Demand for sanitation in Salvador, Brazil: A hybrid choice approach

Andreia C. Santos [a,*], Jennifer A. Roberts [a], Mauricio L. Barreto [b], Sandy Cairncross [a]

[a] London School of Hygiene and Tropical Medicine, London, United Kingdom
[b] Instituto de Saúde Coletiva, Universidade Federal da Bahia, Brazil

ARTICLE INFO

ABSTRACT

Funds to promote access to water and sanitation in developing countries are scarce and most of the investments come from the national governments and households sources, not international resources. In many of these countries, mainly in middle income countries, households are paying direct taxes to access these services, and understanding what determines their choice and motivation is fundamental to promote access to them. It has been argued that is not enough to supply a wide range of alternatives unless the individuals can recognise their benefits and sustainability.

The objective of this paper is to understand the choice of sanitation technology by residents in the city of Salvador, Brazil. We propose a unique hybrid choice model that incorporates a set of latent attitudinal variables and explains how the demographic factors within a household influence choice. The substantial difference of our hybrid choice model from descriptive frameworks is that it integrates choice and latent variables (such as attitudes and preferences) allowing us to model explicitly the cognitive process that influences sanitation adoption, draw conclusions from cognitive variables associated with individuals' socio-economic and demographic characteristics, and establishes a causal pathway among these variables.

The results show that the attributes of health protection, accessibility, privacy, and house modernisation were what households cared about when opting for flush toilet and sewerage connection, rather than the high cost and consequent household socio-economic status associated with them. The hybrid model is statistically consistent with these findings, and seems to fill the gap between behavioural theory and discrete choice models applied to sanitation.

© 2011 Elsevier Ltd. All rights reserved.

Introduction

Funds to promote access to water and sanitation in developing countries are scarce and most of the investments come from the national governments and households sources, not international resources (Shordt, van Wijk, & Brikke, 2004). In many of these countries, mainly in middle income countries, households are paying direct taxes to access these services, and understanding what determines their choice and motivation is fundamental to promote access to them. It has been argued that is not enough to supply a wide range of alternatives unless the individuals can recognise their benefits and sustainability (Persson, 2002).

Jenkins (1999) demonstrated that the few latrines installed in rural Benin were paid for by the households, and that the determinants of household choice for sanitation were associated with factors other than the health messages disseminated by governmental and non-governmental agencies. Her findings indicated the necessity for deeper investigation of the demand-side aspects, understanding the cultural and historical context, needs, perception and preferences of households, as a means to improving coverage and sustainability.

Few studies have explored the demand for sanitation. Most have applied the conventional contingent valuation approach that continually fails in giving an accurate prediction of consumer behaviour. The method assumes that the maximum amounts people will pay to avoid a loss (willingness to pay) and the minimum compensation necessary for them to accept it (willingness to accept) is equivalent in probabilistic contexts (Diamond & Hausman, 1994; Knetsch & Sinden, 1994). However, experiments have demonstrated that individuals weigh gains and losses differently, and they have greater sensitivity towards losses than towards gains, contradicting the main assumption of this approach and making it poorly suited to providing answers to policy questions facing planners (Kahneman, 2003; Knetsch & Sinden, 1994). On the other hand, the traditional economic demand model is not able to

predict behaviour when the good or service is not directly exchanged in markets and therefore one cannot speak of equilibrium prices and quantities *per se*, as in the case of sanitation. Instead, the emphasis is usually on understanding preferences, where perception and attitude towards sanitation can be included, and historical and cultural aspects can be discussed, to provide additional information on individuals' behaviour and more robust measures to forecast demand in special markets (Phaneuf, Carbone, & Herriges, 2007).

The framework developed by Jenkins and Scott (2007) for the stages of the sanitation adoption decision has contributed to understanding of its determinants and the development of marketing messages that increase demand. Their model provides a way to identify the aspects that influence what the authors called the three stages of adoption: preference to sanitation, intention to install and choice. These stages are derived from behaviour change models, cognitive theories, and psychometric constructions. The study, however, did not ascertain the causal pathway among the variables (preference, intention and choice), and thus, the behavioural link between the attributes of the alternatives and the decisions of individuals is not established. It is important to understand why individuals' decisions vary as conditions change, not only to replicate the results of conditions existing at the time of the survey, and this is only possible if causal relationships among decision variables are understood (Domenich & McFadden, 1975).

We propose a hybrid choice model where the substantial difference from the descriptive framework of Jenkins and Scott is that it integrates choice and latent variables (such as attitudes and preferences) allowing us to explicitly model the cognitive process that influences sanitation adoption, to draw conclusions from cognitive variables associated with individuals' socio-economic and demographic characteristics, and to establish a causal pathway among these variables. Therefore, while we have presented a model capable to make inference regarding behaviour (for sanitation adoption), i.e., a model that is responsive to changing scenarios, Jenkins and Scott's model is essentially descriptive, reflecting the results of conditions existing at the time of the survey. Whereas this difference seems trivial, it is in fact extremely important to design a model that is causal in terms of allowing application to a variety of planning problems without major data collection and calibration costs, that can be transferable from one urban setting to another, allowing reuse without expensive re-estimation in each setting, and that can be efficient, in terms of providing maximum forecasting accuracy per dollar spent on data collection (Domenich & McFadden, 1975).

The objective of this paper is to use that model to understand the choice of sanitation technology by residents in the city of Salvador, Brazil, incorporating a set of latent attitudinal variables that explain how the demographic factors within a household influence choice.

The hybrid choice model (HCM) and a proposed framework of analysis for sanitation

New developments in the area of transportation and marketing have offered a rich dialogue between the understanding of the cognitive process of perception and attitude, and behaviour forecasting (Walker, 2001). *Hybrid Choice Models* permit the incorporation of idiosyncratic aspects of the cognitive processing into the random utility discrete choice models to be addressed. They permit a more realistic representation of behaviour in the choice process, with a better predictive power, producing consistent and efficient estimates of the parameter, and seem to fill the gap between behavioural theory and discrete choice models (Ben-Akiva et al., 1999).

Essentially, the hybrid model emphasises that history influences context: the choice process is context dependent and governed by psychometric constructions, which are constructed throughout the time and vary among individuals, assuming heterogeneity across them. Attitudes are examples of psychometric constructions. In our framework for analysis of sanitation demand, we hypothesise that attitudes of individuals are key drives underlying behaviour, and affect individuals' preferences and their decision making process towards different sanitation options. Attitudes reflect individuals' needs, values, tastes, and capabilities, and are affected by experiences, socio-cultural and historical aspects (which are formed over time) and socio-economic characteristics (McFadden, 2001). Attitudes also explain part of the measured error of the choice process captured by the discrete model, and reflect an *intention* of action towards behaviour, not the behaviour itself. Then, attitudes influence the observed behaviour, which in demand analysis is represented by preferences. Preferences embody the relative desirability of different choices, and can be defined through experiments (stated preferences) or observed choice/behaviour (revealed preferences: type of sewerage system, toilet, etc., actually adopted) – (Ben-Akiva et al., 2002).

Attitudes are latent constructions and are measured using scaling provides techniques and statistics theory for converting responses into measurements (Ben-Akiva et al., 1999). These measurements have been based on the use of multi-item scales (e.g., a semantic differential scale) and data reduction techniques such as factor analysis, principal component analysis and structural equations (Morikawa, Ben-Akiva, & McFadden, 1996). In sanitation, for instance, they can relate to the advantages, qualities, or importance of sewerage or latrine adoption (e.g., importance of cleanliness, or of health benefits). They are then treated as explanatory predictors (non-directly observed) in choice models (McFadden, 2001). The observable variables, such as socio-economic and demographic characteristic of individuals, are also important to explain behaviour, but they are part of the explanation, not the deterministic behaviour. Hence, the traditional systematic segmentation of the population based only on observable variables to determine collective tendencies is substituted by a more comprehensive explanation based on individuals' considerations, criterion, and preferences (Walker, 2006). The cause–effect relationship among these factors is modelled through causal mapping, using indicators to represent a random component of the utility function (Ben-Akiva et al., 1999).

The hybrid framework for sanitation demand is presented in Fig. 1. In this figure, the traditional discrete choice analysis can be seen from the vertical relationship among the socio-economic and demographic characteristics of individuals and attributes of options (explanatory variables) and the decision process (utility maximisation – U), revealed preference. The hybrid model expands the choice process incorporating psychometric constructions in the analysis. The terms in boxes are variables observed directly (or measured by suitable experiments): characteristics of individuals and attributes of the available sanitation, and the indicators of attitudes. The terms in ovals are latent variables which are not observed directly, but are rather inferred from other variables. The full arrows indicate the direction of causal relationship between variables: latent variables are influenced by explanatory variables and are linked to the decision process through observed measures that include indicators, which represent manifestations of latent constructs. These latent constructs (indicators) can be evaluated by an affective value (which varies in sign: positive or negative) and modelled using a structural latent variable model. The dashed arrows are related to the influence of not observed variable on observed ones. Cultural, social and historical aspects of individuals as well as their needs, values, tastes, capabilities and experiences are not directly measured, but rather assumed to be indirectly captured by individuals socio-economic and demographic characteristics and attitudinal latent constructions, resulting from a multi-stage process in which an underlying causal structure explains the
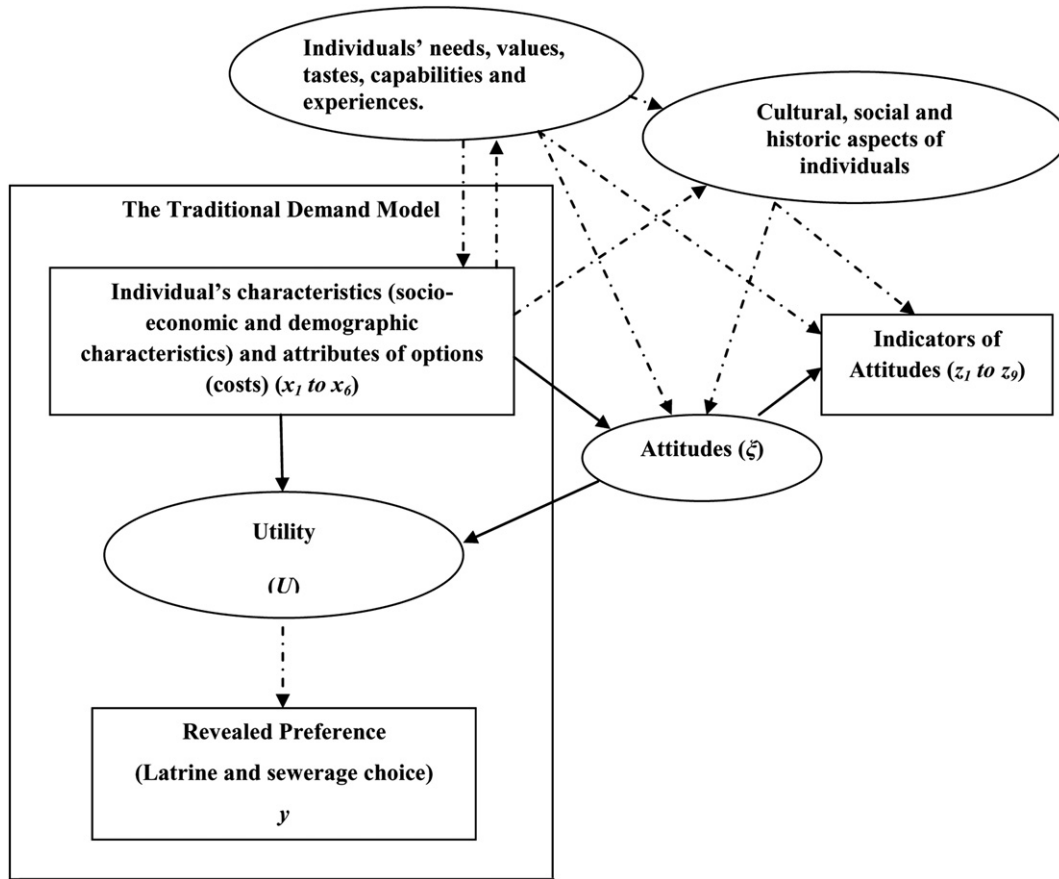
Fig. 1. The hybrid framework for latrine and sewerage choice in Salvador, Brazil. Source: adapted from Ben-Akiva et al. (1999).

influence of various direct and indirect factors on sanitation choice (Ben-Akiva et al., 1999).

The integrated model is composed of two parts: a latent variable model, formed by structural equations, and the discrete choice model, the measurement equations. Structural equations indicate the cause and effect relationships between the observable causal variable to the latent variable. From the values of the latent variables, measurement equations are constructed to demonstrate the distribution of the indicators that are conditioned to the values of latent variables. Measurement equations represent the relationship between observable indicators and the underlying latent variables. This system of equations is then estimated simultaneously or sequentially through any combination of models (Bolduc & Ben-Akiva, 2005). For this analysis we used the mixed logit model presented in the methods section below.

Based on the indirect utility maximisation model, a general mathematical representation of the above relationships can be defined as (Ben-Akiva et al., 1999):

$$U_j = b'x_j + c_j'\xi + v_j \tag{1}$$

Where

$$z = \eta\xi + \omega \tag{2}$$

and,

$$\xi = \delta x + \epsilon \tag{3}$$

then,

$$U = j \text{ if } U_j > U_k; \ \forall k \in J; \ 0, \text{ otherwise} \tag{4}$$

U denotes the indirect utility of option j for the observed outcome y (revealed preference; e.g., type of toilet). The indirect utility is a function of observed and latent variables, and a random error component; z is a vector of observable indicators of ξ; ξ is a vector of individual latent variables, x is a vector of exogenous observable variables that determine ξ (x may or may not be a part of z), b_j and c_j are vectors of unknown parameters to be estimated and η and δ are matrices of unknown parameters to be estimated and $v = (v_1, \ldots, v_j)$, ϵ and ω are measurement errors independent of U, z, and ξ. According to economic consumer theory each individual is able to compare alternatives in a choice set. Theoretically, alternatives are compared using a preference indicator $>$ that compares individuals' options, U_j or U_k. For instance, $U_j > U_k$, that means the individual's revealed preference favours U_j relative to U_k. (Mas-Colell, Whinston, & Green, 1995).

The Bahia Azul sanitation programme in Salvador, Brazil

Starting in 1995, the Bahia Azul sanitation programme had two main periods: between 1995 and 2004 and from 2004 to 2007. The first period included projects directed at improving the collection and final disposal of solid wastes, pollution control, and environmental and sanitary education, but its main objective was to expand the sanitation coverage from 26% to 80% of households. The second stage aimed to cover areas not covered by the first stage of the programme. Evaluation of the effect of this intervention on the health status of the population showed a significant impact on the prevalence and incidence of infectious intestinal diseases (Barreto, 2006; Barreto, Genser, Strina, Teixeira, Assis, Rego et al., 2007).

The sewerage intervention had two components: the first aimed to construct the public sewer system and marine outfalls, the other consisted of thousands of household connections and internal plumbing within households, linking the toilet, bath area, and kitchen sink. This stage should be carried out and paid for each household. The cost of each house connection to Bahia Azul's sewer network was, on average, US$ 20, but an additional surcharge is added every month to the water bill to pay part of the total costs. This additional cost varies, depending on the social class, but all households must pay for it and a State Law determines penalties for the households that do not make regular payments, including the inclusion of their names in financial credit databases (Barreto, 2006).

Salvador has a service, commercial, and industrial economy, with a nominal GDP of US$ 5.8 billion, per capita annual nominal income of US$ 2300, and a Gini index of 0.587, in 2008 (IBGE, 2008). The median income for the 25% poorest is US$ 80 per month and average household expenditure on food, transport and rent is 88% of total income. The educational level is low: 19% of the total population is illiterate, and 72% spent less than nine years at school. About 93% of the total population of Salvador had access to water. Recent figures from the Government of Bahia State showed that in 2008, the sewerage coverage in the city was higher than 80% (Barreto, 2006; IBGE, 2008).

The cost of construction and connecting the internal pipe work to the Bahia Azul sewers was estimated to be US$ 160 at 2006 prices: US$ 140 to construct the internal facilities and US$ 20 to link the internal system to the street sewer. Costs included labour, when households did not do the work themselves, and excluded maintenance. Considering that the median income of the majority of population in Salvador was less than US$ 80 per month, and that most of it was needed for food, transport and rent, a question regarding the success of the Bahia Azul's programme could be raised: would the poor population of the Salvador be able to bear the cost of connection to Bahia Azul's sewers? What are the determinants of households' choice? The answer to this question involves different aspects that are related, not only to the cost of the service, the socio-economic and demographic characteristics of individuals, but also, their perception, beliefs, and attitude towards sanitation, influenced by their historical and cultural background.

Methods

Data collection, timeline of evaluation and study population for the demand study

Data collection for this study was carried out between November 2004 and April 2005, corresponding to the first stage of the programme (1995–2004). The population studied was the same than that evaluated by the epidemiological team (Barreto, 2006; Barreto et al., 2007).

Effective sample used for demand estimation

From the 1128 households selected for the epidemiological study, we obtained information from 721 (64%); 389 (34%) had moved to another neighbourhood or city; and 18 (2%) were revisited, but refused to answer the questionnaire.

Previous qualitative study and questionnaire design

The questionnaire developed by Jenkins (1999) to investigate demand for sanitation in rural Benin was adapted to explore variables related to household attitudes towards toilets with sewerage connection in Salvador. The adaptation of the questionnaire was based on the results from an anthropological study that was carried out during the pre-intervention period to evaluate environmental perceptions and the living conditions of the population (Larrea & Barreto, 2009). It was also based on answers from semi-structured interviews that were conducted with 22 households from different areas in the city.

What was evaluated by the questionnaire?

Attitudes were related to toilets and type of sewerage connection, not type of sewerage connection only. The pilot study of 22 households demonstrated that individuals had difficulty evaluating the connection *per se*, when it is related to their domestic space. For them, an in-house flush toilet connected to a sewerage system (if any) had a different meaning from the sewerage system outside the household. Whereas the in-house system was synonymous with cleanness and health, the external one was mostly related to the community self-respect, house values and smell, this last factor related in some sense to health concerns (Larrea & Barreto, 2009).

Topics included in the survey questionnaire

The survey questionnaire consisted of closed-ended pre-coded questions and some open questions, and incorporated the following topics: characteristics of the house, present defecation practices, and attitudes to the present defecation site for users and non-users. Attitude characterises individuals, their beliefs and intention towards latrines and sewerage systems; it was measured as advantages and disadvantages of these (e.g., little importance to very important) − (Ding & Ng, 2008; Walker, 2001). These measuring scales followed the same structure proposed by (Jenkins, 1999); revealed preferences for latrines and sewerage connection were evaluated in accordance to the existing system the household had at that time, i.e., type of latrine (if any: with or without flush toilet) and sewerage connection (yes or no, and type: Bahia Azul or other); questions about the importance of advantages (drives) and disadvantages (constraint factors) of latrines and sewerage systems, costs of connection and maintenance were also included in the questionnaire.

Definition of adopter and non-adopter

A household that had a flush toilet inside (or outside) the home connected to a system for isolation and disposal of the waste (such as the Bahia Azul sewers, or the mayor system) was considered an adopter. Non-adopters were households with or without an indoor toilet, which lacked a connection to any external isolation and treatment system for the waste (sewage was discharge into the environment, or faeces were deposited in a bag and thrown away). Whether the households had a septic tank connected to a sewerage system or regularly empty the tank not disposing their sewer into the environment (they could have a contract with a company specialised in sewer treatment or other approach), they were also considered an adopter.

Ethics

Two ethics committees approved the research: those of the London School of Hygiene and Tropical Medicine and the Instituto de Saúde Coletiva of the Federal University of Bahia, Brazil. All household heads received a letter informing them of the research objectives and were asked to take part in the study. Those who agreed were asked to sign a consent form.

Table 1
Indicators and scales for attitude latent variables (ξ).

Attitudinal Latent Variable (advantages and disadvantages; scale 2 to 4; 2 = less important and 4 = very important)
Gain prestige from visitors (z_1)
Increasing market value of home (z_2)
Make it easy to defecate (z_3)
Protect from disease (z_4)
Make my house more modern (z_5)
Turns my "image" better to my neighbours (z_6)
It is private (z_7)
It is safe (z_8)
It is comfortable (z_9)

Note: symbols in parenthesis are the same as those used in the equations for determination of the model (see mathematical equations (1) to (4)).

Data entry and analysis

Data were entered using EPI-INFO version 6.0 and analysed using GAUSS version 8 (Aptech, 2006). Mean, standard deviation, and percentages were calculated for descriptive statistics. To test for differences between adopters and non-adopters, Student's t test for independent samples was used to compare the means of a normally distributed variable or the Mann–Whitney U test if the distribution was non-normal. Only the p-values of the tests are presented at a significance level of 95% (i.e., p-value = 0.05).

A simultaneous estimation was performed to assess together the latent model and the discrete choice model. Nine psychometric indicators composed the latent variable model. The indicators for attitude latent variable are presented in the Table 1. Table 2 presents the dependent and independent variables analysed in this model. In short, from equation (1), we have:

$$
\begin{bmatrix} U_{BAZUL} \\ U_{NONE} \\ U_{OTHER} \end{bmatrix} = \begin{bmatrix} b_{11} b_{12} b_{13} b_{14} b_{15} b_{16} \\ b_{21} b_{22} b_{23} b_{24} b_{25} b_{26} \\ b_{31} b_{32} b_{33} b_{34} b_{35} b_{36} \end{bmatrix} \begin{bmatrix} FEMALE \\ AGE \\ INCOME \\ NCHILDREN \\ EDUCATION \\ COST \end{bmatrix}
$$
$$
+ \begin{bmatrix} c_{11} \\ c_{21} \\ c_{31} \end{bmatrix} [\xi_{attitudes}] + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \quad (5)
$$

The final relation in (5) was estimated using maximum simulated likelihood estimation with 125 Halton draws, which is an "intelligent" draw in a simulation method used to reduce the run-times in estimating the integral in some models, when the number of variables increases (Train, 1999).

Goodness-of-fit

To assess the coefficients, p-values were used and the likelihood ratio (LR) statistic was carried out to test the null hypothesis that all the slope coefficients were simultaneously equal to zero (Train & Wilson, 2007).

Results

General characteristics of households and type of sewerage connection

In the household sample (Table 3), 90% of interviewees were female heads of household. It was not by design that the respondents were mainly female. By design, the respondents were the household head, reflecting the data and statements of households (Who is the head of household? Who does take the decisions?). Around 62%, however, were housewives with some "sporadic income" from informal work. The median family monthly income was US$ 158. The sample median age was 32 years, an average of 8.6 years of study and 2 children in the house.

Regarding the type of sewerage system, 69% had already a connection with Bahia Azul's sewers, and 71% had at least a flush toilet; 12% of households had connected to an old municipal sewer system – which partially treated the city sewer and discharged it into the rivers and ocean; currently this system does not exist and it was substitute by the Bahia Azul programme; 10% had access to a storm drain (discharging into the street, lakes, rivers and ocean) or no sanitation at all (Table 3).

The median cost paid by adopters for a sewer connection in 2004 was US$ 28, with a high standard deviation ranging from US$ 0.50 to US$ 524. This volatility was expected because the cost included those households who just paid for the connection, and those who paid for the whole system, including new latrine, flush toilet, and connection (Table 3).

Attitudes towards toilets and sewer connections

Difference between adopters and non-adopters

Attitudes (or intentions) towards latrines and sewerage connections had a few difference between means, when comparing adopters and non-adopters. Among adopters, health protection, accessibility, privacy and security were the well-scored factors followed by comfort, increased value and modernisation of the house, better social image within the neighbourhood and among friends and relatives, and prestige. Non-adopters pointed the house monetary valorisation as the main advantage of sanitation, with accessibility, health protection, house modernisation, and privacy sharing the same score for the second advantage (Table 4). The

Table 2
Dependent and independent variables used in the analysis.

Dependent Variables	Description
BAZUL (U_1)	Toilet connected to Bahia Azul's programme (reference variable)
NONE (U_2)	With or without toilet facilities, but without a sewerage connection
OTHER (U_3)	Other type of connection (mainly the municipal system, not Bahia Azul)
Independent variables	Description
FEMALE (x_1)	Dummy = 1 for female; 0 for male
AGE (x_2)	In years
INCOME (x_3)	Monthly household income in Brazilian currency "Real", 2004
NCHILDREN (x_4)	Number of children in the house
EDUCATION (x_5)	Number of years that household head has studied
COST (x_6)	Cost of sanitation for each household (conditional upon the choice[a]); in Brazilian currency "Real"
ATTITUDE (ξ)	Constructed based in indicators (see Table 1)

Note: symbols in parenthesis are the same as those used in the equations for determination of the model.
[a] Conditional since it depends on the price of the option chosen and not all prices (Persson, 2002).

Table 3
Socio-economic and demographic characteristics, and sanitation choice for a sample of households in Salvador, Brazil (N = 721), in US$.

Female respondent: n (%)	649 (90)
Age of respondent: median (std. dev.)	31.89 (12.23)
Housewife: n (%)	450 (62.4)
Years of education of household head: median (std. dev.)[a]	8.57 (3.83)
Household income (monthly average during 2004): median (std. dev.)[b]	158 (115.75)
Number of children at home: median (std. dev.)[a]	2 (1.37)
Type of sanitation: n (%) *Bahia Azul sewerage*	499 (69.21)
No sanitation	71 (9.85)
Other	151 (20.94)
Flush toilet: n (%)	514 (71.29)
Cost of sanitation: median (std. dev.)	28 (61.86)

[a] N = 719.
[b] N = 716.

differences in attitude scores between the two groups were not statistically significant.

Health was rated better by adopter than non-adopters of this service. The anthropological study pointed out that the population in Salvador could identify the link between lack of sanitation and disease, although no specific link was made with diarrhoeal diseases (Barreto, 2006). Therefore, it was expected that households who had access to this service could better perceive its benefits.

The hybrid choice model with latent variables

The result of the full hybrid choice model is shown in Table 5. Aiming to compare the contribution of latent variables in explaining the choice model, we estimated two models: a hybrid choice model without latent variables and a model with those variables. The results showed that the latent variables clearly improved the magnitude and significance of coefficients, and the difference between the two analyses was extremely significant in terms of the resulting gain (more robust coefficients) and specification of the mixed choice model (likelihood ratio).

The coefficients presented in Table 5 are the log-odds of the effect of each x and ξ (explanatory variables) on U (or y, the dependent variable), always comparable to Bahia Azul, the reference variable. The easiest and simplest method of interpretation of these results is through the sign of parameter estimates and their statistical significance: a positive sign suggests the likelihood of the response increases with the level or presence of x and ξ, other things being equal; a negative sign has the converse interpretation (Liao, 1994).

Flush toilets with Bahia Azul sewerage were more popular than all other sanitation options. Analysing the socio-economic and demographic characteristics of the household we find that they were preferred by all age groups, income and educational levels. The more children in the household, the lower was the probability that this option would be chosen. However, this option still prevails over *Other* when the number of children is included in the analysis. Females prefer Bahia Azul to no system at all, but males prefer *Other* to Bahia Azul sewers. As expected, the higher the cost of other options more households will opt for Bahia Azul sewerage. Positive attitude makes Bahia Azul sewerage preferable to no sanitation, but not to *Other*.

The sign of the coefficient for all socio-economic and demographic variables is the same in both models, with and without latent variables. The model with latent variables increased not only the magnitude and significance of the coefficients, but also the goodness-of-fit of the general model. The Likelihood Ratio chi-square improved when these variables were added.

The magnitude and significance of coefficients reflect the reliability of the correlation between the explanatory and dependent variables and indicates the strength of variables in explaining the model. The estimated coefficients demonstrated primarily that the latent variable attitude played an important role in the household's decision to connect to a sewerage system. The attributes of health protection, accessibility, privacy and security related to flush toilets and the sewerage system, *not only* the toilet and connection themselves reflected in their costs and household socio-economic status, were what the households also cared about when determine their choice. All coefficients were significantly different from zero when latent variables were included in the model. This suggests that these variables captured factors that influenced the household decision in adopting sanitation, and thus reduced the random error term.

Discussion

The objective of this paper was to identify the determinants of demand for sanitation in Salvador, using a framework of analysis that allows the causal relationship among the socio-psychological latent constructs, the socio-economic and demographic characteristics of households, and the attributes of the available sanitation options, to be tested and demonstrated. The hybrid model is statistically consistent with this objective and seems to fill the gap between behavioural theory and discrete choice models applied to sanitation.

In economics, causal inference is treated as a problem of missing data where the choice is revealed from the observed data. It means

Table 4
Attitudes to advantages of toilets with sewerage connection among a sample of adopters and non-adopter households in Salvador, Brazil.[a]

Advantages	Adopter[b]		Non-Adopter		Difference between means (p-value)
	Mean	Std. Dev.[c]	Mean	Std. Dev.[c]	
Gain prestige	3.39	0.639	3.53	0.624	−0.14 (0.575)
House monetary valorisation	3.55	0.553	3.82	0.393	−0.27 (0.433)
Easy access	3.64	0.502	3.71	0.470	−0.07(0.148)
Health protection	3.70	0.487	3.71	0.470	−0.01(0.274)
Make my house more modern	3.48	0.583	3.71	0.470	−0.23(0.768)
Make my social image better	3.43	0.642	3.53	0.514	−0.10 (0.512)
Privacy	3.62	0.516	3.71	0.470	−0.09 (0.547)
Security	3.62	0.508	3.76	0.437	−0.14 (0.184)
Make my house more comfortable	3.58	0.531	3.71	0.470	−0.13(0.241)

[a] Scores range from 2 (most negative) to 4 (most positive) for each advantage.
[b] N = 718; Adopters = 647; Non-adopters = 71.
[c] Std. Dev. = standard deviation.

Table 5
Estimation results for the hybrid model: reference model (without the latent variable) and the latent variables enriched model (Using 125 Halton Draws).

Reference Model (without latent variables)			Latent Variable Enriched Model	
Variables	Coefficients	p-value	Coefficients	p-value
None	−1.54391	0.003	−3.25463	0.000
Other	−1.43529	0.002	−2.87452	0.000
Cost	−0.08354	0.000	−0.58471	0.000
Female: none[a]	−0.90534	0.071	−2.72587	0.000
Female: other	1.54382	0.000	2.03958	0.000
Age: none	−0.37264	0.297	−1.58742	0.000
Age: other	−0.96453	0.000	−1.87413	0.000
Income: none	−0.95482	0.033	−1.00874	0.000
Income: other	−0.35279	0.009	−1.95874	0.001
N children: none	0.85342	0.000	1.88725	0.000
N children: other	−1.85473	0.343	−2.96857	0.001
Education: none	−0.57410	0.000	−1.36874	0.000
Education: other	−1.84537	0.000	−1.97413	0.000
Attitude: none			−3.97425	0.000
Attitude: other			2.87425	0.000
Number of observations: 721			Number of observations: 721	
p-value (model): 0.0000			p-value (model): 0.0000	
LR chi-square: −1632.8			LR chi-square: −1593.2	

[a] It refers to the variable female associated to option none (no sanitation), compared to Bahia Azul option; the same interpretation can be applied to other variables, e.g., Age: other, which refers to the age of respondent associated to the option other.

that as only one of the outcomes can be observed (Y_1 or Y_0), the causal effect (or effects) is defined by their comparison, e.g., $Y_1 > Y_0$. If every individual has two possible outcomes, Y_1 and Y_0, where Y_1 is the outcome if the individual opts for a sewer connection and Y_0 is the outcome if the individual does not, the causal inference becomes a problem of inference with missing data: one of the outcomes is observed in the data (the one chosen), and the other outcome (the one not chosen − a counterfactual) cannot be observed because the individual cannot choose two options at the same time (Maelli & Rubin, 2003; Moffitt, 2003; Rubin, 1974).

From this premise, however, only an exogenous variable (say X) is a causal variable, and the estimated coefficient on X is interpreted as the difference that would occur in Y if a particular individual were exogenously given one more unit of X (Moffitt, 2003). The problem arises when the variable in question whose has effect on Y is potentially endogenous. In this case, "*the action-event is also affected by things we cannot measure, for we can never fully explain and measure the determinants of why individuals take an action or experience an event and others do not*" (Moffitt, 2003, page 449). This would be the case of latent variables (ξ, attitude) that was used in the model to explain choice of sanitation. However, the new advances in statistics and econometrics allow that, through a causal pathway, cause and effect among latent variables and exogenous variables are identified and their effect on the dependent variable is isolated. These advances consist of the integration of a structural model (structural equations: see equations (2) and (3)) to a discrete choice model (see equations (1) and (4)) to assess the covariance or correlation between the latent endogenous variables (caused by exogenous variables) and exogenous ones (not caused) − (Ben-Akiva et al., 1999). The different aspects of the indicators that were used to construct the attitudinal latent variable will be part of another paper as it will require the explanation of other statistical factors such as pathway analysis and structural equations.

The results presented here shown that both, adopters and non-adopters, had a positive attitude towards sanitation in Salvador, and that this attitude influenced their choice of type of system. When comparing the estimated coefficients of attitudes (and their significance) to other estimated coefficients, we can observe that

attitudinal constructions could explain more about sanitation choice than individuals' socio-economic and demographic characteristics and costs. The variable costs, for instance, had the lowest coefficient amongst other variables. It suggests that costs were not the main determinant of households' choice for type of sanitation, rather the attitudes towards sanitation, gender, number of children in the household and age played a much more important role in the choice process.

The main constraint for non-adopters was the number of children in the house: higher the number of children, lower was the chance the household would opt for sanitation. We did not evaluate the variable costs by group of adopters and non-adopters, separately; however findings suggest that policy makers should focus greater attention on, for example, improving subsidy programmes as an incentive for non-adopters with many children in the household if the objective is increasing coverage. For this group, a full subsidy should be considered and partial subsidy should be designed only to the remaining group of non-adopters; those with few children in the household. This plan of subsidy might guarantee a certain level of equity in the distribution of public resources.

The hybrid model established a causal relationship amongst observable and latent variables, identifying the key factors that influenced choice for sanitation in Salvador, allowing policy makers to identify and possibly to draw specific interventions by socio-economic groups. However, the interpretation of the effect of the *indirect* variables affecting behaviour is essentially a descriptive work and need to be discussed in the light of the socio-cultural and historical factors affecting a specific community or society. The attitudinal latent variable was determined by indicators of attitude that represented the advantages of sanitation perceived by the households − health protection, accessibility, privacy, security, comfort, house monetary valorisation and prestige. These indicators were assumed to *indirectly* capture the individuals' needs, values, tastes, capabilities, experiences, cultural, social and historical aspects. These aspects were also directly and indirectly captured by the socio-economic and demographic characteristics of individuals.

According to the anthropological study, in Salvador, access to sanitation was seen as an essential social right, which should be provided by the State for the whole population. Access to sanitation, in this case, had two components: the external availability of the sewerage network, and the internal domestic plumbing. The external availability refers to the government action in making sewerage accessible for the entire population: the fact that some neighbourhood had access to it and others did not, created a high sense of inequality in the community and some social movements to repair this situation (Larrea & Barreto, 2009). The internal availability was understood in part as a household obligation and in part as a government obligation, mainly for the non-adopters group. The households were willing to bear partially the cost of sanitation (Barreto, 2006).

Environmental aspects were also important issues of social concern about sanitation, contrary to the common assumption in industrialised countries that people in developing countries do not care about environmental quality. A Gallup poll conducted in 1992 revealed that 80% of Brazilians had "a great deal" or "a fair amount" of personal concern about environmental effect of poor sanitation. In Mexico, this percentage was 83%, in Uruguay was 82%. These rates were similar to United States rate, which was 85%. A poll of 580 people (243 middle-class people, and 337 poor people) conducted in Salvador indicated that the consequences of the lack of urban sanitation (open-air sewers, polluted beaches, and uncollected garbage) were seen as an important environmental problem. The findings were more remarkable for poor people (40% of responses) than for middle-class people (34% of responses). Other environmental problems mentioned by respondents included destruction of

natural resources, uncontrolled industrialisation, bad policies, and lack of environmental control. Unsanitary conditions remained the most important concern however, even when compared with more general questions like infrequent garbage collection and irregular supply of water (Tesh & Paes-Machado, 2004).

Another explanation is associated to the introduction of a sanitation law, the State Sanitation Law, in the state of Bahia, which compelled the household to pay part of the costs of sanitation (Larrea & Barreto, 2009). The introduction of the law itself could, in part, explain the high coverage of sanitation in the city, even amongst the poorest. For low income classes, sanitation was partially subsidised by the middle and high income classes (Barreto, 2006). Nevertheless, the costs were still high for these poor strata and many of them, in spite of having contracted the service (connected to Bahia Azul), could not pay their bills and so had their credit ratings downgraded (Larrea & Barreto, 2009). Many households, however, preferred this situation to being without a good sewerage system, and this is explained by their high evaluation of this service reflected by the attitudinal constructors. In a previous survey, when households were asked for the "*main thing they wish to change in their neighbourhood*", most answers were related to sanitation, especially sewerage connections (Barreto, 2006).

In rural Benin, Jenkins (1999) found that without strong drives (perception, attitude and other latent construction) for a latrine, the household would be uninterested in a change of sanitation practice. A lack of such drives, not the presence of constraints alone (costs of connection, for instance) was the main reason for non-adoption. The presence of one or more drives had a strong influence on adoption.

In our sample, at about 34% of individuals had moved to another neighbourhood, what can be an indicator of a high percentage of rented houses. We did not conduct any further assessment using this information and therefore this could be considered a potential limitation for the interpretation of our results. However, we believe the Hybrid Choice Model used to estimate demand for sanitation in Salvador is consistent in demonstrating the importance of latent variables underlying revealed preferences, opening the black box of consumer choice, and contributing to our understanding of sanitation demand.

References

Aptech Systems, Inc (Aptech). (2006). *GAUSS 8.0*. Washington: Aptech.

Barreto, M. L. (Ed.). (2006). *Avaliaçao do impacto epidemiológico do programa de saneamento ambiental da Baia de Todos os Santos (Bahia Azul)*. Salvador: Instituto de Saude Coletiva, Federal University of Bahia.

Barreto, M. L., Genser, B., Strina, A., Teixeira, M. G., Assis, A. M. O., Rego, R. F., et al. (2007). Effect of city-wide sanitation programme on reduction in rate of childhood diarrhoea in northeast Brazil: assessment by two cohort studies. *Lancet, 370*(9599), 1622–1628.

Ben-Akiva, M., McFadden, D., Gärling, T., Gopinath, D., Walker, J., Bolduc, D., et al. (1999). Extended framework for modeling choice behavior. *Marketing Letters, 10*, 187–203.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., et al. (2002). Hybrid choice models: progress and challenges. *Marketing Letters, 13*(3), 163–175.

Bolduc, D., & Ben-Akiva, M. (2005). *Hybrid choice models with logit kernel: Applicability to large scale models*. Boston: Elsevier.

Diamond, P. A., & Hausman, J. A. (1994). Contingent valuation: is some number better than no number? *The Journal of Economic Perspectives, 8*(4), 45–64.

Ding, Z., & Ng, F. (2008). A new way of developing semantic differential scales with personal construct theory. *Construction Management and Economics, 26*(11), 1213–1226.

Domencich, T., & McFadden, D. (1975). *Urban travel demand: A behavioral analysis*. New York: North-Holland Publishing Corporation.

Instituto Brasileiro de Geografia e Estatística (IBGE). (2008). *Pesquisa Nacional por Amostra de Domicilios (PNAD)*. Rio de Janeiro: IBGE. Available in. http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2008/default.shtm.

Jenkins, M. W. (1999). *Sanitation promotion in developing countries: Why the latrines of Benin are few and far between*. Davis: University of California. Department of Civil and environmental Engineering.

Jenkins, M. W., & Scott, B. (2007). Behavioral indicators of household decision-making and demand for sanitation and potential gains from social marketing in Ghana. *Social Science & Medicine, 64*, 2427–2442.

Kahneman, D. (2003). Maps of bounded rationality: psychology for behavioral economics. *American Economic Review, 93*(5), 1449–1475.

Knetsch, J. L., & Sinden, J. A. (1994). Willingness to pay and compensation demanded: experimental evidence of an unexpected disparity in measures of value. *The Quarterly Journal of Economics, 99*(3), 507–521.

Larrea, C. K., & Barreto, M. L. (2009). Salud ambiental urbana; aproximaciones antropológicas y epidemiológicas sobre la evaluación de un programa de alcantarillado en contexto de grandes desigualdades sociales. *Quaderns de l'Institut Català d'Antropologia, 22*, 71–99.

Liao, T. F. (1994). *Interpreting probability models: Logit, probit, and other generalized linear models*. Iowa: Sage Publications.

Maelli, F., & Rubin, D. B. (2003). Assumptions allowing the estimation of direct causal effects. *Journal of Econometrics, 112*, 79–87.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. New York: Oxford University Press.

McFadden, D. (2001). Economic choices. *The American Economic Review, 91*, 351–378.

Moffitt, R. (2003). Causal analysis in population research: an economist's perspective. *Population and Development Review, 29*(3), 448–458.

Morikawa, T., Ben-Akiva, M., & McFadden, D. (1996). *Incorporating psychometric data in econometric choice models*. Massachusetts: Massachusetts Institute of technology. Working paper.

Persson, T. H. (2002). Welfare calculations in models of the demand for sanitation. *Applied Economics, 34*, 1509–1518.

Phaneuf, D. J., Carbone, J. C., & Herriges, J. A. (2007). *Non-price equilibria for non-marketed goods*. Washington: Frontiers of Environmental Economics Workshop.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomised and nonrandomised studies. *Journal of Educational Psychology, 66*, 668–701.

Shordt, K., van Wijk, C., & Brikke, F. (2004). *Monitoring millennium development goals for water and sanitation: A review of experiences and challenges*. Delft: IRC International Water and Sanitation Centre.

Tesh, S. N., & Paes-Machado, E. (2004). Sewers, garbage, and environmentalism in Brazil. *Journal of Environment & Development, 13*, 42–72.

Train, K. (1999). *Halton sequences for mixed logit*. Berkley: University of California. Department of Economics. Available in. http://elsa.berkeley.edu/wp/train0899.pdf.

Train, K., & Wilson, W. W. (2007). *Econometric analysis of stated-preference experiments constructed from revealed-preference choices*. Berkley: University of California. Department of Economics. Available in. http://elsa.berkeley.edu/~train/spoffrp.pdf.

Walker, J. L. (2001). *Extended discrete choice models: Integrated framework, flexible error structures, and latent variables*. Massachusetts Institute of Technology, Massachusetts: Department of Civil and Environmental Engineering.

Walker, J. L. (2006). Opening up the black box: enrichment behavioral models of spatial and travel choices. *Journal of Transport Geography, 14*(5), 396–398.