AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Hjortskov, M. (2019). Citizen Expectations and Satisfaction Over Time: Findings From a Large Sample Panel Survey of Public School Parents in Denmark. *American Review of Public Administration*, 49(3), 353-371. <https://doi.org/10.1177/0275074018765822>

Publication metadata

Title:	Citizen Expectations and Satisfaction Over Time: Findings From a Large Sample Panel Survey of Public School Parents in Denmark
Author(s):	Morten Hjortskov
Journal:	<i>American Review of Public Administration</i> , 49(3), 353-371
DOI/Link:	https://doi.org/10.1177/0275074018765822
Document version:	Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Citizen Satisfaction and Expectations over Time: Findings from a Large Sample Panel Survey of Public School Parents in Denmark

Abstract

Expectations are thought to affect how citizens form their attitudes and behavior towards public services. Such attitudes may include citizen satisfaction, where expectations play a fundamental role, and relevant behaviors include choice of services and the decision to voice opinions about them. However, there are few investigations into what drives citizen expectations and even fewer that consider these relationships across time. This article tests whether prior expectations, perceived performance and citizen satisfaction influence future expectations, using a unique dataset that follows individual citizens across two subsequent school satisfaction surveys from 2011 and 2013. The results show that prior expectations have a large and consistent influence on future expectations, as predicted by the literature, whereas the influence from prior perceived performance seems less consistent. Prior satisfaction may also influence future expectations, although the effect is minor. The results indicate that citizens' expectations are rather stable across time and may contain normative values and beliefs about how public services should perform.

Author statement: I would like to thank the 4 anonymous reviewers, Simon Calmar Andersen, Gregg G. Van Ryzin, Oliver James and Christian Bøtcher Jacobsen for thoughtful comments on earlier drafts of this article. I would also like to thank the City of Aarhus for giving access to their satisfaction data.

Introduction

Citizen expectations are an important determinant of attitudes towards public services (Van Ryzin, 2004; James, 2011). If services do not live up to expectations, citizens may be dissatisfied with public services and may react by choosing other services, voting someone else into office, voicing their opinions through protests or complaints or even voting with their feet by moving to another locality (Boyne et al., 2009; Dowding & John, 2012). Moreover, a general push towards measuring performance outcomes has been prominent in public management (Pollitt & Bouckaert, 2011), but many public outcomes are not easy to measure objectively. Public managers and politicians are therefore increasingly giving citizens satisfaction surveys to extract knowledge about performance, satisfaction and expectations (Miller et al., 2009; Van Ryzin, 2013).

However, when using satisfaction as a performance measure, the implicit assumption is that there is a close relationship between satisfaction and performance (Stipak, 1979). Citizen expectations may undercut this performance-satisfaction relationship by creating a standard that performance is compared to before the satisfaction evaluation is formed, and this standard may vary across different citizens (Van Ryzin, 2004, 2006). The possible difference in both expectations and satisfaction will most likely feed back into the performance information created by these satisfaction surveys. It is therefore important for public administration scholars to understand what drives citizen expectations.

Still, little research has examined what influences citizen expectations of public services (James, 2011). Even less research has considered these influences across time, although expectations are thought to update over time when citizens and consumers experience services or products of varying quality (Oliver & Burke, 1999; Higgs et al., 2005). This article therefore tests how citizen expectations update over time by answering the following research question: How do prior expectations, perceived performance and satisfaction affect citizens' future expectations?

Research in consumer satisfaction from the business and marketing literatures has indicated that expectations are updated over time by previous levels of performance in a circular, Bayesian-like fashion – especially when analyzing at the aggregated level (Boulding et al., 1993, p.10; Johnson et al., 1995; Oliver, 2010, p.61). In other words, future expectations are most likely informed by previous experiences with a service and its performance, as well as other contextual factors such as citizen characteristics (Duffy, 2000), word of mouth and communications through the media (Van Ryzin, 2006) and even prior satisfaction (Clow et al., 1997). These possible feedback effects from prior perceptions of performance and satisfaction are even more likely in the public sector because citizens use many services repeatedly over time. The effects of previous performance on later expectations have been investigated before in the literature (see, e.g., Oliver & Burke, 1999; Higgs et al., 2005; James, 2011), but this is not the case with the effect of prior satisfaction. Clow et al. (1997), however, successfully link prior satisfaction to future expectations in three of four investigated service settings. Moreover, there has generally been little focus on citizen expectations and their antecedents in public administration (see James (2011) and Jacobsen et al. (2015) for notable exceptions).

While the possible feedback effect from prior expectations and perceived performance on future expectations would make theoretical sense and match prior research, a feedback effect from prior satisfaction would not. Such a feedback effect would raise practical, theoretical and methodological concerns regarding the use of satisfaction surveys to gather performance information and the consistency of the Expectation-Disconfirmation Model (EDM), which explicitly seeks to take expectations into account when explaining citizen satisfaction (Van Ryzin, 2004). First, in practice governments might try to increase performance or perhaps even try to reduce citizens' expectations in order to create higher satisfaction. But if higher satisfaction at the same time increases expectations for future performance – a feedback effect – then it may be increasingly difficult to meet future

expectations (James, 2011). Such interdependency would also make performance information based on citizen satisfaction more ambiguous and difficult to interpret.

Second, a feedback effect from perceived performance is theoretically meaningful because expectations should be an aggregated measure of the experiences with the service in question – exactly what the theories behind expectation formation and the EDM would state should inform future expectations (Boulding et al., 1993; Zeithaml et al., 1993; Johnson et al., 1995). However, it is less obvious if prior satisfaction at the same time (i.e., controlling for prior perceived performance and expectations) has feedback effects on future expectations. This would be an effect of satisfaction on future expectations that is not informed by perceived performance but instead may rely on emotions, attribution of blame or even cognitive biases and dissonance reduction (Oliver & Winer, 1987; Oliver, 2010). Politicians and practitioners will have a difficult time affecting such influences. Third, a feedback effect from previous satisfaction on future expectations would be a methodological concern, because it would render the EDM endogenous, and it should therefore be viewed as a model that is constantly updated by new experiences with services that might create a spurious relationship between expectations and satisfaction.

This paper attempts to answer how expectations update over time using a two-wave survey design that allows us to follow the same citizen across two consecutive parent satisfaction surveys in Danish public schools (from the years 2011 and 2013). The results show that prior expectations, perceived performance and satisfaction may in fact drive future expectations, but in different ways and with different impact.

The article proceeds as follows. First, the theories of expectation formation and updating and the EDM are reviewed. Second, the research design and analytical choices are discussed. Third, the main results and robustness checks are displayed with a subsequent discussion of the implications of the findings.

Theory

Expectations have been an important topic among satisfaction scholars for a long time (Oliver, 2010). James (2009, p.109, 2011, p.1420) defines expectations as “judgments of what individuals or groups think either will or should happen under particular circumstances”. This definition points at one of the persistent debates about expectations, namely the difference between positive or predictive expectations and normative expectations (Boulding et al., 1993; Niedrich et al., 2005; James, 2009, 2011). While the word “think” implies a prediction of future services and therefore speaks to the positive expectations approach, the word “should” implies normative expectations, meaning citizens’ views on what is reasonable or desirable in the public service. Hence, positive expectations are predictions of how citizens think the service “will” be in the future, whereas normative expectations are rooted in norms and values about how citizens think things “should” be and therefore are considered more stable over time (Prakash, 1984). These expectations concepts are established and used in the public administration literature (e.g., Poister & Thomas, 2011; Jacobsen et al., 2015).

While few studies address normative expectations in the public sector, there has been a substantial increase in studies utilizing positive expectations, especially through the Expectation-Disconfirmation paradigm (James, 2011). This emerging work on positive expectations draws on the extensive literature on rational expectations from economics, where they play an important role for predicting consumer behavior, price and decision-making. In this literature, expectations are a fundamental part of the consumer’s effort to maximize expected utility. The approach of rational expectations assumes that agents use all available relevant information in creating these expectations and that on average the expectations will equal the predictions of the relevant economic model (Muth, 1961).

Rational expectations are often viewed as a continuous Bayesian revision or learning process, where priors are constantly being updated as information is accumulated (Cyert & DeGroot, 1974). On average, citizens should more or less be able to predict the performance of a public service they interact with on a regular basis, because they are constantly able to update their priors and learn from the mistakes they have made in this prediction (Oliver, 1989; Johnson & Fornell, 1991). This process of updating expectations may even take place during service delivery itself (Oliver & Burke, 1999; Higgs et al., 2005).

Researchers have also focused on adaptive expectations, which are also said to be updated along the way by averaging prior expectations and one's most recent personal experiences (Nerlove, 1958). The adaptive expectations theory holds the psychologically attractive view of expectation formation that citizens have an ongoing expectation for the performance of the public service in question which is updated by actual performance and available information. Adaptive expectations are therefore consistent with the psychological theories of anchoring and adjustment (Tversky & Kahneman, 1974; Johnson et al., 1995, p.696).

The primary difference between rational and adaptive expectations is that adaptive expectations are biased towards more recent performance information, depending on how adaptive the expectations are considered to be. The assumption is that adaptive expectations, contrary to rational expectations, typically place greater weight on recent experiences. Perhaps for that reason, adaptive expectations are viewed as better able to predict individual expectations, which can change more quickly and are more exposed to small performance fluctuations, whereas rational expectations typically perform better in the aggregate and in stable markets (Nerlove, 1958; Johnson et al., 1995). Thus, both the rational and the adaptive expectation approaches to the process of positive expectation formation assume that citizens are influenced by previous information about and experiences with a service.

The Expectation-Disconfirmation Model

The importance of expectations in the satisfaction literature stems from the EDM. The EDM was developed in the business and communication literature in the sixties and seventies (Engel et al., 1968; Oliver, 1980) and has travelled to the citizen satisfaction literature in public administration over the last decade (Van Ryzin, 2004, 2006; Andersen & Hjortskov, 2016). Put simply, it states that satisfaction is formed from a citizen's expectations of a service contrasted with their perceived performance of the service. If the service lives up to their expectations (confirms them), the citizen is most likely satisfied – if not, the citizen is most likely dissatisfied. The contrast between citizen expectations and perceived performance is termed disconfirmation, hence the name of the model (Olson & Dover, 1976; Oliver, 1980) (see Figure 1).

The conceptualization of expectations in the EDM literature is most often positive expectations, which resonates well with the cognitive foundations of the model (Oliver, 1993). In the EDM, expectation formation is described within the satisfaction process (Oliver, 1980), and this view has been more or less adapted to public administration with the introduction of the EDM (Van Ryzin, 2004, 2006). In the EDM literature, typical influences in expectation formation are thought to be personal experience, word-of-mouth, social referents, the media, public auditors, image, and the public organizations themselves (LaTour & Peat, 1979; Van Ryzin, 2004).

The key mechanism in the model is therefore the use of expectations as an adopted standard or anchor that can be compared with the performance that is experienced. The EDM holds that since all citizens evaluate a particular service with their own standard as a referent or starting point (expectations for the service), different citizens may evaluate the same service quite differently. Therefore, an investigation of satisfaction without taking expectations into account will be inaccurate according to the EDM (Van Ryzin, 2013).

Updated Expectations and Feedback Effects of Satisfaction

It is easy to forget that the EDM is rarely a one-shot game in the public sector. Many of the services that citizens receive through the public sector in modern society are services they receive or encounter repeatedly over time, such as public schools (as children or as parents), health care (as patients or as relatives) and local roads and parks. These large welfare services are often evaluated continuously through satisfaction surveys, which are also often used for research (Miller et al., 2009). In other words, prior expectations most likely inform future expectations through a process of updating or adaptation (Oliver & Burke, 1999; Higgs et al., 2005), but prior satisfaction and perceived performance might also influence future expectations because of the repeated encounters with public services.

The theory behind the EDM is not blind to the temporal considerations of service delivery and experience in the public sector. Since expectations are thought to be connected to all the other constructs in the model (perceived performance, disconfirmation and satisfaction), the learning and updating mechanism inherent in the expectations construct is incorporated in the model (Bolton & Drew, 1991; Boulding et al., 1993). However, many investigations of the EDM need to assume that expectations are exogenous to prior performance and satisfaction, mainly for methodological reasons. Many of these studies are cross-sectional and are measured at one time point, and they therefore require the measured independent variables to be exactly that: independent. This turns into a strict assumption of exogeneity of expectations, which is used, and widely acknowledged, in many EDM studies in public administration (Van Ryzin, 2004, p.436, 2006, p.600; James, 2009, p.113; Morgeson, 2013, p.292; Jacobsen et al., 2015, p.833). The exogeneity assumption, however, becomes very strong when it is used together with theoretical arguments on how expectations are formed

through various processes of learning, updating and experiencing – chief among those one's previous personal experience with the service.

The marketing and business EDM literature also notes the possible influence of prior perceived performance and satisfaction on subsequent expectations, although it is mostly seen as a phenomenon happening at the aggregate level: *"The role of collective expectations in aggregate satisfaction stems from the fact that they reflect prior levels of performance (e.g., quality) and satisfaction delivered by firms and industries. [...] The satisfaction or dissatisfaction experienced will create expectations for similar levels of quality-related satisfaction in the future"* (Oliver, 2010, p.61). Although it is used as an explanation at the aggregated market level, there is a clear presumption that both prior performance and satisfaction might influence future expectations.

Also, research in public administration has shown that previous performance positively affects future citizen expectations at the individual level. In an observational study, James (2011) found that prior performance as measured by auditors' assessments of local government's performance in England significantly affected citizens' positive expectations and to some extent their normative expectations as well. Moreover, in a subsequent field experiment James found that providing information about the good or bad performance of the local authorities influenced positive expectations but not normative expectations.

Furthermore, satisfaction itself might drive future expectations, an idea expressed as early as 1977 by Hunt: *"It is not a matter of once one is satisfied, he no longer strives, but rather that satisfaction brings raised expectations such that current performance is no longer satisfactory"* (Hunt, 1977, p.26). To this author's knowledge, only one prior study has directly investigated this proposition empirically: Clow et al. (1997). Inspired by Zeithaml et al. (1993), the study develops a model of service expectations consisting of several possible antecedents, such as word-of-mouth, firm image, advertising, and tangibles as well as satisfaction. The authors theorize that past experience with a

service firm consists of both service quality and satisfaction and that prior satisfaction with a service will create positive word-of-mouth about the service and also repeat repurchase behavior and therefore higher expectations in the future (Clow et al. 1997, p.232). The study confirms the relationship in three of four different service contexts.

A feedback effect of satisfaction on future expectations would not only render the EDM endogenous; it would also call into question the cognitive assumptions behind the EDM (Oliver, 1993, p.419). An effect of prior satisfaction on future expectations, controlled for prior perceived performance and expectations, makes it inherently difficult to interpret citizen satisfaction as a pure performance measure. Such an effect would not be informed by the perceived performance but perhaps rely more on personal emotions, equity considerations, attribution of blame or even cognitive biases and dissonance reduction (Oliver, 2010).

Later satisfaction theory has described satisfaction as an affective evaluation, whereas perceived performance has been conceptualized as a more cognitive one (Iacobucci et al., 1994). For example, parents may find the performance of a public school to be mediocre, but they may also generally like the teachers and be disinclined to rate them low on satisfaction. The subjective performance evaluation may still be “...a *hard performance-based judgment*” (Oliver, 1997, p.178), but the satisfaction evaluation will probably also incorporate more emotional aspects. Therefore, an effect of prior satisfaction on future expectations in theory would build more on emotions and perhaps the “warm feeling” of well-being (or, at the other extreme, anger and frustration) associated with those feelings. In other words, citizen satisfaction stripped from performance influences and expectations may contain emotions evoked from the public service encounter (Oliver, 1993).

In sum, it seems that not only prior expectations but also prior perceived performance and satisfaction could potentially positively affect future expectations. These theoretical propositions lead to three different hypotheses, which are illustrated in Figure 1 using the two years of data indicated. The figure

uses the structure of the EDM with performance (Oliver, 2010) as a starting point, albeit with some of the links removed for presentational reasons (see also Van Ryzin, 2006).

[Figure 1 about here]

The first hypothesis, then, is derived as a consequence of the rational and adaptive expectations theories and the updating of expectations based on prior expectations:

H1: Prior expectations affect future expectations positively

The second hypothesis also builds on the theories about the updating of expectations but also on expectation research, which shows that prior performance affects future expectations (e.g., Oliver & Burke, 1999; James, 2011):

H2: Prior performance positively affects future expectations

Only one previous empirical examination of the satisfaction feedback effect has been found (Clow et al., 1997). Since this feedback effect of satisfaction would have large practical, methodological and theoretical consequences for the use of citizen satisfaction, it is important to investigate its magnitude. This leads to the third and final hypothesis:

H3: Prior satisfaction positively affects future expectations

In Figure 1, the black arrows represent Hypotheses 1-3 while the gray arrows represent the links in the EDM, some of which are not directly analyzed in the current study.

Design

Context and Data: Parent Satisfaction with Schools in Denmark

The data in this paper are from Danish public schools in the city of Aarhus, which supplies both primary and lower secondary education. The Danish school system is relatively decentralized: the schools function under municipality governments which determine their funding and the local school regulations, while the Folkeskole Act at the national level provides the overall framework for the schools' activities. Still, there is a high level of discretion at the school level to allow for each school to incorporate their own local context (Ministry of Education, 2017). The case of Danish schools fits well with an investigation of citizen attitudes and expectation formation, since public schools in Denmark, as well as in many other countries, constitute a large portion of the public services that are delivered and have the core features of typical public service delivery: large organizations, complexity, ambiguity, importance and continuous use. Furthermore, public schools are often evaluated through citizen satisfaction surveys in Denmark (Rasmussen et al., 2014) and elsewhere (e.g., Boston Public Schools, 2016; City of New York - Department of Education, 2017).

The satisfaction data are from the City of Aarhus' biannual parent satisfaction survey. The surveys are sent to parents of all children in schools, daycare institutions, after-school programs, and youth clubs. The parents are sent one survey per child per institution, which means that for some children (i.e., for school children who are also enrolled in either an after-school program or a youth club), their

parents actually receive two surveys. For the purpose of this paper, only the school survey data are used.

One particularly attractive attribute of these parent surveys is that they can be linked across years at the individual child level. This means that every second year, for every family, each child enrolled in a public school will have an answer to the satisfaction survey attached to them (given that the parents answered the survey). This creates a panel of satisfaction surveys linked over time at the child level, with two-year gaps in between. The survey data also contain information on which school the child attends and which specific classroom they belong to.

In this paper, survey data from the years 2009, 2011 and 2013 are used. The response rates in the three years were 60.2 percent, 62.6 percent and 65.9 percent, respectively, which corresponds to 17,411, 17,833 and 18,611 answers. There are three questions in 2011 and 2013 that measure the three constructs in the EDM: Satisfaction (“Overall, how satisfied are you with your child’s school?”), perceived performance (“How often do you experience that the school delivers high quality?”) and expectations (“How often do you expect the school to deliver at a high quality?”). This particular expectations question may tap into both positive and normative citizen expectations. However, people often elicit their normative expectations even when asked about their positive expectations (Spreng et al., 1998), and if the question is interpreted as normative it will be a conservative test of the hypotheses, since normative expectations have been shown to be more difficult to change (James, 2011). All three are 5-point Likert questions, with “Very satisfied, Satisfied, Neither dissatisfied nor satisfied, Dissatisfied, Very dissatisfied” as answer possibilities for the satisfaction question, and “Always, Most of the time, Sometimes, Rarely, Never” for the other two questions. In 2009, only the overall satisfaction question was included. Only these single-item measures exist for expectations, perceived performance and overall satisfaction in the surveys, even though several items per construct

is preferable. The internal and external validity challenges that this may create are discussed in the conclusion.

Furthermore, the satisfaction survey data is linked to detailed administrative data on the children's parents from Statistics Denmark from 2011. These rich administrative data can be linked to the survey data through each child's social security number and used as control variables. Moreover, data from the National Tests in Denmark were obtained from the Ministry of Education. The tests are mandatory and they are taken online, where an adaptive algorithm determines their difficulty. The tests used in this article are reading and math from the spring of 2013 (before the 2013 survey), taken by 2nd, 4th, 6th, and 8th grade children (reading) and 3rd and 6th grade children (math), and the results are standardized according to convention (Beuchert & Nandrup, 2014). Lastly, two organizational measures, school budgets measured in kroner per child and change of school leader between 2011 and 2013, were added to the dataset in order to inform the robustness check of relevant influences at this level.

Statistical Models

The statistical methods used in this paper seek to establish whether there is a direct effect of prior expectations, perceived performance and satisfaction on future expectations. To reach this goal, various strategies are used. First, a set of OLS models are presented to observe each of the three variables' influence on future expectations. These models have one central assumption in common: the perceived performance construct captures all relevant performance variation. If not, unobserved performance variation might explain effects of prior expectations, perceived performance and satisfaction on future expectations. This is of course also a theoretical assumption behind the EDM (Andersen & Hjortskov, 2016). Therefore, the second strategy estimates the effects within schools.

This school fixed effects approach controls away any time-invariant variation at the school level by only comparing parent answers within schools. This hopefully captures the common performance within the different schools – some of which may not be captured by the perceived performance construct. Time-invariant variation at the school level could, for example, be the effect of an especially efficient school principal or a certain culture or focus at the school that could affect satisfaction and expectations. The characteristics of the neighborhoods surrounding the schools may also influence parent satisfaction.

In Denmark, children are automatically signed up to the public school in the district where they live. However, parents are free to apply to a public or private school other than their district school, and around 12 percent did so in 2011 (Rambøll, 2011, p.11). Neighborhood characteristics do not typically change much in two years and are therefore partly captured in the school fixed effect (as part of the robustness checks described below, classroom and family fixed effects were also applied to tackle the problem). These unobservables could be correlated with prior expectations, perceived performance and satisfaction, and without the school fixed effects they could end up in the error term of the regression and cause endogeneity bias (Allison, 1990). Equation (1) formalizes the school fixed effects model:

$$E_{ijt} = \beta_0 + \gamma E_{ij(t-1)} + \delta S_{ij(t-1)} + \omega P_{ij(t-1)} + \beta_1 \mathbf{X}_{ij(t-1)} + \alpha_j + \varepsilon_{ij} \quad (1)$$

Here, E is expectations for student i in school j in the year 2013. P is the perceived performance in 2011 with the parameter of interest ω to be estimated and S is the satisfaction in 2011 with the associated parameter of interest δ to be estimated. γ is the estimated effect of the lagged dependent variable expectations in 2011. $\mathbf{X}$ is a vector of individual-level covariates measured in 2011. α is the fixed, cluster-specific parameter at the school level, while ε is the individual error term. In the

analysis, prior satisfaction (S), perceived performance (P) and expectations ($E_{(t-1)}$) are introduced one at a time and in different combinations to illustrate their individual influence.

Equation (1) is estimated with linear fixed effects models with standard errors clustered at the school level. The outcome measure is a single, 5-point Likert-type item (expectations), and because of this the estimations have also been tested with ordinal logit models. The results are essentially the same. It should be noted, however, that these results are obtained without fixed effects, since using these would result in inconsistent estimates due to the incidental parameter problem with relatively few observations within clusters (Rabe-Hesketh & Skrondal, 2012, p.557).

Another point to note is that much of the performance of the public schools may be controlled out with the fixed effects, since parent answers are only compared within schools/classrooms. This also means that the perceived performance variable only contains information from within these clusters. This is of course the purpose of the fixed effects strategy, but the effect of the perceived performance construct will at the same time be stripped from the performance at the school level that is constant across schools. On the other hand, the perceived performance item might contain more information than just performance (Wirtz & Mattila, 2001), which the variable can still pick up on. However, the role of the perceived performance variable in these estimations is as more of a control variable, and the fixed effects models therefore mainly focus on H1 and H3 about prior expectations and satisfaction affecting future expectations.

As part of both of these strategies, prior expectations ($E_{(t-1)}$) are included in some of these OLS and fixed effects regression models. Since an implication of the EDM is that citizens with low expectations should more easily be satisfied given perceived performance, prior expectations could drive both prior satisfaction and future expectations (James, 2009, p.113). This is a case of possible selection on the dependent variable, where an analysis of covariance is a good strategy (Allison, 1990; Morgan & Winship, 2014, p.374). When the lagged dependent variable is introduced in the models,

the interpretation of the effect of prior satisfaction and perceived performance on future expectations is a net effect of these variables on the change in expectations between the two waves of surveys. The key assumption in these analyses is that unobserved variables that affect both satisfaction and expectations do not have direct effects on future expectations, except through prior expectations.

For control variables, a large register dataset on the parents and children is used. Child-level controls include grade, age, gender, number of siblings and change of school. Parent-level controls include education, income, gender of respondent, SES group (for example, unemployed, on maternity leave, early retirement, enrolled in education, etc.), and whether the parent is single. Information on who answered the surveys (mother, father or other) is also included (see Table A.1).

Lastly, as part of the main analyses, a path model is applied. The main advantage in this study of the path model is that it allows for the simultaneous inclusion of several variables and gives estimates for the theoretical connections (paths) between them, including possible mediation. Although mediation is difficult and skepticism regarding whether such estimates can be interpreted causally is common (Green et al., 2010), the path model can inform us how the three central concepts in this article may interrelate across time. Here, the extra wave of the satisfaction survey (2009) from the City of Aarhus is utilized. As explained above, this extra survey only contains the item on satisfaction and not the items on expectations and perceived performance, and therefore it is not used in the OLS and fixed effects models.

Robustness Checks

Four different robustness checks are carried out. First, classroom fixed effects are employed. This does not change equation (1) but merely shifts the meaning of j to classes instead of schools. Classroom fixed effects neutralize possible time-invariant confounders at the classroom level, such

as teacher effects. In addition, children in Danish schools are generally not grouped into specific classes based on their (or their parents') characteristics such as ability, family background, neighborhood characteristics or the parents' satisfaction or expectations (which is generally the case in European school systems). Therefore, the classroom fixed effects might overcome some of the possible selection into schools or neighborhoods, because classroom assignment is formed on a somewhat random basis (Ammermueller & Pischke, 2009).

However, even within classes some confounding factors might exist (Vigdor & Nechyba, 2007). For example, teachers might not be assigned randomly to classes, so if very good teachers are either assigned to or choose to teach high-achieving classes, then this might drive high perceived performance, satisfaction and expectations in these classrooms. This "teacher shopping" might be less of a concern in European school systems (Ammermueller and Pischke 2009), but to accommodate the concern, family fixed effects are employed as the second robustness check. The family fixed effect uses within-family variation to estimate the relationships of interest. This requires at least two children attending school within a family where there are parent answers to both of their school satisfaction surveys. Moreover, it requires variation in the answers to the surveys within the same set of parents over time. In other words, this identification strategy is extremely restrictive and also relies on a smaller and possibly quite selective sample.

On the other hand, the strategy can account for time-invariant unobserved influences on parent perceived performance, satisfaction and expectations common to school children within a family. For example, intra-family views on parenting and upbringing could influence perceived performance, satisfaction and expectations directly and perhaps also indirectly influence how children are treated by teachers at the school and hence affect both satisfaction and expectations through this path. These models are only used as a robustness check, since they involve a tradeoff between high internal

validity by only using the variation within families and lower external validity by restricting the sample to families with more than one child in school.

As a third robustness check, I use a multilevel strategy where the two extra organizational level variables (budget change and change of school principal) are utilized in a random-intercept model and where the prior satisfaction variable is allowed to have varying coefficients across schools in a random-coefficient model (Rabe-Hesketh & Skrondal, 2012). The benefit here is that it is possible to take both the school and the classroom levels into account at the same time while controlling for organizational variables (budgets and change in leadership), which is not possible in the fixed effects models. Moreover, the models also allow different intercepts and coefficients that may vary between schools and classrooms when analyzing expectations.

The fourth and final robustness check concerns the possibility that the educational performance of a child at a school has an effect on prior expectations, perceived performance and satisfaction and on future expectations and hence may be a confounder of the main relationships. To investigate this concern, the results from the national tests are utilized as control variables. As mentioned, the national tests in reading and math are carried out at specific grade levels, and therefore part of this robustness check will be in smaller subgroups of children in 2nd, 3rd, 4th, 6th and 8th grade. To rule out power concerns when investigating subgroups, the tests, which are adaptive to the level of the individual child, are also combined into one measure in one of the specifications under the assumption that they measure the same latent performance dimension.

Results

Appendix A contains an overview of the different variables and their descriptive statistics used in this study. Table 1 displays the simple bivariate correlations between the key variables in the analysis

across the years. The only measure from 2009 is satisfaction, and it is (not surprisingly) correlated with satisfaction in 2011 and 2013. However, it is also correlated with perceived performance in both subsequent surveys and, to a lesser extent, with expectations in both 2011 and 2013. This gives the first indication that satisfaction may affect subsequent expectations.

[Table 1 about here]

Expectations 2011 are correlated with satisfaction in the same survey (2011: .132), in the previous survey (2009: .059) and in the survey in 2013 (.043), albeit with a somewhat lower coefficient in the two latter cases. This also indicates the connection between expectations and prior satisfaction. Expectations also seem to be correlated with perceived performance, both within the same survey (.309 in 2011 and .306 in 2013) and between surveys (.086 and .105). All coefficients in Table 1 are significant at the .001 level.

Table 2 presents the results from the OLS estimations. Model 1 estimates the simple correlation between prior satisfaction and future expectations. This results in an effect of .06, which means that moving one point on the 5-point Likert scale of satisfaction changes .06 points on the 5-point Likert scale of expectations. Although the effect is significant at the .001 level, the substantive impact is not large.

[Table 2 about here]

In Model 2, only Perceived Performance 2011 is entered. There is also a significant effect on Expectations 2013 of roughly the same magnitude as from Satisfaction 2011. Entering the two

variables at the same time in Model 3 seems to attenuate the effect of satisfaction on expectations in 2013. Only the Perceived Performance variable is significant. Model 4 introduces expectations in 2011 as a lagged dependent variable. This also means that the coefficients are to be interpreted differently; they are now effects on changes in expectations from 2011 to 2013 (analysis of covariance). As expected, prior expectations have a large effect on future expectations. The coefficient equals .27 and is significant at the .001 level. Introducing Perceived Performance 2011 in Model 5 does not change this coefficient much, and the coefficient on Perceived Performance is now greatly attenuated and insignificant. In Model 6 it is clear that this is not the case with prior satisfaction. When Prior Satisfaction and Prior Expectations are entered, the coefficient on Prior Expectations again stays largely unaltered, but the coefficient on Satisfaction is still significant at the .05 level and reads .03. Models 7 and 8 contain all of the three variables without covariates (Model 7) and with covariates – including budget – and school principal changes at the school level (Model 8). Again the effect of satisfaction is significant and largely unchanged in magnitude. Neither the introduction of Perceived Performance nor the covariates changes the coefficient much. The Perceived Performance coefficient, on the other hand, is diminished to almost zero in Models 7 and 8 and is insignificant.

The results from Table 2 confirm H1, as prior expectations seem to have a large effect on future expectations. Citizens seem to update their expectations along the way. The evidence concerning H2, however, is mixed. Entered by itself and with Satisfaction (Models 2 and 3), there is a small but significant effect of prior perceived performance on future expectations. However, when entered with Prior Expectations (Model 5) and both Prior Expectations and Satisfaction (Models 7 and 8), there is no effect. It seems that prior expectations account for a substantive part of the effect from prior perceived performance on future expectations. These specifications do not include fixed effects, so Perceived Performance should in theory be able to pick up on any actual performance effects as

discussed in the design section above. Lastly, H3 is confirmed in four of five models containing the Satisfaction variable. Only in Model 3, when Prior Perceived Performance and Satisfaction are entered simultaneously, is the effect of satisfaction insignificant. The effect is, however, quite small. Table 3 presents the results of the school fixed effects models that rule out any time-invariant variation at the school level. This means that the Perceived Performance variable is perhaps stripped from some of the important performance variation at the school level. On the other hand, this is done to see if the satisfaction effect might be an effect of some unmeasured performance dimension at the school level not captured by the Perceived Performance construct. However, the results are basically the same as the OLS estimations in Table 2. This is also the case in Model 8, where the covariates do not include budget or school principal change.

[Table 3 about here]

Robustness Checks

Tables B.1 through B.4 report the robustness checks described in the design section. With a few exceptions, the general results from Tables 2 and 3 hold. Figure 2 presents a selection of the robustness checks (the most inclusive specifications), with the OLS and school fixed effects results from Tables 2 and 3 (Model 8) presented for comparison.

[Figure 2 about here]

First, the effect of prior satisfaction on future expectations seems to hold in all specifications, except the family fixed effects. The effect is still quite small but very much alike in all cases (between .027 and .034 on a 5-point scale). As mentioned, the family fixed effects specification is very demanding of the data and only includes families with more than one child in school. The standard error is therefore slightly higher (.017 vs. around .010 in the other specifications). The slightly higher standard error is also present in the child ability models that include the results from the national tests as controls, especially when using the single tests (see Models 1-7 in Table B.4). Except for the reading test in 2nd and 4th grade, these estimates are insignificant. However, when combining all tests into one measure, which is the model shown in Figure 2, the effect of prior satisfaction is significant at the 5% level with point estimates of .034. In general, this confirms Hypothesis 3, although the substantive impact is not large.

Second, the effect of prior expectations on future expectations is both large and significant in all but one case. The effect is around .23-.27 and significant at the 1% level. The exception is again the family fixed effects, where the effect is almost gone. This may be an effect of the specific subgroup of parents that have more than one child in school ($n = 5104$), but it may also be an effect of the removal of time-invariant variation, such as the personality or values of the parent, that may affect the relationship between prior and future expectations. It is, however, interesting that prior satisfaction almost retains its effect on future expectations while prior expectations do not. In general, it seems that expectations are updated in a Bayesian fashion with a rather fixed core between the two waves of surveys.

Lastly, perceived performance is small and insignificant in all of the models shown in Figure 2. This is (almost) also the case in Tables B.1 through B.4. The exceptions concern the models that do not incorporate prior expectations. Here, there is a tendency for Perceived Performance to be larger (and even significant in the classroom fixed models) and it seems to attenuate the effect of Prior

Satisfaction in the models where only these two variables are entered. This is also the case in the multilevel and the child ability specifications (not shown). In general, Hypothesis 2 does not seem to hold.

A Path Model of the Relationships

As a final way of gaining insight into the relationships between prior citizen expectations, perceived performance, satisfaction and future expectations through the three waves of available data (2009, 2011 and 2013), a path model is employed. As described in the Design section, the path model is estimated with structural equation modeling (structural component) and the model shown is without covariates (including covariates does not change the substantial conclusions). The Perceived Performance and Expectations variables in both 2011 and 2013 have been allowed to have correlated residuals, as this creates a better fit of the model (again, the estimates do not change much if they are not allowed to correlate).

Figure 3 shows the variables included and the estimates of the different paths. All estimates are significant at the 1% level, except the estimate of Satisfaction 2011 on Expectations 2013, which is significant at the 5% level.

[Figure 3 about here]

It is evident from Figure 3 that expectations are interlinked over time. The direct link between Exp2011 and Exp2013 equals .27, which also resembles most of the results from the above regressions. Likewise, Perceived Performance 2011 and 2013 as well as Satisfaction 2011 and 2013 have quite strong direct links between them. Interestingly, the strongest paths are between perceived performance and satisfaction within the same year, which resembles results often obtained in the

satisfaction literature (Van Ryzin, 2006, 2013; Poister & Thomas, 2011), although there is no separate disconfirmation construct available in the current study, unlike in these studies. Furthermore, it seems that expectations' correlation with satisfaction in both 2011 and 2013 is negative, also echoing some of the satisfaction literature (Poister & Thomas, 2011).

Table 4 shows the direct, indirect and total effects in the model. The total effect of Prior Satisfaction in 2009 on Expectations in 2011 is .05. This particular path has not been included in the above regressions, but it is of the same magnitude as the results from the above regressions. The outcome of Expectations 2013 is a bit more informative since we can utilize the full set of variables measured before 2013. The total effect of (Prior) Expectations 2011 is .27, a sizable effect that consistently appears (except in the family fixed effects regressions). Again, this confirms Hypothesis 1. Perceived Performance also has a slight total effect here, but again it is rather small (.02). This speaks in favor of Hypothesis 3, although the substantive impact seems small and most other specifications do not attribute much impact to perceived performance. Satisfaction 2011 has a total effect of .03 on Expectations 2013, and somewhat surprisingly Satisfaction 2009 has almost the same total effect on Expectations 2013 (.03) some four years later. Hypothesis 3, suggesting that prior satisfaction may affect future expectations, receives more support here and it appears that it is not just an effect solely present in the years 2011-2013.

When it comes to the satisfaction outcomes, it is clear that both prior and current perceived performance play a large role. The total effect of Perceived Performance in 2011 and 2013 on Satisfaction in 2013 is .39 and .71 respectively. This is a larger total effect than the one between Satisfaction 2011 and 2013 (.30). It seems that perceived performance does not play a large role in determining citizens' future expectations, but instead correlates quite substantively with future satisfaction. The total effects of both Prior and Current Expectations on Satisfaction in 2013 are negative and quite small (-.04 and -.07 respectively).

One interesting question regarding the satisfaction outcomes is whether the effect of prior satisfaction on future expectations spills over to later satisfaction. This can be calculated using the path coefficients from the model in Figure 3/Table B.5 and multiplying the indirect path from Sat2011 to Sat2013 via Exp2011: $.03 * -.07 = -.002$. This is a very small, negative effect, but it is significant at the 5% level (tested using a nonlinear comparison of point estimates, not shown). It seems that the prior effect of satisfaction on future expectations does spill over to later satisfaction in the form of less satisfaction, but the effect is minuscule. In comparison, the indirect effect of Sat2011 on Sat2013 via Perf2013 is: $.15 * .71 = .107$, which is significant at the 1% level. Together, the two indirect paths add up to the full indirect path of Sat2011, reported in Table 4 under Sat2013 (.105).

Discussion and Conclusion

This article has discussed and investigated how prior expectations, perceived performance and satisfaction may influence future expectations. Since most theories of expectation formation focus on the updating of expectations over time and emphasize the importance of prior experiences in this process, this paper has investigated whether the expectations construct is explained by prior expectations and perceived performance of citizens. Furthermore, it has analyzed whether prior citizen satisfaction may feed back into future expectations, as satisfaction is also mentioned as an antecedent of expectations in the literature.

The proposed relationships were investigated in a number of OLS and fixed effects regressions, multilevel models and a path model. The updating of expectations (H1) has to a large extent been verified and the effect tends to be large. Only in the family fixed effects regressions were they not present. However, the effect of perceived performance on future expectations (H2) is not confirmed. When entered alone or with Satisfaction the effect tends to be significant, albeit quite small. When

entered with Expectations it disappears. This is the picture even in models without the fixed effects (OLS, multilevel and path models) that might attenuate the performance effects. The feedback effect of satisfaction (H3) seems to be present across most models, but it is a quite small effect. Only in the very restrictive family fixed effect regression with all covariates entered (Model 8) does the effect become insignificant. The effect is also not present in some of the robustness checks using child ability tests. However, this may be an effect of lacking power in these subgroups, as the effect emerges as significant again when combining the different ability tests, resulting in more power.

These results have a number of theoretical, methodological and practical implications. However, some limitations should first be mentioned. One limitation is that this analysis covers only two years of data in one city. This may result in less external validity. The trends in expectations, satisfaction and performance might have been special in these two particular years. Although data from Danish schools is generally of high quality, it should be mentioned that the Danish school system, like any school system, has unique features, and the city of Aarhus is the second largest city in Denmark, causing the schools examined to be larger than the average schools in the country. The results from this study should be replicated in other settings to raise its external validity. However, this study is still an improvement over previous analyses, and the path model showed that various effects might also exist from 2009 to 2011.

Another limitation is the use of single items in the analyses. Generally, the use of indices is recommended, since they can minimize measurement error. Furthermore, uncertainty about how citizens interpret the specific expectations question may be raised (positive or normative?). On the other hand, single items, especially on the expectations and satisfaction constructs, are commonly used in the literature, and if citizens interpret the question as normative, as research has shown many do, the results are conservative since normative expectations are likely more difficult to change.

Lastly, in order to make stronger causal claims, experimental work should be carried out to analyze the relationship between prior satisfaction and future expectations. The advantage of the observational data in this study is the large N and the fact that the citizens here have concrete, everyday experiences with the important, specific service they are evaluating over the course of two years. This is difficult to replicate entirely in an experimental setting, but smaller parts can be replicated in survey, laboratory and field experiments, which have become more common in the literature in recent years (e.g., James, 2011; Van Ryzin, 2013). The present study indicates that more of such work could be fruitful.

The implications of the findings in this article are several. First, the theoretical propositions of the classical expectations literature seems to hold: expectations are updated along the way but also seem to maintain a rather fixed core. The large and consistent effect of prior expectations on future expectations implies a certain amount of temporal stability in expectation formation. This speaks in favor of normative expectations when it comes to public services. Perhaps citizens put much more emphasis on their values and beliefs when evaluating public services, especially services that concern their children. If this is the case, it also makes sense that the family fixed effects obliterate the effect of prior expectations – they are to a large extent internal and fixed to the evaluator and therefore do not vary much within parents.

Second, the theoretical EDM can be seen in a somewhat different light. The model is built around the idea that citizen satisfaction is deeply dependent on expectations. However, there seems to be more to the story when prior satisfaction and expectations can influence future expectations. The EDM should, especially when it is used in the public service context, be seen as a model that is updated over time and is dependent on the evaluations made in the previous period in cases where a service is used, and perhaps evaluated, in a repeated fashion.

Third, the results in this paper are somewhat at odds with the cognitive interpretation of the EDM. If satisfaction, controlled for perceived performance, can affect future expectations, then it must be something other than a performance effect. According to the classical EDM, the perceived performance construct is a hard performance-based construct, and in theory it should incorporate every aspect of performance important to a citizen's satisfaction evaluation. Therefore, the effect of prior satisfaction on future expectations must largely consist of influences other than performance. The list may be long, but this paper has mentioned influences such as emotions, dissonance reduction and other cognitive biases.

Fourth, the results suggest a methodological implication. Expectations should generally not be assumed exogenous in the public service context. Perhaps in some cases, like in the case of an entirely new and unfamiliar service, this is a valid assumption – though such cases may, on the other hand, contain other challenges (McGill & Iacobucci, 1992). But in the public service context, many services are provided on a daily basis and citizens have the possibility to adjust their expectations of the services they experience.

Finally, there may be some practical implications of this study. One is that practitioners may want to control for prior satisfaction and analyze satisfaction changes rather than levels to avoid some of these influences. Another practical solution may be to employ normative expectation items in citizen surveys, since these have been found to be less affected by prior perceived performance (James, 2011) and therefore may be less susceptible to the influence of prior satisfaction as well. Although the expectations question used in this paper may be interpreted as normative, a purely normative question may overcome the feedback effect.

Scholars in the citizen satisfaction literature have discussed whether public managers might successfully manipulate expectations with raised satisfaction as a result, albeit with mixed findings (James, 2009, pp.119–120, 2011, p.1432, Van Ryzin, 2004, p.446, 2006, p.609). This study has raised

the concern that higher performance and citizen satisfaction may also raise future expectations that are harder to meet and that the end result might actually therefore be lower, not higher, citizen satisfaction. The results of this study suggest, however, that prior expectations are a large part of future expectations, and that the effects of prior perceived performance and satisfaction are small. Thus, public managers and politicians should continue to focus on raising performance and citizen satisfaction, but the results also indicate that they should communicate clearly what citizens can expect from public services in the future.

Literature

- Allison, P.D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology*. 20 (1). p.pp. 93–114.
- Ammermueller, A. & Pischke, J. (2009). Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*. 27 (3). p.pp. 315–348.
- Andersen, S.C. & Hjortskov, M. (2016). Cognitive Biases in Performance Evaluations. *Journal of Public Administration Research and Theory*. 26 (4). p.pp. 647–662.
- Beuchert, L.V. & Nandrup, A.B. (2014). The Danish National Tests. *Economics Working Papers*. 2014–25. p.pp. 1–51.
- Bolton, R.N. & Drew, J.H. (1991). A longitudinal analysis of the impact of service changes on customer attitudes. *The Journal of Marketing*. p.pp. 1–9.
- Boston Public Schools (2016). *School Climate Survey for Parents 2016*. [Online]. Available from: <http://www.bostonpublicschools.org/Domain/737>. [Accessed: 1 February 2017].
- Boulding, W., Kalra, A., Staelin, R. & Zeithaml, V.A. (1993). A dynamic process model of service quality: From expectations to behavioral intentions. *Journal of Marketing Research*. 30 (1). p.pp. 7–27.
- Boyne, G.A., James, O., John, P. & Petrovsky, N. (2009). Democracy and Government Performance: Holding Incumbents Accountable in English Local Governments. *The Journal of Politics*. 71 (04). p.pp. 1273–1284.
- City of New York - Department of Education (2017). *NYC School Survey - Survey Overview*. [Online]. Available from: <http://schools.nyc.gov/Accountability/tools/survey/default.htm>. [Accessed: 31 January 2017].
- Clow, K.E., Kurtz, D.L., Ozment, J. & Soo Ong, B. (1997). The antecedents of consumer expectations of services: an empirical study across four industries. *Journal of Services Marketing*. 11 (4). p.pp. 230–248.
- Cyert, R.M. & DeGroot, M.H. (1974). Rational expectations and Bayesian analysis. *The Journal of Political Economy*. p.pp. 521–536.
- Dowding, K.M. & John, P. (2012). *Exits, voices and social investment: citizens' reaction to public services*. Cambridge; New York: Cambridge University Press.
- Duffy, B. (2000). *Satisfaction and Expectations: Attitudes to Public Services in Deprived Areas*. [Online]. Rochester, NY: Social Science Research Network. Available from: <https://papers.ssrn.com/abstract=1158936>. [Accessed: 19 October 2017].
- Engel, J.F., Kollat, D.T. & Blackwell, R.D. (1968). *Consumer Behavior*. New York: Holt, Rinehart and Winston.

- Green, D.P., Ha, S.E. & Bullock, J.G. (2010). Enough Already about “Black Box” Experiments: Studying Mediation Is More Difficult than Most Scholars Suppose. *The Annals of the American Academy of Political and Social Science*. 628 (1). p.pp. 200–208.
- Higgs, B., Polonsky, M.J. & Hollick, M. (2005). Measuring expectations: forecast vs. ideal expectations. Does it really matter? *Journal of Retailing and Consumer Services*. 12 (1). p.pp. 49–64.
- Hunt, H.K. (1977). *Consumer Satisfaction and Dissatisfaction: Perspectives and Overview*. Marketing Science Institute.
- Iacobucci, D., Grayson, K. & Ostrom, A. (1994). The Calculus of Service Quality and Customer Satisfaction: Theoretical and Empirical Differentiation and Integration. In: T. A. Swartz, D. E. Bowen, & S. W. Brown (eds.). *Advances in Service Marketing and Management: Research and Practice*. Greenwich, CT: JAI Press, pp. 1–67.
- Jacobsen, R., Snyder, J.W. & Saultz, A. (2015). Understanding Satisfaction with Schools: The Role of Expectations. *Journal of Public Administration Research and Theory*. 25 (3). p.pp. 831–848.
- James, O. (2009). Evaluating the Expectations Disconfirmation and Expectations Anchoring Approaches to Citizen Satisfaction with Local Public Services. *Journal of Public Administration Research and Theory*. 19 (1). p.p. 107.
- James, O. (2011). Managing Citizens’ Expectations of Public Service Performance: Evidence from Observation and Experimentation in Local Government. *Public Administration*. 89 (4). p.pp. 1419–1435.
- Johnson, M.D., Anderson, E.W. & Fornell, C. (1995). Rational and Adaptive Performance Expectations in a Customer Satisfaction Framework. *Journal of Consumer Research*. 21 (4). p.pp. 695–707.
- Johnson, M.D. & Fornell, C. (1991). A framework for comparing customer satisfaction across individuals and product categories. *Journal of economic psychology*. 12 (2). p.pp. 267–286.
- LaTour, S.A. & Peat, N.C. (1979). Conceptual and methodological issues in consumer satisfaction research. *Advances in consumer research*. 6 (1). p.pp. 431–437.
- McGill, A.L. & Iacobucci, D. (1992). The Role of Post-Experience Comparison Standards in the Evaluation of Unfamiliar Services. *NA - Advances in Consumer Research*. 19. p.pp. 570–578.
- Miller, T.I., Kobayashi, M.M. & Hayden, S.E. (2009). *Citizen Surveys for Local Government*. 3rd Ed. International City Management Association.
- Ministry of Education (2017). *About the Folkeskole*. [Online]. 2017. Available from: <http://eng.uvm.dk/primary-and-lower-secondary-education/the-folkeskole/about-the-folkeskole>. [Accessed: 19 October 2017].
- Morgan, S.L. & Winship, C. (2014). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. 2 edition. New York, NY: Cambridge University Press.
- Morgeson, F.V. (2013). Expectations, Disconfirmation, and Citizen Satisfaction with the US Federal Government: Testing and Expanding the Model. *Journal of Public Administration Research and Theory*. 23 (2). p.pp. 289–305.

- Muth, J.F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*. 29 (3). p.pp. 315–335.
- Nerlove, M. (1958). Adaptive expectations and cobweb phenomena. *The Quarterly Journal of Economics*. p.pp. 227–240.
- Niedrich, R.W., Kiryanova, E. & Black, W.C. (2005). The dimensional stability of the standards used in the disconfirmation paradigm. *Journal of Retailing*. 81 (1). p.pp. 49–57.
- Oliver, R.L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of marketing research*. 17 (4). p.pp. 460–469.
- Oliver, R.L. (1993). Cognitive, affective, and attribute bases of the satisfaction response. *Journal of consumer Research*. p.pp. 418–430.
- Oliver, R.L. (1989). Processing of the satisfaction response in consumption: a suggested framework and research propositions. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*. 2 (1). p.pp. 1–16.
- Oliver, R.L. (2010). *Satisfaction: A behavioral perspective on the customer*. Second Edition. Armonk, NY: M.E. Sharpe. Available from: [Accessed: 17 July 2012].
- Oliver, R.L. (1997). *Satisfaction: A behavioral perspective on the customer*. Boston: McGraw-Hill. Available from: [Accessed: 17 July 2012].
- Oliver, R.L. & Burke, R.R. (1999). Expectation Processes in Satisfaction Formation - A Field Study. *Journal of Service Research*. 1 (3). p.pp. 196–214.
- Oliver, R.L. & Winer, R.S. (1987). A framework for the formation and structure of consumer expectations: Review and propositions. *Journal of Economic Psychology*. 8 (4). p.pp. 469–499.
- Olson, J.C. & Dover, P. (1976). Effects of expectation creation and disconfirmation on belief elements of cognitive structure. *Advances in consumer research*. 3. p.pp. 168–175.
- Poister, T.H. & Thomas, J.C. (2011). The Effect of Expectations and Expectancy Confirmation/Disconfirmation on Motorists' Satisfaction with State Highways. *Journal of Public Administration Research and Theory*. 21 (4). p.pp. 601–617.
- Pollitt, C. & Bouckaert, G. (2011). *Public Management Reform: A comparative analysis-new public management, governance, and the Neo-Weberian state*. New York: Oxford University Press. Available from: [Accessed: 28 October 2014].
- Prakash, V. (1984). Validity and reliability of the confirmation of expectations paradigm as a determinant of consumer satisfaction. *Journal of the Academy of Marketing Science*. 12 (4). p.pp. 63–76.
- Rabe-Hesketh, S. & Skrondal, A. (2012). *Multilevel and Longitudinal Modeling Using Stata, Volumes I and II, Third Edition*. 3 edition. College Station, Tex: Stata Press.
- Rambøll (2011). *Evaluering af mere frit skolevalg (2.0)*. København.

- Rasmussen, P.S., Olsen, L. & Brogaard, L. (2014). *Kommuners erfaringer med BTU-konceptet for brugertilfredshedsundersøgelser - Fra beslutning til gennemførelse og anvendelse*. KORA. Available from: [Accessed: 30 October 2015].
- Spreng, R.A., Mackoy, R. & Dröge, C. (1998). Confounds in the measurement of predictive expectations. *Scholarship and professional work from the College of Business*. p.p. 15.
- Stipak, B. (1979). Citizen satisfaction with urban services: Potential misuse as a performance indicator. *Public Administration Review*. 39 (1). p.pp. 46–52.
- Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*. 185 (4157). p.pp. 1124–1131.
- Van Ryzin, G.G. (2004). Expectations, performance, and citizen satisfaction with urban services. *Journal of Policy Analysis and Management*. 23 (3). p.pp. 433–448.
- Van Ryzin, G.G. (2013). Experimental Test of the Expectancy-Disconfirmation Theory of Citizen Satisfaction. *Journal of Policy Analysis and Management*. 32 (3). p.pp. 597–614.
- Van Ryzin, G.G. (2006). Testing the expectancy disconfirmation model of citizen satisfaction with local government. *Journal of Public Administration Research and Theory*. 16 (4). p.pp. 599–611.
- Vigdor, J. & Nechyba, T. (2007). Peer effects in North Carolina public schools. In: L. Woessmann & P. E. Peterson (eds.). *Schools and the equal opportunity problem*. pp. 73–101. Available from: [Accessed: 4 June 2015].
- Wirtz, J. & Mattila, A. (2001). Exploring the role of alternative perceived performance measures and needs-congruency in the consumer satisfaction process. *Journal of Consumer Psychology*. 11 (3). p.pp. 181–192.
- Zeithaml, V.A., Berry, L.L. & Parasuraman, A. (1993). The nature and determinants of customer expectations of service. *Journal of the academy of Marketing Science*. 21 (1). p.pp. 1–12.

Figure 1: The Theoretical Propositions

Note: The figure illustrates the three hypotheses with black arrows and the remaining links from the Expectation-Disconfirmation Model with performance (Oliver 2010) with gray arrows.

Table 1: Pearsons Correlations between Key Variables in the Analysis

	Satisfaction 2009	Satisfaction 2011	Expectations 2011	Perceived performance 2011	Disconfirmation 2011	Satisfaction 2013	Expectations 2013	Perceived performance 2013	Disconfirmation 2013
Satisfaction 2009	1 17360								
Satisfaction 2011	0.358* 8768	1 17783							
Expectations 2011	0.059* 8702	0.132* 17592	1 17617						
Perceived performance 2011	0.298* 8672	0.649* 17527	0.309* 17492	1 17547					
Disconfirmation 2011	0.217* 8655	0.471* 17472	-0.534* 17492	0.640* 17492	1 17492				
Satisfaction 2013	0.261* 5967	0.390* 10637	0.043* 10555	0.336* 10525	0.261* 10495	1 18568			
Expectations 2013	0.054* 5942	0.078* 10600	0.290* 10522	0.105* 10491	-0.140* 10461	0.151* 18430	1 18455		
Perceived performance 2013	0.217* 5914	0.342* 10560	0.086* 10484	0.363* 10453	0.250* 10424	0.660* 18360	0.306* 18346	1 18377	
Disconfirmation 2013	0.150* 5907	0.244* 10552	-0.149* 10476	0.241* 10445	0.332* 10416	0.474* 18329	-0.517* 18346	0.657* 18346	1 18346

Note: Pairwise correlations. * p<0.001. Number of observations below coefficients

Table 2: The Effect of Prior Expectations, Perceived Performance and Satisfaction (2011) on Future Expectations (2013) – OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Satisfaction 2011	0.061** (0.009)		0.013 (0.011)			0.034** (0.009)	0.037** (0.011)	0.032* (0.011)
Perceived Performance 2011		0.092** (0.010)	0.084** (0.013)		0.015 (0.010)		-0.013 (0.012)	-0.010 (0.012)
Expectations 2011				0.270** (0.012)	0.263** (0.012)	0.266** (0.012)	0.267** (0.012)	0.242** (0.012)
Constant	4.359** (0.037)	4.262** (0.039)	4.242** (0.042)	3.383** (0.056)	3.339** (0.060)	3.267** (0.064)	3.270** (0.064)	3.445** (0.176)
Covariates	NO	NO	NO	NO	NO	NO	NO	YES
N	9283	9283	9283	9283	9283	9283	9283	9283

Note: Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.001$, two-tailed tests. Expectations 2013 as dependent variable. Controls include change in school budgets and change in school principal.

Table 3: The Effect of Prior Expectations, Perceived Performance and Satisfaction (2011) on Future Expectations (2013) - School Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Satisfaction 2011	0.057** (0.012)		0.014 (0.013)			0.032* (0.011)	0.038* (0.012)	0.032* (0.013)
Perceived Performance 2011		0.085** (0.012)	0.076** (0.013)		0.015 (0.012)		-0.010 (0.013)	-0.009 (0.013)
Expectations 2011				0.258** (0.012)	0.253** (0.013)	0.253** (0.012)	0.256** (0.012)	0.237** (0.012)
Constant	4.376** (0.048)	4.290** (0.044)	4.267** (0.053)	3.439** (0.055)	3.405** (0.062)	3.328** (0.067)	3.332** (0.068)	2.353** (0.237)
Covariates	NO	NO	NO	NO	NO	NO	NO	YES
N	10329	10329	10329	10329	10329	10329	10329	10329
N (groups)	50	50	50	50	50	50	50	50
Between std. dev.	0.262	0.276	0.274	0.267	0.270	0.266	0.265	0.207
Within std. dev.	0.626	0.624	0.624	0.603	0.603	0.602	0.602	0.595
ICC	0.149	0.163	0.162	0.164	0.167	0.164	0.162	0.108

Note: Standard errors clustered at the school level in parentheses. * $p < 0.05$, ** $p < 0.001$, two-tailed tests. Expectations 2013 as dependent variable. Controls do not include change in school budgets or change in school principal.

Figure 2: Selected Results from Robustness Checks – Effects of Prior Satisfaction, Expectations and Perceived Performance on Future Expectations

Note: Dots represent the point estimates of the three different variables (prior satisfaction, expectations and perceived performance) in the various models listed to the left; spikes represent their 95% confidence intervals in two-tailed tests. OLS and School Fixed Effects presents the results of Model 8 from Tables 2 and 3 respectively. Classroom and Family Fixed Effects presents Model 8 from Tables B.1 and B.2 respectively. Random Effects, Random-intercept and Random-coefficient presents Models 1, 3 and 4 from Table B.3 respectively. The empty multilevel model (2) has been omitted. Child Ability presents Model 7 from Table B.4 with all tests included. Note that the x-axis has been allowed to vary between the three variables for better presentation.

Figure 3: Path Model

Note: Results from a structural equation model with no latent constructs (path model). ‘Exp’, ‘Sat’ and ‘Perf’ are shorthand for expectations, perceived performance and satisfaction respectively. ‘2009’, ‘2011’ and ‘2013’ refer to the waves of the surveys. All path estimates are unstandardized and significant at the 1% level, except the path from Sat2011 to Exp2013 which is significant at the 5% level (stars are omitted for presentational reasons). Residuals from expectations and perceived performance are allowed to correlate in both 2011 and 2013 for a better fit (RMSEA: 0.144→0.05, CFI: 0.99→0.89, SRMR: 0.08→0.02 with no substantive changes in path coefficients as a consequence).

Table 4: Direct, Indirect and Total Effects in the Path Model

Outcome	Independent variable	Direct Effects	Indirect Effects	Total Effects
Exp2011				
	Sat2009	0.054 **		0.054 **
Sat2011				
	Exp2011	-0.073 **		-0.073 **
	Perf2011	0.678 **		0.678 **
	Sat2009	0.192 **	0.195 **	0.387 **
Perf2011				
	Sat2009	0.293 **		0.293 **
Exp2013				
	Exp2011	0.274 **	-0.002 **	0.272 **
	Sat2011	0.026 *		0.026 *
	Perf2011		0.018 **	0.018 **
	Sat2009		0.025 **	0.025 **
Sat2013				

Exp2011		-0.041 **	-0.041 **
Sat2011	0.194 **	0.105 **	0.299 **
Exp2013	-0.069 **		-0.069 **
Perf2011		0.387 **	0.387 **
Perf2013	0.714 **		0.714 **
Sat2009		0.169 **	0.169 **

Perf2013

Exp2011		-0.011 **	-0.011 **
Sat2011	0.149 **		0.149 **
Perf2011	0.259 **	0.101 **	0.360 **
Sat2009		0.133 **	0.133 **

Note: Results from the path model, unstandardized coefficients. * $p < 0.05$, ** $p < 0.001$, two-tailed tests. No control variables included. The residuals of Perf2011 and Exp2011 and Perf2013 and Exp2013 have been allowed to correlate for better fit.

Appendix A – Descriptive Statistics

Table A.1: Descriptive Statistics

		Obs	Mean	Std. Dev.	Min	Max
Child level						
	Expectations 2013	10329	4.6	0.63	1	5
	Satisfaction 2011	10329	4.05	0.81	1	5
	Expectations 2011	10329	4.53	0.68	1	5
	Perceived Performance 2011	10329	3.69	0.74	1	5
	Standardized test result - danish 2. grade	1421	0.15	0.97	-5.45	3.06
	Standardized test result - danish 4. grade	1186	0.22	0.91	-5.33	2.96
	Standardized test result - danish 6. grade	1161	0.18	0.91	-4.42	2.99
	Standardized test result - danish 8. grade	995	0.20	0.87	-4.63	2.52
	Standardized test result - math 3. grade	1334	0.17	0.95	-4.54	3.44
	Standardized test result - math 6. grade	1183	0.19	0.94	-5.54	4.36
	Standardized test result - combined	6097	0.18	0.93	-5.45	3.44
	Child's gender	10329	0.51	0.5	0	1
	Change of school 2011-2013	10329	0.08	0.28	0	1
Gender of respondent 2011						
	Man	10329	0.22	0.42	0	1
	Woman	10329	0.78	0.42	0	1
	Do not wish to answer	10329	0.00	0.03	0	1
Gender of respondent 2013						
	Man	10329	0.23	0.42	0	1
	Woman	10329	0.75	0.43	0	1
	Do not wish to answer	10329	0.02	0.12	0	1
Mother controls						
	Education - High	10329	0.59	0.49	0	1
	Education - Middle	10329	0.32	0.47	0	1
	Education - Low	10329	0.08	0.27	0	1
	Disposable income	10329	263077	118951	-1147162	5989773
	Mother is working	10329	0.87	0.34	0	1
	Mother is single	10329	0.14	0.35	0	1
	Number of Children (mother)	10329	1.59	0.72	1	9
Father controls						
	Education - High	10329	0.48	0.50	0	1
	Education - Middle	10329	0.33	0.47	0	1
	Education - Low	10329	0.07	0.26	0	1
	Disposable income	10329	285163	314652	-4980977	15190888

Father is working	10329	0.74	0.44	0	1
Father is single	10329	0.07	0.25	0	1
School level					
Change in budget per child 2011-2013	9283	-2745	5508	-25899	18272
Change of school principal 2011-2013	10329	0.20	0.40	0	1

Note: School-level variables have not been adjusted for the number of children attending the different schools. 10 schools changed their principal between the surveys of 2011 and 2013. Income is in Danish kroner and can be negative because of taxes and debt. The combined test results measure shown uses the Danish 6th grade results - not the math 6th grade results

Appendix B – Robustness Checks

Table B.1: The Effect of Prior Expectations, Perceived Performance and Satisfaction (2011) on Future Expectations (2013) - Classroom Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Satisfaction 2011	0.054** (0.010)		0.010 (0.011)			0.031* (0.010)	0.032* (0.011)	0.028* (0.011)
Perceived Performance 2011		0.088** (0.011)	0.082** (0.013)		0.018 (0.011)		-0.003 (0.013)	-0.002 (0.013)
Expectations 2011				0.255** (0.012)	0.250** (0.012)	0.252** (0.012)	0.253** (0.012)	0.234** (0.012)
Constant	4.387** (0.039)	4.281** (0.039)	4.263** (0.044)	3.448** (0.053)	3.407** (0.059)	3.339** (0.062)	3.341** (0.063)	2.158* (1.041)
Covariates	NO	NO	NO	NO	NO	NO	NO	YES
N	10329	10329	10329	10329	10329	10329	10329	10329
N (groups)	1160	1160	1160	1160	1160	1160	1160	1160
Between std. dev.	0.367	0.369	0.369	0.352	0.353	0.352	0.351	0.341
Within std. dev.	0.617	0.615	0.615	0.594	0.594	0.594	0.594	0.587
ICC	0.261	0.265	0.264	0.260	0.261	0.260	0.259	0.252

Note: Standard errors clustered at the class level in parentheses. * $p < 0.05$, ** $p < 0.001$, two-tailed tests. Expectations 2013 as dependent variable

Table B.2: The Effect of Prior Expectations, Perceived Performance and Satisfaction (2011) on Future Expectations (2013) - Family Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Satisfaction 2011	0.036* (0.016)		0.029 (0.016)			0.035* (0.016)	0.029 (0.016)	0.027 (0.017)
Perceived Performance 2011		0.034 (0.023)	0.021 (0.023)		0.034 (0.024)		0.021 (0.024)	0.024 (0.024)
Expectations 2011				0.012 (0.027)	0.003 (0.027)	0.007 (0.027)	0.002 (0.027)	0.003 (0.027)
Constant	4.468** (0.066)	4.486** (0.085)	4.417** (0.097)	4.560** (0.121)	4.476** (0.134)	4.438** (0.137)	4.407** (0.143)	4.124** (0.228)
Covariates	NO	NO	NO	NO	NO	NO	NO	YES
N	5104	5104	5104	5104	5104	5104	5104	5104
N (groups)	2437	2437	2437	2437	2437	2437	2437	2437
r² within	0.003	0.002	0.003	0.000	0.002	0.003	0.003	0.022
r² overall	0.007	0.010	0.010	0.082	0.014	0.017	0.012	0.000
r² between	0.008	0.012	0.012	0.115	0.018	0.021	0.015	0.000
Between std. dev.	0.562	0.562	0.561	0.561	0.561	0.560	0.561	0.589
Within std. dev.	0.354	0.354	0.354	0.354	0.354	0.354	0.354	0.353
ICC	0.716	0.716	0.715	0.715	0.715	0.715	0.715	0.736

Note: Standard errors clustered at the family level in parentheses. * $p < 0.05$, ** $p < 0.001$, two-tailed tests. Expectations 2013

as dependent variable. Only covariates varying at the family level included in Model 8.

Table B.3: Multilevel Analysis

	School Random Effects	School and Classroom - Random Intercept (Empty)	School and Classroom - Random Intercept	School and Classroom - Random Coefficient
Satisfaction 2011	0.032* (0.013)		0.031* (0.010)	0.034* (0.011)
Perceived Performance 2011	-0.004 (0.011)		-0.004 (0.011)	-0.005 (0.011)
Expectations 2011	0.242** (0.012)		0.240** (0.010)	0.243** (0.010)
Budget Change 2011-13	0.000 (0.000)		0.000 (0.000)	0.000* (0.000)
School Leader Change 2011-13	-0.007 (0.020)		-0.011 (0.020)	-0.013 (0.020)
Constant	3.445** (0.198)	4.585** (0.016)	3.470** (0.183)	3.463** (0.184)
Covariates	YES	NO	YES	YES
N	9283		9283	9283
N (schools)	43		43	43
N (classrooms)	-		1043	1043
r² within	0.110		-	-
r² overall	0.122		-	-
r² between	0.820		-	-
Random Effects Parameters				
<i>School Level</i>				
<i>Variance of Intercept</i>		0.009 (0.003)	.001 (0.001)	.001 (0.001)
<i>Classroom Level</i>				
<i>Variance of Intercept</i>	-	0.004 (0.002)	0.003 (0.002)	0.237 (0.054)
<i>Variance of Satisfaction 2011</i>	-		-	0.011 (0.003)
<i>Covariance of Satisfaction and Intercept</i>	-		-	-0.051 (0.013)
<i>Variance of Residual</i>	-	0.389 (0.006)	0.346 (0.005)	0.338 (0.005)

Note: Standard errors clustered at the school level in parentheses. * $p < 0.05$, ** $p < 0.001$, two-tailed tests. Expectations 2013 as dependent variable.

Table B.4: Subgroup Analysis: Controlling for Child Ability in Math and Reading National Tests – School Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Reading - 2 nd grade	Reading - 4 th grade	Reading - 6 th grade	Reading - 8 th grade	Math - 3 rd grade	Math - 6 th grade	All tests
Satisfaction 2011	0.071* (0.023)	0.060* (0.031)	0.034 (0.032)	-0.002 (0.034)	0.019 (0.025)	0.034 (0.032)	0.034* (0.015)
Perceived Performance 2011	-0.023 (0.028)	-0.012 (0.033)	-0.023 (0.040)	-0.003 (0.032)	-0.003 (0.028)	-0.012 (0.040)	-0.006 (0.014)
Expectations 2011	0.242** (0.029)	0.216** (0.041)	0.229** (0.033)	0.227** (0.029)	0.270** (0.033)	0.225** (0.032)	0.236** (0.013)
Test result	0.008 (0.014)	0.011 (0.020)	0.002 (0.024)	0.037 (0.035)	-0.011 (0.014)	-0.005 (0.021)	0.004 (0.011)
Constant	3.017** (0.281)	3.486** (0.427)	3.582** (0.500)	3.618** (0.634)	3.584** (0.311)	3.529** (0.497)	3.497** (0.235)
Covariates	YES	YES	YES	YES	YES	YES	YES
N	1421	1186	1161	995	1334	1183	6097
N (groups)	49	49	49	47	49	48	49
r² within	0.190	0.144	0.148	0.121	0.201	0.149	0.113
r² overall	0.214	0.164	0.155	0.136	0.219	0.154	0.127
r² between	0.757	0.642	0.382	0.388	0.634	0.371	0.778

Note: Standard errors clustered at the school level in parentheses. Fixed effects at the school level. * $p < 0.05$, ** $p < 0.001$, two-tailed tests. Expectations 2013 as dependent variable.

Table B.5: The Path Model

Dependent Variable	Independent Variable	Coefficient	Std. Err.
Exp2011			
	Sat2009	0.054**	(0.013)
	Constant	4.309**	(0.053)
Sat2011			
	Exp2011	-0.073**	(0.014)
	Perf2011	0.678**	(0.013)
	Sat2009	0.192**	(0.012)
	Constant	1.050**	(0.077)
Exp2013			
	Exp2011	0.274**	(0.013)
	Sat2011	0.026*	(0.011)

Constant	3.261**	(0.068)
Perf2011		
Sat2009	0.293**	(0.013)
Constant	2.388**	(0.056)
Perf2013		
Sat2011	0.149**	(0.015)
Perf2011	0.259**	(0.016)
Constant	2.094**	(0.051)
Sat2013		
Sat2011	0.194**	(0.012)
Exp2013	-0.069**	(0.015)
Perf2013	0.714**	(0.014)
Constant	0.870**	(0.078)
var(e.Exp2011)		
Constant	0.443**	(0.009)
var(e.Sat2011)		
Constant	0.378**	(0.008)
var(e.Exp2013)		
Constant	0.367**	(0.007)
var(e.Perf2011)		
Constant	0.493**	(0.010)
var(e.Perf2013)		
Constant	0.436**	(0.009)
var(e.Sat2013)		
Constant	0.388**	(0.008)
cov(e.Exp2011,e.Perf2011)		
Constant	0.121**	(0.007)
cov(e.Exp2013,e.Perf2013)		
Constant	0.124**	(0.006)
<i>N</i>		4844
R ² Exp2011		0.004
R ² Sat2011		0.445
R ² Perf2011		0.091
R ² Exp2013		0.087
R ² Perf2013		0.157
R ² Sat2013		0.461
R² Overall		0.135

Note: Unstandardized coefficients. * $p < 0.05$, ** $p < 0.001$, two-tailed tests.
No control variables included. The residuals of Perf2011 and Exp2011 and
Perf2013 and Exp2013 have been allowed to correlate for better fit.