A Brief Final Project Report
Problem-based Learning Using Data Analytics
Econ 451/551, Fall 2016

A Complete Set of Results and Modules can be found online:
https://nepalstudycenter.unm.edu/SustainableResearchLab/Econ451FALL2016.html

Sustainable Development Lab, Nepal Study Center
University of New Mexico

Instructor: Dr. Alok K. Bohara, Professor, Department of Economics
Director, Nepal Study Center

Undergraduate students:
Biraj Rawal, Maylisa May Souriyanyong, Dianna Spilca, & DeShawn Vaughan

Graduate student mentors:
Samrat Kunwar (TA), Soumi Roy Chowdhury, Siobhan Yilmaz, Muhammad Adnan Shahid, and Veeshan Rayamajhee (doctoral students)

BACKGROUND

The purpose of studying pollution in Bhairahawa and the surrounding VDCs is to illuminate the role it has played in the spread of diseases, like diarrhea and cholera. Social conditions such as poor environmental awareness, hygiene practices and inadequate health knowledge were examined carefully in order to convey and understand the role these pivotal elements play in a community's overall health.

DATA

In total, 752 households from all three areas were surveyed using the proportional sampling strategy (see Appendix A of the **Full Report** for the complete analysis.). The whole study area is divided in 3 strata in such a way that the sampling units within the stratum are as homogenous as possible and sampling units between strata are as much heterogeneous as possible. The three strata consist of Siddharthangar municipality, Bagaha VDC and Basantapur VDC. The sampling area consisted of 28 wards that were partially within the Danda catchment area. We stratified the sample by ward to ensure sufficient geographic spread. Further each stratum is subdivided into number

of clusters of wards. There is a total of 28 clusters in the 3 strata, which consists of 14,409 households

The survey protocol included: expert interviews, focus-group discussions, debriefing, pilot survey and the final survey. In preparation for the survey, we started with expert interviews as a quick way to obtain information on the survey area. The expert interviews were conducted with the ward personnel of Siddharthangar municipality; and Bagaha and Basantapur VDCs. These interviews played a vital role in redefining the concrete environmental and health issues in and around Danda River; and the survey questions were modified accordingly. (Please see Appendix A for the design, structure, and sampling strategy)

A focus group discussion was also conducted in PNMHI (Pratiman-Neema Memorial Health Institute, Bhairahawa) with representatives from all three VDC's and municipality. The focus group discussion was conducted on 20 participants that comprised of both male and female participants. The focus group proved valuable in comprehending the problem from an average citizen's perspective. As with the expert interviews, the survey questions were further modified after focus group discussions. After focus group, the next step undertaken was debriefing one-on-one with few sample respondents. Each respondent went through the survey by reading it aloud to themselves and answered all the questions. We recorded the time taken to complete each question and at the end they provided comments on the survey in general. This process played an important role in helping us understand how potential respondents would see the survey and the complications that could arise.

LEARNING DESIGN AND PROCESS

Under the supervision of the main instructor (Professor Alok K. Bohara), the class was developed within the framework of the NSC's Sustainable Development Lab. A select group of undergraduate students (engineering, biology, and economics) were enrolled in this class, and a group of volunteering graduate students was selected to serve as mentors. The group met twice a week. Using the STATA software, the survey data was cleaned up and were broken into several modules: water quality, health incidences, sanitation infrastructure, knowledge and attitude and behavior. Various visualizations techniques were used to unravel data information. The final module was the multivariate analysis examining the impact on health outcome. Students followed the following processes: 1) weekly literature review and reading, 2) in-class data analysis, 3) writing of the module reports. At each stage, graduate mentors interacted with the undergraduate students to help them with the reading, analysis, and writing. The final product was uploaded on a dedicated website: At the end, a set of practical solutions was proposed for implementation, perhaps through a study abroad program.

The primary objective of this pilot project (concept course) was to analyze survey data which was collected in Summer 2016 by the Nepal Study Center at the University of New Mexico with the residents of Bhairahawa. The purpose of this survey was to understand their opinions on river ecosystem, environmental pollution and household water quality in Danda river. In addition,

the survey question also focused on household drinking water practices, and pollution problems in Bhairahawa.

Because some data from the survey is categorical vs. quantitative data, logistic regression models were used. The survey questions are often presented as yes or no questions or opinion/habit based. Binaries were created by only taking the extreme polar end into account as the correct answer. For example, some questions would ask how often would one wash their hands after the use of a toilet. This question came with five possible answer choices. To set up the generic "1" for yes and "0" for no binary, the only answer selected to recode into a "1" would be if they washed their hands every time after bathroom use. All other selections would automatically fall under a "0".

MAIN EMPIRICAL FINDINGS

In the result section, we highlighted the primary descriptive findings of our analysis followed by the empirical method and regression results.

Descriptive Results

While analyzing the water quality of Danda River, it is found that the river is highly polluted with a variety of water pollutants such as Ammonia, PH, Dissolved Oxygen, Phosphate, Nitrate, Arsenic, Coliform, and Fecal Coliform. Thus individuals in proximity to Danda River face a serious issue of water pollution and possible negative health implications. Our data reaffirms that in most of the wards in Siddharthanagar, at least 10% of their population contracted a waterborne disease. Contamination of diseases are also influenced by the behavior and practices of individuals including use of flush toilets, treatment of drinking water, and hand washing behaviors. We see that 40.3% of the households that did not wash their hands after using toilet face a diarrhea outbreak. This rate declined to 16.8% among households that practiced handwashing behavior. Of the other practices, only 8% of people from Siddharthanagar, 4% from Basantapur, and 9% of people from Bagaha treat their drinking water sources. The lack of hygienic practices can be attributed to the low level of knowledge of the households as only 10.1% had heard about different water pollutants such as E.coli.

Causal Analysis: Describing the variables

The likelihood of contracting diarrhea or other illness is altered by a set of independent variables and is analyzed through logistic regressions.

$$Illness = f(Personal\ Behavior, Santiation\ Condition, Drinking\ Water\ Source, Surrounding\ Environment) + u$$

The dependent variables we wanted to test were the contraction of diarrhea by adults and children (variables denoted Children Contracting Diarrhea and Adult Contracting Diarrhea). We treated these as two separate variables because research has shown that children are at a higher risk of contracting these diseases (CDC, WHO, et al.), and thus, we wanted to account for this in our models. Also, due to the nature of the survey we had data on households contracting other

illnesses (Typhoid, Jaundice, etc.). This prompted us to create two more dependent variables to test which we denoted as Children Getting Sick (children getting sick with any illness) and Adult Getting Sick (adults getting sick with any illness).

To see how the likelihood of contracting diseases was effected by outside factors we account for a series of independent variables that are in binary form. The first, *Personal Behavior* was defined as the frequency in which individuals in a household washed their hands. Those who washed their hands frequently were tested against those who wash rarely and didn't wash at all. Next, we accounted for *Sanitations Conditions* which are defined as two separate variables. The first, was in regards to how households treated water; those who filtered or boiled their water were tested against those who did not. The second variable accounted for the availability of a sanitation facility. Households who have access to a flush toilet and septic were tested against those who did not. *Drinking Water Source* was also accounted for because we wanted to observe the difference from getting water from a tube well had in comparison to getting it from public tap water. Finally, we took into account the impact that *Surrounding Environment* may have had on contracting diarrhea or other illness by controlling for which city they lived in and how far the households were from the Danda River (households within a ten-minute distance to the river were tested against households who lived further away).

We ran different 3 sets of models where the variables enter the models sequentially. In Model 1- we analyzed the prevalence of diarrhea resulting from personal behavior and surrounding environment (which city they lived in) as our independent variables. In Model-2 we added other important factors such as water treatment and access to a sanitation facility. Finally, in Model-3 we added in two new factors that tested the impact of getting drinking water from a tube well rather than public tap water and the nearness of a household to the Danda River.

Causal Analysis: Magnitude of effect

The regression results complement our descriptive findings as described above. Our initial analysis states that hand washing leads to a significant reduction in the likelihood of contracting diarrhea among children and adult within a range of 9% - 21%. The magnitude, sign, and significance did not change with our inclusion of water treatment behavioral variables as control factors. Further, we saw that in addition to handwashing, boiling or filtering of water contribute significantly to the reduction in the likelihood of diarrhea by 6-9% among children and adult. Accessibility to flush toilet reduces the diarrhea incidence by 4% to 9%. Finally, we ran a full model with handwashing, water treatment, and sanitation variables. In addition, we included the water source and proximity to Danda river as additional control variables. The results again illuminated the significance of hand washing with the point estimates highlighting the decrease in the occurrence of diarrhea from 7-20%. The significance of other factors varied, but there is still reason to believe that water treatment practices, sanitation facilities, and hygienic practices contribute significantly in the lower likelihood of water borne diseases.

RECOMMENDATIONS

a. Scientific Data Collection

At the heart of this project lies the idea of collecting good scientific data (water, air, bio-diversity) and using it to inform the public and policy makers to help them make better decisions. We propose to implement a monitoring program that collects data on a regular basis (monthly and quarterly) on the following four areas:

- Chemical water quality testing: (Nitrate, pH, Dissolved Oxygen, Biochemical Oxygen Demand (BOD), Ammonia, Turbidity, Phosphate, Temperature, Coliform, and E. Coli.)
- Biological assessment Aquatic insects and their larvae (macroinvertebrates) will be trapped, captured, and identified using a Biotic Index (insect key containing the list of pollution sensitive organism).
- Physical landscape assessment: The riparian corridor will be examined for its physical condition (e.g., erosion, inadequate buffer, vegetation, and invasive species).
- Geomorphic assessment: Stream flow can affect the amount of dissolved oxygen, which in turn can impact the animal in the water. The flow (speed) and the volume of water in the river can reflect the changing environment (e.g., climate change).

Data monitoring protocols conducted by DEMP (Danda Ecological Monitoring Program) should play a key role in data collection to not only monitor pollution rates but help guide restoration initiatives for the river and its surrounding communities in order to foster positive public health outcomes and the progression of health education programs.

Testing chemical levels (pH, coliform, nitrate, dissolved oxygen, biochemical oxygen demand, phosphates, temperature) should be done in conjunction with a study of the local flora and fauna in order to assess the level of pollution and damage to the local ecosystem in and around the Danda River. An emphasis should be placed on disseminating the tasks of the monitoring program through citizen science, utilizing local communities and students from local schools to help collect the data using simple but effective tools and instruction. By utilizing the direction of DEMP, local communities near the Danda River and individuals both nationally and internationally interested in the preservation of environmental habitats with a focus on public health can be guided using the program's specific water quality monitoring program, mirrored after successful programs worldwide that demonstrate a holistic approach to the issue of pollution as a political, social, health, and scientific priority.

b. Environmental Education

The citizen science aspect of the proposal recommends getting school students involved in the data collection process. Hands-on practical projects for students will definitely improve the quality of the environmental education, but it also impacts the attitude of the parents and the other household members. Given the empirical finding regarding the impact of hand-washing on the health outcome in the data, for example, the city officials may also benefit by investing in the public awareness campaigns.

A change in environmental education (EE) could help improve sanitation practices. It is clear that individuals are unaware of the severity of the pollution levels and consequences of such harmful pollutants. For instance, in Siddharthanagar 89% of households had not heard of Coliform, despite 75% of their water being polluted with it. This lack of awareness displays the need for environmental education. Furthermore, Environmental educational programs should be interlaced with the current curriculum at all academic levels and as part of community outreach within each location. Doing so will teach children at a young age the importance of these practice while at the same time getting adults to be familiarized as well.

In a study done in Costa Rica, an EE program was implemented to young students on wildlife conservation. A pretest was assessed to help analyze what knowledge the students and their parents already possessed on the topic. The students on average scored 71% on the pretest (15 of 21 questions) while the parents scored an average of 38% (8 of 21 questions). Within a duration of four weeks, educational coloring books and games were administered to the students. These packets were to be worked alongside their parents. A posttest was given eight months after to analyze both group's retention. The students' retention had been lower with an average score of 67%. The parents' retention had increased with an average score of 52%. Although there was not an increase in both groups, the information administered can still be exercised from parent to child (C. Vaughan et. Al).

This type of implementation in schools can be done for students around the Danda River, and in turn create an increased level of awareness which will ultimately lead to better health.

c. Sanitation & Hygiene

As observed from our data analysis, people who suffered from common water borne disease were the ones who lacked a proper drinking water facility, had improper dumping behavior or similar practices which made them prone to common water borne diseases. The city authorities can use a few policy instruments and regularity measures to influence public behavior (e.g., unlawful dumping), but they can also come up with subsidies and incentive packages to improve sanitation related infrastructure (e.g., levy fees to create a garbage pick system, government subsidies to promote flush toilets). For example, use of ceramic water purifier resulted a 50% drop in diarrheal diseases in Cambodia since its introduction in 2001. It also successfully reduced E. coli by 99.99%. It was created through a joint effort by UNICEF and the Water and Sanitation Program. A similar program can be implemented in Bhairahawa.

d. Public Policy

The most important aspect of this pilot exercise is to promote an *evidence-based decision making* process. To that end, we believe that the sound scientific data collection program as proposed here – DEMP-- will be extremely valuable. In fact, the proposal outlined in this exercise offers a big picture which attempts to bring all the parties together – scientists, the public, and policy makers.

Using a holistic approach allows for analysis to be done across a full spectrum of data, and in doing so in terms of public policy, *data* will serve as the proprietor of change, and not *ideology.* Economist Munir Sheik highlights the importance of using a holistic approach to public policy when analyzing macroeconomic performance. In his analysis, he notes that the United States and Canada both have very large economies, yet when you include factors outside of productivity into the analysis like carbon taxes, and wealth distribution the overall size of one's economy is no longer as impressive. Which illuminates the prevailing need to consider a full spectrum of data in any analysis.

e. Environmental Conservation and Eco-System Services: Bio-park Wildlife Refuge

The city of Siddharthanagar is on the Danda River, which flows through the farmlands, open space, and urban settlements, and goes through one municipality and three or more rural counties. Some segments of the river are pristine and harbor numerous species of birds, reptiles, insects and a healthy riparian ecosystem. In this urban ecosystem the health and wellbeing of the people is interconnected with the riparian ecology and the lifeform it supports. Any loss of biodiversity of this vital river system impacts the eco-tourism and the image of the world heritage site of Lumbini, the birthplace of Buddha. Thus, to promote and ensure a common human-bio interest, we propose for a development of a *bio-park wildlife refuge* along the Danda River.

BIBLIOGRAHPY

Ballantyne, R., Fien, J., & Packer, J. (2001). Program Effectiveness in Facilitating Intergenerational Influence in Environmental Education: Lessons From the Field. *The Journal of Environmental Education, 32*(4), 8-15. doi:10.1080/00958960109598657

Chhapola, V., & Brar, R. (2013, August 26). Impact of an educational intervention on hand hygiene compliance and infection rate in a developing country neonatal intensive care unit. *International Journal of Nursing Practice, 21*(5), 486-492. doi:10.1111/ijn.12283

Knatterud.,G.L., Rockhold, F.W., George, S.L., Barton, F.B., Davis, C.E., Fairweather, W.R., Honohan, T., Mowery, R, O'Neill, R. (1998). Guidelines for quality assurance in multicenter trials: a position paper. Controlled Clinical Trials, 19:477-493.

Luby, S. P., Agboatwalla, M., Painter, J., Altaf, A., Billhimer, W. L., & Hoekstra, R. M. (2004). Effect of Intensive Handwashing Promotion on Childhood Diarrhea in High-Risk Communities in Pakistan. *Jama, 291*(21), 2547. doi:10.1001/jama.291.21.2547

Luby, S. P., Agboatwalla, M., Feikin, D. R., Painter, J., Billhimer, W., Altaf, A., & Hoekstra, R. M. (2005, July 16). Effect of handwashing on child health: A randomised controlled trial. *The Lancet, 366*(9481), 225-233. doi:10.1016/s0140-6736(05)66912-7

Most, .M.M., Craddick, S., Crawford, S., Redican, S., Rhodes, D., Rukenbrod, F., Laws, R. (2003). Dietary quality assurance processes of the DASH-Sodium controlled diet study. Journal of the American Dietetic Association, 103(10): 1339-1346.

Sharek, P. J., Benitz, W. E., Abel, N. J., Freeburn, M. J., Mayer, M. L., & Bergman, D. A. (2002). Effect of an Evidence-Based Hand Washing Policy on Hand Washing Rates and False-Positive Coagulase Negative Staphylococcus Blood and Cerebrospinal Fluid Culture Rates in a Level III NICU. *Journal of Perinatology, 22*(2), 137-143. doi:10.1038/sj.jp.7210661

Vaughan, C., Gack, J., Solorazano, H., & Ray, R. (2003). The Effect of Environmental Education on Schoolchildren, Their Parents, and Community Members: A Study of Intergenerational and Intercommunity Learning. *The Journal of Environmental Education,34*(3), 12-21. doi:10.1080/00958960309603489

Whitney, C.W., Lind, B.K., Wahl, P.W. (1998). Quality assurance and quality control in longitudinal studies. Epidemiologic Reviews, 20(1): 71-80.

World Health Organization. (2007). Combatting Waterborne Disease at the Household Level.

World Health Organization (2009). Country Profiles of Environmental Burden of Disease: Nepal.

World Health Organization (2016). Disability-adjusted life years (DALYS).

Xuan, L. T., & Hoat, L. N. (2013). Handwashing among schoolchildren in an ethnically diverse population in northern rural Vietnam. *Global Health Action, 6*(0). doi:10.3402/gha.v6i0.18869

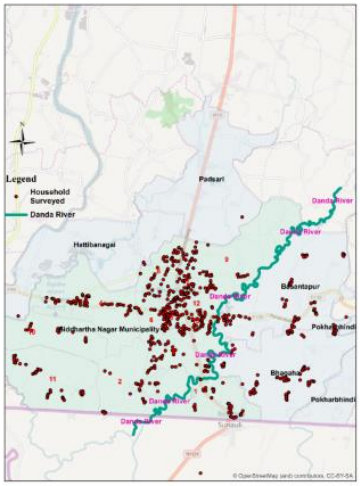
Magiorakos AP, Leens E, Drouvot V, May-Michelangeli L, Reichardt C, Gastmeier P, Wilson K, Tannahill M, McFarlane E, Simon A. Pathways to clean hands: highlights of successful hand hygiene implementation strategies in Europe. Euro Surveill. 2010;15(18):pii=19560.

Sheik, A Munir. (January 2016). Evidence-based policy Development: A framework and its application

Appendix A

SURVEY DESIGN

The survey protocol included expert interview, focus-group discussions, debriefing, pilot survey and the final survey. In preparation for the survey, we started with expert interviews as a quick way to obtain information on the survey area. The expert interviews were conducted with the ward personnels of Siddharthanagar municipality; and Bagaha and Basantapur VDC s. These interviews played a vital role in redefining the concrete environmental and health issues in and around Danda River; and the survey questions were modified accordingly.



A focus group discussion was also conducted in PNMHI (Pratiman-Neema Memorial Health Institute, Bhairahawa) with representatives from all three VDC's and municapility. The focus group discussion was conducted on 20 participants that comprised of both male and female participants. The focus group proved valuable in comprehending the problem from an average citizen's perspective. As with the expert interviews, the survey questions were further modified after focus group discussions. After focus group, the next step undertaken was debriefing one-on-one with few sample respondents. Each respondent went through the survey by reading it aloud to themselves and answered all the questions. We recorded the time taken to complete each question and at the end they provided comments on the survey in general. This process played an important role in helping us understand how potential respondents would see the survey and the complications that could arise.

After the 3 steps mentioned above, the first pre-test survey was conducted on 50 randomly sampled households in five wards of Siddharthanagar; and 3 wards each from Basantapur and Bagaha. We selected 30 households for the municipality and 10 households each for the VDCs. The result from this survey was used to estimate some preliminary result, primarily to understand whether respondents were properly understanding survey questions. While the results did not look too out of ordinary, there were some problems in respondents understanding the choice experiment section for Danda river in the survey. As a result, the survey was slightly modified and another pretest was undertaken with 50 different household samples with the same logistics as the first pretest survey.

After the second pretest survey, the final survey was conducted between June 05 – July 16, 2016 for a total of 752 respondents across Siddharthanagar, Basantapur and Bagaha. The survey was administered in *Nepali* and *Maithili* language by 8 college level students that were trained for 3 days in survey techniques. Enumerators used a scripted introduction to assure that each respondent received the same amount of introductory information. The enumerators were also given a GPS device to record the household coordinates of each respondent. Finally, each household that undertook the survey was presented with a detergent powder as a gift for their time.

Survey Design and Methodology:

The survey design was divided into two stages:

Stage 1:

The whole study area is divided in 3 strata in such a way that the sampling units within the stratum are as homogenous as possible and sampling units between strata are as much heterogenous as possible. The three strata consist of Siddharthanaga municipalit, Bagaha VDC and Basantapur VDC. The sampling area consisted of 28 wards that were partially within the Danda catchment area. We stratified the sample by ward to ensure sufficient geographic spread.

Further each stratum is subdivided into number of clusters of wards. There is a total of 28 clusters in the 3 strata, which consists of 14,409 households.

Strata	Name	No. of households	Number of wards	Number of clusters to be sampled	No. of households to be sampled	Over sampling
1	Siddharthanagar Municipality	12,497	12	12	520	572*
2	Bagaha VDC	707	8	4	30	90*
3	Basantapur VDC	1,205	8	4	50	90*
	Total	**14,409**	**28**	**20**	**600**	**752**

Stage 2:

The clusters are selected from each stratum with probability proportional to size of households so that the total number of clusters is 20 and households is 600.

In total, we surveyed 752 households for the final survey. Enumerators were instructed to take a random route through the wards, stopping at every 5th house for a total of 5 households per enumerators per day. Furthermore, they were asked to interview only household representatives age 18 or older. If a household declined to take the survey, the enumerators would go to interview the corresponding house.