

Technical note on the determinants of water and sanitation in Yemen

Giacomo Sbrana ¹

Department of Economic and Social Affairs (DESA)
United Nations, New York
02 June 2009

1. Introduction

This technical note briefly describes and analyses the main empirical results obtained from an econometric estimation of the main determinants of water and sanitation achievements in Yemen. It has been elaborated as part of the technical backstopping that the Development Policy and Analysis Division of the United Nations Department of Economic and Social Affairs (DPAD/UN-DESA) is currently providing in the framework of the project “Assessing Development Strategies to Achieve the Millennium Development Goals (MDGs) in the Arab region”, in close collaboration with UNDP’s Regional Bureau for Arab States and the World Bank. As such, the econometric estimation used in the note follows closely the specification that is set up in the economy-wide model system that is used in the project to assess MDG related strategies and their macroeconomic trade-offs.² In the following two sections, respectively, the estimation methodology and the analysis of empirical results are presented.

2. Estimation methodology

The main data source for the econometric estimations is the Demographic and Health Survey (DHS) that the Central Statistical Organization of the Ministry of Planning and International Cooperation of the Republic of Yemen conducted for 2003. Two

¹ Associate Expert at the Development Policy and Analysis Division of the United Nations Department of Economic and Social Affairs (DPAD/UN-DESA). Comments are welcome and can be addressed to the author by email (sbrana@un.org). The views and opinions expressed here are those of the author and do not necessarily reflect those of the United Nations or its member states.

² The economy-wide model system is known as MAMS (*Maquette* for MDGs simulation). It is an innovative type of dynamic computable general equilibrium (CGE) model that has been especially designed by the World Bank to analyse the impact of public investments and interventions on MDG achievements, sectoral economic growth, and employment, among other potential areas of analysis. This model has been extended through its application in 18 countries of the Latin American and Caribbean (LAC) region. The model and results from its application for LAC countries can be found in: Marco V. Sánchez, Rob Vos, Enrique Ganuza, Hans Lofgren, and Carolina Díaz-Bonilla (eds), *Public Policies for Human Development. Feasible Financing Strategies for Achieving the MDGs in Latin America and the Caribbean*, London: Palgrave/Macmillan (forthcoming).

specifications were used to conduct the econometric estimations for, respectively, access to improved drinking water and access to improved sanitation facilities. Accordingly, two dummy variables for the left hand side of the two specifications were specified using the following DHS variables:

<i>Main source of drinking water</i>	<i>Type of toilet facility</i>
piped supply	flush toilet connected to sewage
cooperative supply	flush toilet not connected to sewage
artesian well	bucket
regular well	pit
guil (spring)	toilet with tank
covered pond	open air
uncovered pond/ krif	public/street toilet
tanker truck	
bottled water	

Those individuals surveyed who claimed they had piped or cooperative supply, artisan or regular well, or bottled water, were considered to have access to improved drinking water for the construction of the first dummy variable.³ For these individuals, therefore, the

³ According to the Joint Monitoring Program (JMP) for water and sanitation of the World Health Organization and UNICEF, improved water sources include household connections, public standpipes, boreholes, protected dug wells, protected springs, and rainwater collections. Unimproved water sources, on

dummy variable was given a value of one. The remaining sources of water were considered poor enough to be part of an improved drinking water system, so individuals who claimed they had access to these were all given a value of zero for the construction of the dummy variable.

As for sanitation, improved facilities are considered to be a flush toilet connected or not connected to sewage, a pit, or a toilet with tank. Individuals with access to this type of facilities were given a value of one for the construction of the dummy variable, whereas, on the other hand, a value of zero was assigned to all other individuals claiming they have access to one of the other facilities. By these definitions, the DHS indicates that 72% and 52% of Yemen's population had access to, respectively, improved drinking sources and improved sanitation facilities in 2003. It is then clear that the gap to achieving the targets for water and sanitation that are part of MDG7 remains fairly wide.

Due to the binary response variables describing access to water and sanitation, a standard logit regression procedure was employed to carry out the estimations. The following specification was estimated for the dummy variables (y_i) of, respectively, access to improved drinking water and access to improved sanitation facilities:

$$Pr ob(y = 1) = \frac{\exp^{\alpha*wealth+\beta*area+\gamma*spending_pc+\theta*electricity}}{1 + \exp^{\alpha*wealth+\beta*area+\gamma*spending_pc+\theta*electricity}}$$

whereby the probability that the event “access to improved drinking water/access to improved sanitation facilities” occurs is represented; or otherwise:

$$Pr ob(y = 0) = \frac{1}{1 + \exp^{\alpha*wealth+\beta*area+\gamma*spending_pc+\theta*electricity}}$$

and, where:

wealth: represents a wealth index that is calculated for each individual as the fraction of assets and amenities that the individual's household own of a

the other hand, would include unprotected wells, unprotected springs, vendor-provided water or tanker truck-provided water.

bundle of eighteen assets and amenities.⁴ This would be a proxy for household income per capita when this is seen as a determinant of access to drinking water and/or basic sanitation as in the MAMS framework. The choice of using a wealth index based on household assets and amenities is in line with the precedent used in Case, Paxson and Ableidinger (2004)⁵ and in Fortson (2008)⁶ as this index provides an easier and more understandable interpretation of the coefficients.

area: dummy variable that describes if the individual lives in the urban (assuming a value of 1) or in the rural (assuming a value of 0) areas, the inclusion of which intends to control for the important disparities that are observed between the two areas.

spending_pc: this variable represents per-capita public expenditure on “Health and Population” by governorate in 2004. Due to lack of availability of more disaggregated data, population expenditure, which includes water and sanitation, had to be used in tandem with health expenditure. In view of no other alternative, either, data on public expenditure on “Health and Population” were used to construct this variable even though these are extraneous to the DHS. They were provided by the Ministry of Finance and pertain to the final account of the local authority's budget for the 2004 fiscal year. These data and their total summation were also divided by the total population of, respectively, each governorate and the country, in order to express them in per-capita terms. Subsequently, the ratio of each per-capita spending by governorate and the per capita spending for the country was computed. This ratio is a measure of how much each governorate spends on “health and population” compared with the national

⁴ The bundle includes: radio, television, video, refrigerator, dish, electrical cooking stove, water heater, sewing machine, electric fan, washing machine, air conditioner, electrical Hoover, blender, bicycle, motorcycle, car, telephone and flush toilet or pit latrine.

⁵ Case, Anne, Christina H. Paxson and Joseph Ableidinger. 2004. “Orphans in Africa: Parental Death, Poverty, and School Enrollment”. *Demography* 41(3): 483-508.

⁶ Jane G. Fortson. 2008. “The Gradient in Sub-Saharan Africa: Socioeconomic Status and HIV/AIDS”. *Demography* 45(2): 303-322.

spending. The construction of the variable - as explained – was necessary because the use of governorate-specific spending per capita by itself would have been difficult to interpret in a standard logistic regression (especially when the focus is on the marginal effects). Since most of the provision of water and sanitation services is carried out by the government in Yemen, this variable is a good proxy for per capita spending on water and sanitation in the MAMS framework.

electricity: this is a variable that expresses the proportion of people that have access to electricity provided by the government and by governorate.. It can be considered as a proxy of access to public infrastructure along the lines of the MAMS framework.⁷

3. Empirical results

The results obtained after estimating the equation above, separately for water and sanitation, are all reported in table 1. These include the value of the estimated parameter; the z-statistics in brackets; the marginal effects, which measure how much the probability of having access to improved drinking water or access to improved sanitation facilities would change as a result of a variation in each of the determinants; and, the implicit elasticities, which indicate the percentage change of the probability of having access to improved drinking water or access to improved sanitation facilities given a 1% change in each of the determinants. Furthermore, other summary statistics including sample size and robustness of the estimated specifications are part of the table.

According to the empirical results, wealth evidently has a strong and significant positive impact on the probability of having access to improved drinking water and sanitation facilities. Richer individuals are, therefore, expected to have more access to such services. The probability of access to the two services is respectively 3% and 20% higher for those

⁷ The public infrastructure sector in MAMS generally accounts for all other public infrastructure not pertaining to education, health, and water and sanitation. Usually, it includes the provision of public electricity and transport (public roads, bridges, airports, etc.). The 2003 DHS of Yemen, however, contains no information on transport so only electricity is used in the estimations.

individuals that live in the urban areas. This result is expected to the extent that urban areas generally possess better infrastructure than rural areas, especially for sanitation.

From a policy viewpoint, the empirical results also indicate that public spending per capita in health and population would increase the probability of having access to improved drinking water and sanitation facilities by about 1% and 4%. Electricity seems to play an important role, too. The probability of having improved water and sanitation increases by 12% and 18%, respectively, where there is access to public provision of electricity. In terms of the implicit elasticity, electricity turns out to be the second most important determinant - after wealth - which is an indication that public infrastructure may have an important synergy effect on increasing the access to improved water and sanitation.

In general, the impact of the determinants on the probability of having access to sanitation facilities is higher than in the case of the equation estimated for access to drinking water. This should not come as a surprise considering that, as indicated earlier, only little more than half of the population has access to decent sanitation facilities. Coverage of the drinking water supply is much wider, although there are still important gaps to tackle. From a policy viewpoint, the results help to conclude that public spending in improving the access to sanitation facilities would lead to achieve more than public spending in increasing the water supply.

Table 1. Estimation results for water and sanitation^{1/}

	Water equation			Sanitation equation		
	Parameter estimates	Marginal effects	Elasticities	Parameter estimates	Marginal effects	Elasticities
<i>wealth</i>	1.265 (22.94)	0.246	0.077	5.997 (85.68)	0.961	0.554
<i>area</i>	0.137 (5.46)	0.026	0.009	1.141 (43.66)	0.203	0.129
<i>spending_pc</i>	0.056 (2.35*)	0.011	0.015	0.272 (9.62)	0.043	0.104
<i>electricity</i>	0.604 (8.34)	0.117	0.071	1.135 (13.07)	0.182	0.194
Sample:		78,982			83,083	
Mean of dependent variable		72%			52%	
LR chi2(4)		2,362.99			35,027.86	
Prob > chi2		0.0000			0.0000	
Pseudo R2		0.0254			0.3047	

^{1/} The following notes apply to this table: (i) z-statistics are presented in brackets; (ii) the statistical significance is at the 1% in all cases but those where an asterisk has been added; (iii) the marginal effects are defined as $\Delta y/\Delta x$, where Δ denotes change, y is the value of the dependent variable, and x represents the value of the determinant (s); and, the elasticity is computed as follows: $(\Delta y/y)/(\Delta x/x)$.