.

Stanovich, Keith E., and Richard F. West. 2000. 'Advancing the Rationality Debate'. *Behavioral and Brain Sciences* 23 (05): 701–717.

Stradling, Stephen, Michael Carreno, Tom Rye, and Allyson Noble. 2007. 'Passenger Perceptions and the Ideal Urban Bus Journey Experience'. *Transport Policy* 14 (4): 283–92. doi:10.1016/j.tranpol.2007.02.003.

Swärdh, Gunilla Björklund and Jan-Erik. 2016. 'Valuing in-vehicle comfort and crowding reduction in public transport'. Text.Preprint. Accessed July 13. http://swopec.hhs.se/ctswps/abs/ctswps2015_012.htm.

Tirachini, Alejandro, Lijun Sun, Alexander Erath, and Artem Chakirov. 2016. 'Valuation of Sitting and Standing in Metro Trains Using Revealed Preferences'. *Transport Policy* 47 (April): 94–104. doi:10.1016/j.tranpol.2015.12.004.

Trafikanalys, Trafikanalys. 2013. 'Utvärdering av marknadsöppningar i kollektivtrafiken'. 2013:13. Stockholm: Trafikanalys (TRAFA - Transport Analysis). http://www.trafa.se/globalassets/rapporter/rapport_2013_13_utvaerderingar_av_marknadsoeppningar_i_kollektivtrafiken_-_rapport_2013.pdf.

Van de Velde, Didier, Arne Beck, J. van Elburg, and Kai-Henning Terschüren. 2008. 'Contracting in Urban Public Transport'. *Amsterdam: European Commission*. http://www.fedea.net/docs/iv_workshop_transporte/van%20de%20Velde%20-%202014%2005%2003%20Madrid%20(v1.1).pdf.

Wardman, Mark, and Gerard Whelan. 2011. 'Twenty Years of Rail Crowding Valuation Studies: Evidence and Lessons from British Experience'. *Transport Reviews* 31 (3): 379–98. doi:10.1080/01441647.2010.519127.

Yannis, Tyrinopoulos, and Aifadopoulou Georgia. 2008. 'A Complete Methodology for the Quality Control of Passenger Services in the Public Transport Business'. http://www.openstarts.units.it/dspace/handle/10077/5965.