Environment for Development

Discussion Paper Series July 2021 ■ EfD DP 21-07

Satisfaction with Water Services Delivery in South Africa

The Effects of Social Comparison

Nouran Zenelabden and Johane Dikgang

Central America
Research Program in Economics and Environment for Development in Central America Tropical Agricultural Research and Higher Education Center (CATIE)

Chile
Research Nucleus on Environmental and Natural Resource Economics (NENRE)
Universidad de Concepción

China
Environmental Economics Program in China (EEPC)
Peking University

Colombia
The Research Group on Environmental, Natural Resource and Applied Economics Studies (REES-CEDE), Universidad de los Andes, Colombia

Ethiopia
Environment and Climate Research Center (ECRC), Policy Studies Institute, Addis Ababa, Ethiopia

Ghana
The Environment and Natural Resource Research Unit, Institute of Statistical, Social and Economic Research, University of Ghana, Accra

India
Centre for Research on the Economics of Climate, Food, Energy, and Environment, (CECFEE), at Indian Statistical Institute, New Delhi, India

Kenya
School of Economics
University of Nairobi

Nigeria
Resource and Environmental Policy Research Centre, University of Nigeria, Nsukka

South Africa
Environmental Economics Policy Research Unit (EPRU)
University of Cape Town

Sweden
Environmental Economics Unit
University of Gothenburg

Tanzania
Environment for Development Tanzania
University of Dar es Salaam

Uganda
EfD-Mak, School of Economics and Department of Agribusiness and Natural Resource Economics, Makerere University, Kampala

USA (Washington, DC)
Resources for the Future (RFF)

Vietnam
University of Economics
Ho Chi Minh City, Vietnam

Satisfaction with water services delivery in South Africa: the effects of social comparison

Nouran Zenelabden and Johane Dikgang[1]

Abstract

The subjective well-being literature tends to assess the effects of social comparison on public services by downplaying the significance of asymmetric comparison effects. Moreover, the literature tends to use cross-sectional data, and lacks national coverage. This paper investigates the role of social comparisons in determining the satisfaction of South African households with municipal water service delivery. We use a unique balanced-panel dataset from 2015-2017 with national coverage, from Statistics South Africa General Household Surveys. Unobserved heterogeneity is controlled by using the random effects model with a Mundlak term. We find indications of rivalry between households at provincial level, but indications of altruism or risk sharing between closer neighbours. Moreover, we find evidence of both downward and upward comparison, with the latter having the strongest effect. Comparison effects influence households differently. We conclude that, since satisfaction with water service delivery seems to be strongly influenced by psychological and behavioural factors such as social comparison, satisfaction surveys serve a limited purpose as a foundation for public policy, because satisfaction is determined in part by factors that are unrelated to the actual service experienced by households. Our empirical evidence confirms this line of reasoning. The findings are robust for variety of reference groups.

Keywords: comparison, households, municipality, satisfaction, water.

JEL Codes: D60, D62, H41, I31

---

[1] Nouran Zenelabden, E-mail: 2515036@students.wits.ac.za and Johane Dikgang* Corresponding author, E-mail: johane.dikgang@wits.ac.za, School of Economics and Finance, University of the Witwatersrand, Johannesburg, South Africa. We would like to thank Mr. Niel Roux at Statistics South Africa for his cooperation and for granting us access to the dataset required for the purpose of this study. We would also like to thank Ida Frössander and Petra Noreback from the Stockholm School of Economics, Sweden, for their contribution in the conceptualization of this paper and Alain Pholo from Public and Environmental Economics Research Centre (PEERC), University of Johannesburg for his assistance with data. We thank Environment for Development (University of Gothenburg, Sweden) for comments and editorial support.

1.1 Introduction

Past studies in the economics literature have shown that social comparison – satisfaction compared to some benchmark – is an essential determinant of subjective well-being. According to Firebaugh and Tach (2012) there is evidence that people compare their positions to those of their co-workers, neighbours, and others, based on shared demographic characteristics. Fox and Kahneman (1992) found that social comparisons may affect well-being in public domains of life, where information about others may be more readily available; however, in private domains, where information about others is scarce, social comparisons may be concluded from satisfaction judgments.

When social comparisons are found to be an important determinant of satisfaction, it is often taken as a sign that satisfaction is being influenced by psychological and behavioural factors, which for public policymaking are difficult to observe and influence (Deichmann and Lall, 2007). In such instances, it is not clear who public policy should target, or what the welfare effects would be of implementing policies aiming at reducing the influence of comparison effects on satisfaction. Therefore, evidence of significant effects of social comparison on satisfaction might undermine the use of satisfaction surveys as a basis for public policy making (see Frössander and Noreback, 2014).

Most of the existing literature on the relationship between personal characteristics and relative public service satisfaction is derived from developed countries. In developing countries, few studies have evaluated the effect of personal characteristics and the relative magnitude of the effects on people's satisfaction.

Vasquez (2015) and Vasquez and Trudeau (2011), for example, focused on the socioeconomic determinants of individual satisfaction. Other studies have investigated the effect of relative magnitude on satisfaction; for example, Deichmann and Lall (2007) estimated relative service reliability based on nearest geographic neighbour in Bangalore, India, and argued that relative service performance influenced household satisfaction. Furthermore, Vasquez et al. (2011) used the prospect theory to evaluate the influence of user perception on satisfaction with water services in Leon, Nicaragua. However, Deichmann and Lall (2007) and Vasquez et al. (2011) do not provide evidence on a national level; rather, they use region-specific data during a certain time of the year, which may lead to biased estimates. Moreover, they do not investigate the importance of asymmetric (upward and downward) comparisons that might affect household satisfaction differently. Identifying the gaps that remain could assist governments and water utilities in improving public policy, especially in countries with many disparities, such as South Africa.

This paper contributes to and extends the currently limited literature from developing countries. In summary, the contribution of this study is threefold. Firstly, we use data with national coverage, allowing us to capture the heterogeneity in the whole population and disparities in service delivery by different geographical levels (i.e., province, municipality and primary sampling unit (PSU), the latter approximately equivalent to an enumerator area). Secondly, we constructed a unique balanced-panel dataset from 2015-2017, in contrast to the cross-sectional data used in most of the literature. Panel data is ideal since it allows us to control time-fixed effects and individual random effects. Thirdly, in addition to analysing the role of relative magnitude in household satisfaction with water service delivery, we examined the effect of asymmetric comparisons on relative satisfaction in measuring the extent of social comparison on household satisfaction, which depends on households' position relative to the reference group.

We find evidence that social comparisons matter to households' satisfaction with water service delivery. At the provincial level, results display a positive relationship between a household's probability of being satisfied with water service delivery and its water service reliability relative to mean reliability, controlling for households' own service reliability. However, at Municipal and PSU level, the effect is the opposite. This indicates that the effect of social comparisons depends on whether comparisons are made to close neighbours or to more distant others. In addition, when testing for asymmetric comparisons, we find evidence of both *upward* and *downward* comparisons at Municipality and PSU level, whereas at the provincial level we only find evidence of downward comparisons – suggesting that social comparison effects differ between households.

The remainder of the article is organized as follows. Section 2 is the literature review. Section 3 provides the analytical framework. Section 4 presents the data while section 5 presents the results. Section 6 concludes the article.

2. Literature review

2.1 Satisfaction and utility

‘Satisfaction’ means different things to different disciplines. In economics, satisfaction can mean well-being, utility or happiness (Dolan, Peasgood and White, 2008). ‘Utility’, on the other hand, refers to the satisfaction expected from the consumption of goods and services. Consequently, satisfaction of individuals is the basis of the utility function. Utility derived from services can be categorised into two broad perspectives. The first is the service provider's perspective, which may refer to the provider’s ability to satisfy the needs of the user. Second is the user's perspective, which may refer to a psychological feeling of pleasure, satisfaction, well-being and happiness resulting from the use of the service. In this case, utility is a subjective or relative concept.

The increasing interest of economists in evaluating subjective utility or relative satisfaction has its origins in the ‘Easterlin paradox’. In the mid-70s, Richard Easterlin observed that increasing the income of all individuals does not increase the happiness of all, meaning that happiness does not always go upward as income grows (Easterlin, 1974; 1995). Thus far, two non-alternative explanations for the Easterlin paradox have been put forward in the literature. Both are directly related to the nature of the comparisons that an individual makes (Barcena-Martin, Cortes-Aguilar and Moro-Egido, 2016).

One explanation is based on the existence of internal benchmarks, which involves aspirations and dynamic comparisons with the individual’s own experiences at different points in time. The other explanation – which is a focus of this study – is based on external benchmarks, and holds that comparisons are made between peers such as neighbours, co-workers and friends. These social comparisons assume that what matters to people is not the absolute level of their resources, but their level of resources relative to the reference or comparison group (Barcena-Martin, et al., 2016).

2.2 Determinants of well-being

The existing literature has extensively assessed the key factors that affect individual's wellbeing and satisfaction. However, studies on the effect of social comparison on an individual's satisfaction with public services, especially in developing countries, are limited. In this section we present a review of the literature relating to the main co-variates included in our study. We start by assessing the main determinants of an individual's satisfaction with public services. This is followed by a review of the literature that evaluates the effect of social comparison on an individual's satisfaction, and in different domains of life.

Recent studies on public service satisfaction show that an individual's satisfaction is partially determined by service-related factors, and partially by other factors, including socio-economic and demographic characteristics. However, one caveat associated with these studies is that data collectors use different scales and methods for measuring satisfaction, making comparisons between studies cumbersome. In addition, while self-reported satisfaction has been shown to correlate strongly with certain factors, the mixed empirical evidence on causal direction and the lack of experimental data makes it difficult to establish causality (Dolan, Peasgood and White, 2008).

Nevertheless, these studies are important. A picture that emerges from these studies is that socio-economic characteristics and other factors correlate with public-service satisfaction. Even though comparability is limited when it comes to magnitude of estimates, these studies can at least give an indication of what factors are expected to correlate with public service satisfaction, and in what direction. Vasquez and Trudeau's (2011) showed that duration of hours of tap water per day during the rainy season (winter) and dry season (summer) has a positive and significant effect on household satisfaction across all models. In addition, personal characteristics including education, age and income have a negative effect on users' satisfaction. For internal consistency, a satisfaction index was estimated using factor analysis, and the results support logit models.

Vasquez (2015) undertook a similar study in San Lorenzo, Guatemala. The study used the probit model and presented four different models, with a focus on the valuation of user satisfaction for each service-related factor. Results indicate that user satisfaction is based on the frequency of service interruptions and on pressure levels. Customer satisfaction decreases with water supply interruption, while customers with higher water pressure are more satisfied.

2.3 Overview of social comparison studies

In the absence of objective means to evaluate their own situation, individuals compare themselves to others (Festinger, 1954). More specifically, individuals will engage in social comparisons with *similar* and *proximal* others since it is difficult to accurately compare oneself to a very different reference group (Festinger, 1954). In economics, influences from social comparisons were recognized later. The focus here has mainly been on income comparisons and the dominant finding has been a negative correlation between subjective well-being and the income of a reference group to whom the individual compares herself (e.g., Blanchflower and Oswald, 2004; Clark and Oswald, 1996; Ferrer-i-Carbonell, 2005).

There is no general consensus in the literature on who constitutes an individual's reference group, since the actual interaction between the individual and the reference group is difficult to observe and since individuals usually compare themselves to different groups in different contexts (Deichmann and Lall, 2007; Kingdon and Knight, 2007). As such, reference groups could vary between population groups, change throughout an individual's life and depend on the individual's degree of social and geographical isolation (Clark and Senik, 2010; Fafchamps and Shilpi, 2008; Herrera and Roubaud, 2006). In other cases, reference groups seem to stay the same throughout life. For example, Senik (2008) shows that it is common for individuals in post-transition countries to compare their welfare to people they knew before the transition, implying that former classmates or colleagues constitute the relevant reference groups even later in life.

Considering the difficulties in defining reference groups, it is common to take account of both social and geographical factors. For example, besides including racially defined reference groups, Kingdon and Knight (2007) assume reference groups based on cluster and district levels in South Africa and conclude that higher average income of the reference group is associated with higher subjective well-being when the reference group is defined on a local cluster level (close neighbours). Interestingly, this coefficient becomes negative when the reference group is defined on a broader, district, level (including strangers).

Due to lack of knowledge on actual reference groups, most empirical studies simply assume that reference groups are exogenously given. However, there is substantial support in the social psychology literature for the notion that people actively choose their reference standards (Falk and Knell, 2004). In response to this, a limited number of studies in recent years have investigated the endogeneity of reference groups (e.g., Knight, Song and Gunatilaka, 2009). Literature on social comparisons has yielded mixed empirical evidence on the sign of the relationship between subjective well-being and the performance of the reference group.

When reviewing empirical evidence, it therefore becomes important to distinguish between symmetric comparison effects and asymmetric comparison effects. A majority of the literature considers *symmetric comparison effects* where a change in the reference group's mean performance affects all individuals' subjective well-being in the same way. A positive relationship between individuals' own satisfaction and the performance of the reference group, controlling for individuals' absolute performance, is often explained by positive externalities stemming from altruism or risk-sharing within the community, whereas a negative relationship is hypothesized to originate from feelings of rivalry (Kingdon and Knight, 2007).

The *asymmetric comparison effect* approach implies that a change in the reference group's mean performance could influence different individuals in different ways. Hence, the sign of the reference group effect on an individual's own utility might depend on the comparison direction. Blanco-Perez (2012) summarises the asymmetric comparison, stating that upward comparisons could negatively affect individuals' satisfaction due to the 'envy effect'. Alternatively, upward comparison could positively affect individuals' well-being due to the so-called 'information effect', by which individuals see an improvement in the reference group's performance as an indicator of their own future performance improvement. In the case of downward comparisons, there could be a positive effect from the 'prestige' associated with being better off than one's reference group.

However, research has also shown that individuals might also feel 'regret' at being better off, and that downward comparisons could therefore negatively affect their well-being. While the empirical evidence is mixed, it is commonly assumed that an underperforming reference group (upward comparison) affects individuals' well-being more than a reference group that outperforms the individual (downward comparison). However, the effect of upward and downward comparisons on individuals' well-being is still a disputed question.

To this end, we took advantage of recent evidence on the relationship between well-being and public service delivery in developing countries. To the best of our knowledge, there are only two studies (Deichmann and Lall, 2007; Vasquez et al., 2011) that focus on the effects of social comparison on people's satisfaction with water service delivery. Deichmann and Lall (2007) used household survey data for two major cities in India: Bangalore and Jaipur. Their study estimates relative service reliability using the ratio comparison model, the ratio being between the service level for the household and that of its reference group, defined by nearest neighbour and welfare. The study showed that household satisfaction with water service was highly correlated with households' relative water-service position. Households that received equal or worse water services than their peers tended to be more satisfied, controlling for the household's absolute level of service availability.

Despite studies demonstrating that social comparison matters to the satisfaction of individuals, this cannot be generalised to the whole population, as the studies focused on specific regions in the country during a specific time period. Moreover, they did not pay much attention to the extent of the change in reference group impact on household satisfaction, and how this impact differs – which implies asymmetric comparison.

3. Analytical Framework

The reported water provision household satisfaction rating is expected to be based on the utility derived from such a service. A utility theoretical framework is used, in which utility (satisfaction with service delivery, in this case) is a function of actual service performance, as well as socio-demographic characteristics forming expectations (see Deichmann and Lall, 2007). Thus, our framework can be represented by equation 1:

$$U = f(A, E) \quad \ldots\ldots (1)$$

where (U) represents the utility or satisfaction households derive from water provision, (A) represents a vector of characteristics describing actual service quality and reliability performance, and (E) is a vector of individual and community characteristics that determine expectations.

Deichmann and Lall (2007) develop this simple model into a specification of service delivery of water (2), in which satisfaction with water service delivery (Si^*) depends on the household's received performance (Pi) and on a relative measure of the performance received by the household in relation to what the household's reference group receives on average (Pi/P^*). Moreover, satisfaction depends on individual and household characteristics (Ii), as well as on the benefit from consumption of goods and services other than water (Yi):

$$S_i^* = \alpha' P_i + \tau' (P_i/P^*) + \delta' I_i + \gamma' Y_i + \varepsilon_i \quad \ldots\ldots (2)$$

In equation 2, (S_i^*) is household satisfaction with water service delivery, P_i captures household own service performance, while P^* reflects service performance level in the household reference group. (P_i/P^*) represents the ratio of the service level of a household to that of its reference group. A household's satisfaction increases with improvements in its own service performance (α'), as well as receiving better service performance than its peers (τ') when $P_i > P^*$. By contrast, when $P_i < P^*$, the relative performance ratio (P_i/P^*) decreases, and satisfaction is reduced. (I_i) is a set of individuals, households and service-specific characteristics. (Y_i) is the benefits from consumption of other goods. Lastly, (ε_i) is a normally distributed error term with zero mean and variance σ^2.

Our empirical analysis is based on panel data estimation accounts for the unobserved heterogeneity problem that is associated with the use of self-reported data. A common concern with this type of data is the existence of omitted individual characteristics, such as personality traits and attitudes that simultaneously influence the dependent and explanatory variables. Thus, we make some important modifications, as follows:

$$S^*_{i,t} = \alpha' P_{i,t} + +\tau'(P_{i,t} P_t^*) + \delta' I_{i,t} + \gamma' Y_{i,t} + \varphi' T_t + v_i + \eta_{i,t} \quad \ldots\ldots\ldots\ldots\ldots(3)$$

where, $\varepsilon_{it} = v_i + \eta_{it}$. In equation (3), ($S^*_{i,t}$) is satisfaction with water service delivery, ($P_{i,t}$) is the reliability of water services, ($P_{i,t}/P^*_t$) is relative reliability of water services, ($I_{i,}$) is a vector of individual and household characteristics, and ($Y_{i,t}$) is the benefits from consumption of other goods. Our first modification was the inclusion of a fixed time effect (T_t) to account for changes that are constant across individuals but vary across time, such as changes in weather patterns, inflation and elections. Secondly, we included time-invariant unobserved individual effects to correct for the existence of omitted variables, where (v_i) is the time-invariant individual effect, and (η_i) is the disturbance term. Hence ε_{it} (the error term) consists of two parts: one that varies by individual effect (v_i) and another that has zero correlation with the independent variables (η_{it}) as presented above.

It is assumed in equation 4 that any possible correlation between (v_i) and the independent variables ($I_{k,i}$) is zero (v_i), 'an individual random effect'. However, in our model we suspect the presence of correlation between unobserved effects and the independent variables.

Despite previous econometric theories (see Wooldridge, 2010) and previous studies suggesting that fixed-effect is the most appropriate strategy, factoring out this time-invariant unobserved heterogeneity. The drawback of this approach is that results would be based solely on within-group variation, since we included bivariate variables in our model; this would lead to many observations being excluded. Controlling for random effects could thus be more suitable, since this would allow for individual effects in the error term; but assuming these are randomly distributed and not correlated with the control variables, it would not exclude observation due to no within-group variation. Moreover, we formally compared the random effects and fixed effects and conducted the Hausman test, the results of which suggest that we cannot reject the null hypothesis of all panel-level means being zero; therefore, the random-effect model is the more efficient estimation method.

The key assumption in the random-effect model is that the unobserved heterogeneity is not correlated with the control variables. In order for this to hold, we make a second modification, by applying the approach developed by Mundlak (1978). Thus, we run a random-effects model with Mundlak corrections, as commonly used in the literature, to capture the correlation between control-variable means and unobserved individual characteristics. The Mundlak approach assumes there is correlation between the individual random effect (v_i) and a subset of the independent variables ($I_{k,i}$), which may be defined as:

$$V_i = \Sigma_j \lambda_j \overline{z_j}_{,i} + v_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

The individual random effect (V_i) therefore consists of two parts. The first part is correlated with the mean of a subset ($\overline{z_j}_{,i}$) of the independent variables ($I_{k,i}$), where $k>j$. The second part (v_i) has zero correlation with the independent variables. In our model, we choose a subset of independent variables ($z_{j,i}$) which includes the natural logarithm of the household's monthly income, squared years of education of the household head, and the natural logarithm of household size, because unobserved individual characteristics such as attitude correlate with control variables such as income and education.

Our third modification involves the inclusion of a new independent variable: whether the household pays for water, denoted by (D_{it}). Due to contextual factors, rather than estimating the benefit from consuming goods and services other than water, we use this new binary variable specifying whether the household pays for water. Because of the Free Basic Water policy in South Africa (which affords poorer households 6kl of water per month at no cost to them), as well as a culture of non-payment of water services, for most people the ability to consume other goods and services is not affected by water purchases. It is therefore more interesting to see what the effect is on satisfaction of paying for a service that some people get partially free.

The fourth and last modification is the inclusion of a second performance variable. Where Deichmann and Lall (2007) studied the *reliability* of water supply (P_i) when controlling for actual performance, we argue that satisfaction with water service delivery is also influenced by the actual *quality* of the water, which is confirmed by Vásquez et al. (2011). Hence, a second performance variable for quality ($P^{q}_{i,t}$) is included in our specifications. However, we only look at relative performance with regard to the reliability variable, since we assume that other households' water *quality* is less observable to others than their water *reliability*.

In conclusion, we extended upon the Deichmann and Lall model (equation 2) by adding time fixed effects, Mundlak corrections, a binary variable for payment of water, and a second performance variable. Thus, our first model specification (i.e., social comparison) becomes:

$$S^{*}_{i,t} = \alpha' P_{i,t} + \pi' P^{q}_{i,t} + \tau'(P_{i,t}/P^{*}_{t}) + \delta' I_{i,t} + \gamma' \mathrm{D}_{i,t} + \varphi' T_{t} + \Sigma_{j} \lambda j \overline{z_{j,i}} + \mathrm{v}_{i} + \eta_{i,t} \quad \text{..............(5)}$$

where ($S^{*}_{i,t}$) is satisfaction with water service delivery, ($P_{i,t}$) is the reliability of water services, $P^{q}_{i,t}$ is quality of water services, ($P_{i,t}/P^{*}_{t}$) is relative reliability of water services, ($I_{i,t}$) is a vector of individual and household characteristics, ($\mathrm{D}_{i,t}$) is a binary variable indicating whether the household pays for water, (T_{t}) is time fixed effects, and ($\Sigma_{j} \lambda j \overline{z_{j,i}} + \mathrm{v}_{i} + \eta_{i,t}$) is the error term. The parts of the error term not correlated with individual characteristics ($\mathrm{v}_{i} + \eta_{i,}$) are assumed to follow a normal distribution, with mean equal to zero and variance equal to $\sigma 2$.

In order to test for asymmetric (upward and downward) comparison effects, a second model was specified. This specification comes from Ferrer-i-Carbonell's (2005) study, which showed that individuals might feel unhappy or dissatisfied if their income is below average, and vice versa. Therefore, we added two variables, which measure the effect of receiving water service that is reliable than the mean service of the reference group. 'More' is referred to as ($M_{i,t}$) and 'less' is referred to as ($L_{i,t}$), to account for downward and upward comparisons respectively. We defined these variables as follows:

$$If\ P_{i,t} = P^{*}_{t}\ then\ M_{i,t} = 0\ and\ L_{i,t} = 0$$

$$If\ P_{i,t} > P^{*}_{t}\ then\ M_{i,t} = P_{i,t} - P^{*}_{t}\ and\ L_{i,t} = 0$$

$$If\ P_{i,t} < P^{*}_{t}\ then\ M_{i,t} = 0\ and\ L_{i,t} = P^{*}_{t} - P_{i,t} \quad \text{.................. (6)}$$

If the household received reliable water equal to the mean service of the reference group ($P_{i,t} = P^{*}_{t}$), then both downward and upward comparisons are equal to zero, which is defined by ($M_{i,t}$ =0) *and* ($L_{i,t}$=0). If the household received more reliable water relative to the mean service of the reference group ($P_{i,t}$>P_t^*), then more reliable water is defined by ($M_{i,t}$=$P_{i,t}$−P_t^*). If the reliability of water is less than the mean service of the reference group ($P_{i,t}$<P_t^*), then less reliable water is defined by ($L_{i,t}$ =P_t^*−$P_{i,t}$).

Our second model (i.e., asymmetric comparisons), which tests the asymmetric comparison effects, includes these two variables ($M_{i,t}$) and ($L_{i,t}$) instead of the relative reliability variable ($P_{i,t}/P_t^*$) included in equation 6, and becomes the following:

$$S_{i,t}^* = \alpha' P_{i,t} + \pi' P_{i,t}^q + \theta' M_{i,t} + \mu' L_{i,t} + \delta' I_{i,t} + \gamma' D_{i,t} + \varphi' T_t + \Sigma_j \lambda j \overline{z_{j,i}} + v_i + \eta_{i,t} \quad \ldots\ldots\ldots..(7)$$

The vector of individual and household characteristics ($I_{i,t}$) in equations (5) and (7) includes gender, age, years of education, years of education squared, ethnicity, the natural logarithm of household size, the natural logarithm of household income, whether or not the household owns its own home, and whether or not the household lives in a metropolitan area. Moreover, in line with Deichmann and Lall (2007), we also included fixed effects for the main source of drinking water, as well as fixed effects for locality type and province. Together with all the other variables, these variables will be described in detail in the following section.

Finally, like much of the well-being literature, we cannot claim unequivocal causal relationships, for variables such as education, income, and whether the household pays for water (i.e., receives a water utility bill), because we do not explicitly deal with the issue of endogeneity. Unfortunately, an instrumental variables strategy to tackle this problem is not easy to implement because our data comes from a General Household survey where some of the questions are related to subjective measures of life satisfaction and others to objective measures, and because we cannot follow individuals over time. Nonetheless, to mitigate against potential bias due to endogeneity, we implemented the following: i) avoided subjective variables among the regressors as much as possible, and ii) controlled for unobserved heterogeneity across places via a multilevel model as a means to deal with omitted variables.

By adopting a multilevel model over the more conventional single-level models, the estimated coefficients, and more importantly their standard errors, consider clustering at various levels of the population structure (see Ballas and Tranmer, 2011; Browne and Steele, 2009; Hox, 1995). Balls and Tranmer (2011) argue that a multilevel approach accounts for the fact that people in the same household, district, or region are not independent of one another. Another key advantage of this approach is that it is possible to estimate different intercepts for different geographical (e.g., province, district, municipality) and group (e.g., household) levels and also to examine how these intercept estimates may be affected by the introduction of explanatory and control variables.

Rice and Jones (1997) argue that the primary focus in this kind of study is on the estimation of the magnitude of variances at different levels and how these variances relate to explanatory variables, in contrast to more conventional econometric approaches that view higher level effects as nuisance parameters that need to be conditioned or adjusted to achieve desirable properties of fixed part parameters. By modelling higher level effects, endogeneity bias is avoided.

In addition, multilevel models can address potential issues of spatial heterogeneity by adopting a ‘‘random coefficients’’ approach (Snijders and Bosker 1999), provided the relationship between the dependent variable and the explanatory variable(s) can be different in each level and/or geographical region. The causal relationship in happiness equations is not entirely clear (see Oswald, 2007) and there is a possibility that the endogeneity of some of the regressors may have considerable effects on the measured variation at individual and household levels.

4. Data

4.1 Data Description

The empirical analysis in our study is based on data from the General Household Survey (GHS) of Statistics SA, with national coverage and with ‘province’ as the lowest geographical unit of identification. The GHS has been conducted on an annual basis since 2002, and measures the living conditions of South African households. Since 2015, the GHS has used a Master Sample frame, developed in 2013 as a general-purpose sampling frame to be used for all Stats SA household-based surveys. The 2013 Master Sample is based on information collected during the 2011 Census conducted by Stats SA. Our study employs data for the years 2015-2017 for approximately 33 000 dwelling units. GHS data provide a wide range of information about individual and household demographics which allows us to capture all the regional and social disparities among South African households.

Approximately two-thirds of the dwelling units remained in the sample over the whole period of analysis from 2015 to 2017. Of these dwellings, approximately 19 000 received municipal water during these three years, which is a requirement for analysing satisfaction with water service delivery. Since the majority of dwelling units remained in the sample during this period, it is possible to create a balanced panel by using a unique identifier number for each household, similarly to the method of Frössander and Noreback (2014). The main difference between this study and that of Frössander and Noreback is that we made use of the latest data (2015-2017), while they used the 2009-2011 data.

Table 1 describes all the variables used in the regressions, as well as their expected signs. Similar to the method used by Stats SA, the rating of overall water service is used as a proxy for the dependent variable satisfaction (see Stats SA 2016; 2017). However, the score constructed by Stats SA can take on values from 1 to 3, with 1=’poor’, 2=’average’, and 3=’good’. We redesigned this ordinal variable into a binary variable for satisfaction that takes on a value of 1 if the rating is ‘good’, and 0 if the rating is ‘average’ or ‘poor’. We consider users who rate services ‘good’ to be satisfied, and those who rate services ‘average’ or ‘poor’ to be dissatisfied.

Table 1: Description of the variables and expected sign

Variable	Definition	Hypothesized correlation with satisfaction
Dependent variables		
Satisfaction	1=‘Good’, 0=‘Average’ or ‘poor’	
Independent Variables		
Water-related factors		
Perceived water service reliability	Reliability Score, 1-4	(+) effect from receiving better services
Perceived water quality	Quality score, 1-5	(+) effect from receiving better quality of water
Pays for water	1=‘Pays for water’, 0=Otherwise	(-) effect from paying for a service that many don't pay for.
Individual characteristics		
Female	1=‘Female’, 0=‘Male’	(+) effect for female, male (omitted)
HH Age	Age of household head	(+) effect with higher age
HH Education	Dummies no schooling, primary education, secondary education and tertiary education	(-) effect with higher education levels, no schooling (omitted)

Ethnicity	Dummies for Black, Coloured, Asian and White	Black (omitted), Coloured (+), Asian (+), White (+)
Household Characteristics		
HH size	Size of household	(-) effect with the increase in household size
HH income	Household income includes earned income as well as social grants, remittances and private pensions.	(+) effect with income increase
Owns dwelling	1= ‘Owns’, 0=Otherwise	(-) effect with owning dwelling
Lives in metro	1=‘Lives in metro’, 0=Otherwise	(+) effect with living in metro
Fixed *effects*		
Year	Dummies for *2015, 2016 and 2017*	Ambiguous
Geo Type	Dummies for farms urban and traditional	Traditional (omitted), Urban (+) , Farms (-)
Provinces	Dummies for Western Cape, Eastern Cape, Northern Cape, Free State, KwaZulu-Natal, North West, Gauteng, Mpumalanga and Limpopo	Western Cape (omitted), Gauteng (+), other provinces (-)
Drinking water	Dummies for *Piped Water in Dwelling, Piped Water in Yard, Borehole on Site, Rainwater Tank, Neighbour's Tap, Public Tap, Water carrier, Borehole off Site, Flowing Water, Dam/Pool, Well, Spring, Other*	Piped water in dwelling (omitted), other sources (-)

The redesign of the dependent variable is done to get more straightforward interpretations of the regressions. Since an ordinal scale does not reflect mathematically equal steps, it is not obvious that the step from ‘poor’ to ‘average’ is the same as the step from ‘average’ to ‘good’. Addressing this issue by redesigning the ordinal variable into a binary variable is a common method used in similar studies (see Lewis and Pattinasarany, 2009).

Regarding service-related variables, we constructed both a reliability score index and a quality score index to control for perceived water-service performance. The *perceived water service reliability* score can take on values 1-4, and is constructed in line with the household food insecurity access scale for measurement of household food access (see Coates, Swindale and Bilinsky, 2007). Figure 1 below illustrates how the perceived water service reliability score index in terms of interruptions was constructed.

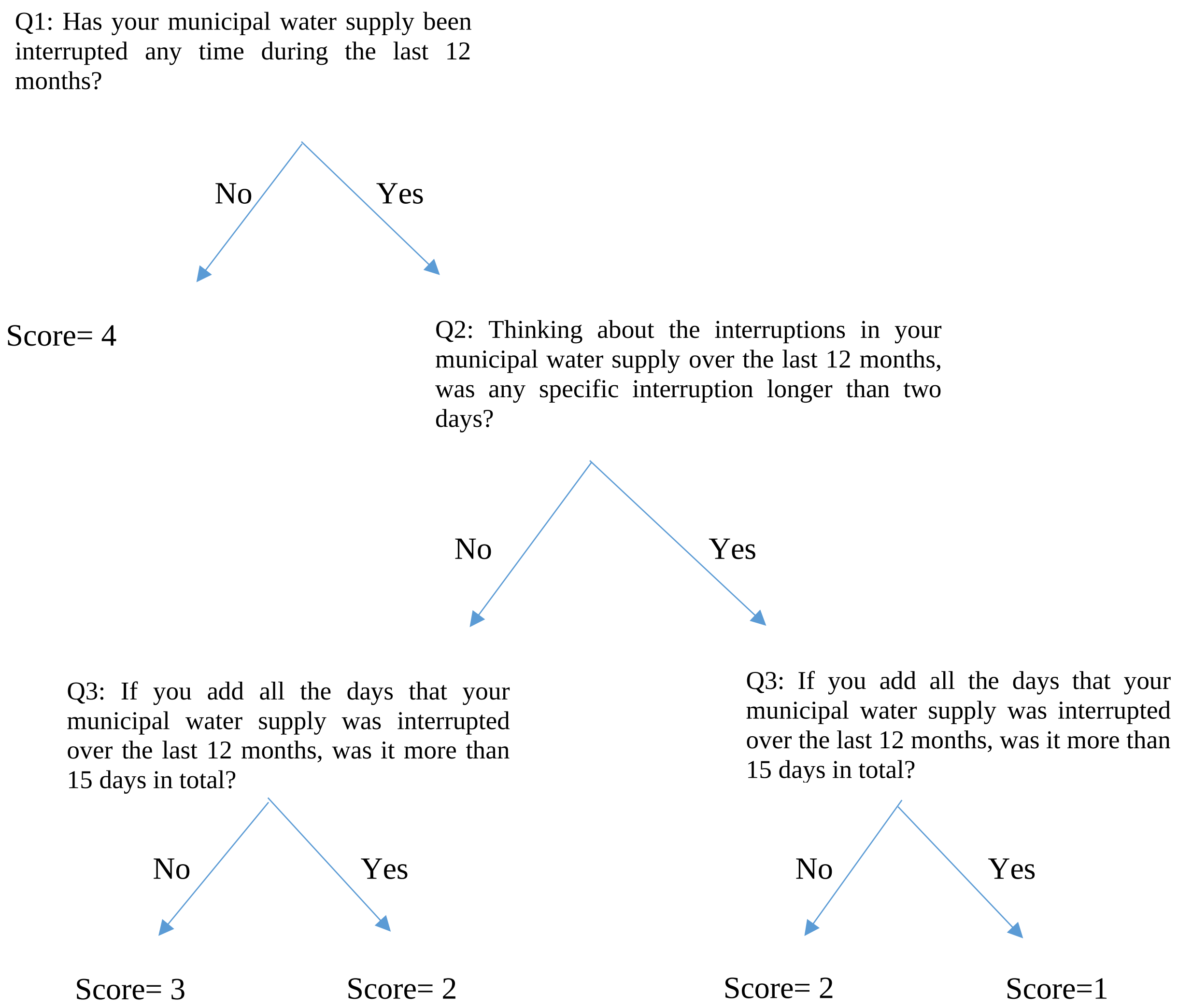

Figure 1: Construction of reliability score[2]

[2] The lowest score (1) was assigned to those reporting interruptions in their water supply during the last twelve months. The second-lowest score (2) is assigned to respondents who had interruptions that either amounted to a total of fifteen days or where at least one interruption lasted for two days. Next, the score of (3) is assigned to respondents who reported interruptions, but none of the interruptions lasted for two days or longer, and total time without water during the last 12 months did not exceed 15 days. The highest score of (4) is assigned to respondents reporting no interruptions at all during the period.

The idea is that binary responses to occurrence and frequency of interruptions may be summarised in a scale to provide a continuous measure for households. Table 2 gives an overview of the distribution of the reliability score and satisfaction in our sample.

Table 2: Distribution of satisfaction and reliability score

Overall Satisfaction	**Satisfied**	**Dissatisfied**		
	57.59%	42.41%		
Perceived water service reliability score	1	2	3	4
	21.48%	10.03%	20.81%	47.68%

Table 2 shows that approximately 58 percent of the households were satisfied with the overall water service. Moreover, nearly half of the households in our sample (around 48 percent) reported the highest score (4), implying that they did not experience any interruption at all. In contrast, around 21 percent of the households reported the lowest score of 1, which suggests that they experienced interruptions frequently in the past 12 months.

In addition to the satisfaction index above, we also construct a perceived water quality index. Perceived water quality is strongly linked with satisfaction levels (taste, odour, colour), hence it is vital to construct this measure.

Q1: Has the water from the main source of drinking had any treatment? Is it…			
			Four 'Yes' = score 5
…safe to drink?	Yes	No	Three 'Yes' and one 'No' = score 4
…clear (has no colour/free from mud)?	Yes	No	Two 'Yes' and two 'No' = score 3
…good tasting?	Yes	No	One 'Yes' and three 'No' = score 2
….free from bad smell?	Yes	No	Four 'No' = score 1

Figure 2: Construction of quality score

A *perceived water quality score* is constructed from four binary questions, on whether the water received was safe to drink; clear (had no colour and was free from mud); good in taste; and free from bad smells. Scores range from 1 to 5. The *relative reliability* is constructed from the reference group ratio, which is defined as a household's perceived water service reliability divided by the mean perceived water service reliability of the household's assumed reference group.

In our analysis, reference groups are defined exogenously, created by sorting respondents into groups according to geographical proximity (province - prv, municipality - Mun and PSU) and ethnicity (*Black*, *Coloured*, *Asian*, *White*). There are a number of reasons for this. First, reference groups are commonly defined both socially and geographically in the literature (see Akerlof and Kranton, 2000; Deichmann and Lall, 2007). In a society such as South Africa, with sharp racial divisions, aspirations may be related to what can be achieved by persons of one's own race. Hence, race identifies the reference group; race-based relativities may be important.

Second, this reference group construction is in line with the paper on South Africa by Kingdon and Knight (2007), who consider ethnic and idiomatic differences to be important factors in the construction of reference groups, due to the legacy of apartheid. Table 3 presents definitions for the reference groups, the number of total reference groups in our sample, and the approximate number of households per reference group in the sample. Note that reference groups for the PSU level are based only on geographical proximity, not ethnicity, given the small (approximate) number of households in each PSU group in our sample.

Table 3: Construction of reference groups

Reference group definition	**Reference groups**	**Households per reference group**
Provincial (i.e., Prov): *Ethnicity* (Black, Coloured, Asian, White) and *geographical proximity* (9 provinces)	36	357
Municipal (i.e., Mun): *Ethnicity* (Black, Coloured, Asian, White) and *geographical proximity* (221 municipalities)	884	15
PSU: *Geographical proximity* (2944 PSUs)	2944	4

We expect the effect of relative water-service delivery to be in line with results by Kingdon and Knight (2007). We thus hypothesised that social comparisons will have different effects on satisfaction, depending on the level at which the reference group is defined. At the lowest geographical level (PSU), we expect to see a negative effect of relative water-service reliability; whereas, at the higher geographical levels (municipal and provincial), where reference groups have more distance between households, we expect to see the opposite, controlling for households' own water service reliability. Regarding the asymmetric comparison effect, the hypothesis effects are in line with the downward and upward comparisons described by Blanco-Perez (2012).

At the provincial and municipal levels, we thus hypothesise positive downward comparisons and negative upward comparisons. However, at the PSU level we expect upward comparisons to be positively correlated with satisfaction, and downward comparisons to have no significant effect. Besides theoretical motivation, this last hypothesis is also anchored in contextual factors. Since clean water is a constitutional right in South Africa, we expect upward comparison to be the most relevant effect. In summary, our study tests the following hypotheses:

4.2 Descriptive Statistics

Table 4 describes all the variables used in the regressions and presents summary statistics. Like the method used by Stats SA, the rating of overall water service is used as a proxy for the dependent variable satisfaction (see Stats SA 2016; 2017). On average, 58 percent of the households are satisfied with the water service delivery. In general, households are happy with the service reliability and the quality of water they receive, which is reflected by an average score of approximately 5, signalling excellent quality. The average score for service reliability is approximately 3 out of 5, which reflects that the majority of households see the reliability of water as 'good'.

Table *4*: Descriptive statistics for selected variables

Variable	**Obs.**	**Mean**	**Std. Dev**
Dependent variables			
Satisfaction	38 546	0.58	0.49
Independent Variables			
Water-related factors			
Perceived water-service reliability	38 546	2.95	1.2
Perceived water quality	38 048	4.78	0.82
Pays for water	38 546	0.42	0.45
Individual characteristics			
Female	38 546	0.45	0.50
Age household head	38 546	49.24	15.54
Primary education	37 935	0.20	0.39
Secondary education	37 935	0.57	0.49
Tertiary education	37 935	0.13	0.34
Ethnicity – Coloured, Asian, White	38 546	0.10, 0.02, 0.06	0.30, 0.145, 0.244
Household Characteristics			
Household size	38 546	3.67	2.33
Household income	35 428	6948.37	9097.28
Owns dwelling	38 449	0.65	0.48
Lives in Metro	38 546	0.41	0.49
Reference groups			
Relative reliability – Prov, Mun, PSU	38 546	1, 1, 1	0.40, 0.37, 0.32
More reliable water service than – Prov, Mun, PSU	38 546	0.42, 0.33, 0.24	0.54, 0.48, 0.43
Less reliable water service than – Prov, Mun, PSU	38 546	0.42, 0.33, 0.25	0.63, 0.54, 0.45
Fixed *effects*			
Urban	38 546	0.73	0.45
Farms	38 546	0.01	0.12
Free State	38 546	0.07	0.26
Eastern Cape	38 546	0.12	0.32
Gauteng	38 546	0.24	0.43
KwaZulu-Natal	38 546	0.16	0.36
Limpopo	38 546	0.1	0.3
Mpumalanga	38 546	0.09	0.28
North West	38 546	0.06	0.24
Northern Cape	38 546	0.05	0.22

It seems that some households have experienced service interruptions, hence their low scores. Perhaps this is not surprising given South Africa's ageing infrastructure, water leaks, and lack of funding for operation and maintenance, which in turn have a negative impact on service quality. Some households in our sample experienced interruptions of up to half an hour, which affected overall service delivery. Another interesting finding is that more than half (58 percent) of the households did not pay for water. Most water utilities give poorer households the first 6kl of water every month. Some utilities opt for universal coverage, extending this free basic government water subsidy to all their customers, irrespective of income level.

The majority of household heads in our sample are black (82 percent), with an average age of approximately 49 years. 57 percent of the household heads have had secondary education. When we look at household characteristics, we find that the average household size in our sample is 4 persons per household. Although average income is R6 900 per month, slightly over half (52 percent) of our sample earn R4 000 per month or less. The fact that over half earn a low income may explain why most households do not pay for water, and signals that most are beneficiaries of the government's free basic water programme.

5. Results

We used average marginal effects rather than marginal effects computed at a fixed point such as the mean. By averaging the marginal effects, we present a more realistic picture (see Bartus, 2005). Since some of the individual and household characteristics are likely to be correlated with each other – for example, household income, education, home ownership, and water utility bill – we tested for multicollinearity (i.e., correlation between predictors, which can adversely affect estimation results). We computed the variance inflation factors (VIF) – a measure of the degree of multicollinearity in our models. A common rule of thumb is that a mean VIF value greater than 10 indicates multicollinearity. All of our specifications generate mean VIF values below 10, averaging 4.93, which suggest absence of a serious multicollinearity problem among the individual and household characteristics.

Furthermore, the estimates of Rho indicate the contribution of the panel component to total variance – approximately 31 percent of total variance – and the reported likelihood ratio test with the null hypothesis that Rho is equal to zero is rejected, which indicates that the probit models are statistically significant. To determine justification of the

Mundlak specification, we test for joint significance of the chosen sub-sample of independent variables believed to correlate with the error term (mean log of household income, mean years of education, and mean log of household size), but the null hypothesis that the estimates are jointly equal to zero cannot be rejected. This suggests that unobserved heterogeneity is uncorrelated with the chosen sub-sample of independent variables; hence, there is no support for the Mundlak specification being superior to the traditional random effects probit specification. Therefore, our analysis relies on the random effects probit specification.

Average marginal effects of relative water service reliability on satisfaction with water services delivery at a provincial, municipal, and PSU level are reported in columns 1,2 and 3 of Table 5. The results show significant estimates for relative reliability at all geographic levels. For the provincial level, results are in line with hypothesis H1, as relative reliability is positive and significant at the 1 per cent level of significance. This implies that higher perceived water-service reliability relative to the reference group is associated with increased probability of being satisfied with water provision. More specifically, a one-unit increase in the reference ratio correlates with an average increase of 6.1 per cent in the probability of being satisfied with water-service delivery, controlling for households' own service reliability.

Table 5: Probit model with random effects

	Provincial level	**Municipal level**	**PSU level**
VARIABLES	**dy/dx**	**dy/dx**	**dy/dx**
Relative reliability	0.061**	-0.032***	-0.120***
	(0.026)	(0.011)	(0.009)
Perceived water reliability	0.095***	0.129***	0.150***
	(0.009)	(0.004)	(0.003)
Perceived water quality	0.093***	0.093***	0.093***
	(0.003)	(0.003)	(0.003)
Female	-0.006	-0.005	-0.004
	(0.005)	(0.005)	(0.005)
Age	0.0002	0.0002	0.0002
	(0.000)	(0.000)	(0.0009)
Age squared	-8.27e-07	-7.93e-07	-1.19e-06

	(8.68e-06)	(8.67e-06)	(8.60e-06)
Primary education	0.016*	0.016*	0.014
	(0.009)	(0.009)	(0.009)
Secondary education	0.014	0.014	0.010
	(0.009)	(0.009)	(0.009)
Tertiary education	-0.003	-0.002	-0.004
	(0.012)	(0.012)	(0.012)
ln Household Size	-0.009**	-0.008**	-0.008**
	(0.004)	(0.004)	(0.004)
ln Household Income	0.005**	0.005**	0.004*
	(0.002)	(0.002)	(0.002)
Owns dwelling	-0.007	-0.007	-0.005
	(0.006)	(0.006)	(0.006)
Live in Metro	0.038***	0.033***	0.026***
	(0.007)	(0.007)	(0.007)
Pays for water	0.028***	0.028***	0.025***
	(0.006)	(0.006)	(0.006)
Coloured	0.005	-0.011	-0.015
	(0.012)	(0.011)	(0.011)
Asian	0.055***	0.036*	0.032*
	(0.020)	(0.019)	(0.019)
White	0.045***	0.027**	0.023*
	(0.014)	(0.013)	(0.013)
Urban	0.044***	0.038***	0.019**
	(0.007)	(0.008)	(0.008)
Farms	-0.052**	-0.055**	-0.075***
	(0.021)	(0.021)	(0.021)
Fixed effects	Yes	Yes	Yes
Rho	.3126	.3131	.3105
Likelihood-ratio test (rho=0)	733.22***	735.78***	719.45***
Observations	34,092	34,092	34,092

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors (shown in parentheses) are adjusted for clustering at the PSU level. Dependent variable is *Satisfaction with Water Service Delivery*. *No schooling*, *Tribal* and *Black* for *Education*, *Geotype* and *Ethnicity* respectively are the reference categories, and therefore omitted from the estimation. Fixed effects include *Year* and *Province*.

The results from the municipal and PSU level display a negative and highly significant estimate of relative reliability, indicating that a one-unit increase of the reference ratio is on average associated with a 3.2 per cent and 12 per cent decrease in the probability of being satisfied with water-service delivery at municipal and PSU level respectively, controlling for households' own service reliability.

As expected, estimates of perceived water-service reliability and perceived water quality are positive and highly significant in the three specifications. While the magnitude of *perceived water quality* remains constant, there is an inverse relationship between geographical level and the magnitude of *perceived water-service reliability.* The implication is that a one-unit increase in the quality score is associated with an average increase of 9.3 per cent in the probability of being satisfied with water service delivery, at all geographical levels; a one-unit increase in the reliability score is associated with an average increase of 9.54 per cent, 12.9 per cent and 15 per cent in the probability of being satisfied for the provincial, municipal and PSU reference groups respectively. The estimates for paying for water are positive and significantly correlated with the probability of being satisfied with water service delivery. The magnitude of *pays for water* remains almost constant on the three specifications.

For individual and household characteristics, the three model specifications reveal estimates similar in magnitude, sign and significance level. We find no significant correlation between satisfaction and *female*, *age*, *age squared* or *owns dwelling*. Furthermore, estimates for years of education are insignificant, except for household head with *primary education,* which is positive and significant at the rather weak 10 per cent significance level at provincial and municipal levels. The implication is that households with primary education are more likely to be satisfied than household heads with no schooling.

Regarding the non-linear variables, estimates of *log of household size* and *log of household income* are significant at the 5 per cent significance level in all three specifications, with positive and negative signs respectively. The implication is that household water satisfaction is likely to increase with an increase in income at all levels, while it is likely to decrease with an increase in the number of household members. These outcomes are in line with our expectation, since an increase in the number of household members leads to an increase in water consumption, which in turn decreases satisfaction, especially if the household experiences water interruption. Also, this might lead to exceeding the 6kl threshold for free basic water.

In addition, the dummies for ethnicity are highly significant for all levels, with the exception of being *Coloured*. Results suggest that being *Asian* or *White* is on average positively and significantly correlated with the probability of being satisfied with water-service delivery. This effect is the greatest in magnitude in the provincial specification. Moreover, households in *metros* are more likely to be satisfied compared to their counterparts in *non-metro* areas, at all levels. Finally, it appears that living in *urban* areas is associated with an increased probability of being satisfied with water services, compared to living in a *traditional rural* area, while households on *farms* are less likely to be satisfied compared to households living in a *traditional* area, for all three specifications.

Table 6: Asymmetric comparison effects with random effects

VARIABLES	**Provincial level** dy/dx	**Municipal level** dy/dx	**PSU level** dy/dx
More water reliability	0.015	-0.012*	-0.059***
	(0.017)	(0.007)	(0.006)
Less water reliability	0.058***	0.026***	0.062***
	(0.017)	(0.007)	(0.006)
Perceived water reliability	0.144***	0.135***	0.156***
	(0.017)	(0.005)	(0.003)
Perceived water quality	0.093***	0.093***	0.092***
	(0.003)	(0.003)	(0.003)
Female	-0.005	-0.005	-0.004
	(0.005)	(0.005)	(0.005)
Age	0.0001	0.0001	0.0001
	(0.0009)	(0.0009)	(0.0009)
Age squared	-1.88e-07	-9.10e-07	-1.27e-06
	(8.64e-06)	(8.67e-06)	(8.60e-06)
Primary education	0.017*	0.016*	0.015
	(0.009)	(0.009)	(0.009)
Secondary education	0.014	0.014	0.011
	(0.009)	(0.009)	(0.009)
Tertiary education	-0.002	-0.000	-0.004
	(0.012)	(0.012)	(0.012)

ln Household Size	-0.009**	-0.009**	-0.008**
	(0.004)	(0.004)	(0.004)
ln Household Income	0.004**	0.004**	0.004*
	(0.002)	(0.002)	(0.002)
Owns dwelling	-0.006	-0.006	-0.004
	(0.005)	(0.005)	(0.005)
Lives in metro	0.036***	0.030***	0.024***
	(0.007)	(0.007)	(0.007)
Pays for water	0.029***	0.027***	0.022***
	(0.006)	(0.006)	(0.006)
Coloured	-0.013	-0.011	-0.016
	(0.013)	(0.012)	(0.010)
Asian	0.033	0.035*	0.031
	(0.022)	(0.019)	(0.018)
White	0.027	0.026*	0.022*
	(0.016)	(0.013)	(0.013)
Urban	0.045***	0.035***	0.016*
	(0.008)	(0.008)	(0.008)
Farms	-0.053**	-0.058***	-0.078***
	(0.021)	(0.021)	(0.021)
Fixed effects	Yes	Yes	Yes
Rho	.3085	.3130	.3105
Likelihood-ratio test (rho=0)	711.61***	734.18***	718.40***
Observations	34,092	34,092	34,092

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors (shown in parentheses) are adjusted for clustering at the PSU level. Dependent variable is *Satisfaction with Water Service Delivery*. *No schooling, Tribal* and *Black* for *Education, Geotype* and *Ethnicity* respectively are the reference categories, and therefore omitted from the estimation. Fixed effects include *Year* and *Province.*

Results for asymmetry in social comparisons at the provincial, municipal and PSU level are presented in Table 6. At the provincial level, we find a positive and significant effect of downward comparisons and an insignificant effect of upward comparisons, controlling for households' own water service reliability. Since there's no evidence of an upward comparison effect, we cannot establish that households with less service reliability are more likely to be satisfied.

Furthermore, at the municipal and PSU specifications, estimates for the downward comparison effect are negative and significant, controlling for household's own water-service reliability. The implication is that, on average, receiving more reliable water service than the reference group is associated with a decrease in the probability of being satisfied with water-service delivery. However, the effect of upward comparisons is positive and highly significant, indicating that, on average, receiving less reliable water services than one's reference group correlates with an increase in the probability of being satisfied.

We see from the above that the magnitude of downward comparison and upward comparison effects at municipal level is relatively small compared to PSU level, whereas the magnitude for both asymmetric variables is approximately similar in magnitude at the PSU level (1.2 per cent and 2.6 per cent for municipal, compared to 5.9 per cent and 6.2 per cent for PSU). Contrary to our expectation, the asymmetric comparison effects at the municipal level are noted in comparison with the results at the PSU level, rather than the provincial level.

These results indicate that, when the reference group is defined at the lowest geographical level (PSU), we observe the opposite effect of relative water service reliability on satisfaction, such that an increase in the frequency of interruptions relative to the PSU reference group correlates with a higher average probability of satisfaction, controlling for households' own reliability of water service. Hence, at this level we find signs of altruism or risk-sharing within the reference groups. As outlined in the literature review, a possible explanation for this result is that households receiving more reliable water service than their PSU reference group experience feelings of regret. On the other hand, it could also be the case that households receiving less reliable water service than the reference group benefit from an information signal of how their future water service delivery will be.

Conducting the asymmetric comparison effects regressions, reported in Table 6, we find signs of both regret (downward comparisons) and information signalling (upward comparisons), with the latter being the strongest effect. Given that both effects exist and differ in magnitude, we conclude that comparison effects do seem to affect households differently at municipal and PSU levels. However, what is commonly interpreted as an altruistic effect could in fact also be a sign of households with below-mean services benefitting from using water from neighbours with more reliable water. Alternatively, it could be that households with above-mean services suffer from neighbours who expect to share it, which would reflect both positive and negative externalities from risk-sharing. This could also result in households receiving above-mean services having to pay not only for their own water consumption, but also for that of their less well-off neighbours.

In the discussion on social comparisons, it is also important to remember that the performance variable Perceived Water Service Reliability is affecting results through both a direct effect and an indirect effect via relative reliability. An improvement in actual reliability would hence be correlated with satisfaction in two ways, and differently at province, municipality and PSU level. At the provincial level, both Perceived Water Service Reliability and Relative Reliability would increase its positive effect on satisfaction by this improvement. At the municipal and PSU level, however, an improvement of actual reliability would affect Perceived Water Service Reliability and Relative Reliability positively, but the implied correlation with satisfaction would be more ambiguous, due to the negative estimate of Relative Reliability.

6. Conclusion

This paper aims to analyse the factors affecting household satisfaction with water service delivery in South Africa, to what extent social comparison can affect satisfaction, and how this effect varies between different definitions of reference groups. By using a longitudinal dataset and running a multilevel model, we are able to explore not so much the cause-effect relationships, but the within-individual changes, as well as inter-individual differences over time.

We find significant effects of water-related factors and household and individual characteristics on households' satisfaction with water service delivery. In addition, social comparisons significantly influence households' satisfaction with municipal water service delivery for the three reference group definitions.

Interestingly, the sign of the effect varies depending on the reference group definition. When the reference group is defined by the largest geographical unit (province) as well as by ethnicity, we find a positive relationship between households' probability of being satisfied with water service delivery and its relative water service reliability, controlling for households' own water service reliability. When the reference group instead is defined by the smallest geographical units (municipality and PSU groups), results suggest a negative relationship between households' satisfaction and their water service reliability relative to the reference group.

The main implication is therefore that households' evaluation of their relative water service reliability depends on whether comparisons are made to close neighbours or to more distant others. Further, when investigating the asymmetry of comparison effects, we find evidence of both *upward* and *downward* comparisons at municipal and PSU levels and *downward* comparisons at the provincial level, suggesting that comparison effects impact households differently. We conclude that satisfaction surveys serve a limited purpose as a basis for public service assessment because psychological and behavioural factors such as comparison effects are found to be significant for the probability of being satisfied. While these factors, unrelated to experience by the actual service user, are difficult for policymakers to influence, citizen feedback through satisfaction studies could still have an intrinsic value in a society with an escalating trend of public protests and governmental distrust.

Studies that model social comparison in public services are still scarce in developing countries, and further research is needed in several areas. First, it is important to investigate whether investigations on satisfaction and social comparison effects should be extended to other public services. Second, it would be interesting if further research investigated the extent to which water-service satisfaction is influenced by the public service provider rather than the actual public service. Lastly, satisfaction surveys in this area could include questions about who households compare themselves to, as well as how households perceive the reliability of service delivery to other households. This suggests that reference groups will be endogenously defined in the collected data. This would allow more variation in the design of reference groups, as well as allowing more accurate measurement of the perceived relative service received.

References

Akerlof, G. and Kranton, R. (2000). Economics and identity. *Quarterly Journal of Economics*, *115*, 715-753.

Ballas, D. and Tranmer, M. (2011). Happy People or Happy Places? A Multilevel Modeling Approach to the Analysis of Happiness and Well-Being. *International Regional Science Review, 35*(1), 70–102.

Bartus, T. (2005). Estimation of marginal effects using Margeff. *The Stata Journal: Promoting Communications on Statistics and Stata, 5* (3), 309-329.

Bárcena-Martín, E., Cortés-Aguilar, A. and Moro-Egido, A. I. (2016). Social comparisons on subjective well-being: The role of social and cultural capital. *Journal of Happiness Studies, 18* (4), 1121-1145.

Blanchflower, D. and Oswald, A. (2004a). Well-being over time in Britain and the USA. *Journal of Public Economics, 87* (7-8), 1359-1386.

Blanchflower, D. and Oswald, A. (2004b). Money, sex and happiness: An empirical study. *Scandinavian Journal of Economics, 106* (3), 393-415.

Blanco-Perez, C. (2012). Rethinking the relative income hypothesis. Dissertation. *SOEPpaper No. 501.*

Browne, W. J., and Steele, F. A. (2009). Editorial: Recent Advances in Multilevel Modelling Methodology and Applications. *Journal of the Royal Statistical Society* Series A 172:535–36.

Clark, A. and Senik, C. (2010). Who compares to whom: The anatomy of income comparisons in Europe. *The Economic Journal*, *120* (544), 573-594.

Clark, A. E. and Oswald, A. J. (1996). Satisfaction and comparison income. *Journal of Public Economics, 61* (3), 359-381.

Coates, J., Swindale A. and Bilinsky P. (2007). *Household Food Insecurity Access Scale (HFIAS) for Measurement of Food Access: Indicator Guide.* Washington DC: The Food and Nutrition Technical Assistance Project, Academy for Educational Development.

Deichmann, U. and Lall, S. V. (2007). Citizen feedback and delivery of urban services. *World Development, 35* (4), 649-662.

Dolan, P., Peasgood, T. and White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology, 29* (1), 94-122.

Easterlin, R. (1974). Does economic growth improve the human lot? Some empirical evidence. In R. David and M. Reder (Eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz*. New York: Academic Press.

Easterlin, R. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization*, *27* (1), 35-47.

Fafchamps, M. and Shilpi, F. (2008). Subjective welfare, isolation, and relative consumption. *Journal of Development Economics, 86* (1), 43-60.

Falk, A. and Knell, M. (2004). Choosing the Joneses: Endogenous goals and reference standards. *The Scandinavian Journal of Economics*, *106* (3), 417-435.

Ferrer-i-Carbonell, A. (2005). Income and well-being: An empirical analysis of the comparison income effect. *Journal of Public Economics*, *85* (5-6), 997-1019.

Festinger, L. (1954). A theory of social comparison process. *Human Relations*, *7*, 117-140.

Firebaugh, G. and Tach, L. (2012). Income, Age, and Happiness in America. In: Marsden, P. (Ed). *Social Trends in American Life*. Princeton: Princeton University Press, 267-287.

Fox, C.R. and Kahneman, D. (1992). Correlations, causes, and heuristics in surveys of life satisfaction. *Social Indicators Research*, *27*, 221-234.

Frössander, I. and Noreback, P. (2014). Satisfaction with water service delivery in South Africa (unpublished Master's dissertation). Stockholm School of Economics, Sweden.

Gravelle, H. and Sutton, M. (2009). Income, relative income, and self-reported health in Britain 1979-2000. *Health Economics, 18* (2), 125-145.

Herrera, J., Razafindrakoto, M. and Roubaud, F. (2006). The determinants of subjective poverty: A comparative analysis in Madagascar and Peru. Développement, Institutions et Mondialisation (DIAL). Working Paper No. DT/2006/01.

Hox, J. J. (1995). Applied Multilevel Analysis. Amsterdam: TT.

Kingdon, G. G. and Knight, J. (2007). Community, comparisons and subjective well-being in a divided society. *Journal of Economic Behavior and Organization*, *64* (1), 69-90.

Knight, J., Song, L., and Gunatilaka, R. (2009). Subjective well-being and its determinants in rural China. *China Economic Review, 20* (4), 635–649.

Lewis, B. D. and Pattinasarany, D. (2009). Determining citizen satisfaction with local public education in Indonesia: The significance of actual service quality and governance conditions. *Growth and Change*, *40* (1), 85-115.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, *46* (1), 69-85.

Oswald, A. (2007). Commentary: Human Well-Being and Causality in Social Epidemiology. *International Journal of Epidemiology 36*, 1253–54.

Rice, N., and Jones, A. (1997). Multilevel models and health economics. *Health Economics, 6,* 561–575.

Senik, C. (2008). Ambition and jealousy: Income interactions in the 'old' Europe versus the 'new' Europe and the United States. *Economica*, *75* (299), 495-513.

Snijders, T. A. B., and Bosker, R. J. (1999). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. London: Sage.

Stats SA - Statistics South Africa (2015). *General Household Survey 2017 Statistical Release – P0318*. Pretoria: Statistics South Africa.

Stats SA (2016). *General Household Survey 2016 Statistical Release – P0318*. Pretoria: Statistics South Africa.

Stats SA (2017). *General Household Survey 2017 Statistical Release – P0318*. Pretoria: Statistics South Africa.

Vásquez, W. F. (2015). A validity assessment of consumer satisfaction to measure performance of water services in Guatemala. *Journal of Water, Sanitation and Hygiene for Development, 5* (2), 301-309.

Vásquez, W. F. and Trudeau, J. (2011). External and internal consistency of user evaluations. *International Journal of Public Administration, 34*(14), 918-925.

Vásquez, W. F., Trudeau, J. and Franceschi, D. (2011). Can user perception influence the quality of water services? Evidence from León, Nicaragua. *International Review of Administrative Sciences, 77* (3), 481-503.

Wooldridge, J. M. (2010). Econometric Analysis of Cross Section and Panel Data, 2nd edition. The MIT Press.