

Supporting Information for

Socio-hydrological Modelling to Assess Reliability of Urban Water System under Formal-Informal Supply Dynamics

Rakhshinda Bano^{1,2}, Mehdi Khiadani¹ and Steven Burian³

^{1.} School of Engineering, Edith Cowan University, 270 Joondalup Drive, Joondalup, Perth, WA 6027, Australia

^{2.} US Pakistan Centre for Advanced Studies in Water, Mehran University of Engineering and Technology, Jamshoro, 76062, Sindh Pakistan.

^{3.} Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, UT 84112, USA.

Contents of this file :

Figures S1 to S8

Text S1

Tables S1

Introduction

The document provides the System Dynamics (SD) model sectors representing the relationships between input and output variables and preliminary data analysis completed to obtain the logistic regression coefficients to be used in the SD model. Other input data being used is primarily obtained from secondary sources including government department websites <http://www.pbs.gov.pk/content/pakistan-social-and-living-standards-measurement>, published articles, theses and reports and grey literature and newspapers. Data on current supply, distribution proportion to different sectors, losses and budget were directly obtained from WASA Hyderabad in hard copies. Additional dataset used in the study are available at Edith Cowan University's data repository and can be accessed at <http://dx.doi.org/10.25958/5efd3803d64fa>.

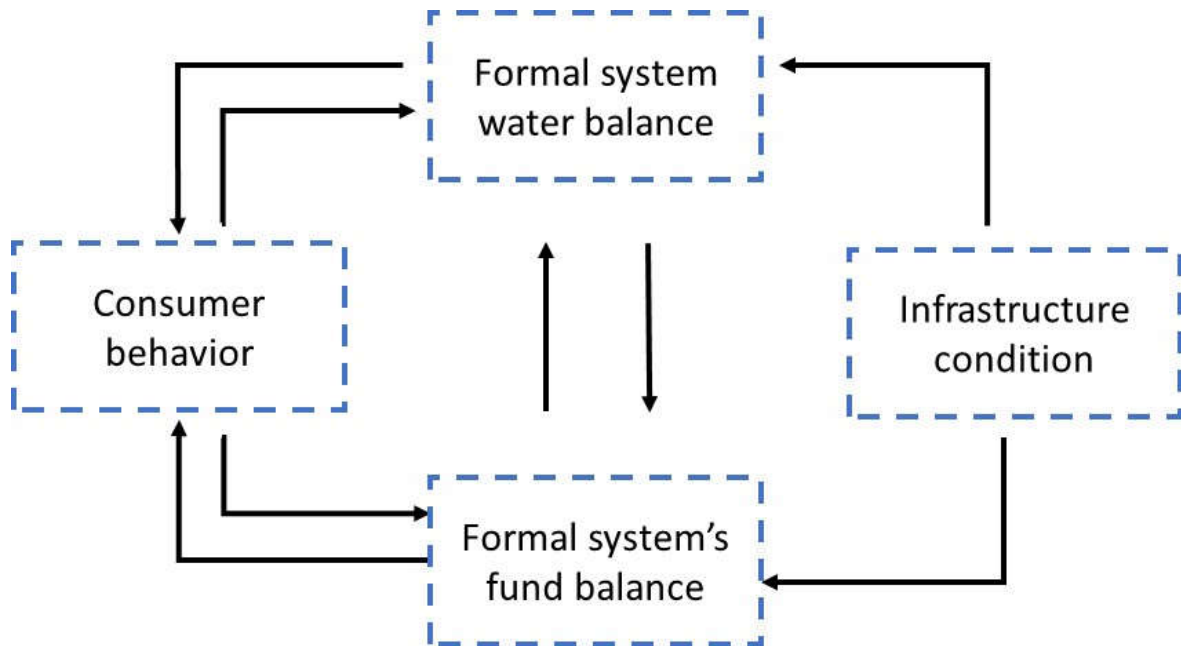


Figure S1. System components of the socio-hydrological processes captured in the study.

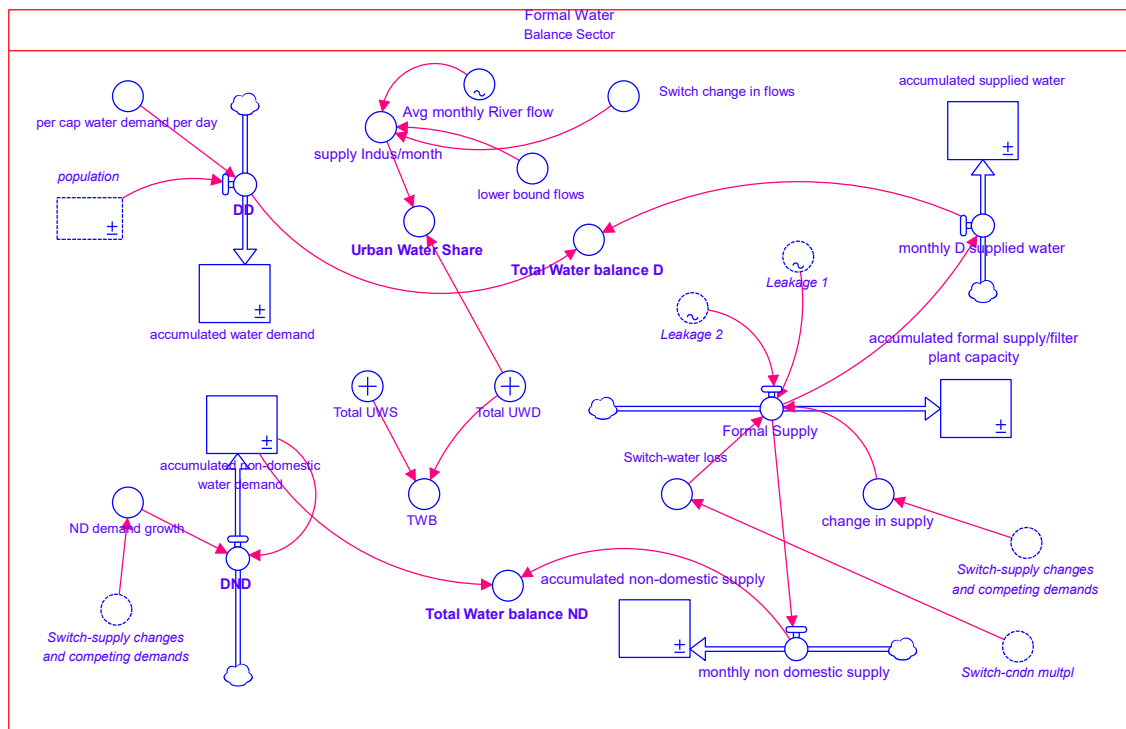


Figure S2. Formal water balance sector: Domestic and non- domestic demand sums up to make total water demand (TWD).

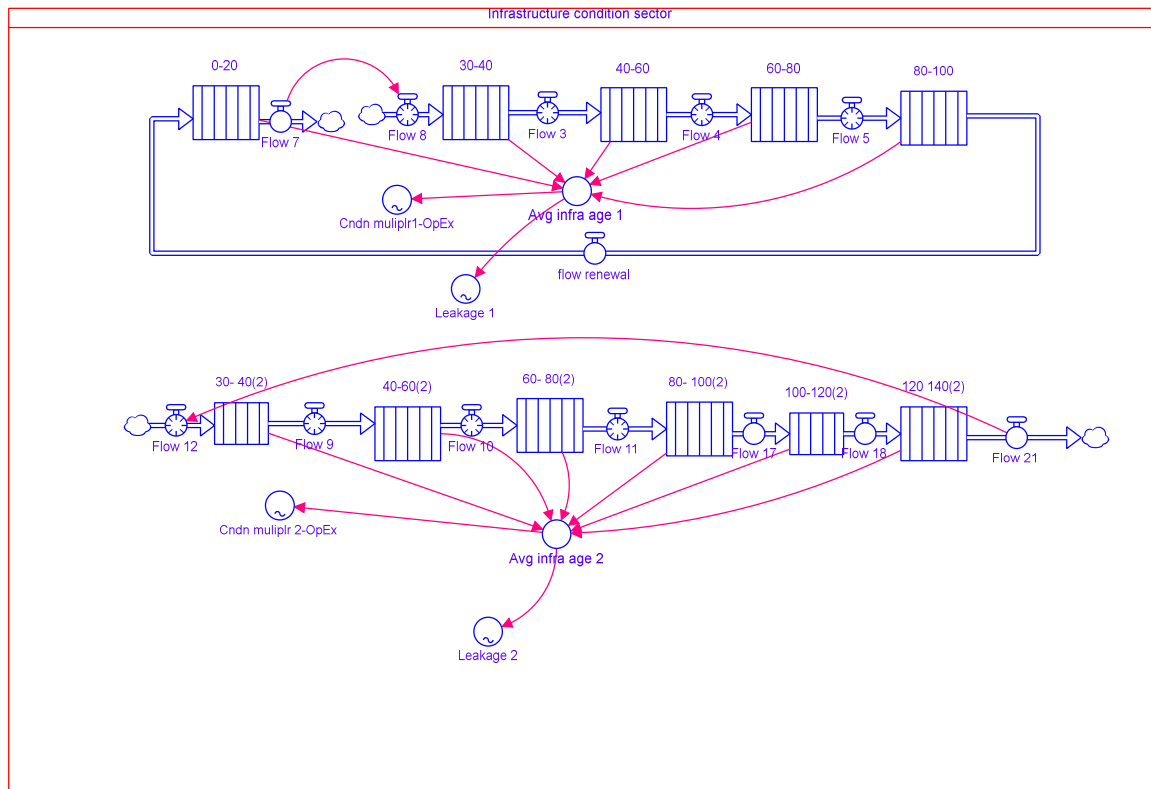


Figure S3. Infrastructure condition sector.

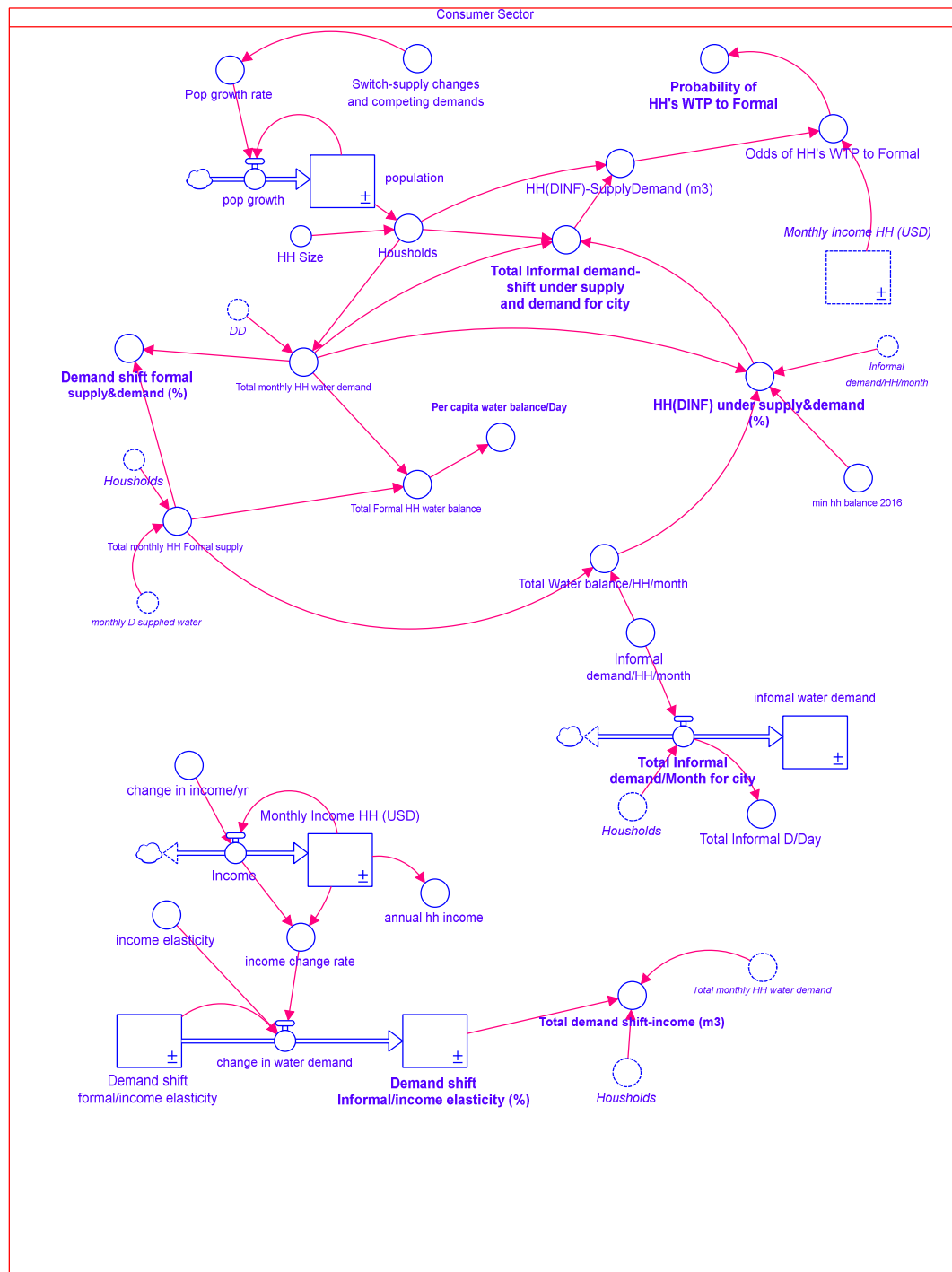


Figure S4. Consumer sector.

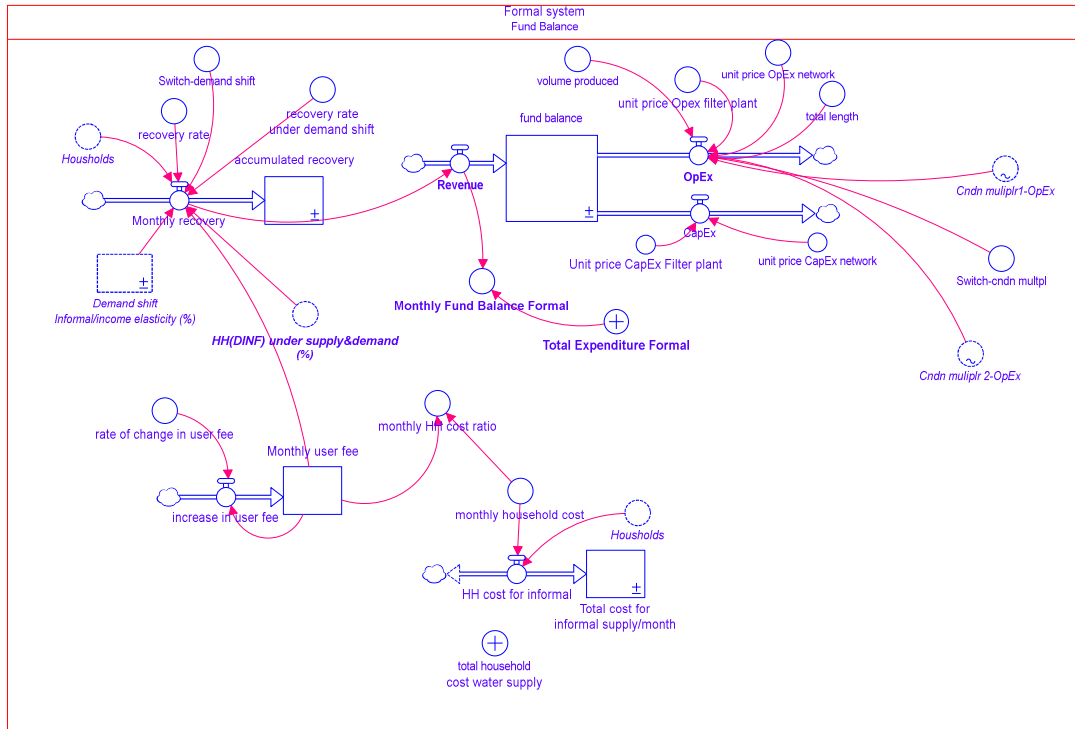


Figure S5. Financial balance sector.

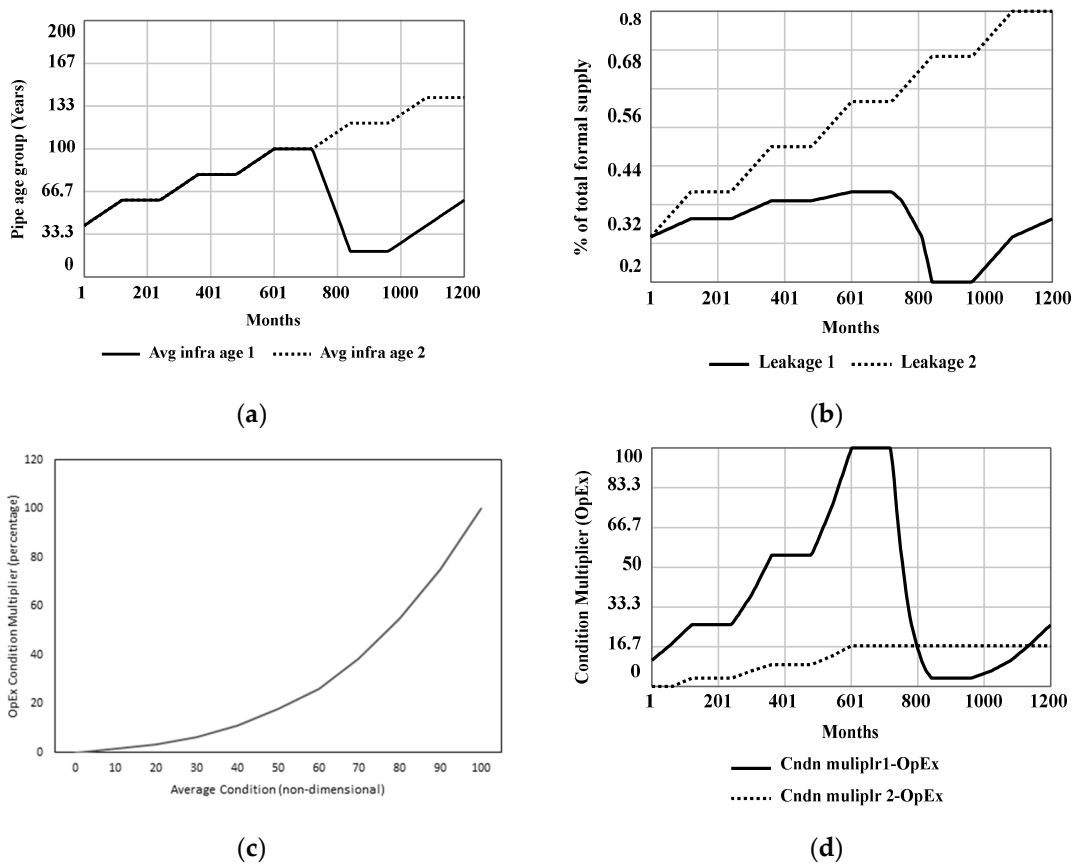


Figure S6. Infrastructure ageing: Infrastructure condition is presented by the (a) infrastructure ageing and (b) associated water losses. (c) Condition multiplier OpEx is used to inflate the unit price OpEx

for the network resulting in an (d) operational expenditure under two scenarios tested in the model. OpEx will be lower under low infrastructure maintenance, with no funds available and resulting in degradation of the infrastructure, represented with increasing water losses upto 70% in the model. This scenario represents the current condition of the urban utilities in developing countries as reported by Kalhoro (2017) for the city of Hyderabad.

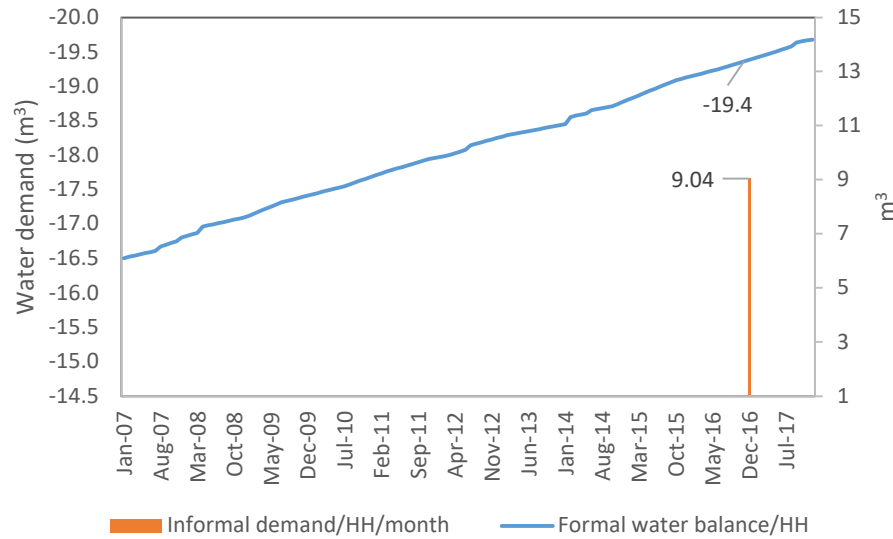


Figure S7. Informal demand versus formal supply balance per household. Based on the assumption of informal demand to be residual, the estimated for the year 2016 by the model is 19.4m³ while it was observed to 9.04 m³ by (Imad, 2017).

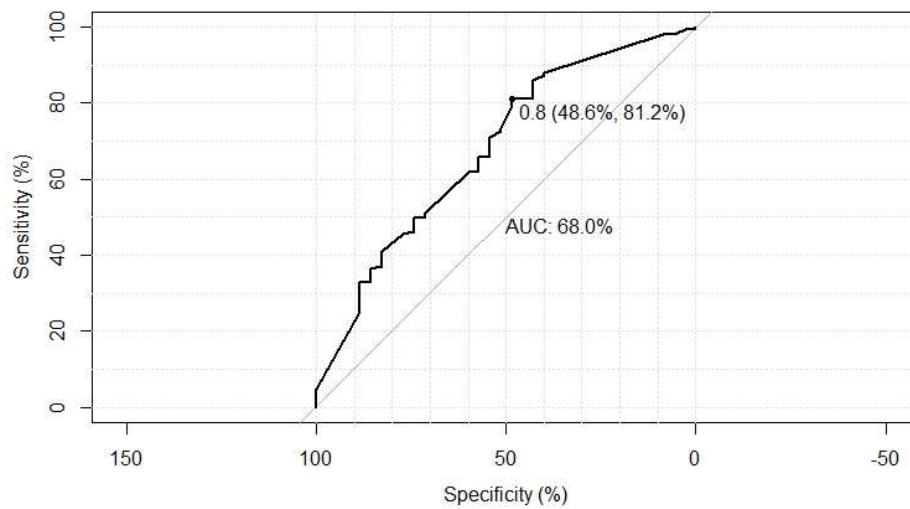


Figure S8. The Area under the ROC Curve - Logistic Regression analysis. The Area under the ROC Curve (AUC) evaluates how well the logistic regression model has classified the positive and negative outcomes at all possible cutoff values. The area under the curve (AUC) is observed to 68%, which means model's predictive performance is average. The higher the value, the better the model's predictive performance is. If the value is 0.5 or less then the performance is considered poor.

Text S 1: Informal demand and WTP to formal (Logistic Regression Analysis)

Logistic regression analysis was used to measure the relationship between consumers' WTP or water fee recovery rate as response variable and informal demand at household level and HH income

as predictor variables. Given the binary nature of the response variable, logistic regression was chosen over linear regression. Data analysis involved filtering the data for missing values. Out of 380 total sample size, 368 responses were finalized for the analysis followed by data splitting into training and testing with a split ratio of 75% and 25% respectively. The logistic regression coefficients mentioned in **Error! Reference source not found.** show that informal demand is negatively affecting the response variable. Though the relationship is not significant with a p value of 0.52. On the other hand, income is positively affecting the response variable with a p value of 0.00095. model's validity is measured through area under ROC (as mention in **Error! Reference source not found.**), the model's predictive performance is also average which shows that though informal demand determines a decline in recovery rate, the responses obtained are not enough to display a strong relationship.

Table S1. Logistic Regression Coefficients.

	<i>Estimate</i>	<i>Std.Error</i>	<i>Z value</i>	<i>Pr(> z)</i>
(Intercept)	0.271815	0.3579386	0.759	0.447619
Informal Demand (Litres)	-0.00022	0.0003511	-0.63	0.528823
Income 2	1.82584	0.5527355	3.303	0.000956***
Income 3	1.480839	0.525791	2.816	0.004856**
Income 4	1.827286	0.6809808	2.683	0.007290**

Table S2: Calibrated input data values

Year	Population (GR-1.93%)	Non-domestic water demand (Gallons) (GR 1.5%)
1998	1204434	-
1999	1227680	-
2000	1251374	-
2001	1275525	-
2002	1300143	-
2003	1325236	-
2004	1350813	-
2005	1376883	-
2006	1403457	-
2007	1430544	37885887
2008	1458154	38454175
2009	1486296	39030988
2010	1514981	39616452
2011	1544221	40210699
2012	1574024	40813860
2013	1604403	41426068
2014	1635368	42047459
2015	1666930	42678171
2016	1699102	43318343
2017	1731895	43968118

References

- Imad, H.U. *Consumer's Willingness to Pay for Municipally Supplied Water: A Case Study of Hyderabad City*; Mehran University of Engineering and Technology: Jamshoro, Pakistan, 2017.
- Kalhor, M.I. *Report of Commission of Inquiry*; Supreme Court of Pakistan: Islamabad, Pakistan, 2017; p. 157.