Predictors of Drinking Water Boiling and Bottled Water Consumption in Rural China: A Hierarchical Modeling Approach

Alasdair Cohen, Qi Zhang, Qing Luo, Yong Tao, John M. Colford, Jr., and Isha Ray

Department of Environmental Science, Policy and Management, University of California at Berkeley, Berkeley, California United States of America

School of Public Health, University of California at Berkeley, Berkeley, California United States of America

National Center for Rural Water Supply Technical Guidance, Chinese Center for Disease Control and Prevention, Beijing, China

Energy and Resources Group, University of California at Berkeley, Berkeley, California United States of America

Supporting Information

ABSTRACT: Approximately two billion people drink unsafe water. Boiling is the most commonly used household water treatment (HWT) method globally and in China. HWT can make water safer, but sustained adoption is rare and bottled water consumption is growing. To successfully promote HWT, an understanding of associated socioeconomic factors is critical. We collected survey data and water samples from 450 rural households in Guangxi Province, China. Covariates were grouped into blocks to hierarchically construct modified Poisson models and estimate risk ratios (RR) associated with boiling methods, bottled water, and untreated water. Female-headed households were most likely to boil (RR = 1.36, p < 0.01), and among boilers those using electric kettles rather than pots had higher income proxies (e.g., per capita TV ownership RR = 1.42, p < 0.01). Higher-income households with younger, literate, and male heads were more likely to purchase (frequently contaminated) bottled water, or use electric kettles if they boiled. Our findings show that boiling is not an undifferentiated practice, but one with different methods of varying effectiveness, environmental impact, and adoption across socioeconomic strata. Our results can inform programs to promote safer and more efficient boiling using electric kettles, and suggest that if rural China’s economy continues to grow then bottled water use will increase.

1. INTRODUCTION AND OBJECTIVES

Over the last few decades, substantial progress has been made in access to water, sanitation, and hygiene (WaSH) in many low and medium income countries (LMIC). However, approximately 2 billion people still lack access to safe drinking water.1 Diarrheal, typhoid, and paratyphoid deaths attributed to unsafe water have declined by 18% (age-standardized) from 2005 to 2015,2 but only an estimated 16% of 188 countries currently meet the Sustainable Development Goal Target 6.1 of universal access to safe and affordable water.3

The high costs of providing centralized drinking water treatment and distribution in areas with low population densities and/or challenging topography make the lack of access to safe drinking water in LMICs a predominantly rural problem.4 In response, point-of-use household water treatment (HWT) technologies such as filtration, chlorination, and solar or ultraviolet disinfection are often recommended to allow households to treat their drinking water. Yet, despite decades of extensive efforts to develop and promote an assortment of HWT products, achieving the sustained adoption and consistent use of HWT remains extremely challenging.5,6

The most commonly used HWT method by far is boiling, with an estimated 1.2 billion users globally.7–9 Boiling is straightforward to use, does not substantially change the taste of the water, and can provide complete pathogen inactivation (regardless of pathogen types or water turbidity).10

Field studies have repeatedly demonstrated boiling’s effectiveness with regard to microbiological water quality.11–14 Boiling also has significant drawbacks: after it cools boiled water is susceptible to recontamination,15 boiling with biomass often produces household air pollution (HAP), exposure to which causes cardiovascular and respiratory disease,16,17 and the fuels used for boiling can be expensive or time-consuming to collect.18

Many studies have compared the effectiveness of different variants of HWT technologies; for example, chlorination with and without a coagulant, biosand versus ceramic filters, ceramic filters with and without colloidal silver.19–24 Boiling, however, is treated as an undifferentiated HWT practice, rather than as a “technology” with multiple variants. Our 2013–2014 HWT study in rural China,25 where the majority of the population boils drinking water, was the first published research we are aware of that disaggregated boiling methods into boiling with
electric kettles and open pots. Between the two boiling methods, the use of electric kettles was associated with lower counts and concentrations of Thermotolerant Coliforms.

The socio-cultural and behavioral determinants of HWT adoption remain understudied and, therefore, poorly understood. Findings from the limited research on the behavioral and attitudinal factors of HWT adoption indicate that reasons for adoption are highly context-specific and often not based on an understanding of germ theory, perceived health risks, or health knowledge. The technology adoption literature suggests that technologies or methods that are user-friendly, easy to comprehend, easy to demonstrate, and have an advantage over the status quo are more likely to be adopted. For boiling water, electric kettles are far more convenient than metal pots heated over biomass-fueled fires. Electric kettles automatically shut themselves off after the water has boiled, whereas boiling with open pots requires more vigilance from the user. It is also easier to boil smaller quantities of water with an electric kettle, and the risk of recontamination upon cooling is much lower in a closed kettle.

In our initial study, we also found that 34% of households regularly drank bottled water, 40.3% \((n = 139)\) of which were found to be contaminated with TTC. Contamination of bottled, or “packaged”, water is not a problem unique to China. Bottled water is neither considered a form of HWT nor an “improved” drinking water source, but many LMIC, such as Indonesia, India, and Thailand, are experiencing double-digit annual growth rates in bottled water consumption (2010–2015 compound annual growth rate = 12.2%, 10.9%, and 12.5%, respectively). The consumption of bottled water in China is also growing rapidly (at 14.3%), and in 2013 China surpassed the U.S. to become the world’s largest market for bottled water, though its per capita consumption is still below the global average. In effect, bottled water has become a competitor to HWT.

Our first objective in this paper was to understand which types of households continued to drink untreated water, despite the high prevalence of boiling overall. In light of the apparent advantages of boiling with electric kettles, our second aim was to understand why, in areas where electricity access is near-universal and electric kettles are not prohibitively expensive, did some households boil with pots while others used kettles? Or, for policy purposes, in LMIC settings in which the government wished to promote drinking water boiling with electric kettles, which types of households might be most likely to adopt kettles and which would face the greatest barriers to adoption? Finally, in response to the growth in bottled water consumption in rural China (and in other middle-income countries), and taking into account the microbiological contamination present in bottled water (that we had identified), we wished to understand which types of households purchased bottled water, and why.

As far as we are aware, this is the first study in the WaSH sector focused on the demographic, socioeconomic, and behavioral predictors of boiling drinking water and of the adoption of different technologies of boiling. Given the environmental health implications of hundreds of millions of daily water boiling events via the combustion of biomass, and the potential environmental impacts of the rapid growth of the bottled water market, China is a particularly suitable setting for investigating these questions.

2. MATERIALS AND METHODS

2.1. Site Selection, Survey Instruments, and Data Collection. In 2013–2014, we carried out a cross-sectional study in rural areas of the Guangxi Zhuang Autonomous Region (Guangxi Province), one of China’s poorest provinces. In collaboration with the National Center for Rural Water Supply Technical Guidance (NCRWSTG) and the Chinese Center for Disease Control and Prevention (CCDC) in Guangxi Province, we selected two relatively low-income counties and used a population-weighted, multistage, geographically stratified, cross-sectional sampling design to randomly select 15 study villages: eight in County A and seven in County B.

We could not find existing data on the predictors of HWT in rural China, so we cast a wide net for our initial data collection by using the Multidimensional Poverty Assessment Tool (MPAT) household survey. MPAT is a thematic indicator based on field-tested survey questions related to key sectors of basic needs and rural livelihoods. As shown in Supporting Information (SI) Figures S1 and S2, MPAT survey responses are aggregated into subcomponents, which are in turn aggregated into 10 components (e.g., Food & Nutrition Security, Health & Healthcare, Sanitation & Hygiene) using a 10–100 scale, where 100 is the optimal value. A tested Chinese-language version was already available and we created additional survey items (piloted and double-blind translated) to collect additional data on HWT, fuel use, and water-related beliefs and behaviors.

Survey data and drinking water samples were collected from 450 households across the 15 villages (30 households/village) during the summer/rainy season (with follow-up data collection from 120 households during the winter/dry season). Our study was powered to measure the proportion of households boiling their drinking water. To control for clustering, we used data from provincial CCDC officials, as well as pilot data, to estimate the intraclass correlation coefficient for our power calculations. Completed surveys were subjected to a three-stage quality control process. Household drinking water samples were assayed for Thermotolerant Coliforms (TTC), an indicator of fecal contamination. Additional information is reported in Cohen et al.

2.2. Selection of Model Covariates: Water-Related. Rather than conducting statistical tests to identify significant bivariate associations between the dependent variables (DV), i.e., HWT methods, and all potentially associated independent variables (IVs) from our extensive household surveys, the choice of water-related covariates known to be associated with HWT from other studies was determined a priori. In addition to drinking water related covariates often controlled for in WaSH studies, e.g., household size and demographics, drinking water sources), we hypothesized that household income would be a key predictor of HWT use. We were unable to estimate household income directly, and the government-collected income data we had were already aggregated at the village level. We therefore used several proxies for household income and wealth (e.g., per capita TV ownership) as well as proxies for access to services (e.g., time needed to reach the nearest health clinic). These data were also used to assess the robustness of the village-level income data. Based on previous behavior-related HWT research, we also asked respondents how many of their relatives and neighbors likely boiled.
water, and the reasons they did or did not prefer boiled or bottled water.

2.3. Selection of Model Covariates: MPAT-derived. To identify survey questions with a high probability of being associated with, and potentially contributing to, HWT or bottled water use, we used the open-source MPAT Excel Spreadsheet to calculate MPAT subcomponent and component values for each household (using the default, standardized, valuations and weightings; details in SI, page S4). Using these semicontinuous variables, we employed a three-stage process, using forward and backward stepwise logistic regression at each stage, to identify variables strongly associated with the use of boiling or bottled water.

For the first stage, stepwise logistic regression was used with the MPAT component values (as IVs) and a probability threshold of 0.2 (i.e., the minimum p-value for component inclusion in the final model). Backward stepwise logistic regression was also used with a probability threshold of 0.2 for removal from the model. MPAT components that had no association with the boiling or bottled water DVs were not included in the next step. In step two, the MPAT subcomponent results belonging to those MPAT components identified in the first stage were subjected to stepwise logistic regression (forward and backward), but using a more restrictive probability threshold of 0.15. This second step helped identify which MPAT subcomponents contained survey questions significantly associated with the DVs. Lastly, a probability threshold of 0.15 was used (forward and backward stepwise logistic regression again) to identify potentially relevant survey items/questions associated with the MPAT subcomponents identified in step two. The results for each of these three steps, and the associated survey questions, are provided in the SI (Tables S1 and S2). For the survey items so identified, provided there was a potentially viable causal link with the outcomes, covariates were created.

All potential covariates are provided in SI (Table S3), along with a simplified hierarchical conceptual framework (after Genser et al.41) that shows the factors hypothesized to impact HWT and their potential causal interlinkages (SI Figure S3).

2.4. Hierarchical Model Construction and Sensitivity Analyses. To model the association of potential predictors on HWT use, we created three binary DVs: one for boiling (n = 215) versus untreated (n = 75); one comparing boiling with electric kettles (n = 122) to boiling with pots (n = 93); and one to compare bottled water use (n = 157) with boiling (n = 215). A number of survey items were used to cross-validate these HWT classifications (defined as using a given HWT method, or not treating drinking water, most of the time). Bottled water is not usually considered a HWT method, but given the relatively widespread use of bottled water in rural China, and considering that most bottled water users heat their water before consuming it, we considered bottled water as a form of HWT and as an alternative to boiling.

In order to better understand which factors may motivate or otherwise be associated with boiling water practices, bottled water use, or drinking untreated water, we used a hierarchical approach to build our models. Specifically, we used modified Poisson regression with a log link and cluster-robust standard errors (SE) to estimate risk ratios (RR) for potential predictors of HWT use (i.e., we used RR to estimate the likelihood of using a given HWT method associated with hypothesized predictor variables). As discussed elsewhere,25,41 when analyzing cross-sectional data with a generalized linear model for Poisson regression with a log link for binary outcomes, the exponentiated coefficients are RR (rather than incidence-rate ratios); compared to logistic regression, these models are more robust to omitted covariates.

When hierarchical model construction is used in the fields of Education, Psychology, and Public Health, the convention is to start with distal hierarchical blocks (e.g., socioeconomic status), and then iteratively add more proximal blocks.41,46 Following the behavioral literature cited above, we hypothesized that water-related perceptions and water-source characteristics would have a disproportionate impact on the decision to use or not to use HWT. Therefore, for the analyses presented here, we loaded the models with the variables theorized to be most proximal first, and then built out to include the most distal variables last. We first analyzed the covariates in each thematic block in isolation of the other covariates in that block, then together in isolation of the other blocks, and then incrementally built the models by adding thematic blocks, starting with the most proximal theorized predictors. If we found that some variables introduced too much collinearity we removed them from their respective blocks before the full models were finalized (e.g., this was the case for the dummy variables for whether or not respondents believed their neighbors or relatives boiled water). For continuous covariates, we used functional form assessments to help determine the likelihood that transformations were appropriate.

Full models (all blocks) were adjusted by removing covariates that did not contribute sufficiently to the outcomes, as evidenced by small effect sizes and/or wide, nonsignificant, 95% confidence intervals (CIs), and that were also nonsignificant when analyzed in isolation. Some theoretically relevant variables (e.g., household size) were retained as controls even if the associated CIs were large and nonsignificant. For model adjustments within and between blocks, we used likelihood-ratio tests when possible; when the number of available observations for the full and restricted models differed, we used Wald tests instead. If we found that either of the binary covariates for head of the household’s gender or marital status was significantly associated with a given DV, we used interaction terms (see SI Table S3).

Compared to traditional modeling approaches, we believe this approach provides a more responsible and transparent means of conducting exploratory analyses when using observational data. Model diagnostics and sensitivity analyses were also conducted, including comparisons with other models (e.g., multilevel and single-level logit models) and bootstrapping was used to evaluate effect SE sensitivity. For all analyses, missing data were ignored.

2.5. Ethics and Reporting. The study was approved by the Committee for the Protection of Human Subjects at the University of California Berkeley (protocol ID: 2012-05-4368) and by the Ethics Review Board at the NCRWSTG, CCDC; all participants provided consent. Statistical analyses were conducted using STATA (v13.1, StataCorp, College Station, TX). This paper was prepared using the STROBE47 reporting guidelines (see SI, pages S2 and S3, for a completed checklist).

3. RESULTS

3.1. Household Survey and MPAT indicator results. The 2012 reported annual income for County A was RMB 4425 (USD 702), and RMB 6912 (USD 1097) for County B. For those MPAT components for which sufficient data was available, the distributions were relatively normal; in addition,
Table 1. Descriptive Statistics for Model Variables by HWT Method

<table>
<thead>
<tr>
<th>Water and Behavior-Related</th>
<th>boil: electric kettle</th>
<th>boil: pot</th>
<th>bottled</th>
<th>untreated</th>
<th>row mean (n for row)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH believes DWQ is good/very good: %</td>
<td>41.2</td>
<td>28.7</td>
<td>46.7</td>
<td>64.0</td>
<td>44.6 (190)</td>
</tr>
<tr>
<td>HH has improved drinking water source: %</td>
<td>55.7</td>
<td>55.1</td>
<td>39.7</td>
<td>39.2</td>
<td>47.6 (211)</td>
</tr>
<tr>
<td>HH believes most/all nearby relatives boil: %</td>
<td>55.3</td>
<td>75.8</td>
<td>44.2</td>
<td>6.5</td>
<td>50.1 (125)</td>
</tr>
<tr>
<td>HH believes most/all neighbors boil: %</td>
<td>54.6</td>
<td>75.4</td>
<td>45.3</td>
<td>9.1</td>
<td>50.8 (125)</td>
</tr>
</tbody>
</table>

Access to Health Services
- minutes to clinic for basic care: mean 12.4, 15.6, 9.1, 8.8, 11.3 (440)
- minutes to clinic for advanced care: mean 28.0, 30.9, 22.7, 23.8, 26.0 (443)
- HH can afford professional care: % 51.6, 43.0, 60.5, 69.3, 55.9 (250)

Economic Indicators
- number of TVs in HH/HH population: mean 0.60, 0.56, 0.55, 0.71, 0.60 (440)
- village-average price for 19L W bottle: mean 7.68, 8.00, 7.82, 6.55, 7.60 (447)
- home can withstand severe weather: % 93.2, 86.1, 81.9, 88.9, 87.0 (369)
- HH uses safe fuel for cooking and heating: % 81.5, 40.9, 81.1, 62.7, 69.8 (300)

Demographic Indicators
- head of the HH’s age: mean 52.0, 56.7, 49.8, 53.1, 52.4 (446)
- male-headed HHs: % 88.5, 69.6, 83.3, 93.3, 83.6 (372)
- married head of HH: % 95.0, 83.3, 90.1, 94.4, 90.7 (392)
- head of the HH is a single female: % 2.52, 11.2, 6.67, 1.39, 5.88 (24)
- head of the HH is a single male: % 2.52, 5.62, 3.33, 4.17, 3.72 (16)
- head of the HH is a married F or F&M: % 9.24, 18.0, 10.7, 5.56, 10.9 (47)
- head of the HH is a married male: % 85.7, 65.2, 79.3, 88.9, 79.8 (343)
- head of the HH is literate: % 66.7, 46.2, 74.7, 76.0, 66.8 (294)
- HH population (live in HH > 9 months): mean 4.16, 3.20, 4.22, 3.80, 3.92 (447)

HH believes DWQ is good/very good are in the group that does not treat their water. Using the data summarized in Table 1 to calculate RRs without any adjustment, we found that a household believed that its drinking water quality was good or very good, it was 22% less likely to boil its water (adjusted RR = 0.78, 0.64–0.94, p < 0.01); the full-model estimate changed only slightly compared to the first block (RR = 0.75; Table 2), and is also quite close to the unadjusted RR of 0.73 (presented in the previous section). For households with relatively poor access to basic health care (i.e., longer durations to reach health clinics), the likelihood of boiling was also higher. Single, female headed households were 36% more likely to boil than drink untreated water (adjusted RR = 1.36, 1.12–1.66, p < 0.01). Among all female headed households, 43% (n = 22) were widowed and, as might be expected, their mean age (60 years, SD = 12.3) was higher than that of married female headed households (55.3, SD = 14.1), and higher than that of male headed households (52.1, SD = 12.1) (see SI, Tables S10–S11). Of the female headed households boiling with pots specifically, 45% (n = 9) were widowed and none were literate (SI, Tables S12–S13).

In addition to the estimated RRs, we used the predicted mean probabilities from the models to graph observations against the predicted probability for using a given HWT method. This approach provided a convenient means to visualize trends and interaction effects. For example, we see in Figure 1 (Panel A) that, as access to basic health care worsens, the likelihood of boiling (versus drinking untreated water) increases; however, when we disaggregate based on the head of the household’s gender (Panel B), we see that female and joint male–female headed households are likely to boil their water regardless of the distance to the nearest health clinic.
<table>
<thead>
<tr>
<th>Covariate Block</th>
<th>Risk Ratio (95% Confidence Interval)</th>
<th>Model</th>
<th>Boil vs Untreated</th>
<th>Electric Kettles vs Pots</th>
<th>Bottled vs Boil</th>
<th>Model Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Water-Related</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water Quality</td>
<td>0.75** (0.61–0.91)</td>
<td>0.78* (0.64–0.95)</td>
<td>0.79* (0.65–0.97)</td>
<td>0.78** (0.64–0.94)</td>
<td>1.12 (0.81–1.56)</td>
<td>-259.2</td>
</tr>
<tr>
<td></td>
<td>1.11 (0.91–1.34)</td>
<td>1.09 (0.92–1.29)</td>
<td>1.09 (0.93–1.27)</td>
<td>1.10 (0.93–1.29)</td>
<td>0.99 (0.64–1.54)</td>
<td>N/A</td>
</tr>
<tr>
<td>Health-Related</td>
<td>1.005** (1.002–1.008)</td>
<td>1.005** (1.002–1.009)</td>
<td>1.005** (1.002–1.008)</td>
<td>0.997 (0.974–1.021)</td>
<td>0.98*** (0.97–0.99)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>0.90 (0.79–1.01)</td>
<td>0.91 (0.79–1.05)</td>
<td>0.91 (0.79–1.04)</td>
<td>1.12 (0.80–1.56)</td>
<td>1.04 (0.57–1.89)</td>
<td>0.49*** (0.35–0.67)</td>
</tr>
<tr>
<td>Economic Indicators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TVs in HH/HH population</td>
<td>0.86 (0.70–1.05)</td>
<td>0.83d (0.68–1.02)</td>
<td>1.42** (1.16–1.74)</td>
<td>N/A</td>
<td>N/A</td>
<td>1.06 (0.91–1.25)</td>
</tr>
<tr>
<td>Village-average price for 19L W bottle</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>1.06 (0.91–1.25)</td>
<td>1.06 (0.91–1.25)</td>
</tr>
<tr>
<td></td>
<td>1.12 (0.80–1.54)</td>
<td>1.12 (0.80–1.56)</td>
<td>1.12 (0.80–1.56)</td>
<td>1.04 (0.57–1.89)</td>
<td>1.04 (0.57–1.89)</td>
<td>0.49*** (0.35–0.67)</td>
</tr>
<tr>
<td>Demographic Indicators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head of the HH’s age</td>
<td>1.003 (0.997–1.010)</td>
<td>0.989* (0.980–0.998)</td>
<td>0.985*** (0.978–0.993)</td>
<td>N/A</td>
<td>N/A</td>
<td>1.06 (0.91–1.25)</td>
</tr>
<tr>
<td>HH head is male</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>1.06 (0.91–1.25)</td>
</tr>
<tr>
<td>HH head is married</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>1.06 (0.91–1.25)</td>
</tr>
<tr>
<td>HH head is married F or F&amp;M</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>1.06 (0.91–1.25)</td>
</tr>
<tr>
<td>HH head is single male</td>
<td>1.08 (0.87–1.35)</td>
<td>1.04 (0.74–1.75)</td>
<td>0.995 (0.65–1.53)</td>
<td>1.03 (0.73–2.08)</td>
<td>0.93 (0.63–1.39)</td>
<td>0.071</td>
</tr>
<tr>
<td>HH head is single female</td>
<td>0.95 (0.72–1.26)</td>
<td>1.06 (0.74–1.75)</td>
<td>0.995 (0.65–1.53)</td>
<td>1.04 (0.73–2.08)</td>
<td>0.93 (0.63–1.39)</td>
<td>0.071</td>
</tr>
<tr>
<td>HH head is literate</td>
<td>1.36** (1.12–1.66)</td>
<td>1.08 (0.87–1.35)</td>
<td>0.95 (0.72–1.26)</td>
<td>1.04 (0.73–2.08)</td>
<td>0.93 (0.63–1.39)</td>
<td>0.071</td>
</tr>
<tr>
<td>HH head is literate</td>
<td>1.05 (0.88–1.26)</td>
<td>1.06 (0.74–1.75)</td>
<td>0.995 (0.65–1.53)</td>
<td>1.04 (0.73–2.08)</td>
<td>0.93 (0.63–1.39)</td>
<td>0.071</td>
</tr>
<tr>
<td>HH head is literate</td>
<td>1.13*** (1.07–1.19)</td>
<td>1.16 (0.83–1.53)</td>
<td>1.26 (1.02–1.58)</td>
<td>1.16 (0.83–1.53)</td>
<td>1.26 (1.02–1.58)</td>
<td>0.076</td>
</tr>
<tr>
<td>HH population (live in HH &gt; 9 months)</td>
<td>0.997 (0.967–1.028)</td>
<td>1.03 (0.96–1.10)</td>
<td>1.03 (0.96–1.10)</td>
<td>1.03 (0.96–1.10)</td>
<td>1.03 (0.96–1.10)</td>
<td>0.099</td>
</tr>
<tr>
<td>Model Indicators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log pseudolikelihood</td>
<td>−259.2</td>
<td>−254.14</td>
<td>−237.51</td>
<td>−230.82</td>
<td>−150.63</td>
<td>−246.04</td>
</tr>
<tr>
<td>n</td>
<td>273</td>
<td>269</td>
<td>252</td>
<td>246</td>
<td>177</td>
<td>331</td>
</tr>
</tbody>
</table>

*Minutes to clinic for advanced care. **Notes: HH = household | * p < 0.05; ** p < 0.01; *** p < 0.001 | c p = 0.071 | d p = 0.076. e This association is almost entirely explained by the significantly higher proportion of well-constructed homes in County B (100%), compared to County A (67.5%).
Looking to the model comparing boiling with electric kettles to pots, as rates of TV ownership (a proxy for household wealth and/or income) increase, so too does the likelihood of boiling with electric kettles. In addition, larger household populations and younger heads of household were both associated with a higher likelihood of using kettles. As shown in Figure 1 (Panels C, D), among those who boil, across the spectrum of head of household age, literate heads of household and male heads of households appear to be more likely to use electric kettles as compared to illiterate or female heads of household who have a relatively higher likelihood of boiling with pots.

With regard to bottled water use versus boiling, younger heads of household are more likely to use bottled water (RR = 0.985 for each one-year increase), as are literate heads of household. As shown in Figure 1 (Panel E), for almost all ages, literate heads of household are more likely to use bottled water versus boiling. In addition, we see that improved access to advanced health services is associated with higher probabilities of using bottled water, and again this can be conditioned further on literacy Figure 1 (Panel F). It is noteworthy that there are similar trends in Figure 1 Panels C and E, such that younger literate heads of household appear more likely to use bottled water than to boil, and more likely to use electric kettles than pots if they do boil.

### 3.3. Key Demographic, Socioeconomic, And Water-Quality Results

The trends reflected in the descriptive statistics for head of the household’s age and HWT use (Table 1) remain relatively constant even after controlling for other covariates (Table 2) such that older heads of household appear more likely to boil generally, and to boil with pots rather than electric kettles. The perception that most other households boil their drinking water was also strongly associated with a given household’s likelihood of boiling in most villages (SI, Figures S8–S9).
While the resolution of reported village income data is crude, as shown in Table 3, our comparison of mean values for proxies of income, wealth, and access to services over lower, middle, and upper levels of village income suggest that, overall, the data appear to provide an accurate indication of household economic status (see SI, page S21, for additional details).

As reflected in our model results and Figure 2, there is an association between HWT and reported village incomes such that the poorest households (or, more conservatively, households in the poorest villages) are most likely to boil their water overall, and as incomes increase a larger and larger proportion of those boiling their water do so with electric kettles. The difference in bottled water use between the first and second village income groups in Figure 2 may be partially explained by the comparatively high costs of bottled water (for 19L bottles) in the lowest income group (mean = RMB 9.6) compared to the second lowest group (mean = RMB 7.17) and the overall mean cost of bottled water (RMB 7.6) (additional details in SI, page S22).

While we see that some socioeconomic factors appear to predict HWT use, one would expect that the comparative effectiveness of the different HWT methods for pathogen inactivation would be similar across income groups. Because this study was not powered to detect subgroup differences in TTC concentrations by HWT method and village income levels, in Figure 3 we stratified by village income thirds (rather than fifths as in Figure 2). As can be seen, households drinking untreated water have the highest associated exposure to TTC in all income groups, and, with the exception of bottled water in the middle income group, those using electric kettles appear to have the lowest exposure. While the relatively small number of observations in each strata, and correspondingly wide (and often overlapping) confidence intervals, limit interpretation, it is noteworthy that the overall level of contamination appears to be lowest in the upper income village group.

4. DISCUSSION

4.1. The Potential Benefits of Promoting Boiling with Electric Kettles. The HAP produced by boiling with solid fuels is, arguably, the primary drawback of boiling as a HWT method. Among the primary global health risks, unsafe water is ranked 14th, and HAP eighth.3 With the advent of near-universal electricity access, boiling water has become a technological choice as opposed to an undifferentiated practice. Though boiling with biomass is likely responsible for a relatively small proportion of total HAP exposure, given the widespread use of boiling in rural China, an increase in the proportion of households boiling with safer fuels or electricity could have a substantial positive impact on indoor and outdoor air quality.

The analyses presented here shed light on which households might be most likely to switch to a safer method of boiling with electric kettles, and which might require more targeted promotion approaches. Broadly speaking, based on these results we hypothesize that if an effective promotional messaging strategy were to be developed and used, many Chinese households currently boiling with pots would start using electric kettles because they are a fast, easy, and convenient means of continuing their pre-existing boiling behavior. However, because we found that older heads of household, and female headed and poorer households in particular, were more likely to boil using pots and biomass, the one-time cost of an electric kettle and/or the associated electricity costs could be significant barriers to adoption. Thus, any would-be electric kettle promotion program might need to
offer subsidies or other financial incentives in order to elicit a desired level of behavior change among such households.

4.2. Bottled Water and the Costs of Convenience. Our results showed that households with relatively higher incomes and younger, more literate, and often male heads of household are more likely to purchase bottled water generally, or to use electric kettles if they boiled (see also SI, Tables S15–S16). For such households, therefore, there is a competition of sorts between low-cost, efficient, and convenient boiling with electric kettles and relatively high-cost, but highly convenient, bottled water, which we found to be frequently contaminated.

When asked why they use bottled water, 46% of households reported doing so because bottled water was “convenient” followed by 21% who preferred bottled water because they believed it was safe. The other two reasons given were affordability and bottled water’s “good taste” (10% each—see SI, Table S17). The CCDC officials (as well as UNICEF-China staff) with whom we worked had observed this trend in their other work, and agreed that bottled water’s growing popularity in rural China was due largely to its convenience. On this point, it is noteworthy that most bottled water dispensers in China (for 19L bottles) have a built-in heating element. Our study did not address this directly, but published studies on bottled water inactivation, the bottled water dispenser/heating reservoirs not address this directly, but published studies on bottled water (for 19L bottles) have a built-in heating element. Our study did not address this directly, but published studies on bottled water inactivation, the bottled water dispenser/heating reservoirs themselves could also be sources of contamination.

Perceptions of convenience appear to be associated with access and cost, and one would expect higher income villages to better support economies of scale in the supply of a wide range of goods and services including bottled water, as compared to lower income villages. In addition to the data on bottled water cost presented above, after dividing all 15 villages into two groups, lower incomes (reported mean annual income < RMB $100) and higher incomes (>RMB $100), we found that the price of a 19L bottle of water was significantly lower (RMB 6.96, SD = 2.32) in the higher income villages than in the lower income villages (RMB 7.99, SD = 2.29) (two-sided t test, \( p = 0.0072 \)). We found no relationship, however, between the price of the bottled water and its microbiological quality (SI, Figure S11).

Perhaps equally informative are the reasons given for not using bottled water. Most such households reported that bottled water is too expensive (38% overall), with smaller proportions responding that bottled water is not convenient to purchase (14% overall), or is not safe (14%) (SI, Table S18). Though not reported on explicitly, we suspect that the status value of bottled water may be partly responsible for its growing use in poor areas of rural China. Overall, our findings suggest that bottled water use is not primarily driven by a desire for safer water, but rather by people’s perceptions of its convenience.

4.3. Limitations. Our study had some limitations which could moderate our conclusions. Our analyses were hampered in part by the relative lack of similar research on boiling in general, and boiling in China in particular. Few studies have focused on the cost dimension of boiling,\(^{18,49}\) and we are aware of only one boiling-focused study\(^{48}\) that made any mention of electric kettles (albeit in the paper’s SI section and in reference to only two households). Additionally, due to censorship of some MPAT survey questions (and the entire MPAT Village Survey), we were unable to fully calculate all the MPAT subcomponents and components.

The distributional assumptions inherent in all such parametric models apply with respect to our analyses here as well; however, we contend that a hierarchical model building approach provides a more transparent and responsible method for model construction compared to traditional regression approaches. Bayesian Network Analysis (and the use of directed acyclic graphs) offers another method for identifying and analyzing potential dependencies among model covariates and for understanding how some covariate effects on DVs of interest may be mediated by other covariates.

In our study the bottled water samples were only tested for indicators of microbiological contamination. With regard to chemical contamination in bottled water, available research (mostly in Chinese language journals) suggests that this is also an issue of concern in China.\(^{51,52}\)

4.4. Policy Implications. This is the first study we are aware of to focus on the socioeconomic predictors of household water treatment and bottled water use in rural China. Among those who boil their water, we found that poorer households with older (and often illiterate) heads of household were more likely to boil with pots and solid fuels, while relatively wealthier households with younger (often literate) heads of household were more likely to boil with electric kettles. Female-headed households had a higher preference for boiling, but primarily with pots and solid-fuels.

For many Chinese, the boiling of drinking water is not necessarily considered a form of “treatment”; rather, there is a widespread cultural preference for drinking water that has been boiled (even if it is no longer hot). In 2010, residential fuel combustion for cooking, boiling, and heating in China was estimated to contribute to 32% of the country’s total outdoor air pollution burden;\(^{53}\) and as much as 80% of these emissions were attributed to the incomplete combustion of wood and agricultural refuse in rural households.\(^{54}\) An estimated 1.6 million premature deaths per year are now attributed to air pollution exposure in China.\(^{55}\) In light of the well-documented challenges of promoting new HWT methods,\(^{26}\) the most practical way to expand access to microbiologically safe drinking water and reduce HAP in rural areas which lack safe centralized supply, but have reliable electricity access, may be to build upon existing preferences for boiled water and promote an expanded use of electric kettles.

Consumption of bottled water in China nearly doubled from 2010 to 2015.\(^{33}\) We found that households with younger, more literate, heads of household, and higher income proxies, were more likely to buy bottled water, mainly for reasons of convenience. Our study also indicates that cost and difficulty in accessing bottled water are the primary barriers to increased bottled water consumption. Assuming the socioeconomic situation in rural China continues to improve as predicted,\(^{56}\) this strongly suggests that the apparent preference we found for bottled water among younger and more affluent households will drive continued growth in the use of bottled water in the coming years. Given the variation in the microbiological quality of bottled water in China, along with the energy requirements and resources used to package, transport, and deliver bottled water,\(^{57}\) this raises concerns for both human and environmental health.

Finally, our findings demonstrate that boiling should not be considered an undifferentiated HWT practice, but one with different methods of varying effectiveness, environmental impact, and adoption rates across socioeconomic strata.
We thank Gemei Zhong, and the Guangxi CCDC staff, as well as Zhenbo Yang at UNICEF-China, and Benjamin Arnold, the ERG Water Group, Alan Hubbard, Thomas Rath, Vincent Resh, Jeff Romek, and Kirk Smith. Funding and support for this research were provided by the U.S. Environmental Protection Agency (www.epa.gov, STAR Fellowship #91744201-0) and the International Fund for Agricultural Development, a specialized United Nations agency (www.ifad.org). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Funding and support for data collection and water sample analyses were provided by the Chinese Center for Disease Control and Prevention (www.chinacdc.cn).


