

Research Paper

A decision support tool for rural water supply planning

Joseph Cook, Jake Wagner and Gunnar Newell

ABSTRACT

Over a dozen studies have examined how households who travel to collect water (about one-quarter of humanity) make choices about where and how much to collect. There is little evidence, however, that these studies have informed rural water supply planning in anything but a qualitative way. In this paper, we describe a new web-based decision support tool that planners or community members can use to simulate scenarios such as (1) price, quality, or placement changes of existing sources, (2) the closure of an existing source, or (3) the addition of a new source. We describe the analytical structure of the model and then demonstrate its possibilities using data from a recent study in rural Meru County, Kenya. We discuss some limits of the current model, and encourage readers and practitioners to explore it and suggest ways in which it could be improved or used most effectively.

Key words | decision support, development economics, rural water supply, water supply engineering

Joseph Cook (corresponding author)

Jake Wagner

Gunnar Newell

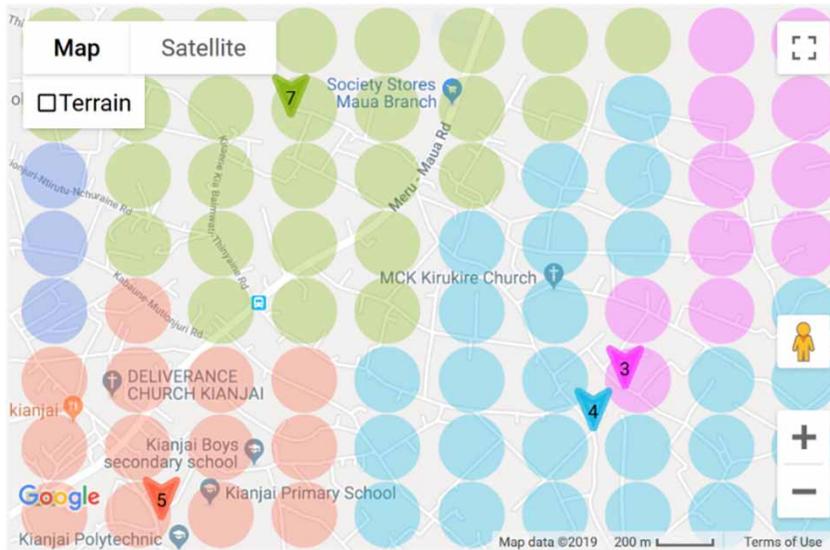
Washington State University,
School of Economic Sciences,
Pullman, WA,
USA

E-mail: joe.cook@wsu.edu

HIGHLIGHTS

- Describes a new open-source rural water decision planning tool.
- Tool focuses on the 'demand' side of simulating household choices and behavior.
- Simulates the impact of changing prices, water quality, or distances to water points.
- Proof of concept shown for a site in rural Meru County, Kenya.
- Tool can inform policy on pricing, maintenance, and placement.

GRAPHICAL ABSTRACT



INTRODUCTION

Approximately one-quarter of households in the world do not have access to improved drinking water services on premises. The issue is concentrated in rural areas: 40% of rural households overall and 75% of rural Sub-Saharan African households travel to collect water (JMP 2019). Continuing four decades of large-scale investments in rural water supply, governments and non-profit organizations continue to invest heavily in improving this situation by giving households access to more and higher quality water sources like protected springs and drilled boreholes. ‘Demand-led’ approaches of the 1990s recommended that communities be given a voice in choosing the right type of technology and be provided with a sense of ownership over the built facilities. Evidence suggests that these policies have improved the sustained functionality of water points (Whittington *et al.* 2009), but problems remain. Approximately one in three handpumps is estimated to be out of service at any one time (Rural Water Supply Network 2013), and user fees and cash on hand continue to be problems (Koehler *et al.* 2015).

These rural households often face a complicated decision in procuring water supply. They may have access

to several sources which vary in quality (e.g. protected deep boreholes vs. polluted surface sources), distance from home, hours of availability, and financial price (either per trip or per month). An urban household with a piped connection must choose how much water to consume from the tap and whether to treat that water, but a rural household makes a three-part decision: where to collect, how much water to collect, and whether the water should be treated.

A large number of empirical studies have examined these types of choices, though the vast majority are in urban or small-town settings where the relevant decision is whether to supplement the unreliable piped supply with non-network water (see Briscoe *et al.* 1981; Mu *et al.* 1990; Persson 2002; Strand & Walker 2005; Larson *et al.* 2006; Nauges & Strand 2007; Basani *et al.* 2008; Cheesman *et al.* 2008; Nauges & Van Den Berg 2009; Boone *et al.* 2011; Kremer *et al.* 2011; Onjala *et al.* 2014; Uwera & Stage 2015; Gross & Elshiewy 2019). Relatively few have examined choices in rural settings (White *et al.* 1972; Briscoe *et al.* 1981; Mu *et al.* 1990; Kremer *et al.* 2011; Gross & Elshiewy 2019; Wagner *et al.* 2019), and only two (Gross & Elshiewy

2019; Wagner *et al.* 2019) have examined rural households' decisions of *how much* water to collect, what economists call 'water demand'. A large number of studies have used stated preference methods (i.e. contingent valuation) to estimate households' willingness to pay for improvements in access or quality (see vanHoutven *et al.* (2017) for a recent review). Although these studies can be used to estimate the percentage of people who will use a source at a given price, they are generally silent on quantity responses.

Despite this body of research, however, it appears that the results have not affected water supply planning on the ground in anything but a qualitative way. Water supply planners or community members evaluating options are routinely faced with difficult questions that this research could inform. Should we increase the per-container fee to improve cost-recovery? If we do, what fraction of customers will revert back to using surface sources? How much less water will an average household collect? Where should we locate new water points? How many new water points should we install? What will happen to household water use if one of the water points fails? If we convert an open spring to a protected source and improve quality, what fraction of the population will use it?

These studies could be used to make predictions for questions like these, but we are unaware of a prediction tool or decision-support system freely and currently available for rural water planners or communities to use. Hopkins *et al.* (2004) use the results of a contingent valuation study in Rwanda to parameterize two location models (the p-median and location set-covering model) to optimize how many new water points to build and where to place them given a cost-recovery constraint or a minimum-service level (maximum distance) constraint. ESRI's ArcGIS software also includes a 'location-allocation tool' that uses similar approaches to choose optimal locations of facilities, which include water points. Hopkins (2015) builds on this work, applying the optimization model to rural Mozambique and adding a model component to estimate (and optimize) the net social benefits to users by valuing time-savings. In both cases, the model emphasizes mathematical optimization given constraints, though the authors make clear that the model could and should be used as an input into a complex, community-driven planning process.

At the same time, the advent of water point mapping (AkvoFlow, mWater, and Water Point Data Exchange (WPDx)) has dramatically increased the amount of geospatial water point data available to planners. WaterAid's Water Point Mapper tool (waterpointmapper.org) is a good example of how water point data are currently being used to produce maps that can be used for monitoring and planning based on minimum-service criteria. The tool does not allow users to interactively explore how changing parameter assumptions or modeling various scenarios would change various outcomes of interest.

In this paper, we describe a decision support tool for rural water supply that planners or community members can use to simulate scenarios such as (1) price, quality, or placement changes of existing sources, (2) the closure of an existing source, or (3) the addition of a new source. The tool is freely available on the web (ruralwaterdecisions.org) and draws from empirical research on household water source choice and demand studies. We plan to continue refining the tool, and welcome input from possible users for the types of questions and interfaces they would find valuable. The objective of this paper is to describe the basic analytical structure of the model and show how the model might be used in a case study from rural Meru County, Kenya. We close with a discussion of the analytical limits of the current version of the model and a call for more empirical research on rural water demand to inform future versions of the model.

METHODS

Framework

Our model assumes that the decisions of (1) how much water to collect and (2) which sources to collect from happen sequentially but in an inter-related way. Households first make the decision of how much to collect and then decide how to allocate that 'demand' across sources available to them. This approach is called the 'linked demand' model and originated in the literature studying how people make decisions about visiting recreation sites for fishing, boating, and hunting (Bockstael *et al.* 1987). Other studies have modeled these inter-related decisions

by first predicting which source will be chosen, then predicting demand/quantity conditional on the household choosing that source (Nauges & Strand 2007; Gross & Elshiewy 2019). We prefer the linked demand approach because it, similar to the ‘almost ideal demand system’ approach (Deaton & Muellbauer 1980; Coulibaly *et al.* 2014), allows households to use more than one source at a time (Elliott *et al.* 2017).

As a simple example, the model first predicts how much water (in 20 L jerricans) that a household will collect in total over a week, which we will discuss in more detail shortly. Suppose this total weekly demand is 50 jerricans (1,000 L). We then predict how the household will allocate those 50 jerricans across the sources available to them by predicting the probability that a household will use that source, again described more below. Suppose our example household had four publicly constructed boreholes and a river within 3 km walk. Suppose that each of the four boreholes charged a user fee. Based on the distance between the household and each source, the source’s price, and the source’s quality, we predict the probability that the household will collect from that source. Our model might predict that the household has an 80% probability of choosing a borehole which is closest to them and happens to also be the cheapest, a 10% probability of collecting from a second borehole which is somewhat farther away but has a higher quality, and a 10% probability of collecting from the free, poor-quality river. Suppose the other two boreholes are so distant from the household that our model predicts a zero probability of the household using them. Demand at each source is calculated by multiplying total demand (50 jerricans) by the probability of use. We would predict that the household would collect 40 jerricans from the first, closest borehole (50×0.80), 5 jerricans from the other borehole, and 5 jerricans from the river (50×0.10). In practice, we limit households to only collecting from their top three sources (ranked by the indirect utility function). This is a crude way to prevent our model from (unrealistically) predicting positive collection by each household from each source in the study site. By aggregating household demand at each source, we can calculate statistics at the source level: how much water will be collected (and revenue raised) from the first borehole by all households in the vicinity.

Parameters and model calculations

The probability of a household collecting from each source is derived from several empirical studies that rely on ‘random utility’ theory in economics (McFadden 1974) to explain a household’s choice of water source. These studies have generally found that price, distance to source, quality, and reliability are all important determinants (Briscoe *et al.* 1981; Mu *et al.* 1990; Persson 2002; Strand & Walker 2005; Larson *et al.* 2006; Nauges & Strand 2007; Basani *et al.* 2008; Cheesman *et al.* 2008; Nauges & Van Den Berg 2009; Boone *et al.* 2011; Kremer *et al.* 2011; Onjala *et al.* 2014; Uwera & Stage 2015; Gross & Elshiewy 2019). Our model focuses on how three key source attributes affect the probability of choosing a source: price, distance, and quality. Following random utility theory, we assume household i ’s indirect utility of collecting from source j is given by: $V_{ij} = -0.11 \times \text{Price} - 0.52 \times \text{Distance} + 0.1 \times (\text{Good Quality}) - 0.1 \times (\text{Poor Quality})$. The first two attributes are continuous measures, but quality is assumed to be discrete: either ‘poor’, ‘fair’, or ‘good’. Households’ preferences are represented by the coefficients on each attribute. The first two coefficient are negative: households prefer shorter distances and lower prices in choosing a source. The third coefficient is positive and the fourth is negative, indicating that households prefer sources of good quality and dislike sources of poor quality. The magnitude of each coefficient indicates how sensitive households are to that attribute: if the price coefficient increases in magnitude, it indicates that households are more sensitive to price. We leave the interpretation of these quality categories to model users, but in general we would map these categories to the JMP ladder as follows: ‘poor’ quality would include surface and ‘unimproved’ sources (i.e. unprotected wells or springs); ‘good’ quality would include ‘safely managed’ sources and possibly ‘basic’ sources; and ‘fair’ quality would include ‘limited’ and possibly ‘basic’ sources. Although some studies have included the source’s availability or potential for interpersonal conflict as attributes, we omit them here for parsimony to focus on the three most important attributes. We also omit wait times, which we discuss below.

We use base case parameters that are taken from a combination of Wagner *et al.* (2019) and our judgment about the existing literature. The user can, however, adjust the

parameters of the utility function. For example, if she believes households are more sensitive to price increases than our base case assumption, she can change the value from the default price parameter of -0.11 to a higher absolute value, i.e. -0.3 . We implement this in the tool with an input field, where the range is bounded by the high- and low-estimates from the empirical literature (Table 1). Implicitly, we only model collection from public sources (i.e. we do not model the household's decision between using their private source and collecting from a public source). We do, however, adjust demand calculations to reflect the prevalence of private sources in the study site.

There are only two existing studies examining total household water demand among rural, unconnected households (Gross & Elshiewy 2019; Wagner et al. 2019). We follow the latter in modeling household demand as a function of only two parameters: household size and 'choice set quality'. The choice set quality is a parameter linking the source choice and water demand decisions: it is unique for each household and is the sum of the utility obtained from collecting each available source weighted by the probability that the household collects from that source (Hanemann 1982; Creel & Loomis 1992). Intuitively, a household with more high-quality, low-cost sources close to them will have a higher 'choice set quality' and thus collect more water per capita than a household with only one poor-quality source. (Choice set quality is implicitly a function of the attributes of available sources and household preferences over those attributes via the utility function.) We again allow the user to vary this parameter; the bounds are given by the 95% confidence interval found in

Wagner et al. (2019), currently the only study to use this approach. The model assumes an average household size for the entire study site since data on individual household sizes would typically be unavailable to a planner. We use a default household size of 5 members based loosely on rural Kenya. Formally, the household demand equation is given by: total weekly demand (20 L containers) = $0.77 + 12.52 * (\text{choice set quality}) + 7.28 * (\text{household size})$. We assume that increasing household size by one member increases water demand by 7.28 20 L jerricans per week with a household-level intercept of 0.77 jerricans per week (average choice set quality is 2.16). This corresponds to approximately 37 liters per capita per day for a family of 5.

We omit for now variables on income or wealth from the household demand equation. These variables are typically included in models of demand for piped water and were included in the two rural water demand studies cited above. These analyses are typically identified based on variation in income or wealth across households in the same study site, but again we expect that most users would not have detailed household income information. We do not believe that there is currently enough information in the existing literature to know how water demand varies across study sites that vary in average income or wealth levels. The model can easily be adapted should this information become available.

CASE STUDY

We demonstrate the decision-support tool by analyzing a sample dataset from Meru County, Kenya. This data comes from a recent water source choice and demand study (Wagner et al. 2019). The site for that study was chosen purposefully because it had a relatively large number of existing water source options available to households, giving us enough observable variation to estimate household preferences for distance, price, and quality. The sample dataset is available for download at ruralwaterdecisions.org. In what follows, we describe the results of three simulations analyzed using the decision-support tool.

Before discussing the simulations, we present results from a benchmark scenario. Figure 1 displays the set of available sources in our study site, along with color-coded

Table 1 | Base case parameters used

Parameter	Base case value [range]
<i>Source choice (Probability of choosing source)</i>	
Price (per 20 L jerrican)	-0.11 [$-1, 0$]
Distance (km)	-0.52 [$-2, 0$]
Good quality	0.10 [$0, 1$]
Poor quality	-0.10 [$-1, 0$]
<i>Household demand (20 L jerricans per week)</i>	
Household size	7.28
Choice set quality	12.52

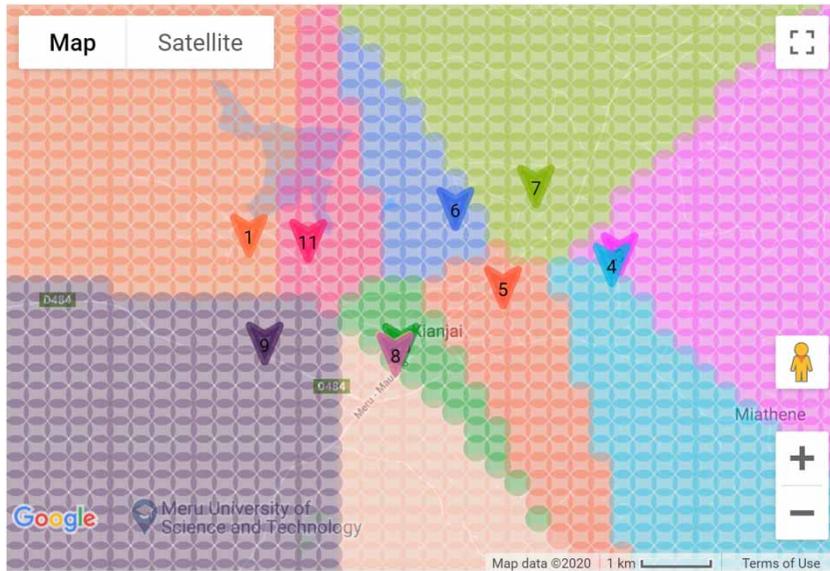


Figure 1 | Benchmark scenario (full study site; dashed box shows the subregion used for simulations).

market segments. Households residing inside the bounds of each market segment are predicted to primarily use the water source of the corresponding color. For example, households located in the light blue-shaded region on the right of the map are predicted to primarily collect water from Source 4. The dashed box shows the region in the northeast of our study site that will be the focus of the following simulations.

The simulations are used to test how changes in source attributes affect both source- and community-level water demand statistics. There are four sources located within the subregion of interest (Figure 2). Each of these sources

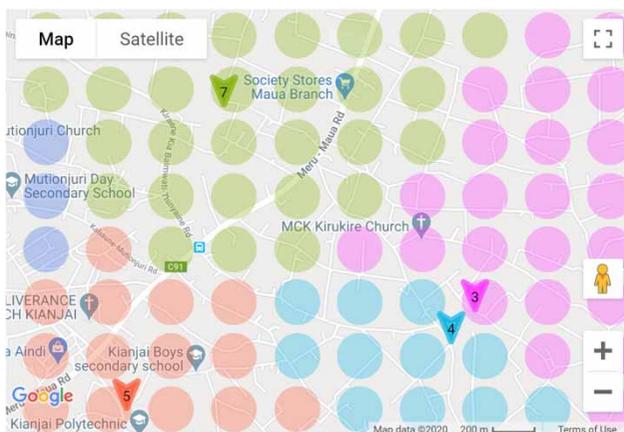


Figure 2 | Benchmark scenario (subregion).

has benchmark attributes (Table 2). In each of the three simulations, we alter the source attributes of one of the sources located in the subregion. Simulation 1 increases the quality at Source 3. Simulation 2 increases the price at Source 7. In Simulation 3, we assume that Source 5 has fallen into disrepair and has been closed.

For each simulation, we generated maps that depict color-coded market segments (Figures 3–5). We also generated a set of demand statistics for the benchmark scenario and each of the simulations (Table 3). For brevity, we will compare the results of each simulation with those of the benchmark case, but cross-simulation comparisons are also insightful.

Simulation A tests the effects of a change in source quality on households' choice of primary source and on demand. Note that when the quality of Source 3 increases from 'poor' to 'good', households that were previously using nearby Source 4 ('fair' quality) switch to using Source 3 (Table 3). While households have stopped using Source 4 as their primary source, they still collect from Source 4, though less frequently. The new market share captured by Source 3 due to its improved quality increases demand from 7,476 to 9,368 jerricans at Source 3.

In Simulation B, we investigate the effect of increasing the price at Source 7 from 2 to 5 Ksh. We see that many households that were previously collecting primarily from

Table 2 | Source attributes – baseline scenario and three simulations

	Attributes	Benchmark		Simulation A	Simulation B	Simulation C
Source 3	Price	1 Ksh		1 Ksh	1 Ksh	1 Ksh
	Quality	Poor	→	Good	Poor	Poor
Source 4	Price	2 Ksh		2 Ksh	2 Ksh	2 Ksh
	Quality	Fair		Fair	Fair	Fair
Source 5	Price	2.5 Ksh		2.5 Ksh	2.5 Ksh	n/a, closed
	Quality	Good		Good	Good	n/a, closed
Source 7	Price	2 Ksh		2 Ksh	5 Ksh	2 Ksh
	Quality	Good		Good	Good	Good

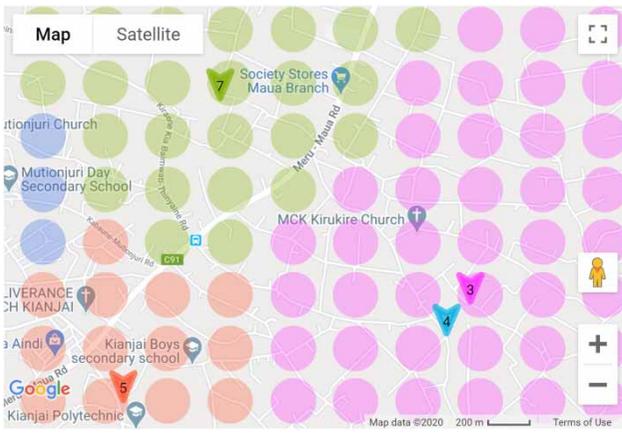


Figure 3 | Improve water quality at Source 3.



Figure 5 | Source 5 (in white) breaks down.

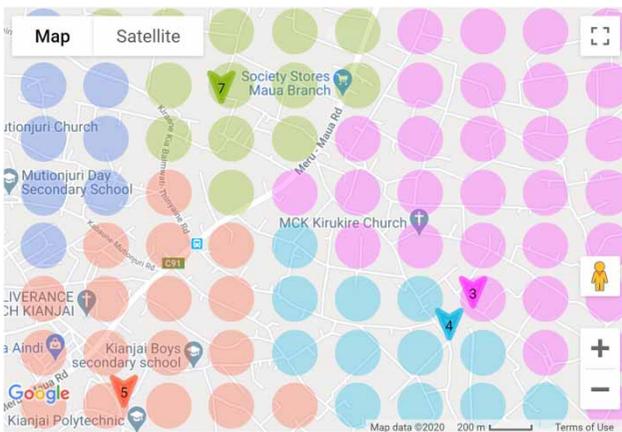


Figure 4 | Increase the price of Source 7.

Source 7 have shifted their primary collection to Sources 4, 5, and 6. Due to the price increase, water collected at Source 7 drops from 6,830 to 4,384 jerricans, but revenues

increase because the fall in demand is offset by the increased price.

Finally, in Simulation C, we consider how the closure of Source 5 affects source choice and demand. We see that when Source 5 closes, Sources 4, 6, and 7 capture most of its demand. Households, however, must now travel much farther to collect water, resulting in the lowest total demand across the three simulations.

DISCUSSION

We begin by acknowledging some of the limitations of the current version of the model which we hope to address in the future. First, households spend time not only in traveling to collect water but also by waiting in queues at the source. The model does not incorporate waiting times, which is

Table 3 | Weekly demand statistics (during the dry season) – baseline and three simulations

		Benchmark	Increase quality at Source 3	Increase price at Source 7	Close Source 5
Study site	Fraction of households using 'good' quality sources	0.32	0.56	0.27	0.26
	Average weekly water demand (20 L jerricans)	58.3	58.5	57.9	57.0
	Average liters per capita per day (L)	33.3	33.4	33.1	32.6
Source 3	Number of primary users (households)	165	342	213	165
	Total demand (20 L jerricans)	7,476	9,368	7,874	7,946
	Total revenue (Ksh)	7,476	9,368	7,874	7,946
Source 4	Number of primary users (households)	127	0	128	158
	Total demand (20 L jerricans)	7,175	6,584	7,455	7,571
	Total revenue (Ksh)	14,350	13,168	14,910	15,143
Source 5	Number of primary users (households)	76	58	78	n/a
	Total demand (20 L jerricans)	5,079	4,620	5,864	n/a
	Total revenue (Ksh)	12,698	11,549	14,659	n/a
Source 7	Number of primary users (households)	173	140	77	181
	Total demand (20 L jerricans)	6,830	6,534	4,384	8,124
	Total revenue (Ksh)	13,660	13,068	21,918	16,249

equivalent to a strong assumption that waiting times are zero at all sources or are equal at all sources (and thus affect the probability the same). Further versions of the model could include the ability for the user to input fixed wait times that vary by source. A more complex dynamic model could allow wait times to increase at a given source as more users are predicted to use that source, feeding back into the decision to use it.

Second, having water delivered to the household by a private vendor is a common feature in many communities, including our study site in Kenya, but is omitted from the model. Similarly, the model currently allows the user to enter the fraction of households with a private, piped connection and then assumes that these households collect water exclusively from their piped connection and do not contribute to demand from sources away from home. Future versions could allow a user to add vending as a water supply option in a community at a fixed price per jerrican and quality. The model would then add vending as a spatially-indeterminate 'source' with zero distance. We could also allow households with piped connections to supplement with sources away from home. Finally, one could use the model from the point of view of water vendors

themselves to predict demand for their services and calculate the optimal location (and quality) to source their water from.

In many places, including some parts of our study site, the groundwater table is shallow enough for households to invest capital costs in digging or drilling wells on their property. This reduces their water collection times and volumetric water prices to zero, and quality can vary depending on the groundwater and the action households take to treat the well water. Incorporating this investment decision would require adding a time dimension to the model beyond the simple monthly time step in the current version. Households would compare the capital costs of investing in the well against the expected benefits over its useful life (years, if not decades), which are in turn a function of the quality of water sources away from home both currently *and* predicted into the future.

Our model is silent on water availability. It is possible that our demand calculations could not be met by available supply, or that aggregate demand over time could lower water tables and dry up some groundwater wells. We are not hydrologists, but would welcome a collaboration to link our demand model with a groundwater model or surface water models (i.e. Soil and Water Assessment Tool

(SWAT) and Variable Infiltration Capacity (VIC)) that could incorporate changes in precipitation (from downscaled climate models) or changes in landscape, soil conservation practices, and/or forest cover.

The companion paper that focused on parameter estimation (Wagner *et al.* 2019) used a bootstrapping approach to simulate confidence intervals around estimated price elasticities of aggregate demand. The decision planning tool discussed here does not include any similar sort of sensitivity analysis, though it could be added in future iterations.

Our tool measure distances using straight line-of-sight, though we know walkers will follow roads and paths to collect water. Where the road network is digitized, it is possible to use spatial tools like ArcGIS or GoogleMaps API to automatically calculate the shortest network route. Unfortunately, in our study site, as in many rural areas, the network that households use consists mainly of dirt roads and paths that are not available digitally. Ho *et al.* (2014) painstakingly digitized this type of network in rural Mozambique and found that although straight-line distances underestimate actual distances by 23%, they are a very good proxy measure. In other words, although straight-line measurements are an underestimate, they are a consistent underestimate such that relative comparisons should be relatively unchanged. We hope approaches that use machine learning to deduce paths from satellite imagery will become operational soon.

Finally, our parameterization of water quality is admittedly simplistic: water can be either 'good', 'fair', or 'poor'. Water quality is a multi-dimensional attribute that includes taste, color/turbidity, and salinity, all of which are perceptible to users. It also includes microbial contamination, which is not readily apparent to users but may be reflected in users' perceptions of the health risks of using that source based on the past experience. Microbial contamination can be measured objectively, but it is users' perceptions of those risks that drive decisions of source choice and demand.

Although researchers have conducted a large number of randomized trials measuring willingness to pay for point-of-use water treatment devices like filters or chlorine tablets (e.g. Ashraf *et al.* 2010), there are many fewer estimates of how *perceived* water quality (Somanathan 2010; Jeuland *et al.* 2015) impacts source choice decisions. In general, we

hope this note illustrates the point that more sophisticated demand-side water supply planning tools could be helpful to the sector, but are at the moment informed by only a handful of studies. More research in this area would be helpful.

We would encourage practitioners who are interested in learning more about the tool to visit ruralwaterdecision-s.org to explore how it works, and to contact us with feedback and recommendations for future model iterations. We believe the model could be useful for policy and practice in a few ways. First, we think it could inform pricing policy that aims to improve funds available for the maintenance of existing water points, whether that maintenance is managed at the village level or at a higher government (i.e. district or province) level. Second, it could help engineers estimate the total water demand at sources. Third, it could shed light on national and global efforts to monitor progress toward Sustainable Development Goals (SDGs) and national standards. Suppose, for example, that a national standard mandates that all households must have an improved water point within 500 m of their home. Suppose that a specific household has an improved water point 400 m away, satisfying the standard. But suppose that because that water point requires a very high user fee, the household chooses to source water from a river 100 m away for free. The household's access situation meets the standard, but they are still consuming unsafe water.

One can imagine practitioners and researchers forming a constructive feedback loop using the model as a core. As a first step, an implementer can use the model in a given location with information only on water source locations and attributes to predict household behavior. This ex-ante prediction could then be compared with water point consumption data (if metered). If the model performs poorly, additional data collection from intercept surveys at water points or short household interviews could provide data for researchers to generate site-specific preference parameter estimates that improve model fit for future use at that location. When such studies are done well and made publicly available (ideally in peer-reviewed outlets), they could be added as parameter value choices in the decision-support tool. This would also add to the public good of improving our ability to accurately forecast rural

water choices and speed progress towards achieving and maintaining the SDGs.

ACKNOWLEDGEMENTS

This work builds on research into how households choose water sources in rural Africa, funded in 2014, by Environment for Development (EfD) with generous support from the Swedish International Development Cooperation Agency (SIDA). We thank Roger Madrigal for helpful comments on an earlier draft.

DATA AVAILABILITY STATEMENT

All relevant data are available from ruralwaterdecisions.org.

REFERENCES

- Ashraf, N., Berry, J. & Shapiro, J. M. 2010 Can higher prices stimulate product use? Evidence from a field experiment in Zambia? *American Economic Review* **100** (5), 2383–2413.
- Basani, M., Isham, J. & Reilly, B. 2008 The determinants of water connection and water consumption: empirical evidence from a Cambodian Household Survey. *World Development* **36** (5), 953–968.
- Bockstael, N. E., Hanemann, W. M. & Kling, C. L. 1987 Estimating the value of water quality improvements in a recreational demand framework. *Water Resources Research* **23** (5), 951–960.
- Boone, C., Glick, P. & Sahn, D. E. 2011 Household water supply choice and time allocated to water collection: evidence from Madagascar. *Journal of Development Studies* **47** (12), 1826–1850.
- Briscoe, J., Chakraborty, M. & Ahmed, S. 1981 How Bengali villagers choose sources of domestic water. *Water Supply and Management* **5**, 165–181.
- Cheesman, J., Bennett, J. & Son, T. V. H. 2008 Estimating household water demand using revealed and contingent behaviors: Evidence from Vietnam. *Water Resour. Res.* **44**, W11428. doi:10.1029/2007WR006265.
- Coulibaly, L., Jakus, P. M. & Keith, J. E. 2014 Modeling water demand when households have multiple sources of water. *Water Resources Research* **50** (7), 6002–6014. <http://doi.wiley.com/10.1002/2013WR015090>.
- Creel, M. & Loomis, J. 1992 Recreation value of water to wetlands in the San Joaquin Valley: linked multinomial logit and count data trip frequency models. *Water Resources Research* **28** (10), 2597–2606.
- Deaton, A. & Muellbauer, J. 1980 An almost ideal demand system. *American Economic Review* **70** (3), 312–326.
- Elliott, M., MacDonald, M. C., Chan, T., Kearton, A., Shields, K. F., Bartram, J. K. & Hadwen, W. L. 2017 Multiple household water sources and their use in remote communities with evidence from Pacific Island Countries. *Water Resources Research* **53**, 9106–9117. <https://doi.org/10.1002/2017WR021047>.
- Gross, E. & Elshiewy, O. 2019 Choice and quantity demand for improved and unimproved public water sources in rural areas: evidence from Benin. *Journal of Rural Studies*. <https://doi.org/10.1016/j.jrurstud.2019.02.021>.
- Hanemann, W. M. 1982 Applied welfare analysis with qualitative response models. Department of Agricultural & Resource Economics, UCB CUDARE Working Papers.
- Ho, J. C., Russel, K. C. & Davis, J. 2014 The challenge of global water access monitoring: evaluating straight-line distance versus self-reported travel time among rural households in Mozambique. *Journal of Water and Health* **12** (1), 173–183.
- Hopkins, O. S. 2015 A regional approach to optimizing the location of rural handpumps. *Journal of Water, Sanitation and Hygiene for Development* **5** (3), 493.
- Hopkins, O. S., Asce, M., Lauria, D. T., Asce, M. & Kolb, A. 2004 Demand-based planning of rural water systems in developing countries. *Journal of Water Resources Planning and Management* **130** (1), 44–52.
- Jeuland, M., Orgill, J., Shaheed, A., Revell, G. & Brown, J. 2015 A matter of good taste: investigating preferences for in-house water treatment in peri-urban communities in Cambodia. *Environment and Development Economics* **21**, 291–317.
- Joint Monitoring Program (WHO and UNICEF) 2019 *Progress on Household Drinking Water, Sanitation and Hygiene 2000–2017: Special Focus on Inequalities*.
- Koehler, J., Thomson, P. & Hope, R. 2015 Pump-priming payments for sustainable water services in rural Africa. *World Development* **74**, 397–411. <http://dx.doi.org/10.1016/j.worlddev.2015.05.020>.
- Kremer, M., Leino, J., Miguel, E. & Zwane, A. P. 2011 Spring cleaning: rural water impacts, valuation, and property rights institutions. *Quarterly Journal of Economics* **126** (1), 145–205.
- Larson, B., Minten, B. & Razafindralambo, R. 2006 Unravelling the linkages between the millennium development goals for poverty, education, access to water and household water use in developing countries: evidence from Madagascar. *Journal of Development Studies* **42** (1), 22–40.
- McFadden, D. 1974 Conditional logit analysis of qualitative choice behavior. In: *Frontiers in Econometrics* (P. Zarembka, ed.). Academic Press, New York, pp. 105–142.
- Mu, X., Whittington, D. & Briscoe, J. 1990 Modelling village water demand: a discrete choice approach. *Water Resources Research* **26** (4), 521–529.

- Nauges, C. & Strand, J. 2007 Estimation of non-tap water demand in Central American cities. *Resource and Energy Economics* **29** (3), 165–182.
- Nauges, C. & Van Den Berg, C. 2009 Demand for piped and non-piped water supply services: evidence from southwest Sri Lanka. *Environmental and Resource Economics* **42** (4), 535–549.
- Onjala, J., Ndiritu, S. W. & Stage, J. 2014 Environment for development risk perception, choice of drinking water, and water treatment evidence from Kenyan towns. *Journal of Water Sanitation and Hygiene for Development* **4** (2), 268–280.
- Persson, T. H. 2002 Household choice of drinking water in the Phillipines. *Asian Economic Journal* **16** (4), 303–316.
- Rural Water Supply Network 2013 *Experiences and Ideas From RWSN's Sustainable Groundwater Community* (November 2012).
- Somanathan, E. 2010 Effects of information on environmental quality in developing countries. *Review of Environmental Economics and Policy* **4** (2), 275–292.
- Strand, J. & Walker, I. 2005 Water markets and demand in Central American cities. *Environment and Development Economics* **10** (3), 313–335.
- Uwera, C. & Stage, J. 2015 Water demand by unconnected households in Rwanda. *Water Economics and Policy* **1** (1), 145002.
- VanHoutven, G. L., Pattanayak, S. K., Usmani, F. & Yang, J. C. 2017 What are households willing to pay for improved water access? Results from a meta-analysis. *Ecological Economics* **136**, 126–135.
- Wagner, J., Cook, J. & Kimuyu, P. 2019 Household demand for water in rural Kenya. *Environmental and Resource Economics* **74** (4), 1563–1584.
- White, G. F., Bradley, D. J. & White, A. U. 1972 *Drawers of Water: Domestic Water Use in East Africa*.
- Whittington, D., Davis, J., Prokopy, L., Komives, K., Thorsten, R., Lukacs, H., Bakalian, A. & Wakeman, W. 2009 How well is demand-driven, community management model for rural water supply systems doing? Evidence from Bolivia, Peru and Ghana. *Water Policy* **11**, 696–718.

First received 27 January 2020; accepted in revised form 26 May 2020. Available online 3 July 2020