

Determinants of Adapted Improved Sanitation in Rural and Urban of Bangladesh and Pakistan

ABSTRACT

This study focuses on the socioeconomic and demographic determinants of improved sanitation adaptation in South Asia, mainly in Bangladesh and Pakistan. For the purpose of mass population study, cross-sectional datasets were obtained for both Bangladesh and Pakistan. The survey is based on a two-stage stratified sample of households. This study hypothesized the status of sanitation into three categorically distributed dependent variables which are “improved sanitation facility”, “unimproved sanitation facility” and “no sanitation facility”. According to the World Health Organization, these variables are supported to categorize the situation of sanitation by Joint Monitoring Program categorization of toilet facilities. The “improved sanitation facilities” includes Flush or pour flushed toilets connected to sewers, Flush or pour flushed toilets or latrines connected to pits or septic tanks, and Ventilated improved pit (VIP) latrines, Pit latrines with slab, Composting toilets including twin-pit latrines and container-based systems. The “unimproved sanitation facilities” includes Pit latrines without slab, Bucket latrines and Hanging latrines. And the “no facility” category includes open defecation (bush/field/desert/road/beach etc.). Wealth Index of the household, Gender of head of the household, marital status of the respondent, Education of household head, Locality of the household, Province/ division are the categorical explanatory variables used in this study; family size of the household and age of the head of the household are two numerical explanatory variables that are taken as important demographic factors. Two candidate models were constructed for predicting no facility, unimproved sanitation and improved sanitation respectively. In the sample, only 1.26% households of Bangladesh do not have any toilet facility while Pakistan has 10.1% households without any toilet facility. The analysis shows agreement among all indicators, for all three component models, and for two data sets. There is no clear trend for model performance improvement by year. All the comparison results indicate the clear improvement of the ordinal logistic model over the multinomial regression model. The AIC, BIC, log likelihood statistics suggest that the highest education of the family members is more influential factor for averting status of a household in Pakistan. The analysis shows agreement among all indicators, for all three component models, and for two data sets. There is no clear trend for model performance improvement by year. All the comparison results indicate the clear improvement of the ordinal logistic model over the multinomial regression model.

Keywords: Bangladesh, Pakistan, adaptation, sanitation, ordinal logistic regression, multinomial logistic regression.

1. INTRODUCTION

In this century of SDGs, sanitation has key importance to attain sustainability in every other aspect of human health and productivity. Sanitation mainly refers to the conditions of public health related to clean drinking water and adequate treatment and disposal of human excreta and sewage; preventing human contact with feces is a vital part of sanitation. Water and sanitation are essential for local development, particularly for sectors such as health, agriculture, social and economic

development, education, and the environment. (MMR Sarker, H Yamashita 2019) Underdeveloped and developing countries are still susceptible in the sense of sanitation where challenges and barriers come in the case of improved sanitation facilities adaptation. In this regard, Bangladesh has made appraisable progress in the past few decades. After the liberation war in 1971, the country was held opened to be attacked by water contaminated diseases due to inadequate knowledge about sanitation in the general people. The nation fought back against contaminating diseases and their epidemics by being educated about the importance of sanitation and applying good sanitation practices in daily life. Though Bangladesh had been recovering excellently from its savage situation of sanitation, anomaly still remains after five long decades from the country's birth. What are the reasons for not adopting good sanitation in life of the people of Bangladesh? Well the answer depends on the socioeconomic and demographic factors of the person and the household. As it is a matter of concern, studies should be conducted to investigate the situation and formulating efficient policies that help people to adopt the sanitation facilities are crying needs of the time being. This study aims i) to determine whether the socioeconomic and demographic factors influence the adaptation of improved sanitation facilities; and ii) to compare the sanitation situation of Bangladesh and Pakistan.

1.1. BACKGROUND AND MOTIVATION

Over 82% of the global population has access to a latrine or toilet that meets the World Health Organization/United Nations International Children's Emergency Fund (WHO/UNICEF) definition of a 'basic' sanitation system (WHO/UNICEF, 2019). In terms of access to sanitation, Bangladesh has progressed tremendously. In Bangladesh, over 99% of the population has access to some form of latrine or toilet (WHO/UNICEF, 2019) and 92.7% of the population use on-site containment methods such as pit latrines (WHO/UNICEF, 2017). Open defecation has dropped from 18% in 2000 to 1% in 2015 (estimated by JMP), and in general, city people have access to some kind of latrine, but the main problem concerned is what happens to the management of human waste after the use of such facilities. As of 2018, the country had nearly ended open defecation. The significant accomplishments of improving sanitation are largely attributed to the leadership of the GoB and initiatives such as Community-Led Total Sanitation, which was developed and launched in Bangladesh in the early 2000s (CRI, 2017). According to a GoB Circular from 2007, Union-level WATSAN committees are responsible for a range of WASH activities including supporting and participating in DPHE activities for awareness raising, coordinating the activities of different stakeholders in the WASH sector, implementing WASH projects, and participating in data collection activities for WASH sector (IRF-FSM 2017, p.5). Within urban communities in Bangladesh, slum dwellers are poorly served. Only 13 percent of households in the country's slums have access to their own sanitation facilities (World Bank, 2018). It is common for 10 households or more to share a single sanitation facility (World Bank, 2018). Shared sanitation facilities are considered more unhygienic compared to private ones. As a result, children living in urban slums are more likely to be stunted compared to those living in other urban areas (World Bank, 2018). WHO (2020d) stated that good WASH along with proper waste disposal practices can help to prevent the COVID-19 transmission at the community level. Adequate WASH practices are crucial to protect human health during any infectious disease outbreaks, including the COVID-19 (Chand et al. 2020; Das et al. 2020b; Mushi and Shao 2020). It is reported that the amount of household and infectious medical wastes is increased in many countries including Bangladesh during the pandemic (Rume and Islam 2020; Rahman et al. 2020). Accompanying the growth in the urban population is the growth of populations in slums, and the United Nations estimates that 10 million people in Bangladesh live in slums (UNDP 2020). Since 1990, the country's population has experienced a 20% gain in water access and 29% gain in sanitation access (World Bank, 2018). Based on estimates, 87% of the population drinks from an improved water source, and about 63% uses improved sanitation facilities. Although there has been significant improvement with respect to access to improved water sources, most of the population currently live in environments plagued by inadequate WASH, which leads to negative health outcomes, and the country records over 76 million diarrhoea episodes each year (Zohura et al. 2020). In rural areas, only about 3 percent of households had piped water connections in 2017 (WB, 2020). Also, 41.7 percent of people's water source is contaminated with fecal bacteria and 61 percent is infected with fecal bacteria by the time it is consumed (UNICEF,

2018).

In Pakistan, 53,000 Pakistani children under five die annually from diarrhoea due to poor water and sanitation (WASH, 2020). Children suffering from repeated episodes of diarrhoea are likely to fall behind in school or drop out altogether. Furthermore, it can also cause stunting that currently affects almost 44 percent of children in Pakistan. While there has been a considerable improvement in the statistics over the last decade, 13 per cent of Pakistan's population -- 22 million people -- still practice open defecation; only 60% have access to basic hygiene facilities (WASH, 2019). The lack of proper toilet facilities is a part of 41 million people's lives in Pakistan. According to The United Nations International Emergency Children's Fund (UNICEF), the lack of toilets leaves people with no choice but to practice open defecation, which can lead to the spread of diseases among communities. Pakistan is the third-largest country where people practice open defecation. UNICEF is working with the government to help build toilet facilities for communities that need them to ultimately improve sanitation in Pakistan. These facilities are especially important for girls to protect them against assault, which happens often during open defecation. The World Bank (2018) recommends shifting the emphasis from improving access to sanitation to improving the quality and safety of WASH infrastructure and the safe management of human waste. Poor water safety, poor hygiene and poorly sited and constructed pit latrines, as well as poor drainage, can lead to drinking water contamination and waterborne diseases even in villages with access to improved water and sanitation, as illustrated by two case studies from KP (Nawab, Esser & Baig, 2017). Drinking water is not often treated, for example, 93.8% of household in the Punjab do not use any water treatment method (Government of Punjab, 2015). The alarming state of water supply and sanitation in Pakistan is creating major health hazards for the entire population and young children are particularly at risk, says a new World Bank report (2018).

2. METHODOLOGY

2.1. Area of Study

This study concentrates on two countries in the south asia region, Bangladesh and Pakistan. Bangladesh consists of eight administrative divisions: Barishal, Chattogram, Dhaka, Mymensingh, Khulna, Rajshahi, Rangpur, and Sylhet. These divisions allow the country as a whole to be separated into rural and urban areas. Pakistan consists of four provinces of Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan; two regions including Azad Jammu and Kashmir (AJK) and Gilgit Baltistan (GB); Islamabad Capital Territory (ICT); and former Federally Administrated Tribal Areas (FATA).

2.2. Data Source

This study focuses on the socioeconomic and demographic determinants of improved sanitation adaptation in South Asia, mainly in Bangladesh and Pakistan. We obtained cross-sectional data from the recent version of Demographic and Health Survey in Bangladesh and Pakistan conducted by the DHS program of USAID.

The sample for the 2017-18 BDHS is nationally representative and covers the entire population residing in non-institutional dwelling units in the country. The survey used a list of enumeration areas (EAs) of the 2011 Population and Housing Census of the People's Republic of Bangladesh, provided by the Bangladesh Bureau of Statistics (BBS), as a sampling frame (BBS 2011).

The sampling frame used for the 2017-18 PDHS is a complete list of enumeration blocks (EBs) created for the Pakistan Population and Housing Census 2017, which was conducted from March to May 2017. The Pakistan Bureau of Statistics (PBS) supported the sample design of the survey and worked in close coordination with NIPS. The 2017-18 PDHS represents the population of Pakistan including Azad Jammu and Kashmir (AJK) and the former Federally Administrated Tribal Areas (FATA), which were not included in the 2012-13 PDHS. The results of the 2017-18 PDHS are representative at the national level and for the urban and rural areas separately. The survey estimates are also representative for the four provinces of Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan; for two regions including AJK and Gilgit Baltistan (GB); for Islamabad Capital Territory (ICT); and for FATA. In total, there is 13 second-level survey domains.

2.3. Sample design

According to BDHS 2017-18, primary sampling unit (PSU) of the survey is an EA with an average of about 120 households. The survey is based on a two-stage stratified sample of households. In the first stage, 675 EAs were selected with probability proportional to EA size, with 250 EAs in urban areas and 425 in rural areas. In the first stage, the sample was drawn by BBS, following the specifications provided by the DHS team. A complete household listing operation was then carried out in all selected EAs to provide a sampling frame for the second-stage selection of households. In the second stage of sampling, a systematic sample of 30 households on average per EA was selected to provide statistically reliable estimates of key demographic and health variables for the country as a whole, for urban and rural areas separately, and for each of the eight divisions. In accord with this design, 20,250 residential households were selected. Completed interviews were expected from about 20,100 ever-married women age 15–49. In addition, in a subsample of one-fourth of the households (about 7–8 households per EA), all ever-married women age 50 and older, never-married women age 18 and older, and all men age 18 and older were weighed and their height measured. In the same households, blood pressure and blood glucose testing were conducted for all adult men and women age 18 and older. The 2017-18 PDHS followed a stratified two-stage sample design. The stratification was achieved by separating each of the eight regions into urban and rural areas. In total, 16 sampling strata were created. Samples were selected independently in every stratum through a two-stage selection process. Implicit stratification and proportional allocation were achieved at each of the lower administrative levels by sorting the sampling frame within each sampling stratum before sample selection, according to administrative units at different levels, and by using a probability-proportional-to-size selection at the first stage of sampling. The first stage involved selecting sample points (clusters) consisting of EBs. EBs were drawn with a probability proportional to their size, which is the number of households residing in the EB at the time of the census. A total of 580 clusters were selected. The second stage involved systematic sampling of households. A household listing operation was undertaken in all of the selected clusters, and a fixed number of 28 households per cluster was selected with an equal probability systematic selection process, for a total sample size of approximately 16,240 households. The household selection was carried out centrally at the NIPS data processing office. The survey teams only interviewed the pre-selected households. To prevent bias, no replacements and no changes to the pre-selected households were allowed at the implementing stages. Due to non-proportional sample allocation, the sample was not self-weighting. Weighting factors have been calculated, added to the data file, and applied so that results are representative at the national level for Pakistan (including FATA and ICT Islamabad) and separately for Azad Jammu and Kashmir and Gilgit Baltistan. The 2017-18 PDHS included all ever-married women age 15-49. Those who were either permanent residents of the selected households or visitors who stayed in the households the night before the survey were eligible to be interviewed. The survey of men was conducted in one-third of the sample households, and all ever-married men age 15-49 in these households were included. The survey was successfully carried out in 561 clusters, after dropping 19 clusters due to security concerns during the fieldwork. These clusters were in Balochistan (1), FATA (2), Gilgit Baltistan (6), Khyber Pakhtunkhwa (4), Azad Jammu and Kashmir (1), Punjab (2), Sindh (1) and ICT Islamabad (2 restricted areas).

2.4. Definition of Variables

DEPENDENT VARIABLE

This study hypothesized the status of sanitation into three categorically distributed dependent variables which are “improved sanitation facility”, “unimproved sanitation facility” and “no sanitation facility”. According to the WHO, these variables are supported to categorize the situation of sanitation by Joint Monitoring Program categorization of toilet facilities. The “improved sanitation facilities” includes Flush or pour flushed toilets connected to sewers, Flush or pour flushed toilets or latrines connected to pits or septic tanks, Ventilated improved pit (VIP) latrines, Pit latrines with

slab, Composting toilets including twin-pit latrines and container-based systems. The “unimproved sanitation facilities” includes Pit latrines without slab, Bucket latrines and Hanging latrines. And the “no facility” category includes open defecation (bush/field/desert/road/beach etc.).

EXPLANATORY VARIABLES

i) Wealth Index of the household

The economic situation in the RDHS was assessed using the wealth index. The index was established based on asset holding by the household (such as land, radio, television, car, etc.) and housing features (energy used, the number of rooms, water source, and toilet type). The wealth index was made applying principal components analysis, which is a complex quantity of the snowballing living condition of a family; this places independent families on a constant range of related wealth (Sintondji et al., 2017).

ii) Family size of the household

The total number of members of the household is counted. In the case of having guest, the guest is counted if s/he was living in that household from the day before the interview was conducted.

iii) Gender of head of the household

Gender of head of the household was counted by male=1 and female=2.

iv) Marital status of the respondent

Civil status of the householder is expressed in never in union=1, married=2, living with partner=3, divorced/separated=4, widower=5.

v) Education of household head

Education of head of the household is categorized into no education=1, incomplete primary=2, complete primary=3, incomplete secondary=4, complete secondary=5, higher education=6.

vi) Locality of the household

Locality or regional residence of the household is in either urban=1 or in rural=0.

vii) Province/ division

Bangladesh is divided into 8 geographical divisions and these divisions are exclusively categorized into Barisal=1, Chittagong=2, Dhaka=3, Khulna=4, Mymensing=5, Rajshahi=6, Rangpur=7, Sylhet=8.

Provinces of Pakistan are categorized into Punjab=1, Sindh=2, KPK=3, Balochistan=4, GB=5, ICT=6, AJK=7, FATA=8.

2.5. Model Specification

Ordinal logistic regression is used to model the relationship between an ordered multilevel dependent variable and independent variables. In the modeling, values of the dependent variable have a natural order or ranking. When the response categories are ordered, in ordinal logistic regression model, the event being modeled not only is having an outcome in a particular category but also preserves information about response categories which are ordered. Ordinal logistic regression models, also known as proportional odds models, utilizing proportional odds, have the following general form shown in

$$\ln\left(\frac{\gamma_i^{(j)}}{1 - \gamma_i^{(j)}}\right) = \ln\left(\frac{P(Y_i \leq j | x_1, x_2, \dots, x_p)}{1 - P(Y_i \leq j | x_1, x_2, \dots, x_p)}\right)$$

$$= \tau_j - (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$$

where Y is response variable with k ordered categories;

j= 1,2 ,....., k-1;

$\gamma_i^{(j)}$ is cumulative probability $P(Y_i \leq j) = P(Y_i = 1) + P(Y_i = 2) + \dots + P(Y_i = j)$ for

$j=1, 2, \dots, k-1$. Note $\gamma_i^{(k)} = P(Y \leq k) = 1$, so it should not be modeled;

Y_i is dependent observations which are statistically independent $i=1, 2, \dots, n$;

x_1, x_2, \dots, x_p are p explanatory variables;

$\beta_1, \beta_2, \dots, \beta_p$ correspond to the regression coefficients for the respective independent variables;

τ_j is the cut-off points between categories.

An ordinal variable is a variable that is categorical and order, here is toilet facility based sanitation status of respondents. In order logit, an underlying score is estimated as a linear function of the independent variables and a set of cutoff level. The probability of observing outcome corresponds to the probability that the estimated linear function, plus random error, is within the range of the threshold level estimated for the outcome. This study also used the multinomial logistic regression model to determine whether the explanatory variables such as wealth index, gender, education, civil status, size of family members, regional residences (rural and urban) and province were associated with access to upgraded sanitation facilities.

Here, the probability of a certain sanitation status from “improved”, “unimproved” and “no facility” sanitation is fitted in the above models depending on the socioeconomic and demographic variables taken. The explanatory variables have their own sub-categories including dichotomous category and nominal categories.

Multinomial logit models do not consider proportional odds and ignore ordered response categories. For k possible outcomes, running $k-1$ independent binary logistic regression models in which one outcome, say k , are chosen as a baseline category and then the other $k-1$ outcomes are separately regressed against the reference outcome. The general form is followed by the following equation:

$$\ln \left(\frac{P(Y_i=j)}{P(Y_i=k)} \right) = \beta_j x_i; \text{ where } j= 1, 2, 3, \dots, k-1$$

Any category of response variable can be chosen to be the baseline or reference category, the model will fit equally well, achieving the same likelihood and producing the same fitted values, only the values and interpretation of the parameters will change, Schafer (2006). In our situation we used the category with worldwide important category so we selected category of (0-no sanitation facility). This means, the comparison will be against the households whom did not been having improved sanitation facility in the year of 2017-18.

2.6. Model Evaluation and Comparison Methods

To compare models, I consider the Akaike’s Information Criterion (AIC) (Akaike, 1974) and the Bayesian Information Criterion (BIC) (Schwartz, 1978) based on the ML method. AIC, a penalized log likelihood criterion is defined by

$$AIC = -2 \ell + 2K \dots \dots \dots (7)$$

Where ℓ is the log likelihood and k is the number of parameters.

$$SC = -2 \ell + K \ln (n) \dots \dots \dots (8)$$

The restriction of ordinal regression originates from the proportional odds assumption even though ordinal regression takes care of ordinal relationship between levels of the dependent variable (Salman and Salem, 2012). The proportional odds assumption is that β is independent of j . In other words, the effects of independent variables, β_s , are constant between different levels of the dependent variable. The proportional odds assumption can be tested by using a likelihood ratio score test to determine whether allowing the effects of independent variables to change will result in significant improvements in model fitness [16]. If the proportional odds assumption is not met, there are still several options, such as using the partial proportional odds model [20]. Our models meet the proportional odds assumption, possibly because of the large sample size and continuous latent response. The proportional odds cumulative-logit model acts well with its connection to the idea of a

continuous latent response.

3. RESULTS AND DISCUSSION

3.1. Descriptive Statistics of Bangladesh and Pakistan according to BDHS 2017-18 and PKHS 2017-18

Descriptive statistics of the study sample are presented in Table 1 shows the frequency, percentage, cumulative percentage and valid percentage of the dependent and independent nominal and continuous variables of BDHS 2017-18 and PKHS 2017-18 data.

TABLE 1: DESCRIPTIVE STATISTICS OF BDHS2017-18 AND PKHS 2017-18

Variables		Frequency		Percent		Valid Percent		Cumulative Percent	
		BD	PK	BD	PK	BD	PK	BD	PK
Sanitation facility	Improved	12250	11842	62.96	81.44	62.96	81.44	62.96	81.44
	Unimproved	6961	1230	35.78	8.46	35.78	8.46	98.04	89.9
	No facility	246	1467	1.26	10.1	1.26	10.1	100	100
	Total	19457	14540	100	100	100	100		
Wealth Index of the household	Poorest	4077	2837	21	19.5	21	19.5	21	19.5
	Poorer	3839	3104	19.7	21.3	19.7	21.3	40.7	40.9
	Middle	3641	2815	18.7	19.4	18.7	19.4	59.4	60.2
	Richer	3798	2720	19.5	18.7	19.5	18.7	78.9	78.9
	Richest	4102	3064	21.1	21.1	21.1	21.1	100	100
	Total	19457	14540	100	100	100	100		
Gender of head of the household	Male	16464	12858	84.6	88.4	84.6	88.4	84.6	88.4

	Rangpur	2467	1697	12.7	11.7	12.7	11.7	89.3	94.3
	Sylhet	2084	832	10.7	5.7	10.7	5.7	100	100
	Total	19457	14540	100	100	100	100		
Locality of the household	Urban	7103	7272	36.5	50	36.5	50	36.5	50
	Rural	12354	7268	63.5	50	63.5	50	100	100
	Total	19457	14540	100	100	100	100		

In the sample, only 1.26% households of Bangladesh do not have any toilet facility while Pakistan has 10.1% households without any toilet facility.

3.2. Development and Evaluation of Ordered Logistic Regression Model

An ordinal logistic regression model is constructed through an explicit enumeration of all available explanatory variables available and the “best” fitted model is selected based on model selection. The same data and model selection methods are applied for the selection of a multiple linear regression model for the purpose of comparison. In this study, both in sample fitness and external validation will be performed for both regression models and the indicators described in earlier sections will be used to evaluate model quality.

TABLE 2. ORDINAL LOGISTIC REGRESSION PARAMETERS WITH BDHS 2017-18 AND PKHS 2017-18 DATA

Bangladesh	Estimate	Pakistan	Estimate	Bangladesh	Estimate	Pakistan	Estimate
[SANITATION_BD = .00]	-2.126*** (0.166)	[SANITATION_PK = .00]	-1.873*** 0.185	Unmarried	0 ^a	Unmarried	0 ^a
[SANITATION_BD = 1.00]	2.084*** (0.157)	[SANITATIO_PK = 1.00]	-1.141*** (0.184)	Married	0.047 (0.157)	Married	-0.024 (0.145)
FAMILY_SIZE	0.014 (0.009)	FAMILY_SIZE	-0.022*** 0.006	Widowed	0.27** (0.132)	Widowed	-0.07 (0.167)
AGE_HEAD	0.009*** (0.001)	AGE	0.014*** 0.002	Divorced/ Separated	0.006 (0.194)	Divorced/ Separated	0.252 (0.322)
Poorest	0 ^a	Poorest	0 ^a	Male	0 ^a	Male	0 ^a
Poorer	0.719*** (0.047)	Poorer	-0.362*** 0.07	Female	0.122** (0.057)	Female	-0.004 (0.078)
Middle	1.618*** (0.052)	Middle	-0.487*** (0.076)	Barisal	0 ^a	Punjab	0 ^a
Richer	2.204*** (0.059)	Richer	-0.277** 0.085	Chittagong	0.002 (0.068)	Sindh	0.209** (0.067)

Richest	2.396*** (0.069)	Richest	-0.129 0.095	Dhaka	-0.503*** (0.067)	KPK	-0.07 (0.066)
No Education	0 ^a	No Education	0 ^a	Khulna	0.432*** (0.071)	Balochistan	0.125 (0.079)
Incomplete Primary	0.239*** (0.045)	Incomplete Primary	-0.067 0.101	Mymensingh	-0.187** (0.068)	GB	0.504*** (0.101)
Complete Primary	0.396*** (0.058)	Complete Primary	-0.122 0.079	Rajshahi	0.233** (0.068)	ICT	0.955*** (0.104)
Incomplete Secondary	0.482*** (0.050)	Incomplete Secondary	-0.188** 0.07	Rangpur	0.141** (0.067)	AJK	1.682*** (0.116)
Complete Secondary	0.554** (0.096)	Complete Secondary	-0.031 0.074	Sylhet	-0.072 (0.071)	FATA	-0.432*** (0.088)
Higher	1.031** (0.072)	Higher	-0.068 0.073	Urban	0 ^a	Urban	0 ^a
Don't Know	0.587 (0.554)	Don't Know	16.511*** (0.00)	Rural	0.406*** (0.041)	Rural	-0.145** (0.052)

Standard error in the parentheses. Asterisks ** and *** indicate statistical significance at the 5% and 1% levels, respectively.

The different measures of sanitation status are regressed for explanatory variables in which added in order of demographic factors and traditional SES. There are strong associations between age, gender, household size, and wealth index and sanitation status of respondents, which controls for demographic factors in Bangladesh and Pakistan.

3.3. Development and Evaluation of Multinomial Logistic Regression Model

TABLE 3. MULTINOMIAL LOGISTIC REGRESSION PARAMETERS WITH BDHS 2017-18 AND PKHS 2017-18 DATA

Bangladesh	Estimate	Pakistan	Estimate	Bangladesh	Estimate	Pakistan	Estimate
[SANITATION_BD = .00]	28.55733*** (0.930)	[SANITATION_PK = .00]	0.057*** (0.925)	Unmarried	0 ^a	Unmarried	0 ^a
[SANITATION_BD = 1.00]	29.92547*** (0.926)	[SANITATIO_PK = 1.00]	11.61777*** (0.982)	Married	0.047 (0.609)	Married	-.832 (0.095)
FAMILY_SIZE	0.014 (0.00)	FAMILY_SIZE	.044*** (0.000)	Widowed	0.27** (0.627)	Widowed	-.737 (0.114)

AGE_HEAD	0.009*** (0.600)	AGE	0.003*** (0.245)	Divorced/ Separated	0.006 (0.403)	Divorced/ Separated	-0.746 (0.117)
Poorest	0 ^a	Poorest	0 ^a	Male	0 ^a	Male	0 ^a
Poorer	0.719*** (0.942)	Poorer	-.158*** (0.100)	Female	0.122** (0.062)	Female	-0.004 (0.078)
Middle	1.618*** (0.946)	Middle	-.399*** (0.000)	Barisal	0 ^a	Punjab	0 ^a
Richer	2.204*** (0.947)	Richer	-.163** (0.133)	Chittagong	0.002 (0.297)	Sindh	0.209*** (0.067)
Richest	2.396*** (0.995)	Richest	.503 (0.00)	Dhaka	-0.503*** (0.604)	KPK	1.027*** (0.000)
No Education	0 ^a	No Education	0 ^a	Khulna	0.432*** (0.356)	Balochistan	0.420*** (0.000)
Incomplete Primary	0.239*** (0.961)	Incomplete Primary	0.374 (0.536)	Mymensingh	-0.187** (0.251)	GB	0.603*** (0.000)
Complete Primary	0.396*** (0.962)	Complete Primary	-.378 (0.657)	Khulna	0.432*** (0.356)	Balochistan	0.420*** (0.000)
Incomplete Secondary	0.482*** (0.962)	Incomplete Secondary	-.279** (0.341)	Rajshahi	0.233** (0.041)	ICT	0.839*** (0.000)
Complete Secondary	0.554** (0.096)	Complete Secondary	-0.031 (0.580)	Rangpur	0.141** (0.004)	AJK	1.459*** (0.000)
Higher	1.031** (0.995)	Higher	-.341 (0.575)	Sylhet	-0.072 (0.000)	FATA	2.911*** (0.000)
Don't Know	0.587 (0.554)	Don't Know	16.511*** (0.00)	Urban	0 ^a	Urban	0 ^a
				Rural	0.406*** (0.000)	Rural	0.250*** (0.000)

The different measures of sanitation status are regressed for explanatory variables in which added in order of demographic factors and traditional SES. There are strong associations between age, gender, household size, and wealth index and sanitation status of respondents, which controls for demographic factors for both Bangladesh and Pakistan.

3.4. Model Comparison Statistics with BDHS 2017-18 and PKHS 2017-18 Data

Two candidate models were constructed for predicting no facility, unimproved sanitation and improved sanitation respectively, with BDHS 2017-18 and PKHS 2017-18 data. To illustrate the model performance, the models were first evaluated and compared by an in-sample validation method with previously introduced measurements in Table 4.

TABLE 4. MODEL COMPARISON STATISTICS WITH BDHS 2017-18 DATA.

Model Statistic	Multinomial Logistic Regression Model			Ordered Logistic Regression Model		
	Improved	Unimproved	No facilities	Improved	Unimproved	No facilities
Sample size	12250	6961	246	12250	6961	246
Sum of absolute residuals	1,656	1,906	2,380	1,433	1,494	2991
Sum of residual squares	2,446	2,722	4,090	1,933	1,986	2,200
Bias	1.087	1.090	1.128	1.026	1.023	3,580
Accuracy	1.117	1.112	1.180	1.103	1.081	1.078
AIC	25188.152			25188.152		
BIC	25203.904			25203.904		
LL	25184.152			25184.152		

The AIC, BIC, log likelihood statistics suggest that the highest education of the family members is more influential factor for averting status of a household in Bangladesh.

TABLE 5. MODEL COMPARISON STATISTICS WITH PKHS 2017-18 DATA.

Model Statistic	Multinomial Logistic Regression Model			Ordered Logistic Regression Model		
	Improved	Unimproved	No facilities	Improved	Unimproved	No facilities
Sample size						
Sum of absolute residuals	1568	1832	2316	1360	1480	2207
Sum of residual squares	2328	2642	4072	1830	1950	3677
Bias	1.096	1.099	1.147	1.023	1.022	1.097

Accuracy	1.126	1.119	1.202	1.099	1.081	1.2
AIC	16407.804			16407.804		
BIC	16422.972			16422.972		
LL	15610.669			15610.669		

The AIC, BIC, log likelihood statistics suggest that the highest education of the family members is more influential factor for averting status of a household in Pakistan.

The analysis shows agreement among all indicators, for all three component models, and for two data sets. There is no clear trend for model performance improvement by year. All the comparison results indicate the clear improvement of the ordinal logistic model over the multinomial regression model. Ordinal logistic models will not predict out-of-range estimations which are not controlled by the multinomial regression model.

4. CONCLUSION

This paper proposes and demonstrates an ordinal logistic regression model for predicting sanitation status in Bangladesh and Pakistan. The model is preferred for its ability to handle the ordinal nature of sanitation component judgments, its explanatory power of the regression analysis, and its accurate prediction power. In this study, both ordinal logistic regression and multinomial regression models have been generated for predicting three main sanitation component ratings. The multinomial logistic model demonstrated in this research can be easily applied with element-level data when it becomes available. In addition to assessing model performance, both in-sample and external validation analysis were performed for all eight evaluation criteria. Finally, it is determined that the ordinal logistic regression method is a better approach than the multiple linear regression method for forecasting sanitation status. It has the inherent advantage of always making meaningful predictions and its predictions are closer to the observations.

5. REFERENCES

- [1] Akaike, Hirotugu. "A new look at the statistical model identification." *IEEE transactions on automatic control* 19.6 (1974): 716-723.
- [2] Chand, Sudham, et al. "COVID-19-associated critical illness—report of the first 300 patients admitted to intensive care units at a New York City Medical Center." *Journal of intensive care medicine* 35.10 (2020): 963-970.
- [3] Centre for Research and Information (CRI) (2017) Water and Sanitation in Bangladesh: A Clean Future. Available at: <http://cri.org.bd/publication/water-&-sanitation-16/#p=1>.
- [4] Das, Manob, Arijit Das, and Ashis Mandal. "Examining the impact of lockdown (due to COVID-19) on Domestic Violence (DV): An evidences from India." *Asian journal of psychiatry* 54 (2020): 102335.
- [5] Islam, SK Shaheenur, et al. "Bovine tuberculosis prevalence and risk factors in selected districts of Bangladesh." *PloS one* 15.11 (2020): e0241717.
- [6] Nawab, Bahadar, Kjell B. Esser, and Shams Ali Baig. "Impact of pit latrines on drinking water contaminations in Khyber Pakhtunkhwa, Pakistan." *Environmental Forensics* 18.4 (2017): 296-306.
- [7] Rahman, Md Arafatur, et al. "Data-driven dynamic clustering framework for mitigating the adverse economic impact of Covid-19 lockdown practices." *Sustainable cities and society* 62 (2020): 102372.

- [8] Sintondji, Luc O., et al. "Socio-demographic characteristics of households as determinants of access to water, hygiene and sanitation in So-Ava, Benin." *Journal of Environmental Science and Public Health* 1.4 (2017): 253-267.
- [9] Schwartz, Alan, and Louis L. Wilde. "Intervening in markets on the basis of imperfect information: A legal and economic analysis." *U. Pa. L. Rev.* 127 (1978): 630.
- [10] Sarker, M. M. R., and Yamashita H., 2019. A spatial panel modeling of water and sanitation insecurity and policy implications in South Asia: an evidence from Bangladesh. In Y. N., Araral E. and K. E. Seetha Ram, eds. *Water Insecurity and Sanitation in Asia*. <https://www.adb.org/sites/default/files/publication/544131/adbi-water-insecurity-and-sanitation-asia.pdf#page=214>. Asian Development Bank Institute, Tokyo. pp. 199-220.
- [11] SAS Institute, *PROC LOGISITC Proportional Odds Test and Fitting A Partial Proportional Odds Model*, 2015, <http://support.sas.com/kb/22/954.html>. 2015.
- [12] United Nations. 2019. "Sustainable Development Goal 8: Promote Sustained, Inclusive and Sustainable Economic Growth, Full and Productive Employment and Decent Work for all." <https://sustainabledevelopment.un.org/sdg8>.
- [13] WASH in HCF (2017) Outcomes of National Workshop on WASH in Health Care Facilities in Bangladesh: Call for Joint Action. Workshop Report. 13 December 2017. Available at: https://www.washinhcf.org/documents/BangladeshWinHCFWorkshop_Dec2017.pdf
- [14] WHO/UNICEF JMP. Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines. WHO/UNICEF Joint Monitoring Program. Geneva: World Health Organization and UNICEF.
- [15] WHO/UNICEF. 2017. "Bangladesh- Population Using Sewer, Septic and Latrines (2017)." <https://washdata.org/data/household#!/dashboard>.
- [16] WHO/UNICEF. 2019. "Progress on Household Drinking Water, Sanitation and Hygiene 2000–2017." New York. https://www.who.int/water_sanitation_health/publications/jmp-2019-full-report.pdf?ua=1.
- [17] WHO/UNICEF JMP (2018) Bangladesh Sanitation Data. Data available at: <http://washdata.org>.
- [18] World Bank. 2018. "Precarious Progress: A Diagnostic of Water Supply, Sanitation, Hygiene, and Poverty in Bangladesh." WASH Poverty Diagnostic, Washington DC.
- [19] World Bank. 2019. Health, Safety and Dignity of Sanitation Workers: An Initial Assessment. Washington DC: World Bank. doi:10.1596/32640.
- [20] World Health Organization (2018) Guidelines on sanitation and health. Geneva. Available at: <http://apps.who.int/iris/bitstream/handle/10665/274939/9789241514705-eng.pdf?ua=1>
- [21] WSUP. 2016. "Dignity for Pit and Septic Tank Emptiers: WSUP Co-Convenes Dhaka Conference to Highlight Need for Improved Working Conditions." <https://www.wsup.com/blog/dignity-for-pitand-septic-tank-emptiers/>
- [22] Wash, O., et al. "Middle East respiratory syndrome coronavirus (MERS-CoV)." (2019).
- [23] Zohura, Fatema, et al. "Effect of a water, sanitation and hygiene program on handwashing with soap among household members of diarrhoea patients in healthcare facilities in Bangladesh: a cluster-randomised controlled trial of the CHOBI7 mobile health program." *Tropical Medicine & International Health* 25.8 (2020): 1008-1015.