Folder Title: Project Management Records - Synergistic Health Effects from Water Supply and Sanitation Interventions - (RPO # 677-25) - 1v

Folder ID: 1572444

Date: 1/1/1992 – 12/31/1993

ISAD(G) Reference Code: WB IBRD/IDA DEC-03-77

Series: Research Project Files Maintained by the Research Administrator

Sub-Fonds: Records of the Office of the Vice President, Development Economics and Chief Economist and later Senior Vice President, Development Economics and Chief Economist (DECVP)

Fonds: Records of the Office of the Chief Economist

Digitized: 11/15/2017

THE WORLD BANK
Washington, D.C.

Synergistic Health Effects from Water Supply and Sanitation Interventions 677-25

[redacted]

The World Bank Group
Archives

1572444

R1999-150 Other #: 2 Box # 149641B

Project Management Records - Synergistic Health Effects from Water Supply and Sanitation Inverventions - (RPO # 677-25) - 1v

Completion

4/16

RECEIVED

APR 13 1993

RESEARCH ADVISORY STAFF

Greg
Clare —
Julio ✓
Edith —
(last)

677-25

6/30
papers to JK - evaluations
covers only to file.
e

Water & Sanitation Division			
Routing Slip		Date: 4-12	
G. INGRAM			
			S-3-033
	Approp Disp.		Note & Return
	Approval		Note & Send On
	Clearance		Per our Convers.
	Comment		Per Your Request
	For Action		Prepare Reply
X	Information		Recommendation
	Initial		Signature
	Note & File		URGENT
From: John Briscoe		Room S-11-043 Ext. 35557	

Remarks:

LL-IN-1 NOTE

DATE: 09-Apr-1993 05:56pm

TO: Gregory Ingram (GREGORY INGRAM)

FROM: John Briscoe, TWUWS (JOHN BRISCOE)

EXT.: 35557

UBJECT: Research Grant

Greg:

I am sending you copies of two papers produced by myself and Dr. James Vanderslice under the grant provided to us.

I hope that you will be pleased with the outcome of this effort, which I think has produced two papers which are significant contributions both in terms of methodology and policy content.

As you can see from the letter from Dr. Vanderslice, the research specifically addressed the methodological issues brought to our attention in the review process. As you can see, both papers are being submitted to professional journals. I'll keep you apprised on the outcome.

Once again many thanks for your efficient help on this.

John

CC: Louis Pouliquen (LOUIS POULIQUEN)

The Sequencing of Environmental Interventions in Developing Countries: Interactions and their Implications

A Report Submitted to the World Bank

April 1, 1993

James VanDerslice, Ph.D.
Carolina Population Center
University of North Carolina at Chapel hill

John Briscoe, Ph.D.
The World Bank

Does Environmental Sanitation Modify the Protective Effects of Breastfeeding Against Diarrheal Disease?

Results from the Cebu Longitudinal Health and Nutrition Survey

A Report Submitted to the World Bank

Final Draft

April 6, 1993

James VanDerslice, Ph.D.
Carolina Population Center
University of North Carolina at Chapel Hill

John Briscoe, Ph.D.
The World Bank

Budget

ALL-IN-1 NOTE

DATE: 28-May-1992 06:07pm

TO: Louis Pouliquen (LOUIS POULIQUEN)

FROM: Gregory Ingram, RAD (GREGORY INGRAM)

EXT.: 31052

SUBJECT: Undercommitment of Research Support Budget Accounts

According to our records, as of May 26, 1992, the Research Support Budget accounts listed below show undercommitments of its FY92 budgets. The amounts left uncommitted and undisbursed by the fiscal year deadlines are lost both to the projects, and to the Research Support Budget.

We notified the project supervisors of the outstanding amounts and fiscal year deadlines for commitments last Friday, May 22, but uncommitted amounts are still large.

Would you please encourage your staff to input the necessary commitments, before the deadlines below.

New consultant requests or extensions	May 29
New contractual services requests	June 19
New travel requests for FY92 departures	June 22

If you have any questions about the management of RSB accounts, please feel free to call Vilma Mataac, the RSB Budget Officer at extension 31030.

Div.	RPO #	Amount uncommitted	Principal supervisor
INUWS	67725	$10,000.00	J. Briscoe

CC: Vilma Mataac (VILMA MATAAC)

F

THE WORLD BANK/IFC/MIGA

OFFICE MEMORANDUM

DATE: June 12, 1992

TO: John Briscoe, Chief, Policy Unit, INUWS

FROM: Gregory K. Ingram, Administrator, Research Advisory Staff

EXTENSION: 31052

SUBJECT: Research project, "Synergistic Health Effects from Water Supply and Sanitation Interventions," RPO 677-25

Thank you for your memorandum of June 10. Accordingly, we will reduce your FY93 allocation by $2,000. Thus the FY93 allocation for this RPO will be $15,000, and the total authorized budget will be $25,000.

cc: L. Pouliquen, INUDR; C. Carnemark, INUWS
S. Shah, C. Else, V. Mataac, E. Thomas, RAD

THE WORLD BANK/IFC/MIGA

OFFICE MEMORANDUM

RECEIVED

JUN 12 1992

RESEARCH ADVISORY STAFF

DATE: June 10, 1992

TO: Mr. Gregory Ingram, Administrator, Research Advisory Staff

FROM: John Briscoe, Chief, Policy Unit, INUWS JB.

EXT: 35557

SUBJECT: Research Project 677-25: Proportion of $4,000 increment to the RPO

Further to your memo of June 5, 1992, we will use $2000 of the $4000 increment to the RPO to increase Dr. Guilkey's input so that the econometric issues may be more fully addressed.

DR:dr

wp51\files\677-25.mem

entered
4/1/92

FORM 1524
(3-88)

THE WORLD BANK/IFC
BUDGET TRANSFER REQUEST

Log No. (For CTRMI Use Only) BTr # 6807

To **Management Information Division, CTR**	Requested by (Name and Signature) Clara Else Cl Else			
Person to Contact Clara Else, RAD	Ext. 31049	Dept./Div. Nos. 623/99	Room No. S-3023	Date 4/1/92

AMOUNTS

	FROM					TO				
EXPENSE CATEGORY	DEPT. NO.	DIV. NO.	EXP. CODE	LEAVE BLANK	AMOUNT $	DEPT. NO.	DIV. NO.	EXP. CODE	LEAVE BLANK	AMOUNT $
1. Regular Salaries			AA	-				AA	+	
2. Temporary			AF	-				AF	+	
3. Overtime			AG	-				AG	+	
4. Staff Travel			CA	-				CA	+	
5. Consultant Travel			CB	-				CB	+	
6. Consultant Fees	671	99	EA	-	$10,000	677	25 ✓	EA	+	$10,000 ✓
7. Contractual Services			EB	-				EB	+	
8. Communications			HC	-				HC	+	
9. Internal Computing			XA	-				XA	+	
10. Translation Services			XB	-				XB	+	
11. Office Machines-Purchase			KA	-				KA	+	
12. Office Machines-Rental			KB	-				KB	+	
13. Office Machines-Maintenance			KC	-				KC	+	
14. Office Machines-Depreciation			KD	-				KD	+	
15. Office Machines-Software/ Supplies			KS	-				KS	+	
16. Office Supplies			LA	-				LA	+	
17. Contingency			VA	-				VA	+	

STAFFYEARS

	FROM			TO		
Staff Category	Staffyears	Dept. No.	Div. No.	Staffyears	Dept. No.	Div. No.
Professional (Grades 18 and above)						
Others (Grades 11 to 17)						
Special						

POSITIONS

FROM				TO			
Dept. No.	Div. No.	Position No.	Title	Dept. No.	Div. No.	Position No.	Title

Remarks
FY92 for RPO 677-25 ✓ "Synergistic Health Effects from Water Supply and Sanitation Interventions"

Briscoe

cc: V. Mataac

WHITE & YELLOW - CTRMI; PINK - Budget Officer; GOLDENROD - Originator

INSTRUCTIONS

General

The Budget Transfer Request Form is designed to systematically collect the information necessary to properly revise the original authorized budget. The columns and lines of this form are designed to accomodate the most common transactions. Should exceptional cases arise, please use either blank lines, erase preprinted lines, use remarks section or use an additional form to transmit the necessary information for your transfer request. Budget transfers should be requested for amounts of $1,000 and above for any one expense category.

Responsibility for Completing the Form

The completion of the form is not a means for seeking CTR approval to transfer the budget but a means of informing CTR about a budget reallocation. The authority to transfer, as explained below, lies with the department or vice-presidency, depending on internal delegation of authority arrangements within the VP-Unit (or VP complex). This form should be filled out by an organizational unit requesting transfers of its approved and distributed budgeted amounts, staff years and positions. If the transfer is between two departments, consultation with the receiving department is always necessary. All forms should be copied to the Chief Administrative Officer in the VP's front office and to the budget officers in all other departments. The authority to sign the budget transfer forms and its distribution may vary for each Vice-Presidency according to specific instructions issued for this purpose by the Chief Administrative Officer and the responsible budget officers. Forms for budget transfers from the Contingency Accounts to the various organizational units will be filled out by the responsible CTR/PBD Officer.

Amount Transfer

As a general rule, vice-presidents and department heads have the authority to transfer discretionary funds from one expense category to another without CTR or PBD approval, except that the representation and hospitality budget may not be increased.

The expense categories which are printed on the form are only those most frequently requested. A transfer, however, may involve other expense categories. The complete list of expense categories which are budgeted in the various department and the corresponding codes is provided below for information:

Expense Category	Code	Expense Category	Code
Regular Salaries	AA	Local Staff Costs	AL
Temporary	AF	Office Occupancy	GA
Overtime	AG	Office Machines:	
Staff Travel	CA	Purchase	KA
Consultant Travel	CB	Rental	KB
Representation	DA	Maintenance	KC
Hospitality	FB	Depreciation	KD
Consultant Fees	EA	Software/Supplies	KS
Contractual Services	EB	Office Supplies	LA
Communications	HC	Other Miscellaneous	MF
Internal Computing	XA		
Translation Services	XB		

General Contingency (960-30) VA(099-30)* *Code for IFC

If a budget transfer involves allocated expenses (benefits or overhead), the requestor should consult the Budget Implementation Guide for the proper expense name and code.

Staffyear Transfer

The information requested under this category for professionals (grades 18 and above), others (grades 11 to 17) and special will be used to make adjustments in the administrative budget staffyears and will be reflected in the Report of Expenses and Budgets. Staffyear information is necessary when transfers are made from the Contingency Accounts because such transfers constitute additions to the authorized staffyears.

Position Transfer

Normally, a position transfer involves also transfers of amounts and staffyears. A request for a position transfer is necessary when the transfer is from one department to another, that is, whenever two departamental codes are involved; i.e. the transfer is from 122-05 to 126-10. Position transfers within a department (from one division to another) are made automatically through the Personnel Action Form and, therefore, no budget transfer is necessary.

Recodings of position titles and changes of position numbers of occupied positions which are usually for intradepartamental changes should not be requested through this form. These changes are not consideed budget transfers and are to be communicated to the Personnel Officer who will, if necessary, inform CTR appropriately.

Remarks

The remarks section should be used for a brief description of the transfer or for explanations necessary to clarify the entries made in the other sections of the form.

THE WORLD BANK/IFC/MIGA

OFFICE MEMORANDUM

DATE: March 17, 1992

TO: John Briscoe, INUWS

FROM: Gregory K. Ingram, Administrator, Research Advisory Staff

EXTENSION: 31052

SUBJECT: Research proposal, "Synergistic Health Effects from Water Supply and Sanitation Interventions"

Thank you for your memorandum of March 12, responding to the suggestions concerning further development of the analytical framework of your model, and indicating your fiscal year allocation of the budget for the above project.

The following account and financial authorizations are issued for your project:

i)	Identification code (RPO #)	677-25
ii)	FY92 authorization	$10,000
iii)	FY93 authorization	$17,000
iv)	Total authorization	$27,000

Expenses related to the project may be charged to the above account number up to the total amount authorized. Please use the account number as an identification code in all documents relating to the project.

As some additional work has been added to early stages of the project, I have extended the completion date from the original, September 30, 1992, to December 31, 1992. Please note, that also will be the last date to make financial commitments against the account. The account will remain open for four months after the completion date to allow for disbursements against outstanding commitments.

You are required to file a completion report on the project no later than two months after its completion date; that will be February 26, 1993. Forms for this purpose may be obtained from this office.

cc: V. Rajagopalan, OSPVP; L. Pouliquen, INUDR; C. Carnemark, INUWS;
M. Bellinger, ORGHD; R. Salandy-DeFour, ACTAS; C. Ramirez, ACTAB;
S. Shah, C. Else, E. Thomas, V. Mataac, J. Kipnis, RAD;
Research Committee members

The World Bank Group
Archives
& Records Management

Record Removal Notice

File Title Project Management Records - Synergistic Health Effects from Water Supply and Sanitation Interventions - (RPO # 677-25) - 1v	**Barcode No.** 1572444

Document Date 12 March, 1992	**Document Type** Memorandum	

Correspondents / Participants
To: Gregory Ingram
From: John Briscoe

Subject / Title
Research Proposal

Exception(s)
Personal Information

Additional Comments

The item(s) identified above has/have been removed in accordance with The World Bank Policy on Access to Information or other disclosure policies of the World Bank Group.

Withdrawn by	**Date**
Sherrine M. Thompson	November 11, 2017

Archives 01 (March 2017)

677-25

RPO#	DEP/DIV	Title	P.I.	PCR DAtE
677-25	TWUWS	Synergistic Health Effects from * Water Supply & Sanitation Interventions	Briscoe	N.A.

Output

James VanDerslice and John Briscoe: "The Sequencing of Environmental Interventions in Developing Countries: Interactions and their Implications." April 1, 1993

__________ "Does Environmental Sanitation Modifiy the Protective Effects of Breast Feeding Against Diarrheal Disease?" April 6, 1993.

* Evaluation Group: Health

The memo from Briscoe to Ingram dated April 9, 1993, in project file, mentions the submission of the two papers listed above.
No mention is made of a PCR, since it seems to be a "research grant"

P.I. PCR DAtE

Water & Sanitation Division			
Routing Slip		**Date:** Mar. 3, 1995	
Gregory Ingram, RAD		N-9-033	
	Approp Disp.		**Note & Return**
	Approval		**Note & Send On**
	Clearance		**Per our Convers.**
	Comment		**Per Your Request**
	For Action		**Prepare Reply**
	Information		**Recommendation**
	Initial		**Signature**
	Note & File		**URGENT**
From: **John Briscoe** **Division Chief** **TWUWS**		**Room S-4-117** **Ext. 35557**	

Clara Else
FYI
g

ok ~~Edith~~ / Anupa -
Log but not ask
for a PCR if it's
really only $10K.
Thanks, Clara

4/13 Logged; spoke to CE; project
funds were $27K; to be determined
if request for PCR will be made.
(677-25)
Ⓔ

With the compliments of

John Briscoe

The World Bank
1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.

Greg:
Two products from the
$10,000 you gave a couple
of years ago.
Many thanks.
John

Drinking-water quality, sanitation, and breast-feeding: their interactive effects on infant health

J. VanDerslice,[1] B. Popkin,[2] & J. Briscoe[3]

The promotion of proper infant feeding practices and the improvement of environmental sanitation have been two important strategies in the effort to reduce diarrhoeal morbidity among infants. Breast-feeding protects infants by decreasing their exposure to water- and foodborne pathogens and by improving their resistance to infection; good sanitation isolates faecal material from the human environment, reducing exposures to enteric pathogens. Taken together, breast-feeding and good sanitation form a set of sequential barriers that protect infants from diarrhoeal pathogens. As a result, breast-feeding may be most important if the sanitation barrier is not in place. This issue is explored using data from a prospective study of 2355 urban Filipino infants during the first 6 months of life. Longitudinal multivariate analyses are used to estimate the effects of full breast-feeding and mixed feeding on diarrhoeal disease at different levels of sanitation. Breast-feeding provides significant protection against diarrhoeal disease for infants in all environments. Administration of even small portions of contaminated water supplements to fully breast-fed infants nearly doubles their risk of diarrhoea. Mixed-fed and weaned infants consume much greater quantities of supplemental liquids, and as a result, the protective effect of full breast-feeding is greatest when drinking-water is contaminated. Similarly, full breast-feeding has stronger protective effects among infants living in crowded, highly contaminated settings.

Introduction

Breast-feeding is an extremely effective means of protecting young infants from diarrhoeal disease. Infants who are not breast-fed have a two-to-three times greater risk of diarrhoea than breast-fed infants and a three-to-five times greater risk than those who are exclusively breast-fed (*1–7*). Other studies have documented even stronger protective effects (*8–11*).

There are two mechanisms through which breast-feeding protects infants from enteric infections. First, it reduces or eliminates exposure to food- and waterborne pathogens. Weaning foods and breast-milk substitutes pose a particular risk since bacterial pathogens can readily multiply in these foods if they are stored at ambient temperatures after preparation (*12–19*). Second, mature breast milk contains several compounds, e.g., secretory IgA, which can improve the infant's ability to resist infection (*20–23*). A number of studies have found significant associations between pathogen-specific antibody levels in breast milk and the risk or severity of diarrhoea caused by that pathogen (*24–26*).

Reducing the level of environmental contamination similarly reduces the risk of diarrhoea. Good sanitation protects infants by creating a series of barriers to keep enteric pathogens out of their environment; excreta disposal facilities isolate human wastes; improved water supplies protect drinking-water from faecal contamination; and handwashing and personal hygiene reduce the transmission of enteric pathogens in the home.

As a result, poor sanitation may pose more of a risk to those who are particularly vulnerable, i.e., non-breast-fed infants. Weaning foods and breast-milk substitutes are more likely to be contaminated in areas where water supply, sanitation, and hygiene are lacking. Furthermore, families living under these conditions often have fewer economic resources and thus are less apt to prepare foods freshly for each meal, adequately reheat previously prepared foods, or store foods under refrigeration. Consequently, mixed-fed and weaned infants living in poor sanitary conditions probably face considerably higher exposures to foodborne pathogens than similarly fed infants in less contaminated environments. Thus, exclusive breast-feeding may provide greater protection to infants living in highly contaminated environments.

---

[1] Assistant Professor, University of Texas-Houston, School of Public Health, MPH Program at El Paso, 901 Education Bldg, U.T.E.P., El Paso, Texas 79968–0642, USA. Requests for reprints should be sent to this author.

[2] Professor, Department of Nutrition, and Fellow, Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

[3] Chief, Water and Sanitation Division, World Bank, Washington, DC, USA.

Reprint No. **5506**

Put simply, breast-feeding and sanitation can be thought of as a set of sequential "barriers" protecting the infant from enteric pathogens. Accordingly, the breast-feeding barrier is most important when the sanitation barrier is not in place. Also, the sanitation barrier is most important when the breast-feeding barrier is absent. The hypotheses that are explored in this article follow from this simple model:

- *Hypothesis 1:* The protective effect of breast-feeding is greatest where sanitary conditions are poor.
- *Hypothesis 2:* The protective effect of good sanitary conditions is greatest among those not breast-fed.

Only two published studies have addressed this issue. Using retrospective data gathered from 1262 women in the Malaysian Family Life Survey, Habicht et al. found that any breast-feeding was associated with lower infant mortality and that this protective effect was significantly stronger among households lacking toilet facilities or piped water (*27*). In a similar analysis of the same data, Butz et al. assessed the effect of the duration of breast-feeding on infant mortality at various levels of sanitation (*28*). Exclusive and supplemented breast-feeding had the strongest protective effects when toilets and piped water were absent, and somewhat smaller effects in households that had either of these facilities. Breast-feeding had the smallest protective effect against infant mortality in households that had both piped water and a toilet. These results support hypothesis 1.

In a related study, Clemens et al. examined the relationship between breast-feeding and the risk of severe cholera among Bangladeshi children under 3 years of age (*4*). The protective effect of any breast-feeding against severe cholera infection was stronger for infants living near (presumably contaminated) rivers, compared with those who did not. Similarly, a stronger association was observed among infants whose families did not have a latrine. These results also concur with hypothesis 1, with breast-feeding having a stronger protective effect in poor sanitary conditions; however, the protective effect also appeared to be greater for families who had a tube-well in their compound. The effects of breast-feeding under good or poor sanitary conditions were not statistically different.

In the present study we have used prospective (rather than retrospective) data from a large, representative sample of infants, with detailed information on feeding patterns, environmental sanitation conditions, and diarrhoeal morbidity to address the following questions:

- Is the protective effect of breast-feeding against diarrhoeal morbidity greatest where water quality and sanitary conditions are poor?
- Is the protective effect of improved water quality and sanitary conditions greatest when breast-feeding is not practised?
- What specific aspects of environmental sanitation are particularly important in protecting children who are not breast-fed?

The answers to these questions have profound implications for resource allocation decisions in infant health and environmental sanitation programmes.

Methods

Data collection

Study design and sample. The investigation used data collected by the Cebu Longitudinal Health and Nutrition Survey (CLHNS) in a prospective study of 3080 children living in urban, periurban, and rural areas of metropolitan Cebu city, Philippines. The survey consisted of 14 interviews of mothers conducted during the third trimester of pregnancy, soon after birth, and every 2 months thereafter until the child was 2 years of age. The sample used in the present analysis comes from a 12-month cohort of all births in 17 randomly selected urban and periurban *barangays* (communities). All the women were informed of the purpose of the study, the types of questions they would be asked, and that participation in the study was completely voluntary. Of the 2555 women recruited, 2355 had single, live births between April 1983 and May 1984 and agreed to participate in the study.

Data for the first 6 months of life were used since during this time the infant's immune system is developing and full breast-feeding is prevalent. Attrition, because of migration out of the area, death, and refusal to participate reduced the sample size to 1963 at the end of 6 months. This loss to follow-up did not appear to result in a selectivity bias (*29, 30*). More information on the survey design and content have been published previously (*30*).

Diarrhoeal disease. At each bimonthly interview the infant's mother or care-giver was asked whether the infant had experienced any episode of diarrhoea in the 7 days prior to the interview day. The local term for diarrhoea used in the questionnaire (*kalibang*) denotes frequent, watery stools. In a separate study on the validity of retrospective morbidity data conducted in the study area, mothers' reports of diarrhoea, based on frequent or loose stools, had a sensitivity of

95–97% and a specificity of 80% compared with diagnoses made at health clinics and hospitals (*31*).

Feeding practices. Infant feeding encompasses a complex set of behaviours, and apparently small differences in behaviour can have a large impact on an infant's exposure to pathogenic organisms and susceptibility to infection. A complete 24-hour dietary recall was taken at each interview, including the amounts of all foods consumed by the child, method of preparation for broad categories of foods, and whether the infant was suckled. For the descriptive statistics, infants were classified at each time period as either exclusively breast-fed, breast-fed and given only non-nutritive liquids (NNL), mixed-fed, or completely weaned, based on the 24-hour recall. Exclusively breast-fed infants received only breast milk. The breast-fed + NNL infants were primarily breast-fed, but also given liquids lacking caloric content, such as teas, brews, and plain water. Mixed-fed infants were those receiving nutritive foods and/or liquids in addition to breast milk. Completely weaned infants did not breast-feed at all. These definitions are generally consistent with other recommendations that have appeared (*32*).[a]

For the bivariate and multivariate analyses, the exclusively breast-fed and breast-fed + NNL infants were combined into a "fully breast-fed" category, indicating that the infants received all nutrition through breast-feeding, and that they were not exposed to potentially contaminated weaning foods. These categories were based on the reported feeding practices 8 days prior to the interview, minimizing the possibility that the reported practice was a result of, rather than a determinant of, diarrhoea in the week before the interview.

In view of the large number of infants in this study, the variety of foods consumed, and the considerable variation in food contamination levels due to microbial multiplication, it was not feasible to test weaning foods directly for bacterial contamination. The effect of consuming contaminated weaning foods is captured primarily by the infant feeding variables: fully breast-fed infants were not exposed to potentially contaminated foods, while the mixed-fed and non-breast-fed infants were exposed. In addition, a variable representing poor food storage practices was constructed, indicating that the child consumed either a breast-milk substitute or a semisolid food that had been stored without refrigeration for more than an hour after preparation. Exposure to waterborne pathogens in non-nutritive liquids was captured by a water quality variable (see below).

Environmental sanitation. The study infants faced a variety of environmental conditions. To capture this complex array of sanitation-related exposure, the following environmental factors were considered: drinking-water quality; access to water; type of excreta disposal facility; presence of excreta in the household's yard; and the sanitation conditions in the household's neighbourhood.

The water sources used by the household were identified during the baseline survey and verified at each bimonthly survey. Between two and five water samples were collected from each drinking-water source over the course of a year; water sources with more variable quality (such as open dug wells) were sampled more frequently. The samples were analysed for the presence of faecal coliforms (FC) using membrane filtration (*33*). A total of 9% of the samples analysed (154/1650) were not used because they exhibited uncharacteristic colonies or heavy background growth. Faecal coliform concentrations were estimated by dividing the number of dark blue colonies observed by the volume of sample filtered.

Several aspects of sanitation were measured in this study. During the baseline survey the household's excreta disposal facility was identified, the yard was inspected for the presence of faecal material, and the respondent was asked whether animals were allowed in the house. To assess the level of community sanitation, an experienced sanitary engineer carried out a series of environmental assessments. The seventeen primary sampling units (*barangays*) were first divided into 41 homogeneous areas, roughly equivalent to neighbourhoods. Each area was rated using structured observations of housing density, type of settlement (e.g., squatter, periurban, etc.), presence of observable faecal material, predominant types of excreta disposal facilities, and drainage. The same individual conducted all assessments, and each area was surveyed twice over the course of the study to check for internal consistency.

Other risk factors. Detailed demographic and socioeconomic data were gathered during the baseline survey, while data on infant growth and the use of preventive health care services were collected at each bimonthy interview. Rainfall data were derived from daily rainfall measurements at 13 stations around the study area.

Model specification

Multivariable models of diarrhoeal disease in children typically include a wide variety of risk factors such as infant feeding practices, water availability

---

[a] *Indicators for assessing breast-feeding practices: report of an informal meeting, 11–21 June 1991, Geneva, Switzerland.* Unpublished document WHO/CDD/SER/91.14, 1991.

and quality, excreta disposal, age, sex, mother's education and household income. While each of these factors may indeed be associated with diarrhoeal disease, the mechanisms through which they affect a child's risk of diarrhoea are quite different. Factors such as drinking-water quality and sanitation are measures of exposure to enteric pathogens, while nutritional status may affect the child's susceptibility to infection.

Socioeconomic factors, in contrast, do not directly affect the risk of diarrhoea, but rather, influence family behaviours which alter the child's exposure to pathogens or susceptibility to infection. For example, in the study population the more educated mothers had a greater tendency to boil their infant's drinking-water, thus reducing their child's exposure to pathogens. Although educational level and water boiling are both associated with diarrhoeal disease, water boiling has a direct effect on the child's exposure to pathogens, while education level has an indirect effect through its influence on the mother's child-care practices. This conceptual model, first proposed by Mosley & Chen (*34*), has been adapted by the Cebu Study Team to investigate the determinants of child health and growth (*29, 30*).

In this model, diarrhoea was specified as a function of factors that directly affect a child's risk of diarrhoea, i.e., exposure to pathogens and susceptibility to infection. Diarrhoea was measured by a dichotomous variable indicating that according to the mother the child experienced an episode of diarrhoea in the 7 days prior to the interview. Breast-feeding practices were measured by two dummy variables, indicating that the infant was fully breast-fed or mixed-fed. Non-breast-fed infants were used as the comparison group. Exposure to contaminated drinking-water was measured using $\log_{10}$(daily dose of faecal coliforms). The dose for each child was estimated by multiplying the predicted concentration of faecal coliforms in their water source 2 weeks before the interview by the amount of water the child consumed. This dose was adjusted for water boiling since this eliminates the risk posed by contaminated drinking-water (*35*). The level of water service was used as a proxy for water use and water-related hygiene. Having water piped into the house has been shown to increase dramatically the amount used. Lack of a water connection was used as a measure of low water usage and poorer household hygiene.

Families may face exposures directly related to the use of excreta disposal facilities. In a random subsample of the study households, private excreta disposal facilities were kept cleaner than public facilities (J. DeClerque, unpublished data, 1985). Thus, private facilities may pose less of a risk to the users than public facilities or informal defecation areas such as vacant lots or the banks of drainage canals. A dummy variable was used to indicate that the household did not have a private excreta disposal facility.

The presence of faecal material in the yard was used as an overall indicator of household sanitation practices. Such material may be from adults or children who do not use toilet facilities or from domestic animals whose faeces may contain enteric pathogens capable of infecting humans (*12*). The presence of animals in a house and the number of other children in the household were used as measures of exposure in the home.

The purpose of an excreta disposal facility is to isolate human wastes so that any pathogens in them cannot infect others. A child whose family uses a toilet or latrine is less likely to come into contact with faecal material than one whose family defecates indiscriminately in areas near the house. However, even children from households with good excreta disposal practices may face considerable exposure if their neighbours defecate indiscriminately. Thus a child's exposure to faecal material is affected not only by his/her family's excreta disposal practices, but also by the practices of the community as a whole. We refer to this aspect of excreta disposal as community sanitation. A measure of exposure to faecal material due to poor community sanitation was developed from the neighbourhood environmental assessments. Of the 41 areas surveyed, 11 were high-density neighbourhoods where excreta were commonly observed. The infants residing in these areas were considered to be exposed to poor community sanitation.

A number of other factors were controlled in this analysis and are discussed in detail elsewhere (*29, 30*). These include the consumption of foods stored without refrigeration, the infant's weight at the previous interview, expressed in units of standard deviations from the sample mean, the use of preventive health care services during the previous 2 months, age and age-squared terms to account for otherwise unmeasured factors that vary with time (e.g., the development of the infant's immune system), the gender of the infant, and the number of days when it rained over the previous 2 weeks.

Estimation methods

A number of concerns motivated the development of the estimation methods used in this study. First, it was important to ensure that the timing of events was consistent with biological plausibility. We therefore used precise lags of each time-dependent exposure variable, so that the measure of exposure precedes the onset of diarrhoea.

A second concern was the potential effect of unobserved factors. While a large number of exposure and susceptibility factors are included in the model, there are other factors that may directly affect an infant's risk of diarrhoea. For example, children differ in their inherent ability to resist infection. Similarly, there may be small yet important differences in child-care practices that are not represented by the available data. Many of these factors are difficult, if not impossible, to observe. Such unmeasured differences between children are referred to as unobserved heterogeneity.

Biases may occur if the unobserved factors associated with the infant's risk of diarrhoea are also associated with the exposure factors under study. Several of the important risk factors considered in this model are directly affected by the mother's behaviours, such as how she feeds her child and what source of drinking-water she uses. Furthermore, these behaviours may be influenced by the mother's perception of the child's risk of diarrhoea. For example, if the mother perceives her infant to be "sickly", or to be facing some particular health threat, she may be more apt to prolong exclusive breast-feeding. If her perceptions are correct, then her child-care practices may be associated with some of the unobserved factors affecting the infant's risk of diarrhoea. As a result, the exposure factor (breast-feeding) could be correlated with unobserved factors that are themselves risk factors for diarrhoea. In this way, unobserved heterogeneity may confound the relationship between breast-feeding and diarrhoea, leading to a biased estimate of the protective effects of breast-feeding (*36*).

In multivariable models, confounding is controlled by including potential confounders in the model. Since in the present case the potential confounders were not observed, they cannot be explicitly included. An alternative approach is the use of instrumental variables (*37*). In the conceptual model presented above, socioeconomic factors were postulated to affect diarrhoea through their influence on health-related behaviours. In the instrumental variable approach, a statistical model is constructed for each of the health-related behaviours, describing that behaviour as a function of its socioeconomic determinants. For example, breast-feeding is modelled as a function of the parent's education, household income and assets, prices of infant foods and cooking fuel, and demographic factors. These models are then used to generate a predicted value for the particular behaviour for each of the households.

The predicted values for the behavioural risk factors are used to represent these factors in the diarrhoea model. Since these predicted values are simply combinations of socioeconomic factors, they should not be associated with the unobserved factors that affect the mother's behaviours or the infant's risk of diarrhoea. As such, when the predicted values are used in the diarrhoea model, the resulting effect estimates will be unbiased (consistent) estimates of the true effect (*37*). A more detailed description of the creation and use of instrumental variables has appeared elsewhere (*29, 30*).

A "random effects" probit model was used to describe the probability of diarrhoea as a function of the risk factors discussed above. This model specifies that the error term is made up of two components, a standard disturbance that is uncorrelated between children and time periods, and an error term that is different for each child but does not vary with time. This error term represents the unobserved heterogeneity, i.e., the unobserved factors specific to each child that affect his/her risk of diarrhoea. Interaction terms and stratified models were used to estimate the effects of breast-feeding for different levels of sanitation. The parameters were estimated using a maximum likelihood procedure (*38*). Simulations were used to assess the risk of diarrhoea associated with different feeding patterns under good and poor environmental conditions.

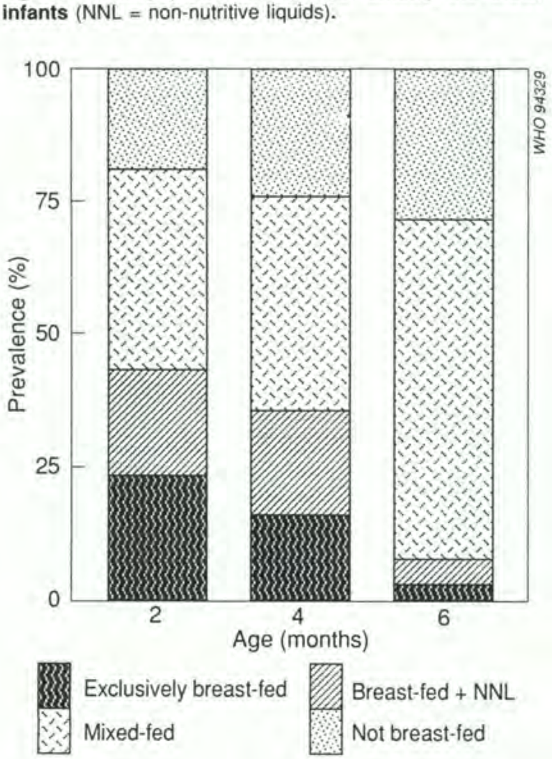
Results

Feeding practices, sanitation and diarrhoeal disease in Cebu

Breast-feeding and exposure to potentially contaminated foods. Infant feeding practices in urban Cebu are typical of those in many developing areas. Very few mothers follow the recommended practice of exclusive breast-feeding for the first 4–6 months. By 2 months of age, 38% of the infants were mixed-fed and 19% were completely weaned (Fig. 1). Another 20% received non-nutritive liquids in addition to breast milk. The proportion of infants who were fully breast-fed dropped to less than 10% by the end of the first 6 months (*39*).

Breast-milk substitutes and semisolid foods seem to pose the greatest risk of bacterial multiplication (*13, 15, 18*). During the first 6 months of life virtually all of the weaned children were given breast-milk substitutes, and these foods accounted for a major part of the infant's diet (88% of the total weight consumed at 2 months; 75% at 6 months). About half of the mixed-fed infants consumed breast-milk substitutes at 2 months of age, and the amounts consumed were, on average, less than half that of the weaned group. Approximately a quarter of those given the substitutes were exposed to milk products that had been stored without refrigeration for over an hour. By 4 months of age, semisolid foods were

Fig. 1. **Feeding practices, by age, among the study infants** (NNL = non-nutritive liquids).



being given to almost half of the mixed-fed and weaned infants, and improper storage was common.

Environmental sanitation. Unlike the situation in many developing countries, water was readily available to almost all the study households. While only 10% of the households had an in-house water connection, less than 3% had to walk more than 5 minutes to fetch water. Furthermore, the quality of the drinking-water available to the majority of households was moderately good to very good. Over half of the households (56%) used deep boreholes fitted with hand pumps or electric pumps, and 29% used a municipal piped supply serving most of the urban areas. Faecal coliform concentrations in these supplies were quite low (geometric mean level = 0.01 colony-forming units (CFU) per 100 ml). The remaining 10% of the households used hand-dug wells, which were frequently contaminated (geometric mean level = 195 CFU per 100 ml).

In contrast to the availability of water, there was no sewerage in any of the study areas. While 54% of the households used flush or pour-flush toilets, very few households had adequate means for disposing of the wastewater, which was commonly discharged into cesspools or open drainage canals. A total of 23% of the households used latrines, and the remaining 23% did not use any facility, defecating into open pits, empty lots, or on the seashore. Toilets and latrines were rarely used to dispose of infants' faeces; the majority of mothers (61%) reported depositing stools under the house, in a vacant lot, or in other places readily accessible to animals or children. Overall, faecal material was readily observed at more than one-third of the households. A total of 30% of the households were rated as being in areas of poor community sanitation, i.e., neighbourhoods with high housing density where excreta were frequently observed.

Infant feeding and sanitation. Many aspects of environmental sanitation and infant feeding were influenced by the same set of underlying socioeconomic factors. For example, household income was inversely related to the duration of breast-feeding and directly related to having an in-house water connection or a private excreta disposal facility. As a result, mothers in households with the highest levels of water supply and sanitation were more likely to wean their children early (Table 1). At 2 months of age, twice as many infants from households with a private excreta disposal facility or in-house water connections were fully weaned compared with households lacking these facilities, and only about half as many were exclusively breast-fed. In contrast, there was little association between feeding practices and the other environmental factors.

Diarrhoeal morbidity and breast-feeding. The proportion of children who experienced diarrhoea in the week preceding the interview rose from 7.2% to 20.4% over the first 6 months (Table 2). There was a clear relationship between feeding practices and diarrhoeal disease (Fig. 2). At 2 months of age the prevalence of diarrhoea among the non-breast-fed infants was nearly three times greater than that among the breast-fed infants. At 2–4 months of age there was a substantial increase in diarrhoeal prevalence among the breast-fed infants, while there was a very sharp rise among the mixed-fed and non-breast-fed aged 4–6 months. After 6 months of age there was little difference in diarrhoeal prevalence between the mixed-fed and non-breast-fed groups.

The mean diarrhoeal prevalence among the infants aged 2–6 months of age is shown in Table 3 for each feeding group stratified by four sanitation variables. For each feeding category, infants living under "poor" sanitary conditions had higher prevalences of diarrhoea, with one exception — fully breast-fed infants in areas of poor community sanitation had about the same prevalence of diarrhoea as

Table 1: **Feeding patterns of the study infants, by sanitation factors, at 2 and 6 months of age**

Feeding practice at:[a]	Private excreta disposal (%) Yes	Private excreta disposal (%) No	In-house water connection (%) Yes	In-house water connection (%) No	Excreta in yard (%) Yes	Excreta in yard (%) No	Poor community sanitation (%) Yes	Poor community sanitation (%) No
2 months of age								
Exclusively BF	16.5	28.7	13.4	24.3	24.9	22.0	24.8	22.5
BF + NNL	18.1	22.0	14.2	20.9	21.6	19.4	22.3	19.3
Mixed-fed	40.7	35.7	35.8	38.2	34.8	39.6	33.4	39.8
Weaned	24.8	13.6	36.6	16.6	18.7	19.0	19.5	18.4
6 months of age								
Exclusively BF	1.7	3.5	1.4	2.9	3.3	2.4	2.2	3.0
BF + NNL	3.0	6.2	2.4	5.0	5.2	4.6	5.0	4.7
Mixed-fed	59.1	67.8	42.3	66.4	64.0	63.5	62.6	64.4
Weaned	36.2	22.4	53.9	25.7	27.5	29.5	30.2	27.9

[a] BF = breast-fed; BF + NNL = breast-fed and given non-nutritive liquids.

those in good sanitation areas. Among weaned infants, diarrhoeal prevalence in each of the poor sanitation strata was consistently high (ranging from 19.5% to 21.5%). Infants from households with in-house water connections had the lowest prevalence of diarrhoea.

Statistical model results

Main effects model. In the multivariable model, both breast-feeding and environmental sanitation were important determinants of diarrhoeal disease during the first 6 months (Table 4, model 1). The protective effect of full breast-feeding relative to no breast-feeding was large and statistically significant. Mixed-feeding had a somewhat smaller, yet statistically significant effect.

Poor environmental conditions were strongly associated with the risk of diarrhoeal disease. Consumption of contaminated water significantly increased the risk of diarrhoea, independently of the type of feeding. In addition, three sanitation variables were significantly associated with diarrhoeal disease. Lack of a private excreta disposal facility and the presence of excreta in the yard had the strongest effects; lack of in-house water was somewhat less important. Poor community sanitation was only marginally significant, and this coefficient was small relative to the other sanitation factors. Other measures of household sanitation, (i.e., having animals in the house and inadequate food storage practices) were not associated with the risk of diarrhoea.

Table 2: **Prevalence of diarrhoea in the 7 days preceding the study interview, by the child's age, sex, and feeding practice**

	% prevalence among infants aged: 2 months	4 months	6 months
Overall	7.2	12.7	20.4
Females	7.1	12.3	20.1
Males	7.4	13.1	20.6
Exclusively breast-fed	4.6	13.0	13.2
Breast-fed + NNL[a]	6.2	11.3	10.4
Mixed-fed	5.6	11.6	19.5
Not breast-fed	14.4	15.9	23.7

[a] NNL = non-nutritive liquids.

Fig. 2. **7-Day prevalence of diarrhoea among the study infants, by feeding practice.**

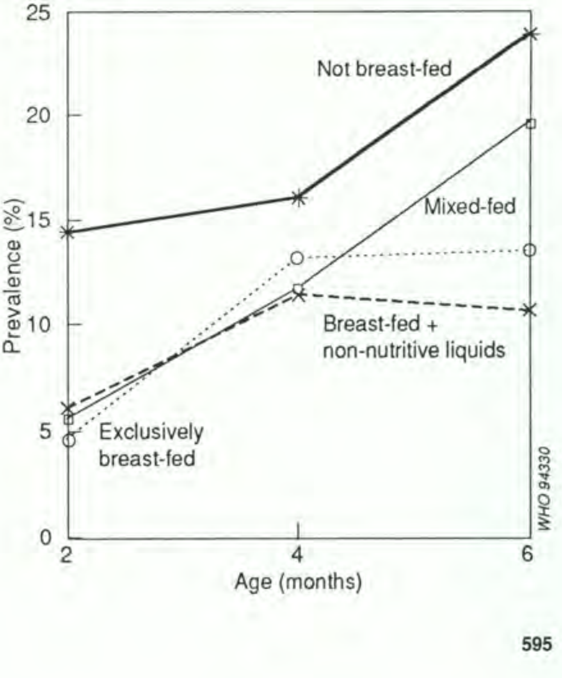


Table 3: **Prevalence of diarrhoea in the 7 days preceding the study interview, and the difference in prevalence relative to fully breast-fed infants aged 2–6 months, by feeding practice and level of sanitation**

	% prevalence of diarrhoea among infants:		
Sanitation variable	Fully breast-fed	Mixed-fed[a]	Not breast-fed[b]
Community sanitation			
Good	8.6	11.6 (3.0)	17.6 (9.0)
Poor	8.1	17.0 (8.9)	20.3 (12.2)
In-house water connection			
Yes	2.0	6.2 (4.2)	13.4 (11.4)
No	8.7	13.6 (4.9)	19.5 (10.8)
Excreta in the yard			
No	7.2	11.6 (4.4)	17.0 (9.8)
Yes	10.5	16.5 (6.0)	21.5 (11.0)
Private excreta disposal			
Yes	6.4	10.2 (3.8)	17.0 (10.6)
No	9.5	15.5 (6.0)	20.5 (11.0)

[a] Figures in parentheses in this column are obtained by subtracting the prevalence of diarrhoea among fully breast-fed infants from that among mixed-fed infants.

[b] Figures in parentheses in this column are obtained by subtracting the prevalence of diarrhoea among fully breast-fed infants from that among infants who were not breast-fed.

Breast-feeding and water contamination. Many of the infants in the study sample were predominantly breast-fed and in addition received NNL supplements. To assess the risk associated with consuming such liquids, the coefficients from the main effects model were used to predict the probability of diarrhoea for each feeding mode at three levels of water quality: no contamination; moderate contamination (10 FC per 100 ml); and high levels of contamination (100 FC per 100 ml). The following mean consumption levels of non-milk liquids for each feeding category were used in this analysis: 39 ml/day for fully breast-fed infants; 119 ml/day for mixed-fed infants; and 209 ml/day for weaned infants. The resulting predicted values are the probabilities that an infant with the given characteristics will experience an episode of diarrhoea over a 7-day period, and represent the average effect for infants aged 2–6 months.

Exclusive breast-feeding and full breast-feeding supplemented with uncontaminated water (breast-fed + NNL) were associated with the lowest risk of diarrhoea (Fig. 3). However, supplementing fully breast-fed infants with even small portions of contaminated water nearly doubled the risk of diarrhoea, from 0.08 to 0.15. Mixed-fed and weaned infants consume much greater quantities of water, and as a result, face much greater risks when their drinking-water is contaminated. Consequently, providing good quality water is most important when the protective effects of breast-feeding are absent.

Similarly, the results can be used to assess the effects of breast-feeding at different levels of water quality. In general, breast-feeding had a strong protective effect; exclusively breast-fed infants had the lowest risk and weaned infants the highest risk. The increase in risk associated with weaning was greatest when drinking-water was highly contaminated. Thus, full breast-feeding will have the strongest protective effect where water quality is poor, i.e., when there is no sanitation barrier to prevent enteric pathogens from entering the water supply.

These results support the two study hypotheses. First, the protective effects of breast-feeding are greatest when drinking-water is contaminated. Second, good water quality has a greater protective effect among infants who are not breast-fed.

Breast-feeding and sanitation. Contamination of infant foods is more likely where sanitary conditions are poor. As a result, eliminating this source of exposure by not giving infants breast-milk substitutes and semisolid foods may be most important under poor sanitary conditions. To assess whether the protective effects of breast-feeding vary with the level of sanitation, the interactions of full breast-feeding with each of the sanitation variables were added sequentially to the main effects model (Table 4, models 2–5). The null hypothesis in each case is that the protective effects of breast-feeding are greatest where sanitary conditions are poor.

Table 4: **Interactions between full breast-feeding with sanitation factors among the study infants aged 2 to 6 months according to the statistical model used** (β and *t* refer to parameter estimate and *t*-statistic, respectively)

	Model 1		Model 2		Model 3		Model 4		Model 5	
Variable	β	*t*	β	*t*	β	*t*	β	*t*	β	*t*
Fully breast-fed	−0.69	−2.5[a]	−0.58	−2.0[a]	−0.66	−2.3[a]	−0.55	−1.6	−0.69	−2.0[a]
Mixed-fed	−0.41	−1.7[b]	−0.42	−1.7[b]	−0.37	−1.5	−0.34	−1.3	−0.41	−1.7[b]
Water quality (log$_{10}$FC dose)[d]	0.26	2.4[a]	0.26	2.4[a]	0.26	2.4[a]	0.26	2.4[a]	0.25	2.3[a]
No in-house water connection	0.19	1.9[b]	0.19	1.9[b]	0.11	0.9	0.19	1.9[b]	0.19	1.9[b]
No private excreta disposal	0.61	3.6[c]	0.60	3.6[c]	0.60	3.5[c]	0.52	2.4[c]	0.61	3.6[c]
Excreta in yard	0.37	2.7[c]	0.38	2.8[c]	0.38	2.8[c]	0.38	2.8[c]	0.36	1.8[b]
Poor community sanitation	0.08	1.6	0.15	2.2[a]	0.08	1.6	0.08	1.6	0.08	1.6
Full breast-feeding interacted with:[e]										
Poor community sanitation			−0.31	−1.5[b]						
No in-house connection					0.75	1.5				
No private excreta disposal.							0.41	0.7		
Excreta in yard									0.02	0.0
Animals in house	0.06	1.4	0.06	1.4	0.06	1.4	0.06	1.4	0.06	1.4
Poor food storage	0.07	0.3	0.05	0.2	0.06	0.2	0.07	0.3	0.07	0.3
No. of children	−0.11	−1.6[b]	−0.11	−1.6[b]	−0.11	−1.7[b]	−0.12	−1.7[b]	−0.11	−1.6[b]
Preventive health care use	0.52	1.8[b]	0.52	1.8[b]	0.52	1.8[b]	0.49	1.7[b]	0.52	1.8[b]
Lagged weight (SD)	0.02	0.8	0.02	0.8	0.02	0.9	0.02	0.8	0.02	0.8
Child's age (weeks)	0.04	1.7[b]	0.04	1.7[b]	0.04	1.7[b]	0.04	1.8[b]	0.04	1.7[b]
Child's age-squared	−0.00	−0.5	−0.00	−0.5	−0.00	−0.6	−0.00	−0.7	−0.00	−0.5
Male child	0.01	0.2	0.01	0.2	0.01	0.2	0.01	0.2	0.01	0.2
Days of rain	0.02	2.5[a]	0.02	2.5[a]	0.02	2.5[a]	0.02	2.5[a]	0.02	2.5[a]
Intercept	−1.33	−3.7[c]	−1.34	−3.7[c]	−1.35	−3.7[c]	−1.39	−3.7[c]	−1.32	−3.6[c]
No. of observations	6 226		6 226		6 226		6 226		6 226	
Log-likelihood	2 313.8		2 313.2		2 313.2		2 313.2		2 313.8	

[a] $P < 0.05$.

[b] $P < 0.10$.

[c] $P < 0.01$.

[d] FC = faecal coliform.

[e] The null hypothesis is that these interactions are negative. Accordingly, one-sided hypothesis tests were conducted.

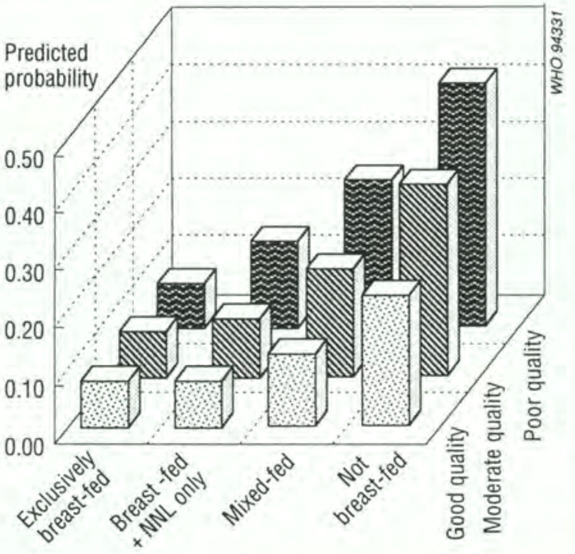
Of the interaction terms, only the poor community sanitation/breast-feeding interaction coefficient was negative and marginally significant (model 2), indicating that, as expected, breast-feeding has a stronger protective effect among infants living in highly contaminated communities. The interaction between no in-house water connection and breast-feeding (model 3) was positive, suggesting that, contrary to our hypothesis, breast-feeding had a stronger protective effect where sanitary conditions were good. The remaining interaction terms were not significant.

It is easier to explore and explain these interactions using stratified models. In the first analysis, households were stratified according to whether they had a water connection in their home. Only a small proportion of the sample had such connections, and the model estimated for these households did not produce stable coefficients.

In the second analysis, households were stratified according to whether they were located in areas with good or poor community sanitation. Separate models were successfully estimated for these two groups (Table 5). While full breast-feeding is a significant protective factor in both areas, the magnitude of the effect in the poor sanitation areas (β = 0.98) is about 1.5 times that in the good sanitation areas (β = 0.61). The difference between these coefficients (0.37) is not statistically significant. Similarly, the effect of mixed feeding was greater for infants living in contaminated communities (0.63 versus 0.40), but the mixed-feeding coefficients were not statistically significant ($P < 0.20$).

The parameters estimated from the stratified models were used to predict the probability of diarrhoea associated with each feeding pattern for areas of good and poor community sanitation (Fig. 4). Poor community sanitation increased the risk of diar-

Fig. 3. **Predicted probabilities of diarrhoea among the study infants, by feeding practice and water quality.** (NNL = non-nutritive liquids).



rhoea among the non-breast-fed infants, but had no effect on the fully breast-fed infants. Thus, full breast-feeding can completely mitigate the risk posed by poor community sanitation.

These results can be interpreted in two ways. Improving community sanitation would clearly have a large impact on weaned infants, and virtually no impact on fully breast-fed infants. The protective effect of good sanitation is therefore greater when infants are not breast-fed, i.e., when the breast-feeding barrier is absent.

Alternatively, the results could indicate that breast-feeding has a stronger protective effect when sanitation is poor. Fully breast-fed infants apparently face much lower risks than weaned infants, and the difference in risks is greater for infants who live in neighbourhoods with poor sanitation. Thus the protective effect of breast-feeding is stronger in the highly contaminated neighbourhoods than in those with good sanitation. In other words, when the sanitation barrier is absent, the protection afforded by breast-feeding becomes even more important.

Discussion

The purpose of our analysis was to test the hypothesis that the protective effect of breast-feeding is greatest where sanitary conditions are poor. While the results do not provide an unambiguous answer, several important conclusions can be drawn.

Full breast-feeding provides significant protection against diarrhoeal disease for infants living in all environments during the first 6 months of life. Exclusive breast-feeding is by far the most protective. Adding even small quantities of contaminated water to the infant's diet can double the risk of diarrhoeal disease. Nevertheless, full breast-feeding (i.e., supplementation with only non-nutritive liquids) is much more protective than mixed-feeding, a result supported by the findings of several previous studies (*1, 7–9*). While mixed-fed infants are at a higher risk than fully breast-fed infants, they are still only half as likely to develop diarrhoea as completely weaned infants. This inverse relationship between the level of breast-feeding and the risk of diarrhoea is probably due to two factors: reduced protection from maternal antibodies in breast milk and an increase in exposure to foodborne pathogens.

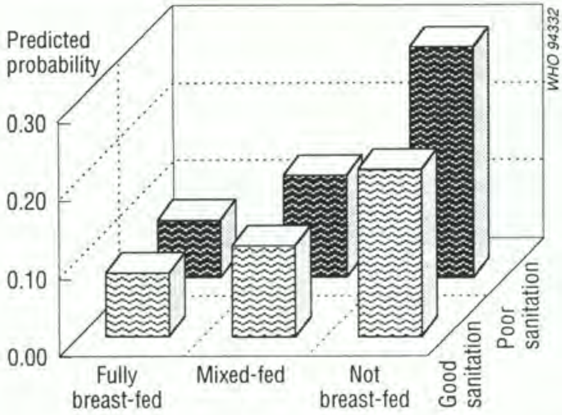
Clearly, under certain circumstances breast-feeding did provide greater protection in poor sanitary conditions. Exclusive breast-feeding had a stronger protective effect when drinking-water was contaminated, and breast-feeding was more protective for infants living in crowded, highly contaminated settings. In contrast, fully breast-fed infants from households lacking water connections or private excreta

Table 5: **Diarrhoea model stratified by community sanitation for infants aged 2–6 months**

	Community sanitation:			
	Good		Poor	
Variable	β	*t*	β	*t*
Fully breast-fed	−0.61	−1.8^{a}	−0.98	−1.9^{a}
Mixed-fed	−0.40	−1.3	−0.63	−1.4
Water quality (log$_{10}$FC dose)	0.29	2.3^{b}	0.16	0.7
No in-house water connection	0.26	2.0^{b}	0.09	0.6
No private excreta disposal	0.42	2.1^{b}	1.07	3.4^{c}
Excreta in yard	0.39	2.3^{b}	0.32	1.3
Animals in house	0.07	1.3	0.04	0.5
Poor food storage	−0.13	−0.4	0.38	0.8
No. of children	−0.07	−0.8	−0.23	−1.8^{a}
Preventive health care use	0.24	0.6	0.88	1.9^{a}
Lagged weight (SD)	0.02	0.5	0.06	1.0
Child's age (weeks)	0.02	0.8	0.06	1.5
Child's age-squared	0.00	0.1	−0.00	−0.8
Male child	0.09	1.4	−0.20	−1.8^{a}
Days of rain	0.03	2.5^{b}	0.01	0.8
Intercept	−1.35	−3.1^{c}	−0.89	−1.3
No. of observations	4 382		1 844	
Log-likelihood	1 560		745	

[a] $P < 0.10$.
[b] $P < 0.05$.
[c] $P < 0.01$.

Fig. 4. **Predicted probabilities of diarrhoea among the study infants, by feeding practice and level of community sanitation.**



disposal facilities did not derive a higher level of protection from full breast-feeding than infants from households that had these facilities. These results are in agreement with the reported effects of breast-feeding and sanitation on infant mortality (*27, 28*) and cholera morbidity (*4*).

The positive interaction between breast-feeding and no in-house water connection was unexpected. Those households with in-house water connections were the wealthiest, and virtually all of these families also owned a private excreta disposal facility. Moreover, only a relatively small proportion of the infants from these households were fully breast-fed (see Table 1). Thus the interaction between an in-house water connection and full breast-feeding may have captured the protective effects of unmeasured child care and hygiene practices among the small group of wealthy mothers who chose to fully breast-feed their infants.

The policy implications of these findings are clear. Full breast-feeding through the first 4–6 months should be encouraged as a means of protecting young infants in all settings from diarrhoeal disease. Furthermore, exclusive breast-feeding is strongly recommended where potable water is not readily available. Also, breast-feeding appears to provide stronger protective effects to infants in crowded, highly contaminated communities. Contamination of supplemental or weaning foods is more likely under these circumstances. In the Cebu region, as in most low-income countries, early supplementation and weaning were most prevalent in these urban squatter and slum areas. Thus, programmes that promote appropriate infant-feeding practices should consider targeting such communities.

In the present study, both water quality and sanitation were important risk factors for diarrhoea. Consumption of contaminated water, lack of private excreta disposal, and the presence of excreta in the yard were associated with the largest increases in risk. As hypothesized, the protection provided by high-quality drinking-water and good sanitation in the community appear to be greatest for non-breast-fed infants. A high-quality water source is particularly important for mixed-fed and weaned infants since they consume the largest quantities of water. Similarly, good community sanitation would benefit completely weaned infants the most since they face significantly higher risks from living in neighbourhoods with poor sanitary conditions. Such highly contaminated environments may well increase the risk of weaning foods being contaminated. Efforts to reduce diarrhoea should focus on reducing exposures, particularly foodborne exposures, through improvements in the choice of supplemental foods, preparation practices and storage methods, and better personal hygiene. Increasing the availability of water and access to excreta disposal facilities can do much to enable families to improve hygienic conditions in the home and to reduce the level of contamination in the community.

Acknowledgements

The analysis we have reported here was funded by a grant from the World Bank Research Program. It is part of a collaborative research project involving the Office of Population Studies and the Water Resources Center, University of San Carlos, Cebu, Philippines; the Nutrition Center of the Philippines; and the Carolina Population Center, University of North Carolina at Chapel Hill, NC, USA. Funding for parts of the project design, data collection, and computerization was provided by the following: National Institutes of Health (contract R01-HD19983A, R01-HD23137, and R01-HD18880R); the Nestlé Coordinating Center for Nutrition Research; Wyeth International; the Ford Foundation; the U.S. National Academy of Sciences; the Carolina Population Center; the U.S. Agency for International Development; and the World Bank.

Résumé

Interaction entre les effets de la qualité de l'eau de boisson, des conditions sanitaires et de l'allaitement au sein sur la santé des nourrissons

Les mesures visant à favoriser une bonne alimentation des nourrissons et l'amélioration des conditions sanitaires sont deux stratégies importantes pour réduire la morbidité diarrhéique chez les

jeunes enfants. L'allaitement au sein protège le nourrisson en lui évitant d'être exposé aux pathogènes présents dans l'eau et la nourriture et en renforçant sa résistance à l'infection; l'amélioration des conditions sanitaires réduit l'exposition aux pathogènes en éliminant les excréments de l'environnement humain. L'allaitement au sein et de bonnes conditions sanitaires constituent donc deux barrières complémentaires contre les pathogènes responsables de la diarrhée. En conséquence, l'allaitement au sein est peut-être encore plus important lorsque les conditions sanitaires laissent à désirer.

Pour vérifier cette hypothèse, on a utilisé les données d'une enquête prospective longitudinale sur la santé et la nutrition menée à Cebu (Philippines). Dans cette étude, 2355 enfants d'un milieu urbain ont été suivis pendant les six premiers mois de la vie. Lors de chaque visite, les épisodes diarrhéiques survenus au cours de la semaine écoulée et le type d'alimentation du nourrisson ont été déterminés en interrogeant la mère. Les conditions environnementales ont été établies par observation directe. On a utilisé un modèle de représentation des effets aléatoires par probits pour estimer les effets de l'allaitement au sein exclusif et d'une alimentation mixte dans différentes conditions sanitaires.

L'allaitement au sein a un effet protecteur marqué contre les maladies diarrhéiques, quel que soit l'environnement. La fréquence des diarrhées était trois fois plus faible chez les enfants nourris exclusivement au sein que chez les enfants sevrés; elle était deux fois plus faible chez les enfants recevant une alimentation mixte. Il existe une forte corrélation entre des conditions d'hygiène médiocres et le risque de maladie diarrhéique. La contamination de l'eau de boisson, l'absence de toilettes privées et la présence d'excréments dans la cour des habitations sont les facteurs qui ont l'effet le plus marqué; l'absence d'adduction d'eau est un peu moins importante à cet égard.

Dans certaines circonstances, l'allaitement au sein offre une plus grande protection lorsque les conditions sanitaires sont mauvaises. Premièrement, l'allaitement au sein exclusif a un effet protecteur plus important lorsque l'eau de boisson est contaminée, car les enfants totalement ou partiellement sevrés consomment des quantités de liquides beaucoup plus importantes. Deuxièmement, l'allaitement au sein offre davantage de protection aux enfants vivant dans un milieu surpeuplé ou fortement contaminé, c'est-à-dire lorsque la "barrière sanitaire" est absente. Par contre, la protection offerte par l'allaitement au sein exclusif est à peu près la même, que le logement possède ou non l'eau courante ou des latrines privées.

Les implications pratiques des résultats de cette étude sont claires. Il faut encourager l'allaitement au sein exclusif pendant les quatre à six premiers mois de la vie, car il protège les nourrissons de la diarrhée, quelle que soit la qualité de l'environnement. En outre, l'allaitement au sein exclusif est fortement recommandé lorsqu'il n'est pas facile de se procurer de l'eau potable. Enfin, les programmes visant à améliorer les habitudes d'alimentation des nourrissons devraient être davantage ciblés sur les communautés surpeuplées vivant dans un milieu fortement contaminé.

L'hypothèse de départ, selon laquelle la protection offerte par une eau de boisson de qualité et de bonnes conditions sanitaires est plus marquée pour les enfants qui ne sont pas nourris au sein, a été confirmée. La qualité de l'eau est particulièrement importante pour les enfants partiellement ou totalement sevrés, car ils en consomment des quantités plus importantes. De même, de bonnes conditions sanitaires dans la communauté sont particulièrement importantes pour les enfants sevrés, car ces derniers sont davantage exposés aux risques liés à un environnement insalubre. Les mesures visant à améliorer l'approvisionnement en eau et l'élimination des excreta peuvent contribuer largement à réduire le niveau de contamination dans la communauté et aider les familles à améliorer les conditions d'hygiène à domicile.

References

1. **Feachem RG, Koblinsky MA.** Interventions for the control of diarrhoeal diseases among young children: promotion of breast-feeding. *Bulletin of the World Health Organization*, 1984, **62**: 271–291.
2. **Ahmed F et al.** Community-based evaluation of the effect of breast-feeding on the risk of microbiologically confirmed or clinically presumptive shigellosis in Bangladeshi children. *Pediatrics*, 1992, **90**: 406–411.
3. **Brown KH et al.** Infant-feeding practices and their relationship with diarrhoeal and other diseases in Huascar (Lima), Peru. *Pediatrics*, 1989, **83**: 31–40.
4. **Clemens JD et al.** Breast feeding and the risk of severe cholera in rural Bangladeshi children. *American journal of epidemiology*, 1990, **131**: 400–411.
5. **Morrow AL et al.** Protection against infection with *Giardia lamblia* by breast-feeding in a cohort of Mexican infants. *Journal of pediatrics*, 1992, **121**: 363–370.
6. **Habicht JP et al.** Does breast-feeding really save lives, or are apparent benefits due to biases? *American journal of epidemiology*, 1986, **123**: 279–290.

7. **Guerrant RL et al.** Prospective study of diarrhoeal illness in Northeastern Brazil: patterns of disease, nutritional impact, etiologies, and risk factors. *Journal of infectious diseases*, 1983, **148**: 986–997.
8. **Mahmood DA et al.** Infant feeding and risk of severe diarrhoea in Basrah city, Iraq: a case–control study. *Bulletin of the World Health Organization*, 1989, **67**: 701–706.
9. **Huttly SRA et al.** The epidemiology of acute diarrhoea in a rural community in Imo State, Nigeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1987, **81**: 865–870.
10. **Victora CG et al.** Infant feeding and deaths due to diarrhea. *American journal of epidemiology*, 1989, **129**: 1032–1041.
11. **Popkin BM et al.** Breast-feeding and diarrheal morbidity. *Pediatrics*, 1990, **86**: 874–882.
12. **Black RE et al.** Incidence and etiology of infantile diarrhea and major routes of transmission in Huascar, Peru. *American journal of epidemiology*, 1989, **129**: 785–799.
13. **Rowland MGM et al.** Bacterial contamination in traditional Gambian weaning foods. *Lancet*, 1978, **1**: 136–138.
14. **Barrel RAE, Rowland MGM.** Commercial milk products and indigenous weaning foods in a rural West African environment: a bacteriological perspective. *Journal of hygiene*, 1980, **84**: 191–202.
15. **Black RE et al.** Contamination of weaning foods and transmission of enterotoxigenic *Escherichia coli* diarrhoea in children in rural Bangladesh. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1982, **76**: 259–264.
16. **Molbak K et al.** Bacterial contamination of stored water and stored food: a potential source of diarrhoeal disease in West Africa. *Epidemiology and infection*, 1989, **102**: 309–316.
17. **Imong SM et al.** The bacterial content of infant weaning foods and water in rural northern Thailand. *Journal of tropical pediatrics*, 1989, **35**: 14–17.
18. **Lloyd-Evans N et al.** Food and water hygiene and diarrhoea in young Gambian children: a limited case–control study. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1984, **78**: 209–211.
19. **Motarjemi Y et al.** Contaminated weaning food: a major risk factor for diarrhoea and associated malnutrition. *Bulletin of the World Health Organization*, 1993, **73**: 79–92.
20. **Worthington-Roberts BS.** Lactation and human milk: nutritional consideration. In: Williams SR, ed. *Nutrition in pregnancy and lactation*, 4th ed. St. Louis, MO, Times Mirror/Mosby College Publishing 1989: 244–322.
21. **Welsh JK, May JT.** Anti-infective properties of breast milk. *Journal of pediatrics*, 1979, **94**: 1–9.
22. **Hanson LA et al.** Characteristics of human milk antibodies and their effect in relation to the epidemiology of breastfeeding and infections in a developing country. *Advances in experimental and medical biology*, 1991, **310**: 1–15.
23. **Garza C et al.** Special properties of human milk. *Clinics in perinatology*, 1987, **14**: 11–32.
24. **Ruiz-Pallacios GM et al.** Protection of breast-fed infants against *Campylobacter* diarrhea by antibodies in human milk. *Pediatrics*, 1990, **116**: 707–713.
25. **Glass RI et al.** Protection against cholera in breast-fed children by antibodies in breast milk. *New England journal of medicine*, 1983, **308**: 1389–1392.
26. **Cruz JR et al.** Breast milk anti-*Escherichia coli* heat-labile toxin IgA antibodies protect against toxin-induced infantile diarrhea. *Acta paediatrica Scandinavica*, 1988, **77**: 658–662.
27. **Habicht JP et al.** Mother's milk and sewage: their interactive effects on infant mortality. *Pediatrics*, 1988, **81**: 456–461.
28. **Butz WP et al.** Environmental factors in the relationship between breastfeeding and infant mortality: the role of sanitation and water in Malaysia. *American journal of epidemiology*, 1984, **119**: 516–525.
29. **Cebu Study Team.** A child health production function estimated from longitudinal data. *Journal of development economics*, 1992, **38**: 323–351.
30. **Cebu Study Team.** Underlying and proximate determinants of child health: the Cebu longitudinal health and nutrition survey. *American journal of epidemiology*, 1991, **133**: 185–201.
31. **Kalter HD et al.** Validation of the diagnosis of childhood morbidity using maternal child health interviews. *International journal of epidemiology*, 1991, **20**: 193–198.
32. **Labbok M, Krasovec K.** Toward consistency in breastfeeding definitions. *Studies in family planning*, 1990, **21**: 226–230.
33. *Standard methods for the examination of water and wastewater*, 15th ed. Washington, DC, American Public Health Association, 1985.
34. **Mosley WH, Chen LC.** An analytical framework for the study of child survival in developing countries. In: Mosley WH, Chen LC, eds. *Child survival, strategies for research*. Cambridge, Cambridge University Press, 1984: 25–45 (*Population and development review*, suppl. to vol. 10).
35. **VanDerslice J, Briscoe J.** All coliforms are not created equal: a comparison of the effects of water source and in-house water contamination on infantile diarrheal disease. *Water resources research*, 1993, **29**(7): 1983–1995.
36. **Briscoe J et al.** People are not passive acceptors of threats to health: endogeneity and its consequences. *International journal of epidemiology*, 1990, **19**: 147–153.
37. **Judge GG et al.** *Introduction to the theory and practice of econometrics*. New York, John Wiley & Sons, 1982.
38. **Avery RB, Hotz VJ.** *Hotztran user's manual*. Old Greenwich, CT, CERA Economic Consultants, Inc., 1985.
39. **Zohoori N et al.** Breast-feeding patterns in the Philippines: a prospective analysis. *Journal of biosocial science*, 1993, **25**: 127–138.

American Journal of Epidemiology
Copyright © 1995 by The Johns Hopkins University School of Hygiene and Public Health
All rights reserved

Vol. 141, No. 2
Printed in U.S.A.

Environmental Interventions in Developing Countries: Interactions and Their Implications

James VanDerslice[1] and John Briscoe[2]

This study assesses the effect of drinking water quality on diarrheal disease in good and poor sanitary conditions using a random sample of 2,355 Filipino infants over the first year of life. The study provides powerful confirmation of the importance of environmental factors on diarrhea: The effects of water quality, household sanitation, and community sanitation are strong, consistent, and statistically significant. The positive impact of improved water quality is greatest for families living under good sanitary conditions, with the effect statistically significant when sanitation is measured at the community level but not significant when sanitation is measured at the household level. Improving drinking water quality would have no effect in neighborhoods with very poor environmental sanitation; however, in areas with better community sanitation, reducing the concentration of fecal coliforms by two orders of magnitude would lead to a 40 percent reduction in diarrhea. Providing private excreta disposal would be expected to reduce diarrhea by 42 percent, while eliminating excreta around the house would lead to a 30 percent reduction in diarrhea. The findings suggest that improvements in both water supply and sanitation are necessary if infant health in developing countries is to be improved. They also imply that it is not epidemiologic but behavioral, institutional, and economic factors that should correctly determine the priority of interventions. *Am J Epidemiol* 1995;141:135–44.

developing countries; diarrhea, infantile; sanitation; water supply

Improving health in developing countries through the provision of water supply and sanitation systems has been an important goal of development agencies for decades. Due to the shortage of economic resources available to address these problems, there has been vigorous debate in recent years regarding the "optimal" sequencing of environmental interventions: Which type of intervention (improving water quality, improving access to water to encourage greater use, or improving excreta disposal) has the greatest impact on health and, as a result, should be carried out first?

A large number of studies have attempted to estimate the health effects from improving water supply and sanitation in developing countries. Many studies have reported large impacts, while others report little or no effect. While many of these studies suffer from methodological shortcomings (1, 2), even among the better designed studies, a wide range of impacts have been observed (3, 4).

While some of this variability may be attributed to differences in study design and analytical methods, there is a more fundamental issue at work as well. Where there are interactions among interventions, the effects of a particular intervention will depend not only on the intervention in question, but also on the other factors with which it interacts, notably the other interventions (3, 5, 6). For example, if the interactions are strong, the health impact from an improved water supply may depend critically on whether sanitation conditions of the community or of the family are good or poor. As discussed in detail below, the existence of such interactions has profound policy implications.

Theoretical models suggest that such interactions do exist between water supply and sanitation interventions, such that the impact from improvements in both water supply and sanitation would be greater than the sum of the effects from improving water supply or sanitation alone (3, 5, 6). However, few empirical studies have compared the effects from single and multiple interventions, with mixed results emerging.

The findings of four studies are consistent with the hypothesized effect. Two case-control studies of diarrheal disease in children (one in the Philippines (7) and one in Malawi (8)) found some evidence that improved water quality did have larger effects among households with better sanitation. However, the sam-

---

Received for publication April 15, 1993, and in final form September 19, 1994.

Abbreviation: FC, fecal coliform.

[1] University of Texas Health Science Center at Houston, School of Public Health, MPH Program at El Paso, 901 Education Building, UTEP, El Paso, TX 79968. (Reprint requests to Dr. James VanDerslice at this address.)

[2] Chief, Water and Sanitation Division, The World Bank, S4–143 1818 H Street NW, Washington, DC 20433.

ple sizes were relatively small, and the differences in effect were not statistically significant. A prospective study in Lesotho (9) found a significant interactive effect of increased water usage and latrine ownership on infant weight gain and linear growth over a 6-month period, even though only four of the 119 study infants fell into the "good water/good latrine" category.

Three other studies, however, have shown interactions opposite from those expected, with improved water quality appearing to have smaller impacts in better sanitary conditions. A study of infant mortality in Malaysia reported significant protective effects of having a piped water supply or a toilet on infant mortality. The effect of having both facilities, however, was "less than would be expected from their separate beneficial effects" (10, p. 525), suggesting that the effect of piped water was smaller in households with good sanitation. In an intervention study in the Philippines (11), cholera was monitored in four communities with different levels of water supply and sanitation. While the communities with improved water supply and/or sanitation experienced significantly less cholera than the control community, the incidence in the community provided both services was only marginally less than the single-intervention communities. A large case-control study conducted in Sri Lanka (12, 13) found some evidence that latrine users derived fewer health benefits from an improved water supply, although the authors cautioned that latrine ownership could be acting as a proxy for economic status.

In summary, although a comprehensive review of the literature concluded that improving water quality appeared to have larger health impacts in areas with better environmental sanitation (2), the results of studies that specifically addressed this issue are inconclusive. A major shortcoming in virtually all of these studies is that sample sizes were usually too small to establish the statistical significance of the main effects, much less that of the interactions.

In this paper we attempt to address this important policy issue by taking advantage of a very large, high-quality data set (emanating from the Cebu Longitudinal Health and Nutrition Survey) and extensive prior work on developing a sound understanding of the determinants of diarrheal disease in the study population (14–18). The study assesses how the impact of improved drinking water quality varies with the level of community and household sanitation.

MATERIALS AND METHODS

Data collection

Household surveys. This analysis uses data from the Cebu Longitudinal Health and Nutrition Survey, a prospective study of children from birth to 2 years of age. The study setting is the island of Cebu, located in the central Visayan region of the Philippines, and includes two large cities, four smaller towns, and the surrounding periurban areas. The region has an estimated population of 1 million. Seventeen of the 155 urban barangays (political districts) were randomly selected, and all women residing in these barangays who gave birth between April 1983 and May 1984 were recruited. Of the 2,555 eligible women, 2,355 had single live births and agreed to participate in the study. The number of children in the study decreased over the study period from 2,220 at 2 months to 1,930 at the end of the first year. Most of the attrition was due to out-migration. There did not appear to be any selection bias associated with loss to follow-up (14).

The household survey consisted of a series of interviews by highly trained local fieldworkers. In the baseline survey conducted during the third trimester, detailed demographic and socioeconomic data were collected, the household's excreta disposal facility and drinking water source were identified, and the yard was inspected for the presence of fecal material. At each bimonthly interview, data were gathered on infant feeding and food preparation practices (including water boiling), the child's height and weight, preventive health care use, and diarrheal morbidity over the previous 7 days. The household's source of drinking water was verified at each survey. More information regarding the survey design and content is available elsewhere (14, 15).

Neighborhood environmental assessments. To assess the level of exposure to fecal material in the family's neighborhood, an experienced sanitary engineer carried out a series of environmental assessments. The 17 barangays were divided into 41 homogeneous areas or "neighborhoods." Each area was rated using structured observations in terms of housing density, type of settlement (e.g., squatter or periurban), presence of observable fecal material, predominant types of excreta disposal facilities, and frequency of flooding. The same individual conducted all assessments, and each area was surveyed twice over the course of the study to check for internal consistency.

Water quality sampling. The water source used by the family for drinking was identified at each bimonthly interview, and sanitary surveys were carried out at each source. Between two and five water samples were collected from each source over the course of a year; water sources with more variable quality (such as open dug wells) were sampled more frequently. The water samples were transported on ice to a laboratory where they were analyzed the following morning. Three serial volumes (1, 10, and 100 ml) were filtered through 45-μm glass fiber filters and incubated on M-FC agar at 44.5°C for 24 hours (19). The concentration of fecal coliforms (FCs) per 100 ml was estimated from the number of dark blue colonies observed on each filter. Nine percent of the samples analyzed (154 of 1,650) were not used due to uncharacteristic colonies or heavy background growth.

Model specification

This analysis uses a model of diarrheal disease in which the explanatory variables are partitioned into proximate biologic and behavioral determinants (e.g., nutritional status, breast-feeding practice) and underlying socioeconomic determinants (e.g., income, education). The latter factors do not directly affect the child's risk of diarrhea but have an indirect effect, influencing the parents' child-care choices such as how long to breast-feed or whether to boil their child's water. Only the proximate factors thought to directly affect the child's risk are included in the model. This approach was first proposed by Mosley and Chen (20) and later adapted by the Cebu Study Team to investigate the determinants of child health and growth (14, 15, 17).

Diarrheal disease. Diarrhea is measured by a dichotomous variable that, if positive, indicates that the child experienced diarrhea in the 7 days before the interview, as reported by the mother. The local term for diarrhea used in the questionnaire (*kalibang*) denotes frequent, watery stools. In a separate study conducted in the study area, the mother's recall of diarrheal morbidity based on the observation of frequent or loose stools had a sensitivity of 95–97 percent and a specificity of 80 percent when compared with diagnoses made at health clinics and hospitals (21).

Drinking water quality. Exposure to contaminated drinking water is measured as the $\log_{10}$ daily dose of fecal coliforms. As described in detail elsewhere (18), the dose for each child was estimated by multiplying the predicted concentration of fecal coliforms in their water source at the time of the interview by the amount of water that the child consumed. This dose was adjusted for water boiling as boiling significantly reduced the risk posed by contaminated drinking water (18).

Sanitation. Three variables are used to measure exposure to feces related to excreta disposal. The first variable measures the presence of fecal material in the yard and is used as an indicator of household sanitation practices.

Individuals may also be at risk when facilities are used but are in poor condition, exposing the user to the feces of previous users. A small observational study showed marked differences in the sanitary condition of 49 private and 70 public toilet facilities used by the study households (J. DeClerque, Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, unpublished data, 1985). Fecal material was evident less frequently in private than in public facilities (0.18 vs. 0.34, $p < 0.08$), and littered paper, used for anal cleansing, was present far less frequently in private facilities (0.16 vs. 0.47, $p < 0.0001$). Accordingly, the second measure of sanitation is whether the excreta disposal facility used by the family was private or public.

The purpose of an excreta disposal facility is to isolate human wastes from the human environment so that pathogens in those wastes are not passed on to other individuals. A child whose family uses a toilet or latrine is less likely to come in contact with fecal material than a child whose family defecates indiscriminately in areas near the house. However, a child's exposure is affected not only by the way in which the family disposes of its excreta; children from households that use toilet facilities may still face considerable exposure if their neighbors do not use such facilities. Thus, a child's exposure is affected not only by his or her family's excreta disposal practices, but also by the practices of the community as a whole.

Housing density can have marked impact on the exposures due to indiscriminate defecation. In sparsely settled rural areas, defecation in fields may pose little risk to the community as a whole, while in crowded urban areas indiscriminate defecation by a small proportion of the population may significantly increase the entire community's exposure to pathogens. Thus, in high-density urban areas, inadequate excreta disposal can have the greatest impact on the transmission of diarrheal disease.

The neighborhood environmental assessments were used to identify neighborhoods facing the greatest risks due to inadequate excreta disposal. Neighborhoods with dense housing, poor drainage, and readily observable fecal material were classified as having "poor community sanitation."

Household hygiene. The level of water service (in-house vs. carried to the house) is used as a proxy for water use and water-related hygiene. The number of other preschool children and household crowding (number of persons/room) are included as measures of the likelihood of person-to-person transmission. The presence of animals in the house is used to measure exposure to pathogens from animal feces.

Feeding practices. Breast feeding may reduce or eliminate exposure to pathogens in contaminated foods (22–24) and may decrease the susceptibility to infection through antibodies present in breast milk (25, 26). Feeding practices 1 week prior to the survey are measured by two dichotomous variables indicating whether the infant was fully breast-fed or mixed-fed. Fully breast-fed infants were those who were exclusively breast-fed or who received nonnutritive liquids (e.g., water, teas, brews) in addition to breast milk. Exposure to pathogens in contaminated water used to make teas and brews was captured by the water quality variable. Mixed-fed infants were those given nutritive supplements such as formulas or gruels in addition to breast milk. Infants who were completely weaned were used as the comparison group.

Susceptibility factors. Children with poor nutritional status may be more susceptible to infection. The infant's weight at the previous interview, expressed in units of standard deviations from the sample mean of that age group, was used as a measure of nutritional status. The use of preventive health care (e.g., immunizations or well-baby checkups) in the past 2 months was also used as a measure of susceptibility.

Biologic factors. As in the other Cebu publications, changes in immunologic development over time were captured by including both age and age squared in the model (14, 15). The child's sex was included as a proxy for differential immunologic development. Rainfall in the past 2 weeks was used to capture seasonal effects such as enhanced survival of enteric pathogens in humid conditions.

Statistical methods

The severity of diarrhea for child i at time t ($D^*_{t,i}$) is specified as a function of the health-related behaviors affecting the child's exposure to pathogens and susceptibility to infection ($Y_{t-1,i}$), the child's nutritional status as measured by growth in the previous time period ($G_{t-1,i}$), and biologic exposure and susceptibility factors ($Z_{t,i}$):

$$D^*_{t,i} = \beta_1 Y_{t-1,i} + \beta_3 Z_{t,i} + \mu_{Di} + \epsilon_{Dt,i}. \tag{1}$$

However, since the severity of diarrhea is difficult to assess, the outcome used in this analysis is simply whether the mother reported the child to have experienced any diarrhea in the week preceding the interview. When the severity of diarrheal episode ($D^*_{t,i}$) exceeds some threshold, it is observed by the mother and reported at the interview:

$$D_{t,i} = \begin{cases} 0 \text{ when } D^*_{t,i} \leq 0 \\ 1 \text{ when } D^*_{t,i} > 0. \end{cases} \tag{2}$$

While several exposure and susceptibility factors are included in this analysis, there are other unobserved factors that affect the child's risk of diarrhea. These include the child's inherent ability to resist infection or particular practices in the child's family that increase or decrease the risk of diarrhea. These "omitted" factors are represented in equation 1 by the random error term μ_{Di}. The second error term ($\epsilon_{Dt,i}$) is a standard random disturbance assumed to be normally distributed and independent across individuals and through time.

Biases may arise if the unobserved factors affecting the child's risk of diarrhea are correlated with other unobserved factors affecting their exposure status. For example, a child with a poor ability to resist infection may be at a greater risk of diarrhea. If the child's mother recognizes this risk, she may be more apt to prolong breast feeding. Thus, the child's "unobserved" ability to resist infection would be correlated with his or her risk of diarrhea and with his or her feeding status. Ignoring this correlation could lead to a biased estimate of the effect of breast feeding on the risk of diarrheal disease. An instrumental variables technique, which has been described in detail elsewhere (14–16), was used to correct for this source of bias.

The data from the six longitudinal surveys covering the first year of life were combined into a data set containing one observation for each child at each point in time. A "random effects" probit model was used to describe the probability of diarrhea as a function of the explanatory variables. The random effects model specifies that the error term is made up of two components, a standard disturbance that is uncorrelated between cross-sections and time periods and an error term that is unique for each cross-section and does not vary with time. The parameters of the probit model were estimated using a maximum likelihood procedure found in the HOTZTRAN software (27).

Unlike logistic regression, the parameter estimates from a probit model cannot be directly interpreted in terms of epidemiologic measures of effect such as risk differences or risk ratios. Rather, simulations must be used to estimate these measures of effect. Once the parameters of the probit model ($\boldsymbol{\beta}$) have been estimated, a predicted probability of diarrhea can be computed given any set of values of the independent variables ($\mathbf{X}$) using the cumulative normal distribution function (ϕ):

Predicted $Pr(\text{diarrhea}|\mathbf{X}) =$

$$P(D = 1|\mathbf{X}) = \phi(\mathbf{X}\cdot\boldsymbol{\beta}) \tag{3}$$

The effect of a given risk factor is assessed by predicting the probability of diarrhea at each level of that risk factor, keeping all other variables constant. For example, the effect of having excreta in the yard is estimated by comparing the predicted probability of diarrhea when this variable is set to zero with the predicted probability when the variable is set to one.

The difference between these predicted probabilities is a measure of the excess risk associated with having excreta in the yard. Similarly, the ratio of these predicted probabilities is a measure of the relative risk. Such risk differences and risk ratios are computed for each individual. The mean risk difference and the mean risk ratio are used as the effect estimates for the study population.

Approximate confidence intervals for the effect measures were constructed by simply repeating the simulation using the end points of the 95 percent confidence interval for the coefficient of interest in place of the point estimate. The confidence intervals for the effect measures include the null value precisely when the confidence interval for the parameter estimate includes zero.

RESULTS

Demographic characteristics

There is wide variation in the social, economic, and demographic characteristics of the study population. Almost half of the households (42 percent) include extended family members, and 46 percent have at least two children other than the study infant. Education levels are quite high; more than 75 percent of the parents had completed primary education, another 5 percent had graduated from high school, and almost 10 percent had some postsecondary education. Most of the households (70 percent) were headed by waged or salaried workers; few were engaged in farming or fishing. Household incomes ranged from 0 to 12,500 pesos per week with a median of 200 pesos (approximately $520/year).

Diarrhea in Cebuano infants

The proportion of children experiencing diarrhea in the week preceding the interview increased from slightly more than 7 percent at 2 months of age to 25 percent at 8 months of age (table 1). For the remainder of the year, the prevalence among males remained virtually constant, while the prevalence among females decreased slightly to 22 percent. Over the course of the year, 39 percent of the infants did not experience diarrhea during any of the weeks preceding the six interviews, while 12 percent of the infants were frequently ill, having had diarrhea during three or more of these 1-week periods.

TABLE 1. Seven-day prevalence of diarrhea by child's age and sex, Cebu, Philippines, 1983–1985

Category	Child's age (months)					
	2	4	6	8	10	12
Overall	7.2	12.7	20.4	25.0	24.4	24.1
Females	7.1	12.3	20.1	24.6	22.9	21.9
Males	7.4	13.1	20.6	25.4	25.7	26.0

Excreta disposal

Excreta disposal was not well managed in the study area. Over half of the households used flush or pour-flush toilets, one third of which were located in the house (table 2). However, because Cebu city does not have a sewage system, on-site septic systems are used for disposal. In some cases, effluent from these cesspools and septic tanks was discharged directly into open canals. Twenty-three percent of the households used latrines, and 4 percent used open pits. Slightly more than 20 percent of the households did not use any facility, with individuals defecating in empty lots, on the seashore, or on the banks of rivers and canals.

Toilets and latrines were rarely used to dispose of infants' feces. The majority of mothers (61 percent) reported depositing these stools in places readily accessible to animals or children (e.g., under the house or in a vacant lot). Fecal material was readily observed at more than one third of the households (table 3).

Community sanitation

Those neighborhoods with very dense housing and fecal material readily observable throughout the area were rated as having very poor community sanitation. Eighteen percent of the sample infants resided in these highly contaminated areas.

TABLE 2. Type of excreta disposal facility used, Cebu, Philippines, 1983–1985

Type of excreta disposal facility	Frequency	%
Private facilities		
Cistern toilet (inside)	127	5.4
Water-sealed toilet (inside)	254	10.8
Cistern toilet (outside)	46	2.0
Water-sealed toilet (outside)	655	27.8
Public facilities		
Water-sealed toilet	194	8.2
Latrine	532	22.6
No facility		
Open pit	100	4.3
No facility used (field, canal, seashore)	438	18.6
Other	8	0.3
Unknown	1	<0.1
Total	2,355	100.0

TABLE 3. Presence of excreta around the house, Cebu, Philippines, 1983–1985

Observation categories	Frequency	%
Heavy excreta visible	178	7.6
Some excreta visible	600	25.5
Very little excreta visible	335	14.2
No excreta visible	1,219	51.8
No observation made	23	1.0
Total	2,355	100.1

While the households in these areas had fewer assets (table 4), education levels were similar. Approximately the same proportion of study households in these two areas used some type of excreta disposal facility; however, a smaller proportion of households in the poor sanitation areas had private excreta disposal facilities. Within the areas of very poor community sanitation, there was significant variability in the level of contamination around individual households; 38 percent of the sample households in these neighborhoods did not have fecal material visible in their yards.

Exposure to contaminated water

Almost all of the households had access to an improved water supply; 56 percent were served by boreholes and 29 percent by the municipal piped supply. The remaining households relied on open dug wells (5 percent) and dug wells fitted with pumps (5 percent). Boreholes and the piped supply usually provided high-quality water—more than 75 percent of the samples taken from these sources produced no fecal coliform colonies. About 10 percent of these samples, however, were quite contaminated, containing more than 100 FCs/100 ml. Dug wells had much higher levels of contamination; only 16 percent of the dug wells produced water with less than 10 FCs/100 ml, and about two thirds of the wells had counts >100 FCs/100 ml.

TABLE 4. Comparison of households in communities with good and very poor sanitation, Cebu, Philippines, 1983–1985

	Very poor community sanitation (18%)		Good community sanitation (82%)	
	Mean	SD*	Mean	SD
Household assets (pesos $\times 10^{-3}$)	7.8	20.1	13.9	55.3
Mother's education (years)	7.6	3.2	7.6	3.3
Private excreta disposal facility	0.35		0.48	
In-house water connection	0.11		0.08	
Fecal material visible outside house	0.62		0.27	

* SD, standard deviation.

More than 75 percent of the infants 2 months of age were given water as part of food or brews. Average total consumption doubled over the first year from 363 ml/day at 2 months of age to 647 ml/day at 12 months. Boiling the infant's water was quite common in this population. More than 90 percent of the mothers reported boiling the water given to their 2-month-old infants, and half still boiled water when their children were 1 year old. As a result, only a small proportion of these infants were exposed to large doses of fecal coliforms (table 5). A slightly greater proportion of children in the good sanitation areas were exposed to contaminated drinking water (19 vs. 12 percent) due to the higher levels of contamination found in the dug wells.

Water availability

Water was readily available to virtually all households. Ten percent of the households had in-house connections to the piped supply or to a borehole fitted with an electric pump, and another 48 percent had a source within 1 minute of their respective houses. Only 3 percent of the families had to walk more than 5 minutes to fetch water.

Effects of water supply and sanitation on diarrhea

The results of the diarrhea models are presented in table 6. The first, a "main effects" model, does not allow for interactions among the environmental variables. The results show that the prevalence of diarrhea is significantly greater where drinking water is contaminated, the household does not have a private excreta disposal facility, there is excreta in the yard, community sanitation is very poor, children are not fully breast-fed, and there has been substantial recent rainfall. The only result that initially seems counterin-

TABLE 5. Distribution of daily FC* doses, adjusted for water boiling, by level of community sanitation, Cebu, Philippines, 1983–1985

Daily FC dose	Very poor community sanitation		Good community sanitation		Overall	
	No.	%	No.	%	No.	%
<1	2,179	92.0	9,029	93.7	11,208	93.4
1–10	130	5.5	500	5.2	630	5.3
>10–100	56	2.4	99	1.0	155	1.3
>100	3	0.1	4	0.0	7	0.1
Total†	2,368	100.0	9,632	99.9	12,000	100.1

* FC, fecal coliform.

† Based on one observation for each child at each time period.

TABLE 6. Parameter estimates† and *t*-statistics from diarrhea model, 2–12 months of age, Cebu, Philippines, 1983–1985

Variable	Main effects		Water quality/sanitation interactions					
	Model 1		Model 2		Model 3		Model 4	
	β	*t*	β	*t*	β	*t*	β	*t*
Water quality ($\log_{10}$ FC‡ dose)	0.13	2.6***	0.19	3.3***	0.10	1.3	0.17	2.4***
No private excreta disposal	0.37	3.3***	0.37	3.3***	0.38	3.4***	0.36	3.3***
Excreta in yard	0.26	2.7***	0.25	2.5***	0.26	2.7***	0.28	2.7***
Very poor community sanitation	0.18	4.3***	0.20	4.8***	0.18	4.3***	0.17	4.2***
Water quality interacted with								
Very poor community sanitation			−0.21	−2.1**				
No private excreta disposal					−0.15	−0.5		
Excreta in yard							−0.20	−0.9
No in-house water connection	0.08	1.2	0.08	1.2	0.08	1.2	0.08	1.2
Animals in house	0.03	1.2	0.03	1.1	0.04	1.2	0.04	1.2
Number of children	−0.10	−2.1**	−0.10	−2.1**	−0.10	−2.1**	−0.10	−2.1**
Crowding (persons/room)	0.00	0.4	0.00	0.4	0.00	0.4	0.00	0.4
Fully breast-fed	−0.49	−2.3**	−0.48	−2.2**	−0.50	−2.3	−0.48	−2.3**
Mixed-fed	−0.12	−0.8	−0.11	−0.7	−0.12	−0.8	−0.12	−0.8
Preventive health care use	−0.11	−0.6	−0.08	−0.4	−0.11	−0.6	−0.11	−0.6
Lagged weight (SD‡)	−0.00	−0.1	−0.00	−0.1	−0.00	−0.1	−0.00	−0.1
Child's age (weeks)	0.05	6.9***	0.05	6.9***	0.05	6.9***	0.05	6.9***
Child's age squared	−0.00	−6.0***	−0.00	−6.0***	−0.00	−6.0***	−0.00	−6.0***
Male child	0.07	1.8*	0.07	1.8*	0.07	1.8*	0.07	1.8*
Days of rain in last 2 weeks	0.02	4.1***	0.02	4.1***	0.02	4.1***	0.02	4.1***
Intercept	−1.63	−6.9***	−1.65	−7.0***	−1.62	−6.9***	−1.64	−6.9***
Log likelihood	5,378.6		5,375.3		5,378.2		5,378.6	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

† Parameter estimates from a probit model cannot be directly interpreted in terms of epidemiologic measures of effect (see Materials and Methods).

‡ FC, fecal coliform; SD, standard deviation.

tuitive is that diarrhea is less prevalent in households with more children. On further consideration, however, the greater risk associated with a larger number of children in the home might be offset because this variable could be measuring the mother's experience and her ability to protect her children from illness.

Subsequent models (2–4 in table 6) incorporate a variety of interactions between the environmental variables. Model 2 shows that the interaction between water quality and very poor community sanitation is highly significant. The coefficient is negative and almost equal in magnitude to the water quality coefficient (0.21 ≈ 0.19). This implies, first, that among the infants living in highly contaminated areas, changes in water quality have little effect on the risk of diarrhea (0.19–0.21 ≈ 0) and, second, that in areas with better community sanitation, water quality is strongly associated ($\beta = 0.19$, $t = 3.3$) with diarrhea. In fact, the effect of water quality estimated in the interaction model is nearly 50 percent greater than estimated in the main effects model (0.19 vs. 0.13).

Model 3 shows that the interaction of water quality with the "no private excreta disposal" variable is negative, suggesting that water quality has a greater effect on diarrhea in households having private excreta disposal facilities. Similarly, model 4 shows that the interaction between water quality and excreta in the yard is negative, suggesting that water quality would have less of an impact in households with excreta in their yard. In these two cases, however, the interaction terms are not statistically significant. It should be noted that the effects of the noninteracted factors remain stable across the four different models.

The effects of contaminated drinking water were calculated using the coefficients from model 2 (table 7). As shown, in areas with good community sanitation, the risk of diarrhea increases substantially as the level of contamination increases. Infants consuming 500 ml of water per day (the average for those infants consuming water) with an average contamination level of 20 FCs/100 ml (= 100 FCs/day) face a 69 percent greater risk of diarrhea than infants consuming potable water (relative risk = 1.69, 95 percent confidence interval 1.25 to 2.11). In contrast, the results in table 7 indicate that there is no relation between water quality and the risk of diarrhea in areas with very poor community sanitation.

The coefficients from model 2 were used to estimate the expected health impacts from improving water quality, level of water service, and sanitation (table 8

TABLE 7. Risk differences and risk ratios for diarrheal disease associated with water contamination, by level of community sanitation, Cebu, Philippines, 1983–1985

Daily FC* dose	Community sanitation							
	Good				Very poor			
	Risk difference	95% CI*	Risk ratio	95% CI	Risk difference	95% CI	Risk ratio	95% CI
1	Reference		Reference		Reference		Reference	
10	0.05	0.02 to 0.08	1.32	1.12 to 1.53	–0.01	–0.05 to 0.05	0.97	0.75 to 1.22
100	0.11	0.04 to 0.18	1.69	1.25 to 2.11	–0.01	–0.10 to 0.09	0.94	0.56 to 1.46
1,000	0.17	0.06 to 0.29	2.12	1.39 to 3.00	–0.02	–0.13 to 0.15	0.91	0.39 to 1.73

* CI, confidence interval; FC, fecal coliform.

† Risk difference = (expected proportion with diarrhea at specified FC dose) – (expected proportion with diarrhea for FC dose of 1).

‡ Risk ratio = $\frac{\text{(expected proportion with diarrhea at specified FC dose)}}{\text{(expected proportion with diarrhea for FC dose of 1)}}$.

TABLE 8. Comparison of the effects of alternative environmental interventions on diarrheal disease in children: mean reductions in the predicted probability of diarrhea, Cebu, Philippines, 1983–1985

Intervention	Reduction in diarrhea for affected families (%)	95% CI*
In-house water connections	12	–7 to 28
Private excreta disposal facilities	42	19 to 60
Removing excreta around the house	30	8 to 45
For families in neighborhoods with good community sanitation, reduce water source FC* concentration from 100 FC/100 ml to:		
10 FC/100 ml	24	11 to 34
1 FC/100 ml	40	20 to 54
For families who use good quality drinking water, improving community sanitation	25	16 to 33

* CI, confidence interval; FC, fecal coliform.

and figure 1). For families who do not have in-house connections, providing such a connection would decrease the prevalence of diarrhea in children of these families by 12 percent. This effect would appear to be relatively small primarily because water was readily available to all houses in the study population.

For families who do not have private or well-maintained excreta disposal facilities, the provision of such facilities is estimated to reduce childhood diarrhea by 42 percent. Similarly, for households with excreta around the house, eliminating the excreta would result in 30 percent less diarrhea among the affected families.

As discussed previously, improving water quality would have no impact on diarrhea for infants living in crowded, highly contaminated neighborhoods. Conversely, improving the level of neighborhood sanitation would have little effect where water quality is poor. Table 8 shows the positive effect of improving water quality where community sanitation is good. Reducing the concentration of fecal coliforms from 100 to 1 per 100 ml would be expected to reduce infantile diarrhea by 40 percent in such families. And for families who use good quality drinking water, the table shows that improving the level of neighborhood sanitation would reduce diarrheal prevalence by 25 percent.

These estimated impacts are broadly consistent with the effects found in literature reviews of the better designed health impact assessments. In these assessments, the median reduction in diarrhea associated with improved water quality was 17 percent, while the median reduction associated with improved sanitation was 36 percent (4).

DISCUSSION

This study provides powerful confirmation of the importance of environmental determinants of infant diarrhea. In all models (see table 8), water quality, household sanitation, and community sanitation have

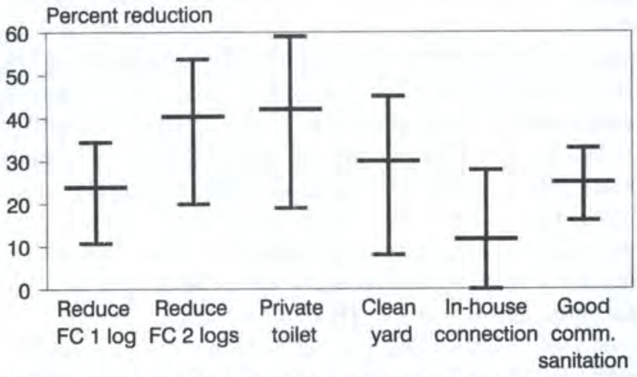


FIGURE 1. Expected reduction in diarrheal prevalence among Cebuano infants for alternative environmental interventions, Cebu, Philippines, 1983–1985.

strong, consistent, and statistically significant effects on diarrhea in infants. The novel aspect of the study, however, is the exploration of interactions among the environmental determinants of diarrhea. Here the findings of the study are both consistent with the predictions of theory and consistent internally. In all cases, the positive impact of improved water quality is greatest for families living under good sanitary conditions. Where the measure of sanitation is at the community level, this relationship is statistically significant. Where the measure is at the household level, the signs of the parameters are as expected but the effects are not statistically significant. The existence and strength of this type of interaction raise three vital policy issues.

First, there is the implication that, where interactions are important, it is impossible to draw any policy conclusions from a study of the health impact of a single intervention (e.g., of improved water quality alone). If a study shows that improving water quality alone has no health impact, does that mean that it is not important to improve water quality? In the presence of strong interactions, the answer is clearly "no," since improving water quality (or any other single risk factor) is a necessary but not sufficient condition for improving health. The conundrum this raises is the following: If the impact study shows that water quality improves health, then this can be used to justify improvements in water quality. If the impact study shows that water quality (alone) does not improve health, then this absence of effect can be attributed to interactions, and improvements in water quality can be argued for with equal force as a "necessary but not sufficient" condition. Since the conclusion from a negative finding is much the same as a conclusion from a positive finding, the finding has little policy relevance.

Second, where interactions are present but not taken into account in an impact study (as is virtually always the case), then too little impact will be attributed to the early intervention (which will pick up only the separate effects) and too much to later interventions (which will pick up the separate *and* joint effects (see 5 for a detailed discussion of this point). In the water and sanitation field, this is particularly important because, with very few exceptions, the "early" intervention is to improve water supply and the "later" intervention is to improve sanitation. The health impacts of sanitation would be overstated accordingly, and the health impacts of water supply would be understated.

Third and finally is the issue of sequencing of interventions. The cost-effective epidemiologic argument would be to first do that intervention for which the separate effect divided by the cost would be the greatest. However, there are insurmountable difficulties in translating this algorithm into practice. Because there are likely to be other interactions involving the potential interventions and contextual factors (e.g., the level of development), it would be necessary to carry out a large number of studies that take account of interactions between interventions in a variety of settings in order to paint a "universal picture" of the impacts of environmental interventions. However, as is obvious from this paper, studies that take these interactions into account are difficult to conduct, require large sample sizes and substantial analytical and financial resources, and are therefore rarely carried out.

In light of these very great conceptual and empirical difficulties, and considering the serious problems faced in financing, operating, and maintaining water and sanitation facilities in developing countries, a rational policymaker would conclude the following: "We know that people in developing countries will not be healthy until they are able to use reasonable amounts of safe, reliable water and until they have adequate excreta disposal facilities. Rather than split hairs about which intervention has the greatest health impact, pay attention to the fundamental issues (elaborated at length elsewhere (28–30)), which are the development of accountable, efficient institutions that can deliver the services people want and value."

ACKNOWLEDGMENTS

This analysis was funded through a grant from The World Bank Research Program. It is part of a collaborative research project involving the Office of Population Studies, directed by Dr. Wilhelm Flieger, and the Water Resources Center, directed by Dr. Herman van Engelen, both of the University of San Carlos, Cebu, Philippines; the Nutrition Center of the Philippines, directed by Dr. Florentino S. Solon; and a group from the Carolina Population Center, University of North Carolina at Chapel Hill (UNC-CH). Dr. Barry Popkin of UNC-CH is the project coordinator. Funding for parts of the project design, data collection, and computerization was provided by the National Institutes of Health (contracts R01-HD19983A, R01-HD23137, and R01-HD18880R), the Nestle's Coordinating Center for Nutrition Research, Wyeth International, the Ford Foundation, the United States National Academy of Sciences, the Carolina Population Center, the United States Agency for International Development, and The World Bank.

We acknowledge the valuable help of Dr. David Guilkey for the development of the statistical models and for his work on the effects of instrumental variables on the estimation of standard errors. We also thank David Fugate, Gina Dahiya, Dr. Linda Adair, Dr. Christine Moe, Arlene Quijada, Rexy Ruiz, and Ruben Quijada for their technical assistance.

REFERENCES

1. Blum D, Feachem RG. Measuring the impact of water supply and sanitation investments on diarrhoeal diseases: problems of methodology. Int J Epidemiol 1983;12:357–65.
2. Esrey SA, Habicht J-P. Epidemiologic evidence for health benefits from improved water and sanitation in developing countries. Epidemiol Rev 1986;8:117–28.
3. Esrey SA, Feachem RG, Hughes JM. Interventions for the control of diarrhoeal diseases among young children: improving water supplies and excreta disposal facilities. Bull World Health Organ 1985;63:757–72.
4. Esrey SA, Potash JB, Roberts L, et al. Health benefits from improvements in water supply and sanitation: survey and analysis of the literature on selected diseases. WASH Technical Report No. 66. Washington, DC: Water and Sanitation for Health Project, 1990.
5. Briscoe J. Intervention studies and the definition of dominant transmission routes. Am J Epidemiol 1984;120:449–55.
6. Briscoe J. A role for water supply and sanitation in the child survival revolution. Bull Pan Am Health Organ 1987;21: 93–105.
7. Baltazar J, Solon F. Disposal of faeces of children under two years old and diarrhoea incidence: a case-control study. Int J Epidemiol 1989;18(suppl):S16-S19.
8. Young B, Briscoe J. Water and health in rural Malawi. Washington, DC: US Agency for International Development, 1986.
9. Esrey SA, Habicht J-P, Casella G. The complementary effect of latrines and increased water usage on the growth of infants in rural Lesotho. Am J Epidemiol 1992;135:659–66.
10. Butz WP, Habicht J-P, DaVanzo J. Environmental factors in the relationship between breastfeeding and infant mortality: the role of sanitation and water in Malaysia. Am J Epidemiol 1984;119:516–25.
11. Azurin JC, Alvero M. Field evaluation of environmental sanitation measures against cholera. Bull World Health Organ 1974;51:19–26.
12. Mertens TE, Cousens SN, Fernando MA, et al. Health impact evaluation of improved water supplies and hygiene practices in Sri Lanka: background and methodology. Trop Med Parasitol 1990;41:79–88.
13. Mertens TE, Fernando MA, Cousens SN, et al. Childhood diarrhoea in Sri Lanka: a case-control study of the impact of improved water sources. Trop Med Parasitol 1990;41:98–104.
14. Cebu Study Team. Underlying and proximate determinants of child health: the Cebu Longitudinal Health and Nutrition Survey. Am J Epidemiol 1991;133:185–201.
15. Briscoe J, Akin J, Guilkey D. People are not passive acceptors of threats to health: endogeneity and its consequences. Int J Epidemiol 1990;19:147–53.
16. Cebu Study Team. A child health production function estimated from longitudinal data. J Dev Econ 1992;38:323–51.
17. Popkin BM, Adair L, Akin J, et al. Breast-feeding and diarrheal morbidity. Pediatrics 1990;86:874–82.
18. VanDerslice J, Briscoe J. All coliforms are not created equal: a comparison of the effects of water source and in-house water contamination on infantile diarrheal disease. Water Res Research 1993;29:1983–95.
19. American Public Health Association. Standard Methods for the Examination of Water and Wastewater. 15th ed. Washington, DC: American Public Health Association, 1985.
20. Mosley WH, Chen LC. An analytical framework for the study of child survival in developing countries. Population and Development Review 1984;10(suppl):25–45.
21. Kalter HD, Gray RH, Black RE, et al. Validation of the diagnosis of childhood morbidity using maternal child health interviews. Int J Epidemiol 1991;20:193–8.
22. Black RE, de Romanã O, Brown KH, et al. Incidence and etiology of infantile diarrhea and major routes of transmission in Huascar, Peru. Am J Epidemiol 1989;129:785–99.
23. Barrell RAE, Rowland MGM. Infant foods as a potential source of diarrhoeal illness in rural West Africa. Trans R Soc Trop Med Hyg 1979;73:85–90.
24. Black RE, Brown KH, Becker S, et al. Contamination of weaning foods and transmission of enterotoxigenic *Escherchia coli* diarrhoea in children in rural Bangladesh. Trans R Soc Trop Med Hyg 1982;76:259–64.
25. Worthington-Roberts BS. Lactation and human milk: nutritional considerations. In: Williams SR, ed. Nutrition in pregnancy and lactation. 4th ed. St. Louis: Times Mirror/Mosby College Publishing, 1989:244–322.
26. Welsh JK, May JT. Anti-infective properties of breast milk. J Pediatr 1979;94:1–9.
27. Avery RB, Hotz VJ. Hotztran user's manual. Old Greenwich, CT: CERA Economic Consultants, Inc., 1985.
28. World Bank. World development report, 1992: development and the environment. Oxford, England: Oxford University Press, 1992.
29. World Bank Water Demand Research Team. The demand for water in rural areas: determinants and policy implications. World Bank Res Obs 1993;8:47–70.
30. Briscoe J. Poverty and water supply: how to move forward. Finance and Development 1992;29:16–19.

Does Environmental Sanitation Modify the Protective Effects of Breastfeeding Against Diarrheal Disease?

Results from the Cebu Longitudinal Health and Nutrition Survey

A Report Submitted to the World Bank

Final Draft

April 6, 1993

James VanDerslice, Ph.D.
Carolina Population Center
University of North Carolina at Chapel Hill

John Briscoe, Ph.D.
The World Bank

ABSTRACT

Full breast-feeding protects infants from diarrheal disease by eliminating exposures to food-borne pathogens and improving resistance to infection. Good sanitation reduces exposure to pathogens by isolating human feces from the human environment. Together breast-feeding and good sanitation form a set of sequential barriers protecting infants from diarrheal pathogens. As a result, breast-feeding may be most important when the sanitation barrier is not in place.

This question was explored using data from the Cebu Longitudinal Health and Nutrition Survey (CLHNS), a prospective study of 2355 urban children from birth to two years of age. A random-effects probit model was used to estimate the effects of full breast-feeding and mixed-feeding while controlling for a wide range of exposure and susceptibility factors.

Breast-feeding provides significant protection against diarrheal disease for infants in all environments during the first six months of life. Fully breast-fed infants have one-third as much diarrhea as weaned infants; mixed-fed are only half as likely to develop diarrhea as completely weaned infants. There is no evidence that mixed-feeding protects infants from diarrhea after six months of age. The interactions of breast-feeding with environmental conditions are as expected for some environmental variables (water quality and neighborhood sanitation) but inconclusive for other environmental variables (water connections and private excreta disposal). More specifically, full breast-feeding has stronger protective effects when drinking water is contaminated and among infants living in crowded, highly contaminated settings. In

contrast, fully breast-fed infants from households lacking water connections or private excreta disposal do not receive any higher level of protection from full breast-feeding than infants from households having these facilities.

INTRODUCTION

Breast-feeding is an extremely effective means of protecting young infants from diarrheal disease. Infants not breast-feeding face two to three times the risk of diarrhea faced by breast-fed infants, and three to five times the risk faced by those exclusively breast-fed (1-8). A few studies have documented even larger protective effects (9-11). Supplementation of breast-feeding with even small amounts of non-milk liquids can significantly increase the infant's risk of diarrhea (3,6,11,12). As a result international health agencies recommend that infants be exclusively breast-fed for the first four to six months, and that breast-feeding be continued throughout the first two years of life (13-16).

There are two mechanisms through which breast-feeding protects infants from enteric infections. First, breast-feeding reduces or eliminates exposure to food- and water-borne pathogens. Weaning foods and breast-milk substitutes pose a particular risk as bacterial pathogens may readily multiply in these foods if they are stored at ambient temperatures after preparation (17-24). Contaminated foods and infant formulas (20,25), preparation of infant foods by an infected individual (26), lack of handwashing before food preparation (27), and storing foods without refrigeration (28,29) have been associated with diarrheal disease in children.

Secondly, breast-feeding can improve the infant's ability to resist infection. Mature breast-milk contains compounds which are thought to protect against enteric infections by inhibiting the colonization and invasion of the small intestine, interfering with bacterial multiplication and viral replication, destroying bacterial cell walls, and stimulating the infant's immune system (30-33). A number of studies have found significant associations between pathogen-specific antibody levels in breast-milk and the risk or severity of diarrhea due to that pathogen (34,35).

Good sanitation protects infants from enteric pathogens by creating a series of barriers to keep pathogens excreted in human waste from being ingested by others. Excreta disposal facilities isolate human wastes from the human environment. Improved water supplies protect drinking water from fecal contamination. Handwashing and good personal hygiene prevent the spread of enteric pathogens in the home.

As a result, poor sanitation may pose more of a risk to those who are particularly vulnerable, namely non-breast-fed infants. Weaning foods and breast-milk substitutes are more likely to be contaminated where water supply, sanitation and hygiene are lacking. Furthermore, families living in these conditions often have fewer economic resources and thus are less apt to freshly prepare foods for each meal, adequately re-heat previously prepared foods, or store foods under refrigeration. Consequently, mixed-fed and weaned infants living in poor sanitary conditions likely face considerably higher exposures to food-borne pathogens than similarly-fed infants in less contaminated environments. Thus, exclusive breast-feeding may provide greater protection among infants living in highly contaminated environments.

Simply put, breast-feeding and sanitation can be thought of as a set of sequential barriers protecting the child from enteric pathogens. Since these "barriers" are sequential, the breast-feeding barrier is most important when the sanitation barrier is not in place. Equally true, the sanitation barrier is most important when the breast-feeding barrier is absent. The hypotheses to be explored in this paper follow from this simple model:

> <u>Hypothesis 1:</u> The protective effect of breast-feeding is greatest where sanitary conditions are poor.
>
> <u>Hypothesis 2:</u> The protective effect of good sanitary conditions is greatest among those not breast-fed.

Only two published studies have addressed this issue. Using retrospective data gathered from 1262 women in the Malaysian Family Life Survey, Habicht et al. (36) found that any breast-feeding was associated with lower infant mortality, and that this protective effect was significantly stronger among households lacking toilet facilities or piped water. In a similar analysis using the same data, Butz et al. (37) assessed the effect of the duration of breast-feeding on infant mortality at various levels of sanitation. Exclusive and supplemented breast-feeding had the strongest protective effects when toilets and piped water were absent, and somewhat smaller effects in households having either one of these facilities. Breast-feeding had the smallest protective effect against infant mortality in households having both piped water and a toilet. This set of results supports Hypothesis 1: The protective effect of breast-feeding is greatest where sanitary conditions are poor.

In a case-control study of Bangladeshi children under three years of age, Clemens et al. (5) examined the relationship between breast-feeding and the risk of severe cholera. The protective effect of any breast-feeding against severe cholera infection was stronger for those infants living near (presumably contaminated) rivers, as compared to those who did not. Similarly, a stronger association was observed among infants whose families did not have a latrine. These results concur with Hypothesis 1 as well, with breast-feeding having a stronger protective effect in poor sanitary conditions. However, the protective effect of breast-feeding also appeared to be greater for families having a tubewell in their compound, a result contrary to the stated hypothesis. It should be noted that the effects of breast-feeding in the good and poor sanitary conditions were not statistically different.

The purpose of this study is to revisit this question using a larger sample from a prospective (rather than retrospective) data set, with detailed information on feeding patterns, environmental sanitation conditions, and diarrheal morbidity. Specifically, three questions are addressed in this study:

1) Is the protective effect of breast-feeding against diarrheal morbidity greatest where sanitary conditions are poor?

2) Is the protective effect of improved sanitary conditions greatest when breast-feeding is not practiced?

3) Which specific aspects of environmental sanitation are particularly important in protecting children who are not breast-fed?

These questions are of particular relevance to policy makers concerned with targeting investments in breast-feeding promotion programs and improvements in environmental sanitation.

METHODS

Data collection

Sample

This analysis uses data from the Cebu Longitudinal Health and Nutrition Survey (CLHNS), a prospective study of children from birth to two years of age. The study setting is the island of Cebu, located in the central Visayan region of the Philippines, and includes two large cities, several towns, and the surrounding peri-urban and rural areas. The region has an estimated population of about one million. This analysis uses data from infants living in urban and peri-urban areas. Seventeen of the 155 urban barangays (political districts) were randomly selected and all women residing in these barangays who gave birth between April 1983 and May 1984 were recruited. Of the 2555 eligible women, 2355 had single live births and agreed to participate in the study. The number of children in the study decreased over the study period from 2220 at two months to 1963 at the end of six months, and to 1761 at the end of two years. Attrition was primarily due to out-migration and deaths. There did not appear to be any selection bias associated with loss-to-follow-up (38).

Household surveys

The household survey consisted of 14 interviews conducted during the third trimester of pregnancy, soon after birth, and every two months until the child was two years of age. During the baseline survey, detailed demographic and socioeconomic data were collected, the household's excreta disposal facility and drinking water source were identified, and the yard was inspected for the presence of fecal material.

Data on infant feeding practices, diarrheal disease, growth, and use of preventive health care services were gathered during each bi-monthy interview, and the source of drinking water was verified. Detailed feeding information was collected for the 24-hour period preceding the interview, and included the number of times the child was suckled, duration of suckling, amount of each supplemental food given, and method of preparation for broad categories of foods (i.e., liquids, breast-milk substitutes, semi-solid foods, solid foods). General feeding practices seven days before the interview were also recorded. More information regarding the survey design and content are available elsewhere (39).

Water quality sampling

Between two and five water samples were collected from each water source over the course of a year; water sources with more variable quality (such as open dug wells) were sampled more frequently. The samples were analyzed for the presence of fecal coliform using membrane filtration (40). Nine percent of the samples analyzed (154 out of 1650) were not used due to uncharacteristic colonies or heavy background growth.

Neighborhood environmental assessments

The purpose of an excreta disposal facility is to isolate human wastes from the human environment so that pathogens in those wastes can not infect others. A child whose family uses a toilet or latrine is less likely to come in contact with fecal material than a child whose family defecates indiscriminately in areas near the house. However, even children from households with good excreta disposal practices may face considerable exposure if their neighbors defecate indiscriminately. Thus a child's exposure to fecal material is affected not only by his family's excreta disposal practices, but also by the practices of the community as a whole. This aspect of excreta disposal is referred to in this paper as community sanitation.

To assess the level of community sanitation an experienced sanitary engineer carried out a series of environmental assessments. The seventeen primary sampling units (barangays) were first divided into 41 homogeneous areas, roughly equivalent to "neighborhoods". Each area was rated using structured observations in terms of housing density, type of settlement (e.g., squatter, peri-urban, etc.), presence of observable fecal material, predominant types of excreta disposal facilities, and drainage. The same individual conducted all assessments, and each area was surveyed twice over the course of the study to check for internal consistency.

Descriptive feeding data

Infant feeding encompasses a complex set of behaviors, and apparently small differences in behavior can have a large impact on an infant's exposure to pathogenic organisms and susceptibility to infection. For the descriptive statistics infants were classified at each time period as either exclusively breast-fed, breast-fed and given only non-nutritive liquids (BF + NNL), mixed-fed, or completely weaned. Exclusively breast-fed infants received only breast-milk. The breast-fed + NNL infants were primarily breast-fed, but also given liquids lacking any caloric content, such as teas, brews, and plain water. Mixed-fed infants were those receiving nutritive foods and/or liquids in addition to breast-milk. Completely weaned infants did not breast-feed at all. These definitions are generally consistent with published recommendations (41,42). Breast-feeding categories for the descriptive statistics were assigned to infants based on the feeding practices reported by the mother for the 24-hour period preceding the interview day.

A complete 24-hour dietary recall was taken at each interview, including the amounts of all foods consumed by the child, the frequency of breast-feeding and minutes of active sucking per feeding. The dietary data were used to compute the amounts of non-nutritive liquids, breast-milk substitutes, semi-solid foods and solid foods consumed by each child.

Modeling diarrheal disease

Conceptual model

Multivariable models of diarrheal disease in children typically include a wide variety of risk factors such as infant feeding practices, water availability and quality, excreta disposal, age, sex, mother's education and household income. While each of these factors may indeed be associated with diarrheal disease, the mechanisms through which they affect a child's risk of diarrhea are quite different. Factors such as drinking water quality and sanitation are measures of exposure to enteric pathogens, while nutritional status may affect the child's susceptibility to infection.

Socioeconomic factors, in contrast, do not directly affect the risk of diarrhea, but rather, influence family behaviors which alter the child's exposure to pathogens or susceptibility to infection. For example, in this population more educated mothers are more apt to boil their infant's drinking water, reducing their child's exposure to pathogens. While education and water boiling are both associated with diarrheal disease, water boiling has a direct effect on the child's exposure to pathogens, while education has an indirect effect through its influence on the mother's child care practices. Breast-feeding, food preparation, and use of preventive health services are other examples of behaviors which affect the child's health and which are influenced by socioeconomic factors such as income and education.

This conceptual model, first proposed by Mosley and Chen (43), was adapted by the Cebu Study Team to investigate the determinants of child health and growth (38,39). In this model, diarrhea is specified as a function of factors directly affecting the child's risk of diarrhea, i.e., their exposure to pathogens and susceptibility to infection. Socioeconomic determinants, such as education, income, and prices, have their effect in the way they influence decisions the parents make regarding the care of their child. These underlying determinants of the child's health are not expected to have a direct effect on morbidity, and are not explicitly included in the model.

Model specification

Diarrhea, the outcome for this study, is measured by a dichotomous variable indicating that, according to the mother, the child experienced some episode diarrhea in the seven days before the interview. The local term for diarrhea used in the questionnaire (kalibang) denotes frequent, watery stools. In a separate study on the validity of retrospective morbidity data conducted in the study area, the mothers' reports of diarrhea based on frequent or loose stools had a sensitivity of 95% to 97%, and a specificity of 80%, when compared to diagnoses made at health clinics and hospitals (44).

For this analysis, the exclusively breast-fed and breast-fed + NNL infants were combined into a "fully breast-fed" category, indicating that the infants received all nutrition through breast-feeding. Exposure to water-borne pathogens in the non-nutritive liquids was captured by a water quality variable (discussed below). These categories were based on the reported feeding practices 8 days prior to the interview, minimizing the possibility that the reported feeding practice was a result of, rather than a determinant of, diarrhea in the week before the interview. Two dummy variables were used to estimate the effects of full breast-feeding and mixed-feeding relative to the completely weaned infants.

Given the large number of infants in this study, and the large number of foods consumed, it was not possible to directly determine levels of contamination in specific foods, nor were detailed food preparation practices recorded for each food item. However, the mother was asked how she generally prepared breast-milk substitutes, semi-solid foods and solid foods consumed by the child the previous day. Foods stored without refrigeration for more than one hour before serving were considered to be subject to bacterial multiplication, and a dichotomous variable was created indicating that the child had consumed at least one such food.

Exposure to contaminated drinking water is measured as the $\log_{10}$ daily dose of fecal coliforms adjusted for water boiling. The dose for each child was estimated by multiplying the predicted concentration of fecal coliforms in their water source at the time of the interview by the amount of water the child consumed. This dose was adjusted for water boiling as boiling was previously found to eliminate the risk posed by contaminated drinking water (45).

The level of water service is used as a proxy for water use and water-related hygiene. Having water piped into the house has been shown to dramatically increase the amount of water used. Lack of a water connection is used as a measure of low water usage and poorer household hygiene.

Families may face exposures directly related to the use of excreta disposal facilities. In a random sub-sample of the study households, private excreta disposal facilities were kept cleaner than public facilities (J. DeClerque, unpublished data). Thus private facilities may pose less of a risk to the users than public facilities or informal "defecation areas" such as vacant lots or on the banks of drainage canals. A dummy variable is used to indicate that the household does not have a private excreta disposal facility.

The presence of fecal material in the yard is used as an overall indicator of household sanitation practices. Fecal material in the yard may be from adults or children who do not use toilet facilities, or from domestic animals whose feces may contain enteric pathogens capable of infecting humans (46). The presence of animals in the house and the number of other children in the household are used as measures of potential exposures in the home.

A measure of exposure to fecal material due to poor community sanitation was developed from the neighborhood environmental assessments. Of the 41 areas surveyed, 11 were high density neighborhoods where excreta was commonly observed. The infants residing in these areas were considered exposed to poor community sanitation.

The infant's weight at the previous interview, expressed in units of standard deviations from the sample mean, was used as a measure of nutritional status. The use of a preventive health care service during the previous two months is used to capture the direct effects of these services, as well as to measure the mothers' awareness of the importance of preventive health care.

Age and age-squared are included to account for otherwise unmeasured factors which vary with time, such as the development of the infant's immune system and changes in the mother's behaviors. The infant's sex is used to capture sex-related differences in child-care practices or differences in immunological development. The number of days of rain in the past two weeks is included to account for behavioral changes related to season, or changes in environmental exposures related to rainfall, such as increased pathogen survival during humid periods.

Estimation methods

While a large number of exposure and susceptibility factors are included in this model, there are still other factors which may directly affect a child's risk of diarrhea. For example, children differ in their inherent ability to resist infection. Similarly, there may be small yet important differences in child care practices which are not represented by the available data. Many of these factors are difficult, if not impossible, to observe. Such unmeasured differences between children are referred to as unobserved heterogeneity.

In statistical models an error term is used to account for the effects of all omitted factors on the outcome variable. For simple linear regressions we assume that these omitted factors have small effects, and that the sum of these effects is, on average, equal to zero (47). Furthermore, this error term is assumed to be uncorrelated with any of the explanatory variables. If the error term is correlated with an explanatory variable, then some of the variation in the dependent variable due to the error term will erroneously be attributed to the effects of that explanatory variable, and the resulting coefficient for that variable will be biased (48).

If the exposure factors are fixed by the researcher, as in a clinical trial, then they should not be correlated with the error term and standard estimation techniques can be used to obtain unbiased effect estimates. Many of the risk factors considered here, however, result from decisions made by the child's mother (e.g., whether to exclusively breast-feed). Furthermore, these decisions may be influenced by her perception of the child's risk of diarrhea. For example, if the mother perceives her infant to be "sickly", she may be more apt to prolong exclusive breast-feeding. If her perceptions are correct, then her child care practices would be associated with the child's inherent healthiness. As a result, in a statistical model of diarrheal disease, "behavioral" risk factors such as breast-feeding could be correlated with the error term representing the (unobserved) "inherent healthiness" of the child. If standard statistical techniques are used, the resulting parameter estimates may be biased (49).

This problem is analogous to confounding. A potential confounder is a factor which is correlated with both the outcome and a risk factor of interest. If the association between the potential confounder and the outcome is not controlled for, then the estimated effect for the risk factor may be larger or smaller than the true effect (50). If unobserved factors affect the child's risk of diarrhea, and are also correlated (through the mothers' actions) with "behavioral" risk factors such as breast-feeding, then these unobserved factors may confound the relationship between the behavioral risk factor and diarrhea. In multivariable models confounding is controlled for by including the potential confounder in the model (50). These unobserved factors, however, can not be explicitly included in the model.

An alternative approach, developed by economists, is the use of instrumental variables (48). In the conceptual model presented above, socioeconomic factors were postulated to affect diarrhea through their influence on health-related behaviors. In the instrumental variable approach, a statistical model is constructed for each of the health-related behaviors, describing that behavior as a function of its socioeconomic determinants. For example, breast-feeding is modeled as a function of the parent's education, household income and assets, prices of infant foods and cooking fuel, and demographic factors. These models are then used to generate a predicted value for that behavior for each of the households.

The predicted values for the behavioral risk factors are used to represent these risk factors in the diarrhea model. Since these predicted values are simply linear combinations of socioeconomic factors, they should not be associated with the child's inherent healthiness, and thus not correlated with the error term. As such, when the predicted values are used in the diarrhea model, the resulting effect estimates will be unbiased (consistent) estimates of the true effect (48). A more detailed description of the creation and use of instrumental variables is presented elsewhere (38,51).

A "random effects" probit model was used to describe the probability of diarrhea as a function of the risk factors discussed above. The "random effects" model specifies that the error term is made up two components, a standard disturbance which is uncorrelated between children and time periods, and an error term which is different for each child but does not vary with time. This latter error term represents the unobserved heterogeneity, that is, the factors specific to each child which affect his or her risk of diarrhea. Separate models were estimated for 2 to 6, and 8 to 24 months of age to reflect changes in feeding patterns on diarrhea and in the effect of breast-feeding in early versus late infancy. The parameters were estimated using a maximum likelihood procedure found in the HOTZTRAN© software (52).

Interaction terms were used to estimate the effects of breast-feeding for different levels of sanitation. Only interactions involving full breast-feeding were used since it was expected to have the strongest differential effects. As an alternative means for assessing the effects of breast-feeding at different levels of sanitation, the sample was stratified by the sanitation variable, and separate models were estimated for each stratum. The advantage of using a stratified model is it is more general; the explanatory variables are free to have different effects on diarrhea in the two strata. The disadvantage of this procedure is that a smaller number of observations is used to estimate the parameters, increasing the variability associated with the parameter estimates.

FEEDING PRACTICES, SANITATION AND DIARRHEAL DISEASE IN CEBU

Socioeconomic and demographic characteristics

There is wide variation in the social, economic, and demographic characteristics of the study population (Table 1). One fourth of the households consist of 3 or fewer people while 22% have 8 or more. Almost half of the households (42%) include extended family members and 46% have at least 2 children other than the study infant. Education levels are quite high in urban Cebu; over 90% of the parents have completed primary education. Fifteen percent of the mothers and 18% of the fathers have graduated from high school, with two-thirds of the graduates having some post-secondary education. Most of the households are headed by workers receiving a wages or salaries (70%). One-fourth are self-employed and 3% are unemployed. Very few households are engaged in farming or fishing. Household incomes range from 0 to 12,500 pesos per week with a median of 200 pesos, which is approximately $520 per year. The top 10% of the households earn over 600 pesos per week, while the lowest 10% earn less than 80.

Breast-feeding and exposure to potentially contaminated foods

Infant feeding practices in urban Cebu are typical of those living in many developing areas. Very few mothers follow the recommended practice of exclusive breast-feeding for the first four to six months. By two months of age 38% of the infants were mixed-fed and 19% were completely weaned (Table 2, Figure 1). Another twenty percent received non-nutritive liquids in addition to breast-milk. At 4 months of age 65% of the infants are given foods or nutritive liquids. After this time the proportion of infants who are fully breast-fed drops to less than 10% by the end of the first six months. Almost half of the infants are completely weaned at one year of age, and only 12% are breast-feeding at all at the time of their second birthday.

The protection afforded by breast-feeding is related to two factors: the consumption of anti-infectious compounds in breast-milk, and the exposure to food-borne pathogens. Exclusively breast-fed infants are the most protected; they consume the largest amounts of breast-milk and are not exposed to contaminated foods or liquids. Infants given only non-nutritive liquids consume relatively small quantities of these liquids (Table 2), however, large numbers of pathogens may be ingested if the water is highly contaminated. Multiplication of bacteria in these non-nutritive teas and brews is unlikely. These infants appear to breast-feed as much as the exclusively breast-fed infants.

Mixed-fed and weaned infants may face high levels of food-borne contamination, particularly if foods are stored without refrigeration. Completely weaned infants consume two to three times as much supplemental foods by mass as the mixed-fed infants during the first year of life. In the second year mean food consumption increases substantially in the mixed-fed group, but remains relatively constant in the weaned group. These differences in consumption levels, however, may have little effect on food-borne exposures. If food is contaminated and stored without refrigeration, increases in bacterial concentrations may range over several orders of magnitude. Given this large range in potential food contamination levels, a three-fold difference the amount of food consumed may be of little practical significance. Furthermore, dose-response studies for enteric pathogens suggest that a three-fold increase in dose may lead to only a small increase in the risk of infection (53,54).

While many nutritive liquids and foods may support bacterial growth, breast-milk substitutes and semi-solid foods seem to pose the greatest risk (24,46,55). Through the first six months virtually all of the weaned children were given breast-milk substitutes (Table 3), and these foods accounted for a major part of the infant's diet (88% of total grams consumed at 2 months, 75% at 6 months). About half of the mixed-fed infants consumed breast-milk substitutes at two months of age, however, the amounts consumed were, on average, less than half that of the weaned group. Approximately one-fourth of those given these milks were consuming milks that had been stored without refrigeration for over an hour. Consumption of breast-milk substitutes by weaned children remained high through the first year of life, but as the infants consumed more at each feeding, the proportion of infants receiving "left-over" milk from a previous feeding dropped substantially.

Semi-solid foods become important by four months of age and continue to constitute a significant part of the diet through the first two years. Weaned children consume only marginally greater amounts of these foods than mixed-fed infants. Poor storage practices for these foods are prevalent through most of the second year.

While few infants under six months of age are given solid foods, almost all are consuming at least small amounts by the end of their first year. By the end of the second year solid foods account for 25% to 30% of the total diet. The proportion of children fed solid foods stored without refrigeration increased over time, possibly reflecting a shift in the childrens' diet towards adult foods.

For weaned infants, breast-milk substitutes pose the greatest risk in the first six months; almost all infants are given these milks, and they constitute the major part of these infants' diets. For the mixed-fed infants, semi-solid foods become a potentially important source of pathogens after four months of age. While the amount of semi-solid foods consumed, and presumably the frequency of consumption, is much smaller than the amount of breast-milk substitutes consumed, semi-solid foods are left unrefrigerated before serving far more frequently, particularly during later infancy. Solid foods increase in importance over the second year as they constitute a greater share of the infant's diet and as the frequency of poor food storage increases.

Environmental sanitation

The study area encompasses a wide range of environmental conditions. There is substantial variation in the level of environmental sanitation from neighborhood to neighborhood, and even between neighboring households. Thus there was no simple means to delineate "good" and "poor" environmental conditions. To capture this complex array of sanitation-related exposures, five environmental factors were considered: drinking water quality, access to water, type of excreta disposal facility, presence of excreta in the household's yard, and a measure of "community sanitation". A more detailed discussion of the environmental conditions is presented in a companion paper (51).

Unlike many areas of the developing world, water is readily available to almost all households. While only 10% of the households had an in-house water connection, less than 3% had to walk more than 5 minutes to fetch water. Furthermore, the majority of households have access to moderately good to very good quality drinking water. Over half of the households (56%) use deep boreholes fitted with handpumps or electric pumps, and 29% use a municipal piped supply serving most of the urban areas. Fecal coliform concentrations in these supplies were quite low (Geometric mean = 0.01 CFU/100 ml). The remaining 10% of the households use hand-dug wells which were frequently contaminated (Geometric mean = 195 CFU/100ml).

In contrast to the availability of water, there was no sewerage in any of the study areas. While 54% of the households used flush or pour-flush toilets, very few households had adequate means for disposing of the wastewater, which was commonly discharged into cesspools or open drainage canals. Twenty-three percent of the households used latrines, and the remaining 23% did not use any facility, defecating into open pits, empty lots or on the seashore. Toilets and latrines were rarely used to dispose of infants' feces; the majority of mothers (61%) reported depositing these stools under the house, in a vacant lot, or in other places readily accessible to animals or children. Overall, fecal material was readily observed at more than one-third of the households.

A child's exposure to fecal material depends not only on his family's defecation practices, but on the practices of neighboring households as well. To capture the community aspects of poor sanitation, each neighborhood was assessed. Thirty percent of the households were rated as being in areas of poor community sanitation, that is, neighborhoods with high housing density where excreta was frequently observed. For the most part these neighborhoods would be considered squatter areas.

Infant feeding and sanitation

Many aspects of environmental sanitation and infant feeding are influenced by the same set of underlying socioeconomic factors. For example household income is inversely related to the duration of breast-feeding and directly related to having an in-house water connection or a private excreta disposal facility. As a result, mothers in households with the highest levels of water supply and sanitation are more likely to wean their children early (Table 4). At two months of age, twice as many infants from households with private excreta disposal facility or in-house water connections were fully weaned as compared to households lacking these facilities, and only about half as many were exclusively breast-fed. These differences in feeding practices persist through the first two years of life. In contrast, there is little association between feeding practices and other environmental factors, such as the presence of excreta in the yard, or with the level of community sanitation. Still, it is striking that among the two month old infants exposed to "poor sanitary conditions", less than half are fully breast-fed, and less than 10% of these high-risk infants continue to be fully breast-fed at six months of age.

Diarrheal morbidity and breast-feeding

The proportion of children experiencing any diarrhea in the week preceding the interview rose from 7% to 20% over the first six months (Table 5), increased marginally over the next six months, and then decreased slightly during the second year. The prevalence among females was slightly less than males, with the greatest differential occurring at around one year of age.

On a simple level, there is a clear relationship between feeding practices and diarrheal disease during the first six months of life (Figure 2). At two months of age the prevalence of diarrhea among fully weaned infants is nearly three times greater than among the breast-fed infants. The gap decreases dramatically over the next two months as the prevalence of diarrhea among the breast-fed infants increases much more rapidly than in the weaned group. Between 4 and 6 months the prevalence remains constant for the fully breast-fed infants, but increases substantially among the mixed-fed and weaned infants. The protective effects of breast-feeding appear to decline over the second six months of life and there is virtually no difference in diarrheal prevalence between weaned and mixed-fed children during the second year of life.

Mean diarrheal prevalence from 2 to 6 months of age is presented in Table 6 for fully breast-fed, mixed-fed and weaned infants, stratified by four sanitation variables. Diarrheal prevalence is greater in the "poor" sanitation group for each feeding category, with one exception; fully breast-fed infants in areas of poor community sanitation have about the same prevalence of diarrhea as those in the good sanitation areas. Among weaned infants, diarrheal prevalence in each of the poor sanitation strata is quite consistent, ranging from 19.5% to 21.5%. Infants in households which have an in-house water connection have the lowest prevalence in each of the feeding groups.

STATISTICAL MODEL RESULTS

Main effects model

In the multivariable model, both breast-feeding and environmental sanitation are important determinants of diarrheal disease during the first six months (Table 7). The protective effect of full breast-feeding relative to no breast-feeding is large and statistically significant. Mixed-feeding has a somewhat smaller, yet statistically significant effect. Breast-feeding was not associated with diarrheal disease after the first six months. As such the remainder of this analysis focuses on the 2 to 6 month time period.

Poor environmental conditions are strongly associated with the risk of diarrheal disease. Consumption of contaminated water significantly increases the risk of diarrhea, independent of the type of feeding. In addition, three sanitation variables were significantly associated with diarrheal disease. Lack of a private excreta disposal facility and the presence of excreta in the yard have the strongest effects; lack of in-house water is somewhat less important. Poor community sanitation is only marginally significant, and this coefficient is small relative to the other sanitation factors. Other measures of household sanitation, (i.e., having animals in the house and inadequate food storage practices) are not associated with diarrhea.

Household assets, mother's education and mother's age were added to the model to test whether the full effects of these underlying socioeconomic factors had been captured by the direct effects of the behavioral exposure variables. This appears to be the case as none of the socioeconomic factors are significantly associated with diarrhea (t-statistics were less than 0.6).

Breast-feeding and water contamination

Many of the infants in this sample were predominantly breast-fed, supplemented by small amounts of water. Consumption of contaminated water, however, was significantly associated with diarrhea. To assess the magnitude of this risk, the coefficients from the main effects model were used to predict the probability of diarrhea for each feeding mode at three levels of water quality: no contamination, moderate contamination (10 FC/100ml), and high levels of contamination (100 FC/100ml). Mean water consumption levels in each feeding category were used in this analysis: 38 ml/day for fully breast-fed infants, 240 for mixed-fed infants, and 934 for weaned infants. The resulting predicted values are the probabilities that an infant with the given characteristics will experience diarrhea over a 7 day period, and represents the average effect for infants 2 to 6 months of age.

Exclusive breast-feeding and full breast-feeding (Breast-fed + non-nutritive liquids) supplemented with uncontaminated water are associated with the lowest risk of diarrhea (Figure 3). However, supplementing fully breast-fed infants with even small portions of contaminated water nearly doubles the risk of diarrhea, from 0.08 to 0.15. Mixed-fed and weaned infants consume much greater quantities of water, and as a result, they face much greater risks when their drinking water is contaminated. Consequently, providing good quality water is most important when the protective effects of breast-feeding are absent.

We can similarly use these results to assess the effects of breast-feeding at different levels of water quality. In general, breast-feeding has a strong protective effect; exclusively breast-fed infants have the lowest risk, weaned infants the highest risk. The increase in risk associated with weaning is greatest when drinking water is highly contaminated. Thus, breast-feeding is most important for infant health when drinking water is contaminated, that is, when there is no "sanitation barrier" keeping enteric pathogens from entering the water supply.

These results support the two stated hypotheses. First, the protective effects of breast-feeding are greatest when water quality is poor. Second, good water quality has a larger protective effect among infants who are not breast-fed.

Breast-feeding and sanitation

Contamination of infant foods is more likely where sanitary conditions are poor. As a result, weaned and mixed-fed infants may face particularly high risks due to poor sanitation. Full breast-feeding, which eliminates exposures to food-borne pathogens, thus may be most important in poor sanitary conditions.

To assess whether the protective effects of breast-feeding are greater in poor sanitary conditions, interactions of full breast-feeding with each of the sanitation variables were added to the main effects model one at a time (Table 8, models 2 - 5). The null hypothesis in each case was that the protective effects of breast-feeding would be greatest where sanitary conditions are poor. The coefficients from the main effects model are included for comparison (model 1).

Of the interaction terms, only the poor community sanitation/breast-feeding interaction coefficient is negative and marginally significant (model 2), indicating that, as expected, breast-feeding has a stronger protective effect among infants living in contaminated communities.

The interaction between no in-house water connection and breast-feeding (model 3) is positive. In this case breast-feeding appears to be more important where sanitary conditions are good. The remaining interaction terms are not significant.

It is easier to explore and explain these interactions using stratified models. In the first analysis, households were stratified by whether they had a water connection in their home. Only a small proportion of the sample had such connections, and the model estimated for these households did not produce stable coefficients.

In the second of these analyses, households were stratified by whether they resided in areas with good or poor community sanitation. Separate models were successfully estimated for these two groups (Table 9). While full breast-feeding is a significant protective factor in both areas, the magnitude of the effect in the poor sanitation areas (0.98) is about 1.5 times that in the good sanitation areas (0.61). The difference between these coefficients (0.37) is not statistically significant. Similarly, the effect of mixed feeding is greater for the infants living in contaminated communities (0.63 vs. 0.40), but the mixed-feeding coefficients are not statistically significant ($p < 0.20$).

These parameter estimates were used to predict the probability of diarrhea associated with each feeding pattern for areas of good and poor community sanitation (Figure 4). Poor community sanitation leads to a large increase in risk among the weaned infants, but has no effect on the fully breast-fed infants. Thus full breast-feeding is able to completely mitigate the risk posed by poor community sanitation.

These results can be interpreted in two ways. Improving community sanitation would clearly have a large impact on weaned infants, and virtually no impact on fully breast-fed infants. This indicates that the protective effect of good sanitation is greater when infants are not breast-fed, that is, when the "breast-feeding barrier" is absent.

Alternatively, these results indicate that breast-feeding has a stronger protective effect when sanitation is poor. Fully breast-fed infants apparently face much lower risks than weaned infants, and difference in risks is greater for infants living in neighborhoods with poor sanitation. Thus in neighborhoods with poor sanitation the protective effect of breast-feeding is stronger than in neighborhoods with good sanitation. In other words, when the "sanitation barrier" is absent, the protection afforded by breast-feeding becomes even more important.

CONCLUSION

Infant feeding practices and diarrheal morbidity patterns in urban Cebu are typical of those in many developing countries. Very few mothers follow the recommended practice of exclusive breast-feeding for the first four to six months. Most give their infants formulas or other drinks within the first two months, and an alarming proportion stop breast-feeding altogether within this time period. The prevalence of diarrhea increases dramatically over the first 8 months, particularly among weaned and mixed-fed infants.

The objective of this study was to assess the protective effects of breast-feeding against diarrheal disease under good and poor sanitary conditions. Three specific questions were posed:

1) Is the protective effect of breast-feeding greatest where sanitary conditions are poor?

2) Is the protective effect of improved sanitary conditions greatest when breast-feeding is not practiced?

3) Which specific sanitary conditions are particularly important in protecting children who are not breast-fed?

The results of this analysis do not provide unambiguous answers to these questions. Nevertheless, several important conclusions can be drawn.

Full breast-feeding provides significant protection against diarrheal disease for infants living in all environments during the first six months of life. Exclusive breast-feeding is by far the most protective. Adding even small quantities of contaminated water to the infant's diet can double the risk of diarrheal disease. Still, full breast-feeding (i.e., supplementation with only non-nutritive liquids) is much more protective than mixed-feeding, a result supported by several previous studies (1,7,9,10). While mixed-fed infants are at a higher risk than fully breast-fed infants, they are still only half as likely to develop diarrhea as completely weaned infants. The protection afforded by mixed-feeding appears to be limited to early infancy, as after six months of age mixed-fed infants are at the same risk of diarrhea as completely weaned infants.

It is clear that in some circumstances breast-feeding does provide greater protection in poor sanitary conditions. First, exclusive breast-feeding has a stronger protective effect when drinking water is contaminated. Second, breast-feeding is more protective for infants living in crowded, highly contaminated settings. In contrast, fully breast-fed infants from households lacking water connections or private excreta disposal do not receive any higher level of protection from full breast-feeding than do infants from households having these facilities. These results are in agreement with the reported effects of breast-feeding and sanitation on infant mortality (36,37) and cholera morbidity (5).

The positive interaction between breast-feeding and no in-house water connection was unexpected. The households who had in-house water connections were the most wealthy and virtually all also owned a private excreta disposal facility. Furthermore, only a relatively small proportion of the infants in these households were fully breast-fed (see table 4). Thus the interaction between no in-house water connection and full breast-feeding may have captured the protective effects of unmeasured child care and hygiene practices of those wealthy mothers who fully breast-fed their infants.

The policy implications are clear. Full breast-feeding through the first four to six months should be encouraged as a means of protecting young infants in all settings from diarrheal disease. Further, exclusive breast-feeding is strongly recommended where potable water is not readily available. However, after six months, promoting a longer duration of partial breast-feeding may do little to decrease diarrheal morbidity but are nonetheless important for other aspects of maternal and child health. Programs encouraging these feeding practices should consider targeting infants in crowded, highly contaminated areas in order to maximize health benefits.

Several sanitation indicators were significant risk factors for diarrheal disease in this study. Consumption of contaminated water, lack of private excreta disposal, and the presence of excreta in the yard were associated with the largest increases in risk. As hypothesized, the protection provided by high quality drinking water and good sanitation in the community appear to be greatest for non-breast-fed infants. A high quality water source is particularly important for mixed-fed and weaned infants as they consume the largest quantities of water. Similarly, good community sanitation would benefit completely weaned infants the most as they face significantly higher risks from living in crowded, highly contaminated communities.

While the protective effects of full breast-feeding are considerable, unsupplemented feeding can not meet the nutritional needs of infants after 4 to 6 months of age, and mixed-feeding does not appear to provide further protection against diarrheal disease after that time. As a result, efforts to reduce diarrhea during the period of highest incidence (6 to 24 months) should focus on reducing exposures, particularly food-borne exposures, through improvements in the choice of supplemental foods, preparation practices and storage methods, and better overall hygiene. Increasing the availability of water and access to excreta disposal facilities can do much to reduce the level of contamination in the community, and enable families to improve hygienic conditions in the home.

ACKNOWLEDGMENTS

This analysis was funded through a grant from The World Bank Research Program. It is part of a collaborative research project involving the Office of Population Studies, directed by Wilhelm Flieger, and the Water Resources Center, directed by Herman van Engelen, both of the University of San Carlos, Cebu, Philippines; the Nutrition Center of the Philippines, directed by Florentino S. Solon; and a group from the Carolina Population Center, University of North Carolina at Chapel Hill (UNC-CH). Barry Popkin of UNC-CH is the project coordinator. Funding for parts of the project design, data collection, and computerization was provided by the National Institutes of Health (Contract R01-HD19983A, R01-HD23137, and R01-HD18880R), the Nestle's Coordinating Center for Nutrition Research, Wyeth International, the Ford Foundation, the U.S. National Academy of Sciences, the Carolina Population Center, the U.S. Agency for International Development, and The World Bank.

REFERENCES

1. Feachem, R. G. and M. A. Koblinsky. 1984. Interventions for the control of diarrhoeal diseass among young children: promotion of breast-feeding. Bulletin of the World Health Organization 62:271-291.
2. Ahmed, F., J. D. Clemens, M. R. Rao, D. A. Sack, M. R. Khan and E. Haque. 1992. Community-based evaluation of the effect of breast-feeding on the risk of microbiologically confirmed or clinically presumptive Shigellosis in Bangladeshi children. Pediatrics 90:406-411.
3. Brown, K. H., R. E. Black, G. Lopez de Romana and H. Creed de Kanashiro. 1989. Infant-Feeding practices and their relationship with diarrheal and other diseases in Huascar (Lima), Peru. Pediatrics 83:31-40.
4. Duffy, L. C., T. E. Byers, M. Riepenhoff-Talty, L. J. La Scolea, M. Zielezny and P. L. Ogra. 1986. The effects of infant feeding on rotavirus-induced gastroenteritis: A prospective study. American Journal of Public Health 76:259-263.
5. Clemens, J. D., D. A. Sack, J. R. Harris, et al. 1990. Breast feeding and the risk of severe cholera in rural Bangladeshi children. American Journal of Epidemiology 131:400-411.
6. Morrow, A. L., R. R. Reves, M. S. West, M. L. Guerrero, G. M. Ruiz-Pallacios and L. K. Pickering. 1992. Protection against infection with Giardia lamblia by breast-feeding in a cohort of Mexican infants. The Journal of Pediatrics 121:363-370.

7. Guerrant, R. L., L. V. Kirchhoff, D. S. Shields, et al. 1983. Prospective study of diarrheal illness in Northeastern Brazil Patterns of disease, nutritional impact, etiologies, and risk factors. Journal of Infectious Diseases 148:986-997.

8. Habicht, J. P., J. DaVanzo and W. P. Butz. 1986. Does breast-feeding really save lives, or are apparent benefits due to biases?. American Journal of Epidemiology 123:279-290.

9. Mahmood, D. A., R. G. Feachem and S. R. A. Huttly. 1989. Infant feeding and risk of severe diarrhoea in Basrah city, Iraq: a case-control study. Bulletin of the World Health Organization 67:701-706.

10. Huttly, S. R. A., D. Blum, B. R. Kirkwood, R. N. Emeh and R. G. Feachem. 1987. The epidemiology of acute diarrhoea in a rural community in Imo State, Nigeria. Transactions of the Royal Society of Tropical Medicine and Hygiene 81:865-870.

11. Victora, C. G., P. G. Smith, J. P. Vaughan, et al. 1989. Infant feeding and deaths due to diarrhea. American Journal of Epidemiology 129:1032-1041.

12. Popkin, B. M., L. Adair, J. Akin, R. E. Black, J. Briscoe and W. Flieger. 1990. Breast-feeding and diarrheal morbidity. Pediatrics 86:874-882.

13. U.S.A.I.D. 1990. Breastfeeding: A Report on A.I.D. Programs. Washington, D.C.: U.S. Agency for International Development.

14. World Health Organization 1992. Programme for Control of Diarrhoeal Diseases: Eighth Programme Report, 1990-1991. Geneva: World Health Organization.

15. Grant, J.P. 1991. The State of the World's Children 1991. Oxford, U.K.: Oxford University Press.

16. Martines, J. C., M. Rea and I. de Zoyza. 1992. Breast feeding in the first six months. British Medical Journal 304:1068-1069.

17. Black, R. E., G. Lopez de Romana, K. H. Brown, N. Bravo, O. G. Bazalar and H. C. Kanashiro. 1989. Incidence and etiology of infantile diarrhea and major routes of transmission in Huascar, Peru. American Journal of Epidemiology 129:785-799.

18. Barrell, R. A. E. and M. G. M. Rowland. 1979. Infant foods as a potential source of diarrhoeal illness in rural West Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene 73:85-90.

19. Barrell, R. A. E. and M. G. M. Rowland. 1980. Commercial milk products and indigenous weaning foods in a rural West African environment: a bacteriological perspective. Journal of Hygiene Cambridge 84:191-202.

20. Black, R. E., K. H. Brown, S. Becker, A. R. M. Abdul Alim and M. H. Merson. 1982. Contamination of weaning foods and transmission of enterotoxigenic Escherchia coli diarrhoea in children in rural Bangladesh. Transactions of the Royal Society of Tropical Medicine and Hygiene 76:259-264.

21. Molbak, K., N. Hojlyng and S. Jepson. 1989. Bacterial contamination of stored water and stored food: a potential source of diarroeal disease in West Africa. Epidemiology and Infection 102:309-316.

22. Imong, S. M., K. Rungruengthanakit, C. Ruangyuttikarn, . Wongsawasdii, D. A. Jackson and R. F. Drewett. 1989. The bacterial content of infant weaning foods and water in rural northern Thailand. Journal of Tropical Pediatrics 35:14-17.

23. van Steenbergen, W., D. A. A. Mossel, J. A. Kusin and A. A. J. Jansen. 1983. Machkos project studies: Agents affecting health of mother and child in a rural area of Kenya. Tropical and Geographical Medicine 35:193-197.

24. Black, R. E., O. de Romano, K. H. Brown, N. Bravo, O. Bazalar and H. Kanashiro. 1989. Incidence and etiology of infantile diarrhea and major routes of transmission in Huascar, Peru. American Journal of Epidemiology 129:785-799.

25. Black, R. E., M. H. Merson, B. Rowe, et al. 1981. Enterotoxigenic Escherichia coli diarrhoea: acquired immunity and transmission in an endemic area. Bulletin of the World Health Organization 59:263-268.

26. Holmberg, S. D., D. E. Kay, R. D. R. Parker, et al. 1984. Foodborne transmission of cholera in Micronesian households. Lancet February 11:325-328.

27. Clemens, J. D. and B. F. Stanton. 1987. An educational intervention for altering water-sanitation behaviors to reduce childhood diarrhea in urban Bangladesh. American Journal of Epidemiology 125:284-291.

28. Ekanem, E. E., C. O. Akitoye and O. T. Adedeji. 1991. Food hygiene behaviour and childhood diarrhoea in Lagos, Nigeria: a case-control study. Journal of Diarrhoeal Disease Research 9:219-226.

29. Wright, C. E., M. El Alamy, H. L. DuPont, et al. 1991. The role of the home environment in infant diarrhea in rural Egypt. American Journal of Epidemiology 134:887-894.

30. Worthington-Roberts, B.S. 1989. Lactation and human milk: Nutritional considerations. In: Nutrition in Pregnancy and Lactation. edited by Williams, S.R. St. Louis: Times Mirror/Mosby College Publishing, pp. 244-322.

31. Welsh, J. K. and J. T. May. 1979. Anti-infective properties of breast milk. J.Ped. 94:1-9.

32. Hanson, L. A., F. Jahil, R. Ashraf, et al. 1991. Characteristics of human milk antibodies and their effect in relation to the epidemiology of breastfeeding and infections in a developing country. Advances in Experimental and Medical Biology 310:1-15.

33. Garza, C., R. . Schanler, N. F. Butte and K. J. Motil. 1987. Special properties of human milk. Clinics in Perinatology 14:11-32.

34. Ruiz-Pallacios, G. M., J. J. Calva, L. K. Pickering, et al. 1990. Protection of breast-fed infants against Campylobacter diarreha by antibodies in human milk. Pediatrics 116:707-713.

35. Glass, R. I., A. sven Svennerholm, B. J. Stoll, et al. 1983. Protection against cholera in breast-fed children by antibodies in breast milk. New England Journal of Medicine 308:1389-1392.

36. Habicht, J. P., J. DaVanzo and W. P. Butz. 1988. Mother's milk and sewage: Their interactive effects on infant mortality. Pediatrics 81:456-461.

37. Butz, W. P., J. P. Habicht and J. DaVanzo. 1984. Environmental factors in the relationship between breastfeeding and infant mortality: The role of sanitation and water in Malaysia. American Journal of Epidemiology 119:516-525.

38. Cebu Study Team 1992. A child health production function estimated from longitudinal data. Journal of Development Economics 38:323-351.

39. Cebu Study Team 1991. Underlying and proximate determinants of child health: The Cebu longitudinal health and nutrition survey. American Journal of Epidemiology 133:185-201.

40. American Public Health Association 1985. Standard Methods for the Examination of Water and Wastewater. Washington, D.C.: American Public Health Association. 15th Ed.

41. Labbok, M. and K. Krasovec. 1990. Toward consistency in breastfeeding definitions. Studies in Family Planning 21:226-230.

42. World Health Organization 1991. Indicators for Assessing Breast-feeding Practices, Report of an Informal Meeting, 11- 21 June 1991, WHO/CDD/SER/91.14. Geneva: WHO.

43. Mosley, W.H. and L.C. Chen. 1984. An analytical framework for the study of child survival in developing countries. In: Child Survival, strategies for research, Population and Development review, supplement to vol. 10. edited by Mosley, W.H. and Chen, L.C. Cambridge: Cambridge University Press, pp. 25-45.

44. Kalter, H. D., R. H. Gray, R. E. Black and S. A. Gultiano. 1991. Validation of the diagnosis of childhood morbidity using maternal child health interviews. International Journal of Epidemiology 20:193-198.

45. VanDerslice, J. and J. Briscoe. 1993. All coliforms are not created equal: A comparison of the effects of water source and in-house water contamination on infantile diarrheal disease. Water Resources Research (In Press)

46. Rowland, M. G. M., R. A. E. Barrell and R. G. Whitehead. 1978. Bacterial contamination in traditional Gambian weaning foods. Lancet January 21:136-138.

47. Alam, N., B. Wojtyniak, F. J. Henry and M. M. Rahaman. 1989. Mothers' personal and domestic hygiene and diarrhoea incidence in young children in rural Bangladesh. International Journal of Epidemiology 18:242-246.

48. Judge, G.G., R.C. Hill, W. Griffiths, H. Lutkepohl and T. Lee. 1982. Introduction to the Theory and Practice of Econometrics. New York: John Wiley and Sons.

49. Briscoe, J., J. Akin and D. Guilkey. 1990. People are not passive acceptors of threats to health: Endogeneity and its consequences. International Journal of Epidemiology 19:147-153.

50. Kleinbaum, D.G., L.L. Kupper and H. Morgenstern. 1982. Epidemiologic Research: Principles and Quantitative Methods. Belmont, CA: Lifetime Learning Publications.

51. VanDerslice, J. and J. Briscoe. 1993. The sequencing of environmental interventions in developing countries. A report prepared for the World Bank. (UnPub)

52. Avery, R.B. and V.J. Hotz. 1985. Hotztran User's Manual. Old Greenwich, Conn.: CERA Economic Consultants, Inc..

53. National Academy of Sciences 1977. Drinking Water and Health, Vol. 1. Washington, D.C.: National Academy of Sciences.

54. Schiff, G.M., G. Stefanovic, E. Young and J. Pennekamp. 1982. Determination of Minimal Infectious Dose of an Enterovirus in Drinking Water. Research Triangle Park, N.C.: U. S. Environmental rotection Agency, Health Effects Research Laboratory.

55. World Health Organization 1983. Minimum evaluation procedure for water supply and sanitation projects. ETS/83.1 CDD/OPR/83.1. Geneva: World Health Organization.

Table 1: Socioeconomic and Demographic Characteristics of the Study Households

Variable	Mean or proportion	s.d.
Child is male	0.53	
Mother's age	25.89	5.84
Father's age	28.71	6.73
Spouse is present	0.94	
Mother's highest grade completed	7.61	3.30
Father's highest grade completed	7.97	3.41
Household has extended family members	0.42	
Household has electricity	0.60	
Household owns radio	0.55	
Household owns television	0.22	
Household owns refrigerator	0.08	
Household income (pesos/wk)	231.18	362.56
Value of assets (pesos x 10^{-3})	12.99	51.49
Distance to nearest road (m)	95.36	214.20
Total rainfall in past 2 weeks (cm)	6.31	4.0
Number of days with rain in past 2 weeks	6.32	2.4

Table 2: Infant Feeding Practices, by Age

	Infant's age					
	2 mo.	4 mo.	6 mo.	12 mo.	18 mo.	24 mo.
Feeding practice (%):						
Exclusively BF[a]	23.2	16.0	2.7	< 0.1		
BF + NNL[b]	20.2	19.3	4.8	< 0.1	< 0.1	
Mixed-fed	38.0	40.6	63.9	56.4	29.6	12.0
Weaned	18.7	24.2	28.6	43.5	70.4	88.0
Mean suckling frequency (times/day):						
Exclusively BF	10.1	9.7	9.7			
BF + NNL	10.3	9.5	9.7			
Mixed-fed	8.0	7.5	7.6	7.5	7.1	6.1
Mean time of active suckling per feed (minutes):						
Exclusively BF	8.8	9.6	11.5			
BF + NNL	9.1	8.8	8.3			
Mixed-fed	8.6	8.4	7.7	7.2	6.3	5.9
Mean intake of all foods and drinks (g/day):						
BF + NNL	33	40	59			
Mixed-fed	311	333	389	683	875	1050
Weaned	1027	1193	1354	1585	1526	1579

---

[a] Exclusively breast-fed

[b] Breast-fed and given non-nutritive liquids

Table 3: Consumption and Preparation of Nutritive Foods, by Age

	Infant's age					
	2 mo.	4 mo.	6 mo.	12 mo.	18 mo.	24 mo.
Proportion receiving breast-milk substitutes (%)						
Mixed-fed	53.3	37.3	18.7	13.5	11.6	15.7
Weaned	97.0	94.1	90.8	79.9	54.2	38.0
Mean intake[a] of breast-milk substitutes (g/day):						
Mixed-fed	388	478	456	412	326	268
Weaned	899	1010	1020	922	737	652
Proportion[a] receiving milks subject to bacterial growth[b] (%):						
Mixed-fed	24.9	13.2	10.4	2.1	0.0	0.0
Weaned	29.1	15.3	11.4	5.7	3.3	1.2
Proportion receiving semi-solid foods (%):						
Mixed-fed	1.0	46.6	83.4	67.3	50.2	35.7
Weaned	3.4	42.3	77.3	69.9	52.7	35.4
Mean intake[a] of semi-solid foods (g/day):						
Mixed-fed	197	71	160	226	227	213
Weaned	81	93	195	291	263	243
Proportion[a] receiving semi-solid foods subject to bacterial growth[b] (%):						
Mixed-fed	50.0	22.8	30.4	28.8	29.5	17.3
Weaned	28.6	19.5	30.1	28.8	28.0	16.6
Proportion receiving solid foods (%):						
Mixed-fed	0.0	12.8	46.7	95.8	98.1	100.0
Weaned	0.0	6.5	34.6	91.7	98.5	99.3
Mean intake[a] of solid foods (g/day):						
Mixed-fed		29	43	131	246	321
Weaned		65	55	160	319	419
Proportion[a] receiving solid foods subject to bacterial growth[b] (%):						
Mixed-fed		14.1	11.6	25.9	40.6	40.9
Weaned		3.4	16.2	22.0	41.5	47.2

[a] Among those consuming this type of food.

[b] Foods stored without refrigeration for more than one hour.

Table 4: Feeding Patterns by Sanitation Factors, at 2 and 6 Months of Age

	Private Excreta disposal		In-house water connection		Excreta in yard		Poor community sanitation	
Feeding practice	Yes	No	Yes	No	Yes	No	Yes	No
2 months:								
Exclusively BF[a]	16.5	28.7	13.4	24.3	24.9	22.0	24.8	22.5
BF + NNL[b]	18.1	22.0	14.2	20.9	21.6	19.4	22.3	19.3
Mixed-fed	40.7	35.7	35.8	38.2	34.8	39.6	33.4	39.8
Weaned	24.8	13.6	36.6	16.6	18.7	19.0	19.5	18.4
6 months:								
Exclusively BF	1.7	3.5	1.4	2.9	3.3	2.4	2.2	3.0
BF + NNL	3.0	6.2	2.4	5.0	5.2	4.6	5.0	4.7
Mixed-fed	59.1	67.8	42.3	66.4	64.0	63.5	62.6	64.4
Weaned	36.2	22.4	53.9	25.7	27.5	29.5	30.2	27.9
12 months:								
Mixed-fed	48.0	63.4	32.1	59.1	59.0	54.8	54.8	57.1
Weaned	52.0	36.5	67.9	40.8	40.7	45.2	45.2	42.8

---

[a] Exclusively breast-fed
[b] Breast-fed and given non-nutritive liquids

Table 5: 7-day Prevalence of Diarrhea, by Child's Age, Sex and Feeding Practice

	Infant's age					
	2 mo.	4 mo.	6 mo.	12 mo.	18 mo.	24 mo.
Overall	7.2	12.7	20.4	24.1	22.9	19.2
Females	7.1	12.3	20.1	21.9	21.1	17.7
Males	7.4	13.1	20.6	26.0	24.5	20.6
Exclusively BF[a]	4.6	13.0	13.2			
BF + NNL[b]	6.2	11.3	10.4			
Mixed-fed	5.6	11.6	19.5	23.6	23.2	20.1
Weaned	14.4	15.9	23.7	24.7	22.8	19.1

[a] Exclusively breast-fed
[b] Breast-fed and given non-nutritive liquids

Table 6: 7-Day Prevalence of Diarrhea (%), and Difference in Prevalence Relative to Fully Breast-fed Infants 2 to 6 Months of Age, by Feeding Practice and Level of Sanitation

	Feeding Practice				
Sanitation Variable	Fully Breast-fed (1)	Mixed-fed (2)	(2-1)	Weaned (3)	(3-1)
Community Sanitation:					
Good	8.6	11.6	(3.0)	17.6	(9.0)
Poor	8.1	17.0	(8.9)	20.3	(12.2)
In-house water connection:					
Yes	2.0	6.2	(4.2)	13.4	(11.4)
No	8.7	13.6	(4.9)	19.5	(10.8)
Excreta in the yard:					
No	7.2	11.6	(4.4)	17.0	(9.8)
Yes	10.5	16.5	(6.0)	21.5	(11.0)
Private excreta disposal facility:					
Yes	6.4	10.2	(3.8)	17.0	(10.6)
No	9.5	15.5	(6.0)	20.5	(11.0)

Table 7: Diarrheal Disease Models, 2 to 6 and 8 to 24 Months of Age

Variable	Age Group			
	2 to 6 Months		8 to 24 Months	
	β	t	β	t
Fully breast-fed	-0.69	-2.5**		
Mixed-fed	-0.41	-1.7*	0.10	0.9
Water quality ($\log_{10}$ FC dose)	0.26	2.4**	0.10	3.1***
No in-house water connection	0.19	1.9*	0.03	0.6
No private excreta disposal	0.61	3.6***	0.18	1.9*
Excreta in yard	0.37	2.7***	0.23	2.6***
Poor community sanitation	0.08	1.6	0.05	1.5
Animals in house	0.06	1.4	0.06	2.5**
Poor food storage	0.07	0.3	0.00	0.0
Number of children	-0.11	-1.6*	-0.09	-1.9*
Preventive health care use	0.52	1.8*	-0.13	-1.3
Lagged weight (std. dev.)	0.02	0.8	-0.04	-1.5
Child's age (weeks)	0.04	1.7*	-0.00	-0.2
Child's age-squared	-0.00	-0.5	-0.00	-0.6
Male child	0.01	0.2	0.14	3.4***
Days of rain	0.02	2.5**	0.01	2.1**
Intercept	-1.33	-3.7***	-0.73	-3.9***
Number of observations	6226		16405	
Log-likelihood	2314		8620	

* - $p<0.10$, ** - $p<0.05$, *** - $p<0.01$

Table 8: Interactions of Full Breast-feeding with Sanitation Factors, 2 to 6 Months of Age

	Model 1		Model 2		Model 3		Model 4		Model 5	
Variable	β	t	β	t	β	t	β	t	β	t
Fully breast-fed	-0.69	-2.5**	-0.58	-2.0**	-0.66	-2.3**	-0.55	-1.6	-0.69	-2.0**
Mixed-fed	-0.41	-1.7*	-0.42	-1.7*	-0.37	-1.5	-0.34	-1.3	-0.41	-1.7*
Water quality (log FC dose)	0.26	2.4**	0.26	2.4**	0.26	2.4**	0.26	2.4**	0.25	2.3**
No in-house water connection	0.19	1.9*	0.19	1.9*	0.11	0.9	0.19	1.9*	0.19	1.9*
No private excreta disposal	0.61	3.6***	0.60	3.6***	0.60	3.5***	0.52	2.4**	0.61	3.6***
Excreta in yard	0.37	2.7***	0.38	2.8***	0.38	2.8***	0.38	2.8***	0.36	1.8*
Poor community sanitation	0.08	1.6	0.15	2.2**	0.08	1.6	0.08	1.6	0.08	1.6
<u>Full breast-feeding interacted with:</u>[a]										
Poor community sanitation			-0.31	-1.5*						
No in-house connection					0.75	1.5				
No private excreta disposal.							0.41	0.7		
Excreta in yard									0.02	0.0
Animals in house	0.06	1.4	0.06	1.4	0.06	1.4	0.06	1.4	0.06	1.4
Poor food storage	0.07	0.3	0.05	0.2	0.06	0.2	0.07	0.3	0.07	0.3
Number of children	-0.11	-1.6*	-0.11	-1.6*	-0.11	-1.7*	-0.12	-1.7*	-0.11	-1.6*
Preventive health care use	0.52	1.8*	0.52	1.8*	0.52	1.8*	0.49	1.7*	0.52	1.8*
Lagged weight (std. dev.)	0.02	0.8	0.02	0.9	0.02	0.8	0.02	0.8	0.02	0.8
Child's age (weeks)	0.04	1.7*	0.04	1.7*	0.04	1.7*	0.04	1.8*	0.04	1.7*
Child's age-squared	-0.00	-0.5	-0.00	-0.5	-0.00	-0.6	-0.00	-0.7	-0.00	-0.5
Male child	0.01	0.2	0.01	0.2	0.01	0.2	0.01	0.2	0.01	0.2
Days of rain	0.02	2.5**	0.02	2.5**	0.02	2.5**	0.02	2.5**	0.02	2.5**
Intercept	-1.33	-3.7***	-1.34	-3.7***	-1.35	-3.7***	-1.39	-3.7***	-1.32	-3.6***
Number of observations	6226		6226		6226		6226		6226	
Log-likelihood	2313.8		2313.2		2313.2		2313.2		2313.8	

---

* - $p<0.10$, ** - $p<0.05$, *** - $p<0.01$

[a] - The null hypothesis is that these interactions are negative. Accordingly, one-sided hypothesis tests were conducted.

Table 9: Diarrhea Model Stratified by Community Sanitation, 2 to 6 Months of Age

Variable	Community Sanitation			
	Good		Poor	
	β	t	β	t
Fully breast-fed	-0.61	-1.8*	-0.98	-1.9*
Mixed-fed	-0.40	-1.3	-0.63	-1.4
Water quality (log FC dose)	0.29	2.3**	0.16	0.7
No in-house water connection	0.26	2.0**	0.09	0.6
No private excreta disposal	0.42	2.1**	1.07	3.4***
Excreta in yard	0.39	2.3**	0.32	1.3
Animals in house	0.07	1.3	0.04	0.5
Poor food storage	-0.13	-0.4	0.38	0.8
Number of children	-0.07	-0.8	-0.23	-1.8*
Preventive health care use	0.24	0.6	0.88	1.9*
Lagged weight (std. dev.)	0.02	0.5	0.06	1.0
Child's age (weeks)	0.02	0.8	0.06	1.5
Child's age-squared	0.00	0.1	-0.00	-0.8
Male child	0.09	1.4	-0.20	-1.8*
Days of rain	0.03	2.5**	0.01	0.8
Intercept	-1.35	-3.1***	-0.89	-1.3
Number of observations	4382		1844	
Log-likelihood	1560		745	

* - p<0.10, ** - p<0.05, *** - p<0.01

Figure 1: Infant Feeding Practices, by Age

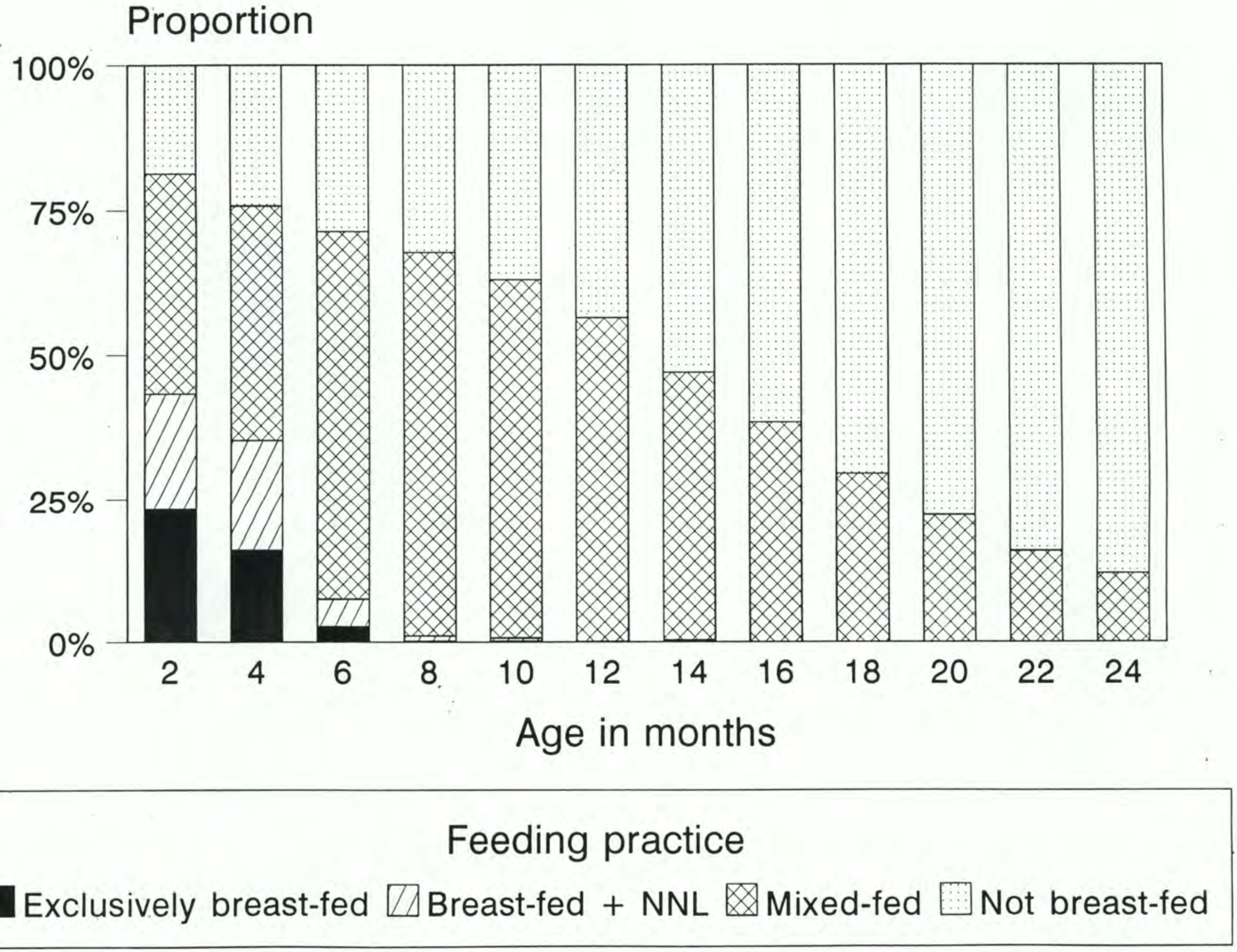
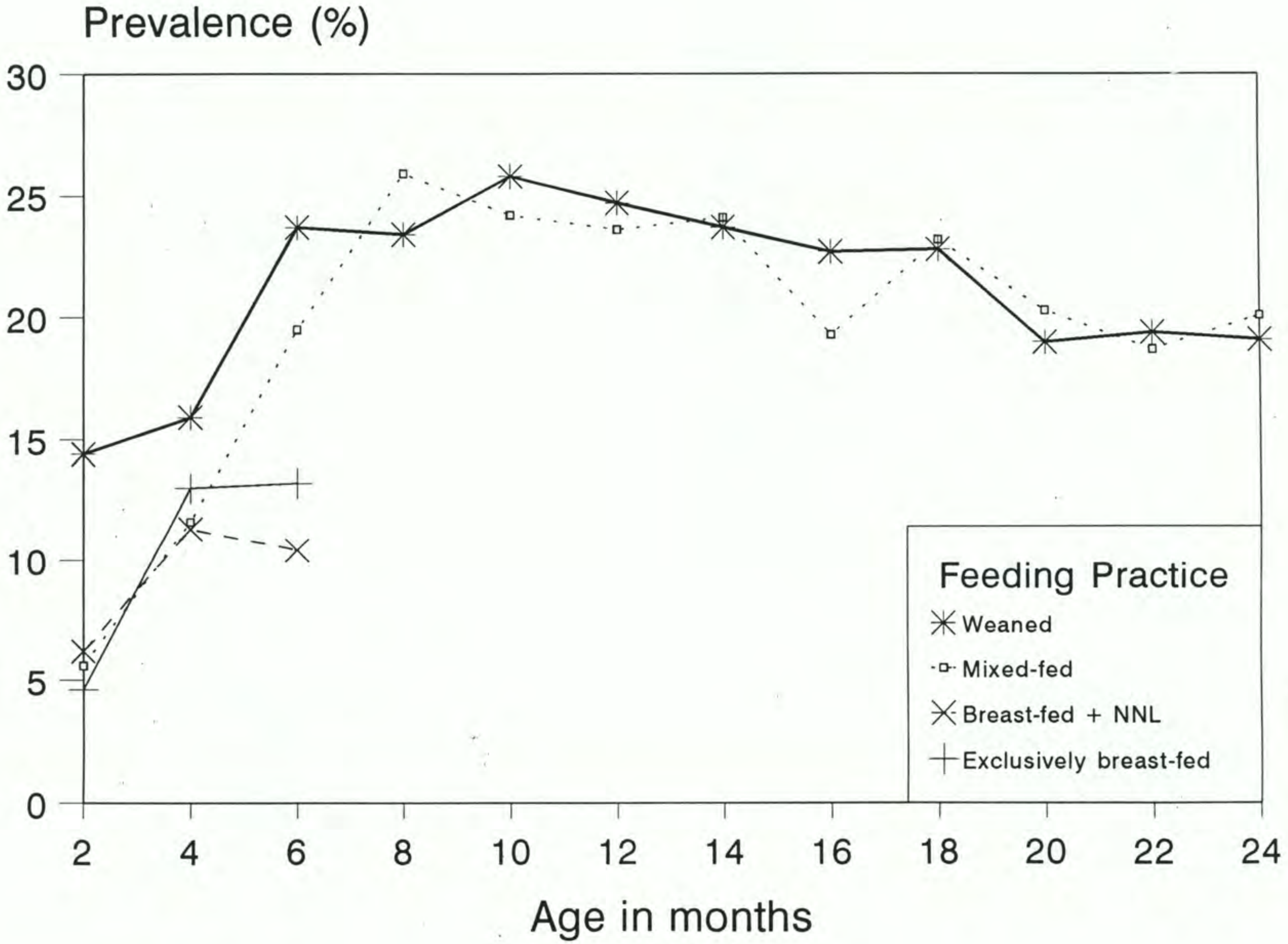
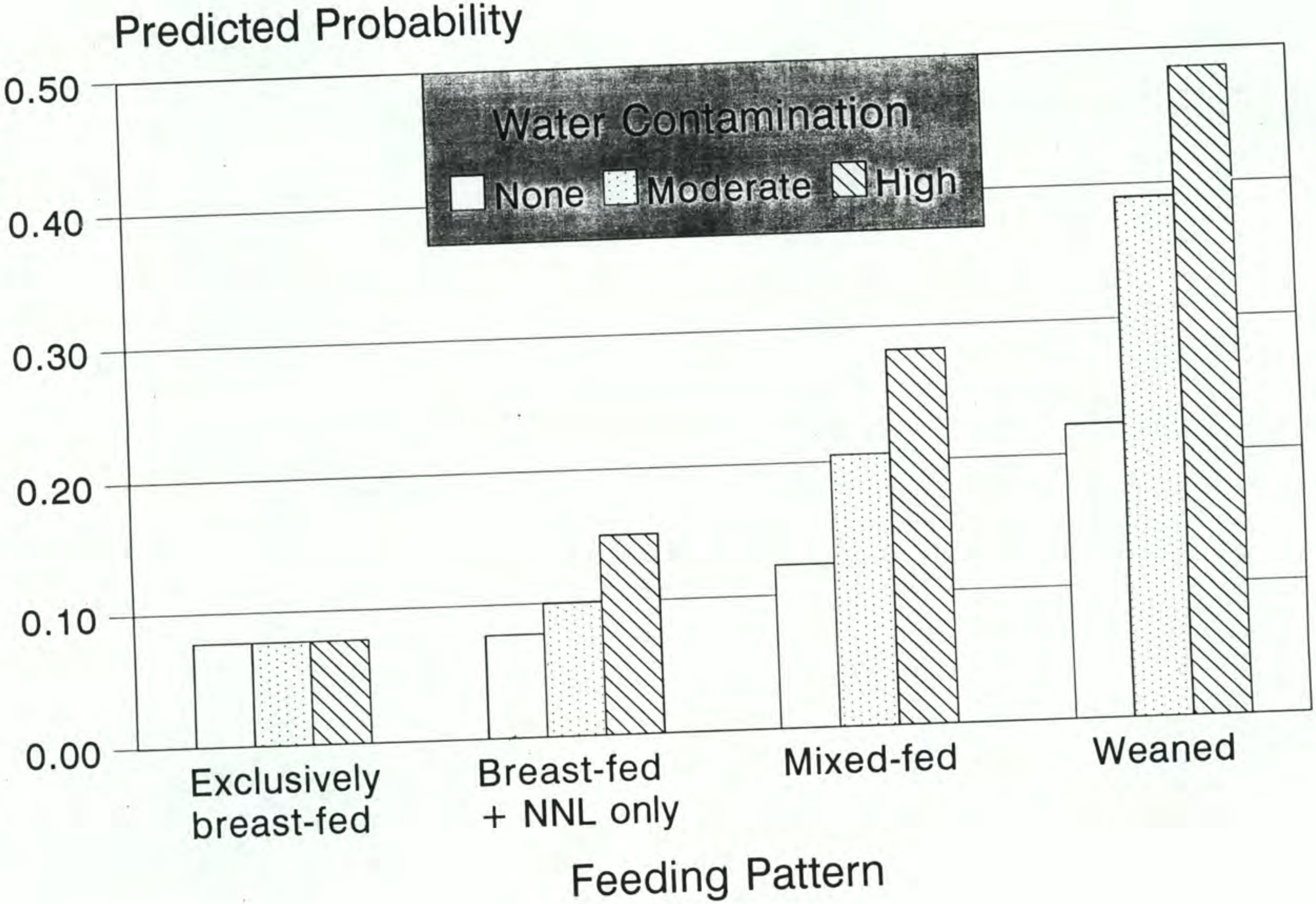


Figure 2: 7-day Prevalence of Diarrhea, by Feeding Practice



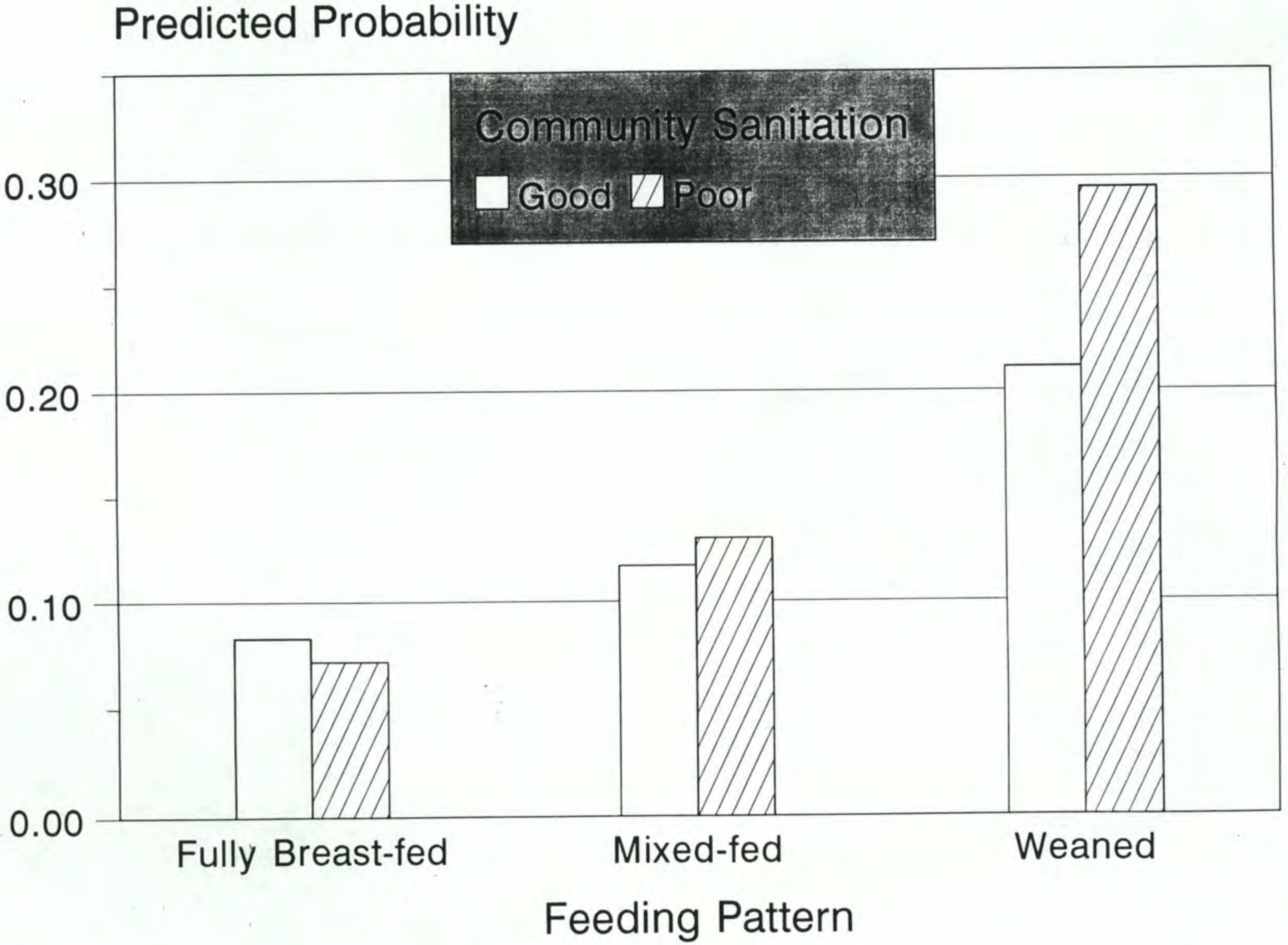
NNL - non-nutritive liquids

Figure 3: Predicted Probabilities of Diarrhea, by Feeding Pattern and Water Quality



NNL - non-nutritive liquids

Figure 4: Predicted Probabilities of Diarrhea, by Feeding Pattern and Level of Community Sanitation



ALL-IN-1 NOTE

DATE: 09-Apr-1993 05:56pm

TO: Gregory Ingram (GREGORY INGRAM)

FROM: John Briscoe, TWUWS (JOHN BRISCOE)

EXT.: 35557

SUBJECT: Research Grant

Greg:

I am sending you copies of two papers produced by myself and Dr. James Vanderslice under the grant provided to us.

I hope that you will be pleased with the outcome of this effort, which I think has produced two papers which are significant contributions both in terms of methodology and policy content.

As you can see from the letter from Dr. Vanderslice, the research specifically addressed the methodological issues brought to our attention in the review process. As you can see, both papers are being submitted to professional journals. I'll keep you apprised on the outcome.

Once again many thanks for your efficient help on this.

John

CC: Louis Pouliquen (LOUIS POULIQUEN)

THE UNIVERSITY OF NORTH CAROLINA
AT
CHAPEL HILL

Carolina Population Center
FAX: 919/966-6638
TELEX: 6501529429 6501529429MCI UW
CABLE: POPCENTER, Chapel Hill, N.C.
TELEPHONE: 919/966-2157

The University of North Carolina at Chapel Hill
CB# 8120, University Square
Chapel Hill, N.C. 27516-3997

April 6, 1993

Dr. John Briscoe
Chief, Water and Sanitation Policy Unit
The World Bank
S11 - 105
1818 H Street, NW
Washington, D.C. 20433

Dear John:

Please find enclosed final versions of the two papers describing synergistic health effects from water supply and sanitation interventions. Please note that the paper on the sequencing of environmental interventions contains an in depth discussion of the potential biases in estimating standard errors when instrumental variables are used in probit models. Monte Carlo simulations conducted by Dr. David Guilkey indicate that the estimated standard errors of the parameters of a probit model were actually greater than the empirical standard deviation of the sample of parameter estimates generated by the simulation. As such, the t-statistics presented in these papers do not appear to be grossly overestimated.

The breast-feeding interactions paper will be submitted to the Bulletin of the World Health Organization, while the paper describing the effects of multiple environmental interventions will be submitted to the American Journal of Epidemiology.

I appreciate the valuable help you have provided in this work.

Sincerely,

Jim VanDerslice, Ph.D.
Research Associate

The Sequencing of Environmental Interventions in Developing Countries: Interactions and their Implications

A Report Submitted to the World Bank

April 1, 1993

James VanDerslice, Ph.D.
Carolina Population Center
University of North Carolina at Chapel hill

John Briscoe, Ph.D.
~~The World Bank~~

ABSTRACT

This study assesses the effect of drinking water quality on diarrheal disease in good and poor sanitary conditions using a random sample of 2,355 Filipino infants over the first year of life. The results from multivariable models show that the harmful effects of contaminated drinking water are much greater in neighborhoods with good community sanitation than in neighborhoods with very poor sanitation. Similar relationships appear to pertain when sanitation is measured at the household level.

The model results were used to compare the impact of alternative environmental interventions. Improving drinking water quality would have no effect in neighborhoods with very poor environmental sanitation, but in areas with better community sanitation, reducing the concentration of fecal coliforms by two orders of magnitude would lead to a 40% reduction in diarrhea. Providing private excreta disposal would be expected to reduce diarrhea by 42% for families who do not have such facilities. Eliminating excreta around the house would lead to a 30% reduction in diarrhea for families whose living areas are contaminated.

The findings raise major policy issues. They suggest that improvements in both water supply and sanitation are necessary if infant health in developing countries is to be improved. They also suggest that health impact evaluations are of little use in setting priorities or deciding on sequences of environmental interventions and imply that it is not epidemiological, but behavioral, institutional and economic factors which should correctly determine the priority and sequencing of interventions.

INTRODUCTION

A large number of studies have attempted to estimate the health effects from improving water supply and sanitation in developing countries. Many studies have reported large impacts, while others report little or no effect. While many of these studies suffer from methodological shortcomings (1,2), even among the better designed studies, a wide range of impacts have been observed (3,4).

While some of this variability may be attributed to differences in study design and analytic methods, there may also be differences in the way that improvements in environmental sanitation affect health in disparate sanitary conditions. Theoretical models suggest that, where adequate water supply and sanitation are lacking, a single environmental intervention may not have a measurable impact on disease transmission (3,5,6). However, once one intervention is functioning and utilized, these models suggest that subsequent interventions will yield substantial health benefits. That is, the effect of any single environmental intervention may depend on the existing sanitary conditions. This implies that there may be important interactions between environmental interventions, and that the health impact from simultaneous improvements in water supply and sanitation may be much greater than the sum of the effects from improving water supply or sanitation alone.

There is little empirical evidence to support or refute this hypothesis. Few studies have attempted to compare the effects from single and multiple interventions, with mixed results emerging.

The findings of four studies are consistent with the hypothesized effect. Two case-control studies of diarrheal disease in children (one in the Philippines (7) and one in Malawi (8)) found some evidence that improved water quality did have larger effects among households with better sanitation. However, the sample sizes were relatively small, and the differences in effect were not statistically significant. A prospective study in Lesotho (9) found a significant interactive effect of increased water usage and latrine ownership on infant weight gain and linear growth over a six month period even though only 4 of the 119 study infants fell into the "good water/good latrine" category.

Three other studies, however, have shown interactions opposite from those predicted, with improved water quality appearing to have <u>smaller</u> impacts in better sanitary conditions. A study of infant mortality in Malaysia (10) reported significant effects of having a piped water supply or a toilet on infant mortality. The effect of having both facilities, however, was "less than would be expected from their separate beneficial effects", suggesting that the effect of piped water was smaller in households with good sanitation. In an intervention study in the Philippines (11), cholera was monitored in four communities with different levels of water supply and sanitation. While the communities with improved water supply and/or sanitation experienced significantly less cholera than the control community, the incidence in the community provided both services was only marginally less than the single intervention communities. A large case-control study conducted in Sri Lanka (12,13) found some evidence that latrine users derived fewer health benefits from an improved water supply, although the authors cautioned that latrine ownership could be acting as a proxy for economic status.

In summary, although a comprehensive recent review of the literature concluded that improving water quality appeared to have larger health impacts in areas with better environmental sanitation (2), the results of studies which specifically addressed this issue are inconclusive. A major shortcoming in virtually all these studies is that sample sizes were usually too small to establish the statistical significance of the main effects, much less that of the interactions.

This paper is an attempt to address this important policy issue by taking advantage of a very large, high quality data set (emanating from the Cebu Longitudinal Health and Nutrition Survey (CLHNS)) and extensive prior work on developing a sound understanding of the determinants of diarrheal disease in the study population (14-18). The study assesses how the impact of improved drinking water quality varies with the level of community and household sanitation.

METHODS

Data Collection

Household surveys

This analysis uses data from the CLHNS, a prospective study of children from birth to two years of age. The study setting is the island of Cebu, located in the central Visayan region of the Philippines, and includes two large cities, four smaller towns, and the surrounding peri-urban areas. The region has an estimated population of 1 million. Seventeen of the 155 urban barangays (political districts) were randomly selected and all women residing in these barangays who gave birth between April 1983 and May 1984 were recruited. Of the 2555 eligible women, 2355 had single live births and agreed to participate in the study. The number of children in the study decreased over the study period from 2220 at two months to 1930 at the end of the first year. Most of the attrition was due to out-migration. There did not appear to be any selection bias associated with loss-to-follow-up (14).

The household survey consisted of a series of interviews by highly trained local fieldworkers. In the baseline survey, conducted during the third trimester, detailed demographic and socioeconomic data were collected, the household's excreta disposal facility and drinking water source were identified, and the yard was inspected for the presence of fecal material. At each bi-monthly interview data was gathered on infant feeding and food preparation practices (including water boiling), the child's height and weight, preventive health care use, and diarrheal morbidity over the previous seven days. The household's source of drinking water was verified at each survey. More information regarding the survey design and content is available elsewhere (14,15).

Neighborhood environmental assessments

The purpose of an excreta disposal facility is to isolate human wastes from the human environment so that pathogens in those wastes are not passed on to other individuals. A child whose family uses a toilet or latrine is less likely to come in contact with fecal material than a child whose family defecates indiscriminately in areas near the house. However, a child's exposure is affected not only by the way in which the family disposed of its excreta: children from households which use toilet facilities may still face considerable exposure if their neighbors do not use such facilities. Thus a child's exposure is affected not only by his family's excreta disposal practices of his family, but also by the practices of the community as a whole.

To assess the level of exposure to fecal material in the neighborhood, an experienced sanitary engineer carried out a series of environmental assessments. The seventeen primary sampling units (barangays) were first divided into 41 homogenous areas, or "neighborhoods". Each area was rated using structured observations in terms of housing density, type of settlement (e.g., squatter or peri-urban), presence of observable fecal material, predominant types of excreta disposal facilities, and frequency of flooding. The same individual conducted all assessments, and each area was surveyed twice over the course of the study to check for internal consistency. Those neighborhoods with very high housing densities and fecal material readily observable throughout the area were rated as having very poor community sanitation.

Water quality sampling

The water source used for drinking was identified at each interview, and sanitary surveys were carried out at each source. Between two and five water samples were collected from each source over the course of a year; water sources with more variable quality (such as open dug wells) were sampled more frequently. The samples were transported on ice to a laboratory where they were analyzed the following morning. Three serial volumes (1, 10 and 100 ml) were filtered through 45-μm glass fiber filters and incubated on M-FC agar at 44.5°C for 24 hours (19). The concentration of fecal coliforms (FC) per 100ml was estimated from the number of dark blue colonies observed on each filter. Nine percent of the samples analyzed (154 out of 1650) were not used due to uncharacteristic colonies or heavy background growth.

Diarrhea model specification

This analysis uses a model of diarrheal disease in which the explanatory variables are partitioned into proximate biological and behavioral determinants (such as nutritional status and breast-feeding practice) and underlying socio-economic determinants (such as income and education). The latter factors do not directly affect the child's risk of diarrhea, but have an indirect effect, influencing the choices parents make regarding the care of their child, such as how long to breast-feed or whether to boil their child's water. This approach was first proposed by Mosley and Chen (20), and later adapted by the Cebu Study Team to investigate the determinants of child health and growth (15).

Diarrheal disease

Diarrhea is measured by a dichotomous variable, which, if positive, indicates that the child experienced diarrhea in the seven days before the interview, as reported by the mother. The local term for diarrhea used in the questionnaire (kalibang) denotes frequent, watery stools. In a separate study conducted in the study area, the mother's recall of diarrheal morbidity based on the observation of frequent or loose stools had a sensitivity of 95% to 97%, and a specificity of 80%, when compared to diagnoses made at health clinics and hospitals (21).

Drinking water quality

Exposure to contaminated drinking water is measured as the $\log_{10}$ daily dose of fecal coliforms. As described in detail elsewhere (18), the dose for each child was estimated by multiplying the predicted concentration of fecal coliforms in their water source at the time of the interview by the amount of water the child consumed. This dose was adjusted for water boiling as boiling significantly reduced the risk posed by contaminated drinking water (18).

Sanitation

Three variables are used to measure exposure to feces related to excreta disposal. The first variable measures the presence of fecal material in the yard and is used as an indicator of household sanitation practices.

Individuals may also be at risk when facilities are used but are in poor condition, exposing the user to the feces of previous users. A small observational study showed marked differences in the sanitary condition of 49 private and 70 public toilet facilities used by the study households (J. DeClerque, unpublished data). Fecal material was evident less frequently in private than public facilities (18% vs. 34%, $p<0.08$), and littered paper, used for anal cleansing, was present far less frequently (16% vs. 47%, $p<0.0001$) in private facilities. Accordingly, the second measure of sanitary condition is whether the facility used was private or public.

Exposure to feces in the neighborhood also depends on the proportion of people who are not using a toilet facility, and on the type of settlement. In sparsely settled rural areas, defecation in fields may pose little risk to the community as a whole, while in crowded urban areas indiscriminate defecation by a small proportion of the population may significantly increase the entire community's exposure to pathogens. Thus it is in high-density urban areas that inadequate excreta disposal can have the greatest impact on the transmission of diarrheal disease. The neighborhood environmental assessments were used to identify neighborhoods where community-level sanitation conditions posed the greatest threat. These were typically squatter settlements with high housing density, poor drainage, and fecal material readily observed in all areas.

Household hygiene

The level of water service (in-house or carried to the house) is used as a proxy for water use and water-related hygiene. The number of other pre-school children and household crowding (number of persons/room) are included as measures of the likelihood of person-to-person transmission. The presence of animals in the house is used to measure exposure to pathogens from animal feces.

Feeding practices

Breast-feeding may reduce or eliminate exposure to pathogens in contaminated foods (22-24), and decrease the susceptibility to infection through antibodies present in breast milk (25,26). Feeding practices one week prior to the survey are measured by two dichotomous variables indicating whether the infant was fully breast-fed or mixed-fed. Fully breast-fed infants were those who were exclusively breast-fed, or who received only non-nutritive liquids (e.g., water, teas, brews) in addition to breast-milk. Exposure to pathogens in contaminated teas and brews was captured by the water quality variable. Mixed-fed infants were those given nutritive supplements, such as formulas or gruels, in addition to breast-milk. Infants who were completely weaned acted as the comparison group.

Susceptibility factors

Children with poor nutritional status may be more susceptible to infection. The infant's weight at the previous interview, expressed in units of standard deviations from the sample mean, is used as a measure of nutritional status. The use of preventive health care (e.g., immunizations or well-baby check-ups) in the past two months is also used as a measure of susceptibility.

Biologic factors

As in the other Cebu publications, changes in immunological development over time are captured by a non-linear term which depends on both age and age-squared (14,15). The child's sex is included as a proxy for differential immunological development. Rainfall in the past two weeks is used to capture seasonal effects, such as enhanced survival of enteric pathogens in humid conditions.

Statistical Methods

Structural models

Detailed descriptions of the statistical methods used in this analysis have been presented elsewhere (14-16). This discussion focuses on the issues to be faced in estimating the effects of environmental variables on diarrheal disease. The severity of diarrhea for child i at time t ($D^*_{t,i}$) is specified as a function of the health-related behaviors affecting the child's exposure to pathogens and susceptibility to infection ($Y_{t-1,i}$), the child's nutritional status as measured by growth in the previous time period ($G_{t-1,i}$), and biologic exposure and susceptibility factors ($Z_{t,i}$):

$$D^*_{t,i} = \beta_1 Y_{t-1,i} + \beta_2 G_{t-1,i} + \beta_3 Z_{t,i} + \mu_{Di} + \epsilon_{Dt,i}. \qquad (1)$$

When the severity of diarrheal episode exceeds some threshold, it is observed by the mother and reported at the interview:

$$D_{t,i} = \begin{cases} 0 & \text{when} \quad D^*_{t,i} \le 0 \\ 1 & \text{when} \quad D^*_{t,i} > 0 \end{cases} \qquad (2)$$

While several exposure and susceptibility factors are included in this analysis, there are other, unobserved factors which affect the child's risk of diarrhea. These include the child's inherent ability to resist infection, or particular practices in that child's family which increase or decrease their risk of diarrhea. These "omitted" factors, referred to as unobserved heterogeneity, persist through time and are represented in equation 1 by the random error term μ_{Di}. This source of variation is assumed to be unique to each individual and normally distributed. The second error term ($\epsilon_{Dt,i}$) is a standard random disturbance assumed to be normally distributed and independent across individuals and through time. This error term represents "random events" which affect the infant's risk of diarrhea.

Socioeconomic factors, such as household income, mother's education and prices, are not included as they do not directly affect the child's risk of diarrhea. Rather, socioeconomic conditions are one set of factors which influence how parents care for their child. These health-related behaviors ($Y_{t,i}$) can be described as a function of past behaviors ($Y_{t-1,i}$), growth ($G_{t-1,i}$), diarrheal morbidity ($D_{t-1,i}$) and respiratory infections ($R_{t-1,i}$) in the previous time period, and existing socioeconomic conditions ($Z_{t,i}$):

$$\begin{aligned} Y_{t,i} = \alpha_1 G_{t-1,i} + \alpha_2 D_{t-1,i} + \alpha_3 R_{t-1,i} + \\ \alpha_4 Y_{t-1,i} + \alpha_5 Z_{t,i} + \mu_{Yi} + \epsilon_{Yt,i}. \end{aligned} \tag{3}$$

This equation also includes two error terms. The first error term, μ_{Yi}, represents unobserved characteristics of the mother or household which affect health-related child care practices in the home. For example, a mother's perception of whether her child is sickly or robust may influence her decision to prolong breast-feeding or boil the child's water. The second error term, $\epsilon_{Yt,i}$, represents purely random disturbances uncorrelated through time or across individuals. Both of these terms are assumed to normally distributed.

A mother's perception of her child's healthiness, captured by μ_Y, may well be associated with the child's actual state of health, captured in part by μ_D. Thus the error terms representing unobserved factors in the two structural equations are correlated:

$$\text{Cov}\ (\mu_{Di}, \mu_{Yi}) \neq 0. \qquad (4)$$

To complete the system which defines how a child's health is determined, similar structural equations can be written describing growth and respiratory infection as functions of health-related behaviors, previous morbidity and growth, underlying socioeconomic and biologic variables, and unobserved heterogeneity. When estimating the parameters of the model, it is important to note the simultaneous nature of the system. Diarrhea, respiratory infections, growth and the health-related behaviors are jointly determined by each other and by the exogenous socioeconomic and biologic factors. In statistical jargon, these are "endogenous" variables.

This approach to the study of diarrheal disease differs from traditional epidemiologic research in several ways. First the model is biologically-based, including only factors thought to have direct effects on the child's risk of diarrhea. It is specifically these exposure and susceptibility factors, such as infant feeding practices, water quality and sanitation, that are of interest to program planners and policy makers. Second, the model acknowledges that many of these factors are not fixed characteristics, but rather, result from choices that the parents make. Furthermore, these choices may be affected by the health of the child, as well as by the parent's unique perception of their child's healthiness. These refinements capture an important behavioral aspect of the processes affecting a child's health, that parents may recognize otherwise unobserved risks to their child's health, and modify their behaviors to reduce those risks (16).

Estimating the parameters

This system of structural equations provides a more realistic model of the processes affecting child health. However, it also reveals potential problems in obtaining unbiased estimates of the health effects of water supply and sanitation. Consider equations 1, 3 and 4 presented above where the equation 3 has been re-written for time period t-1:

$$D^*_{t,i} = \beta_1 Y_{t-1,i} + \beta_2 G_{t-1,i} + \beta_3 Z_{t,i} + \mu_{Di} + \epsilon_{Dt,i}. \qquad (1)$$

$$\begin{aligned} Y_{t-1,i} = {} & \alpha_1 G_{t-2,i} + \alpha_2 D_{t-2,i} + \alpha_3 R_{t-2,i} + \\ & \alpha_4 Y_{t-2,i} + \alpha_5 Z_{t-1,i} + \mu_{Yi} + \epsilon_{Yt-1,i}. \end{aligned} \qquad (3)$$

$$\mathrm{Cov}\ (\mu_{Di}, \mu_{Yi}) \neq 0. \qquad (4)$$

The health-related behaviors at t-1 ($Y_{t-1,i}$) are a function of, and thus correlated with, the unobserved factors affecting these behaviors (μ_{Yi}), such as the parent's perception of whether their child is sickly or robust. Furthermore, these unobserved factors are correlated with unobserved factors directly affecting the child's risk of diarrhea, such as their inherent ability to resist infection (eq. 4). As a result, the behavioral regressors (Y_{t-1}) in equation 1 (the diarrhea structural equation) are correlated with the error component μ_D. If standard techniques are used to estimate the parameters of equation 1, the correlation between the endogenous variable Y_{t-1} and the error term μ_D would lead to a biased estimate of the effect Y on the risk of diarrhea. This is because the variation in the dependent variable due to unobserved factors is incorrectly attributed to Y through its correlation with the error term μ_D.

Instrumental variables

Several methods are available for obtaining consistent parameter estimates for endogenous regressors in structural models. An instrumental variables approach is used in this work. An instrument is a variable which is correlated with the endogenous explanatory variable, but not correlated with the error term. An instrument for an endogenous behavioral risk factor can be created starting with the structural model for that variable (eq. 3). If the endogenous right-hand-side variables are substituted out using their respective structural equations, equation 3 becomes:

$$Y_{t,i} = \delta_1 G_{t-2,i} + \delta_2 D_{t-2,i} + \delta_3 R_{t-2,i} + \delta_4 Y_{t-2,i} + \delta_5 Z_{t,i} + \delta_6 Z_{t-1,i} + \mu_i + \epsilon_{t,i}. \quad (5)$$

The remaining endogenous variables at t-2 can be successively substituted out in a similar manner until only exogenous variables remain:

$$Y_{t,i} = \gamma_1 Z_{t,i} + \gamma_2 Z_{t-1,i} + \ldots + \gamma_s Z_{t-s,i} + \mu_i + \epsilon_{t,i}. \quad (6)$$

This is the reduced-form of structural equation 3. It specifies the behavioral factors as a function of present and past exogenous variables. The reduced-form can be used to create instruments for the endogenous variables. Due to the large number of exogenous variables used, it was infeasible to include all lagged values of the exogenous variables. In general only contemporaneous values of the exogenous variables were used:

$$Y_{t,i} = \gamma\ Z_{t,i} + \mu_i + \epsilon_{t,i}. \tag{7}$$

The exogenous variables, such as education or prices, are not correlated with the unobserved factors μ or the random error ϵ. As such equation 7 can be estimated using standard techniques. The resulting parameter estimates can be used to generate consistent predicted values for the endogenous behavioral variables:

$$\hat{Y}_{t,i} = \hat{\gamma}\ Z_{t,i} \tag{8}$$

These predicted values are used as instruments as they are correlated with the original endogenous variable, and not correlated with the unobserved factors represented by the error term μ. When instrumental variables are used in place of the endogenous variables in the structural models, there is no longer any correlation between the regressors and the error terms, and the resulting parameter estimates are consistent estimates of the "true" effect (27).

Inference using instrumental variables

While the predicted value of an endogenous variable for a given individual is an unbiased (consistent) prediction, there is a degree of uncertainty associated with these predictions as they are derived from <u>estimates</u> of the reduced-form parameters; the true values of these parameters are unknown. However, the uncertainty associated with the predicted values is not taken into account when they are used to estimate the parameters of the diarrhea model (eq. 1): the predicted values are treated as fixed when in fact they are not. As a result, the standard errors of the estimated parameters may be underestimated, and the associated t-statistics may be inflated (28). This would overstate the significance level of a hypothesis test that the parameter is significantly different from zero. While it is theoretically possible to obtain correct standard errors (28), it is infeasible given the large number of instruments used in this model.

A recent Monte Carlo study explored the use of several econometric methods for estimating the effects of endogenous variables on a dichotomous outcome (29). Several simulations were run varying the reduced-form R^2, the structural model R^2, the correlation between the exogenous variables in the reduced-form and structural equation, and the correlation between the error terms of the structural equations. Under each set of conditions a data set was generated from an underlying model, and then used to estimate the parameters of that model. This process was repeated 1,000 times generating 1,000 parameter estimates for each set of conditions.

To assess the bias in the estimated standard errors from ignoring the variability in the instrumental variables, the empirical standard deviation of the 1,000 parameter estimates was compared to the average of the 1,000 estimates of the standard error for that parameter. Contrary to expectation, the average estimated standard error was actually greater than the empirical standard deviation for almost all sets of conditions. The ratio of average standard error estimate to the empirical standard deviation ranged from 0.98 to 1.19.

This appears to result from a decrease in the explanatory power of the model when instruments are used (D. Guilkey, personal communication). The instruments are not perfect representations of the actual values; in cross-sectional studies, the reduced-form R^2s are typically low, ranging from 0.10 to 0.30. The difference between the actual and predicted values can be written as the difference between the reduced-form model (eq. 7) and the equation used to create the predicted values (eq. 8):

$$Y_{t,i} - \hat{Y}_{t,i} = (\gamma\, Z_{t,i} + \mu_i + \epsilon_{t,i}) - \hat{\gamma}\, Z_{t,i} \tag{9}$$

Asymptotically the estimate of γ converges on its true value and there is no difference between the actual and predicted values. However, in finite samples there will be differences between the actual and predicted values for each individual. Thus, when instruments are used in place of the actual values in the structural equation, these differences become a source of random variation. As a result, the level of unexplained variation in the dependent variable is increased, which increases the estimated standard errors of the parameters.

In summary, there are two factors affecting the estimates of the standard errors. Treating the predicted values as fixed ignores a source of variability and underestimates the standard errors. On the other hand, using instruments increases the unexplained variability in the structural model, which increases the estimated standard errors. The Monte Carlo results suggest that when the reduced-form R^2s are low, the latter effect is predominant, and the estimated standard errors will on average be somewhat overstated.

Estimation procedure

The actual procedure is a two-stage process. In the first stage, reduced-form equations are specified and estimated for each of the endogenous variables. The full reduced-form (eq. 6) would include all exogenous variables from the behavioral, growth, diarrhea and respiratory infection structural equations for the present and all previous time periods. However, estimating reduced-forms using all these variables would be infeasible. Accordingly, a smaller group of the exogenous variables thought to be the most important were identified and used for all instruments (Table 1). Other relevant exogenous variables were included on a case-by-case basis. Only contemporaneous values of the exogenous variables were used. Separate reduced-form equations were estimated for each time-period. OLS estimation was used for continuous endogenous variables; tobit and probit estimators were used for censored and dichotomous variables. When endogenous variables were censored due to an endogenous process, the instrument was created using a simultaneous equation approach. For example, food preparation practices were recorded only for those children fed supplemental foods. The first equation modeled whether the child was fed supplements, and the second equation modeled whether food preparation practices were good or poor for those children fed supplements. These two reduced-form equations were jointly estimated allowing the error terms to be correlated. This procedure corrects for the bias due to endogenous sample selectivity (28).

In the second stage, the exogenous variables and instrumental variables from the six longitudinal surveys covering the first year of life were combined into a data set containing one observation for each child at each point in time (approx. 12,000 observations). A "random effects" probit model was used to describe the probability of diarrhea as a function of the explanatory variables. The "random effects" model specifies that the error term is made up of two components, a standard disturbance which is uncorrelated between cross-sections and time periods, and error term which is unique for each cross-section and does not vary with time. The parameters of the probit model were estimated using a maximum likelihood procedure found in the HOTZTRAN© software (30).

Estimating effect measures and confidence intervals

Unlike logistic regression, the parameter estimates from a probit model can not be directly interpreted in terms of epidemiologic measures of effect, such as risk differences or risk ratios. Rather, simulations must be used to estimate these effects. Once the parameters of the probit model (β) have been estimated, a predicted probability of diarrhea can be computed for any set of values of the independent variables ($\mathbf{X}$), using the cumulative normal distribution function (Φ).

$$\text{Predicted Pr(diarrhea}|\mathbf{X}) = P(D=1|\mathbf{X}) = \Phi\ (\mathbf{X}\ \beta) \tag{10}$$

The effect of a given risk factor is assessed by predicting the probability of diarrhea at each level of that risk factor, keeping all other variables constant. For example, the effect of having excreta in the yard is estimated by comparing the predicted probability of diarrhea when this variable is set to zero to the predicted probability when the variable is set to one. The difference between these predicted probabilities is a measure of the excess risk associated with having excreta in the yard. Similarly, the ratio of these predicted probabilities is a measure of the relative risk. Such risk differences and risk ratios are computed for each individual. The mean risk difference and mean risk ratio are used as the effect estimates for the study population.

Approximate confidence intervals for the effect measures were constructed by simply repeating the simulation using the end points of the 95% confidence interval for the coefficient of interest in place of the point estimate. The confidence intervals for the effect measures include the null value only when the confidence interval for the parameter estimate includes zero.

RESULTS

Demographic characteristics

There is wide variation in the social, economic, and demographic characteristics of the study population. Almost half of the households (42%) include extended family members and 46% have at least 2 children other than the study infant. Education levels are quite high; over 75% of the parents had completed primary education, another 5% had graduated from high school, and almost 10% had some post-secondary education. Most of the households (70%) were headed by waged or salaried workers; few were engaged in farming or fishing. Household incomes ranged from 0 to 12,500 pesos per week with a median of 200 pesos (approximately $520/year).

Diarrhea in Cebuano infants

The proportion of children experiencing diarrhea in the week preceding the interview increased from just over 7% at 2 months of age to 25% at eight months of age (Table 2). For the remainder of the year, the prevalence among males remained virtually constant, while the prevalence among females decreased slightly to 22%. Over the course of the year, 39% of the infants did not experience diarrhea during any of the weeks preceding the six interviews, while 12% of the infants were frequently ill, having had diarrhea during 3 or more of these one-week periods.

Excreta disposal

Excreta disposal was not well managed in the study area. Over half of the households used flush or pour-flush toilets, one-third of which were located in the house (Table 3). However, Cebu city does not have a sewerage system so on-site septic systems are used for disposal. In some cases effluent from these cesspools and septic tanks was discharged directly into open canals. 23% of the households used latrines, and 4% used open pits. Just over 20% of the households did not use any facility, defecating in empty lots, on the seashore, or on the banks of rivers and canals.

Toilets and latrines were rarely used to dispose of infants' feces. The majority of mothers (61%) reported depositing these stools in places readily accessible to animals or children (e.g., under the house or in a vacant lot). Fecal material was readily observed at more than one-third of the households (Table 4).

Community sanitation

Those neighborhoods with very high housing densities and fecal material readily observable throughout the area were rated as having very poor community sanitation. Eighteen percent of the sample infants resided in these highly contaminated areas.

While the households in these areas had fewer assets (Table 5), they had similar education levels. Even within the areas of very poor community sanitation there was variability in the level of contamination; 38% of the sample households in these neighborhoods did not have fecal material visible in their yards. While approximately the same proportion of study households in these two areas used some type of excreta disposal facility, a much higher proportion of households in the good sanitation areas had private excreta disposal facilities.

Exposure to contaminated water

Almost all of the households had access to an improved water supply; 56% were served by boreholes and 29% by the municipal piped supply. The remaining households relied on open dug wells (5%) and dug wells fitted with pumps (5%). Boreholes and the piped supply usually provided high quality water -- over 75% of the samples taken from these sources produced no fecal coliform (FC) colonies. About 10% of these samples, however, were quite contaminated, containing more than 100 FC/100ml. Dug wells had much higher levels of contamination; only 16% of the dug wells produced water with less than 10 FC/100ml, and about two-thirds of the wells had counts greater than 100 FC per 100ml.

Over 75% of the infants at two months of age were given water as part of foods or brews. Average total consumption doubled over the first year from 363 ml per day at 2 months to 647 ml per day at 12 months. Boiling the infant's water was quite common in this population. Over 90% of the mothers reported boiling the water given to their 2 month-old infants, and half still boiled water when their child was one year of age.

As a result, only a small proportion of these infants were exposed to large doses of fecal coliforms (Table 6). A slightly greater proportion of children in the good sanitation areas were exposed to contaminated drinking water (19% vs. 12%) due to the higher levels of contamination found in the dug wells.

Water availability

Water was readily available to virtually all households. Ten percent of the households had in-house connections to the piped supply or to a borehole fitted with an electric pump, and another 48% had a source within one minute of their house. Only 3% of the families had to walk more than 5 minutes to fetch water.

Effects of water supply and sanitation on diarrhea

The results of the diarrhea models are presented in Table 7. The first, "main effects model" does not allow for interactions among the environmental variables. The results show that the prevalence of diarrhea is significantly greater where: drinking water is contaminated; the household does not have a private excreta disposal facility; there is excreta in the yard; community sanitation is very poor; children are not fully breast-fed; and there has been substantial recent rainfall. The only result which initially seems counter-intuitive is that diarrhea is lower where there are more children. Upon further consideration, however, the greater risk associated with a larger number of children in the home might be offset because this variable might be measuring the mother's experience and her ability to protect her children from illness.

Subsequent models (2-4 in Table 7) incorporate a variety of interactions between the environmental variables. Model 2 shows that the interaction between water quality and very poor community sanitation is highly significant. The coefficient is negative and almost equal in magnitude to the water quality coefficient ($0.21 \simeq 0.19$). This implies, first, that among the infants living in highly contaminated areas, changes in water quality have little effect on the risk of diarrhea ($0.19 - 0.21 \simeq 0$) and, second, that in areas with better community sanitation, water quality is strongly associated ($\beta=0.19$, $t=3.3$) with diarrhea. In fact, the effect of water quality estimated in the interaction model is nearly 50% greater than estimated in the main effects model (0.19 vs. 0.13).

Model 3 shows that the interaction of water quality with the "no private excreta disposal facility" variable is negative, suggesting that water quality has a greater effect on diarrhea in households having private excreta disposal facilities. Similarly, model 4 shows that the interaction between water quality and excreta in the yard is negative, suggesting that water quality would have less of an impact in households with excreta in their yard. In these two cases, however, the interaction terms are not statistically significant. It should be noted that the effects of the non-interacted factors remain stable across the four different models.

The effects of contaminated drinking water were calculated using the coefficients from model 2 (Table 8). Columns a and b show that in areas with good community sanitation the risk of diarrhea increases substantially as the level of contamination increases. Infants consuming 500 ml of water per day (average for those infants consuming water) with an average contamination level of 20 FC/100ml (= 100 FC/day) face a 69% greater risk of diarrhea than infants consuming potable water (RR=1.69, 95% CI = 1.25 to 2.11). In contrast, columns c and d show that there is no relationship between water quality and the risk of diarrhea in areas with very poor community sanitation.

The coefficients from model 2 were used to estimate the expected health impacts from improving water quality, level of water service, and sanitation (Table 9 and Figure 1). Line (a) of Table 9 pertains to families who do not have in-house connections. Providing in-house connections to such families would decrease the prevalence of diarrhea in children in these families by 12%. This effect would appear to be relatively small primarily because water was readily available close to the houses of all in this study population.

Line (b) shows the effect of providing private (or well-maintained) excreta disposal facilities to those who do not have such facilities. Such an intervention is estimated to reduce childhood diarrhea by 42% in such families. Line (c) similarly shows that eliminating excreta around house would result in 30% less diarrhea among the affected families.

As discussed previously, improving water quality would have no impact on diarrhea for infants living in crowded, highly-contaminated neighborhoods and, conversely, improving the level of neighborhood sanitation would have little effect where water quality is poor. Line (d) of Table 9 shows the positive effect of improving water quality where community sanitation is good. Reducing the concentration of fecal coliforms from 100 to 1 per 100ml would be expected to reduce diarrhea by 40% in such families. And line (e) shows the positive effect of improving neighborhood sanitation where water source quality is good. Improving the level of neighborhood sanitation would reduce diarrheal prevalence by 25% among families which use good quality water.

These estimated impacts are broadly consistent with the effects found in literature reviews of the better-designed health impact assessments, in which the median reduction in diarrhea associated with improved water quality was 17%, while the median reduction associated with improved sanitation was 36% (4).

DISCUSSION

This study provides powerful confirmation of the importance of environmental determinants of infant diarrhea. In all models (see Table 7), water quality, household sanitation and community sanitation have strong, consistent and statistically significant effects on diarrhea in infants. The novel aspect of the study, however, is the detailed exploration of interactions among the environmental determinants of diarrhea. Here the findings of the study are both consistent with the predictions of theory, and internally consistent. In all cases the positive impact of improved water quality is greatest for families living under good sanitary conditions. Where the measure of sanitation is at the community level, this relationship is statistically significant; where the measure is at the household level the signs of the parameters are as expected but the effects are not statistically significant.

These results have important policy implications. By far the most important implication is to confirm that there would be very large health benefits in water supply and sanitation conditions were improved in Cebu. But there are a variety of more subtle implications, too.

The study provides strong evidence, consistent both internally and with theory, regarding the interactions among various environmental interventions. There are several suggestive ways of stating these findings. The first is that, where environmental conditions are poor, only an integrated "package" of environmental interventions will have a marked impact on health. The second is that improvement in one aspect of environmental quality is a necessary but not sufficient condition for health improvements. A third is that, where overall environmental quality is poor, the first intervention, no matter what it may be, will have little direct impact on health. A fourth is that the importance of an initial improvement in environmental quality should be measured not by the direct short-term effect but by the degree to which that intervention makes other, subsequent, interventions effective.

One obvious and curious conclusion is that the information from health impact studies which do not (as most do not) assess interactions is of relatively little use to planners. Consider a case in which a study in a poor community shows that there is no discernible impact when water quality is improved. The simplistic interpretation would be (and frequently has been) that scarce resources should not be used for improving water quality in such a community. The results of the interactions investigated in the Cebu study show that such a conclusion is seriously misleading; it is only through making such apparently ineffective investments that subsequent investments (in, say, sanitation) will have an effect on health. Since the conclusion for policy is the same whether or not a health impact study has a positive finding, the logical conclusion is that the findings of such a study are irrelevant for policy.

The second conclusion of the findings on interactions relates to priorities and sequencing. Because of the interactions it is impossible to set priorities or define a particular sequence of interventions on epidemiological grounds. All that the epidemiology tells us is, in the present context, that improvements in both water supply and sanitation are necessary, and very important. How such a package of improvements is to be implemented has little to do with epidemiology, and much to do with perceived needs, economics and institutions (subjects which the authors have explored at length in other contexts (31,32,33)).

ACKNOWLEDGMENTS

This analysis was funded through a grant from The World Bank Research Program. It is part of a collaborative research project involving the Office of Population Studies, directed by Wilhelm Flieger, and the Water Resources Center, directed by Herman van Engelen, both of the University of San Carlos, Cebu, Philippines; the Nutrition Center of the Philippines, directed by Florentino S. Solon; and a group from the Carolina Population Center, University of North Carolina at Chapel Hill (UNC-CH). Barry Popkin of UNC-CH is the project coordinator. Funding for parts of the project design, data collection, and computerization was provided by the National Institutes of Health (Contract R01-HD19983A, R01-HD23137 and R01-HD18880R), the Nestle's Coordinating Center for Nutrition Research, Wyeth International, the Ford Foundation, the U.S. National Academy of Sciences, the Carolina Population Center, the U.S. Agency for International Development, and The World Bank.

We acknowledge the valuable help of David Guilkey for the development of the statistical models and for his work on the effects of instrumental variables on the estimation of standard errors. We also thank David Fugate, Gina Dahiya, Linda Adair, Christine Moe, Arlene Quijada, Rexy Ruiz, and Ruben Quijada for their technical assistance.

REFERENCES

1. Blum, D. and R. G. Feachem. 1983. Measuring the Impact of Water Supply and Sanitation Investments on Diarrhoeal Diseases: Problems of Methodology. International Journal of Epidemiology 12:357-365.
2. Esrey, S. A. and J. P. Habicht. 1986. Epidemiologic evidence for health benefits from improved water and sanitation in developing countries. Epidemiologic Reviews 8:117-128.
3. Esrey, S. A., R. G. Feachem and J. M. Hughes. 1985. Interventions for the control of diarrhoeal diseases among young children: improving water supplies and excreta disposal facilities. Bulletin of the World Health Organization 63:757-772.
4. Esrey, S.A., J.B. Potash, L. Roberts and C. Shiff. 1990. Health benefits from improvements in water supply and sanitation: Survey and analysis of the literature on selected diseases. WASH Technical Report No. 66. Washington, D.C.: Water and Sanitation for Health Project.
5. Briscoe, J. 1984. Intervention studies and the definition of dominant transmission routes. American Journal of Epidemiology 120:449-455.
6. Briscoe, J. 1987. A role for water supply and sanitation in the child survival revolution. Bulletin of the Pan American Health Organization 21:93-105.
7. Baltazar, J. and F. Solon. 1989. Disposal of faeces of children undre two years old and diarrhoea incidence: A case-control study. International Journal of Epidemiology 18(Suppl.):S16-S19.
8. Young, B. and J. Briscoe. 1986. Water and Health in Rural Malawi. Washington,D.C.: USAID.

9. Esrey, S. A., J. P. Habicht and G. Casella. 1992. The complementary effect of latrines and increased water usage on the growth of infants in rural Lesotho. American Journal of Epidemiology 135:659-666.

10. Butz, W. P., J. P. Habicht and J. DaVanzo. 1984. Environmental factors in the relationship between breastfeeding and infant mortality: The role of sanitation and water in Malaysia. American Journal of Epidemiology 119:516-525.

11. Azurin, J. C. and M. Alvero. 1974. Field evaluation of environmental sanitation measures against cholera. Bulletin of the World Health Organization 51:19-26.

12. Mertens, T. E., S. N. Cousens, M. A. Fernando, et al. 1990. Health impact evaluation of improved water supplies and hygiene practices in Sri Lanka: background and methodology. Tropical Medicine and Parasitology 41:79-88.

13. Mertens, T. E., M. A. Fernando, S. N. Cousens, B. R. Kirkwood, T. F. de C. Marshall and R. G. Feachem. 1990. Childhood diarrhoea in Sri Lanka: a case-control study of the impact of improved water sources. Tropical Medicine and Parasitology 41:98-104.

14. Cebu Study Team 1992. A child health production function estimated from longitudinal data. Journal of Development Economics 38:323-351.

15. Cebu Study Team 1991. Underlying and proximate determinants of child health: The Cebu longitudinal health and nutrition survey. American Journal of Epidemiology 133:185-201.

16. Briscoe, J., J. Akin and D. Guilkey. 1990. People are not passive acceptors of threats to health: Endogeneity and its consequences. International Journal of Epidemiology 19:147-153.

17. Popkin, B. M., L. Adair, J. Akin, R. E. Black, J. Briscoe and W. Flieger. 1990. Breast-feeding and diarrheal morbidity. Pediatrics 86:874-882.

18. VanDerslice, J. and J. Briscoe. 1993. All coliforms are not created equal: A comparison of the effects of water source and in-house water contamination on infantile diarrheal disease. Water Resources Research (In Press)

19. American Public Health Association 1985. Standard Methods for the Examination of Water and Wastewater. Washington, D.C.: American Public Health Association. 15th Ed.

20. Mosley, W.H. and L.C. Chen. 1984. An analytical framework for the study of child survival in developing countries. In: Child Survival, strategies for research, Population and Development review, supplement to vol. 10. edited by Mosley, W.H. and Chen, L.C. Cambridge: Cambridge University Press, pp. 25-45.

21. Kalter, H. D., R. H. Gray, R. E. Black and S. A. Gultiano. 1991. Validation of the diagnosis of childhood morbidity using maternal child health interviews. International Journal of Epidemiology 20:193-198.

22. Black, R. E., O. de Romano, K. H. Brown, N. Bravo, O. Bazalar and H. Kanashiro. 1989. Incidence and etiology of infantile diarrhea and major routes of transmission in Huascar, Peru. American Journal of Epidemiology 129:785-799.

23. Barrell, R. A. E. and M. G. M. Rowland. 1979. Infant foods as a potential source of diarrhoeal illness in rural West Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene 73:85-90.

24. Black, R. E., K. H. Brown, S. Becker, A. R. M. Abdul Alim and M. H. Merson. 1982. Contamination of weaning foods and transmission of enterotoxigenic Escherchia coli diarrhoea in children in rural Bangladesh. Transactions of the Royal Society of Tropical Medicine and Hygiene 76:259-264.

25. Worthington-Roberts, B.S. 1989. Lactation and human milk: Nutritional considerations. In: Nutrition in Pregnancy and Lactation. edited by Williams, S.R. St. Louis: Times Mirror/Mosby College Publishing, pp. 244-322.

26. Welsh, J. K. and J. T. May. 1979. Anti-infective properties of breast milk. Journal of Pediatrics 94:1-9.

27. Judge, G.G., R.C. Hill, W. Griffiths, H. Lutkepohl and T. Lee. 1982. Introduction to the Theory and Practice of Econometrics. New York: John Wiley and Sons.

28. Maddala, G.S. 1983. Limited-dependent and Qualitative Variables in Econometrics. Cambridge: Cambridge University Press.

29. Guilkey, D., T. A. Mroz and L. Taylor. 1992. Estimation and testing in simultaneous equations models with discrete outcomes using cross section data, UNC-CH mimeo. (UnPub)

30. Avery, R.B. and V.J. Hotz. 1985. Hotztran User's Manual. Old Greenwich, Conn.: CERA Economic Consultants, Inc..

31. World Bank. 1992. World Development Report, 1992: Development and the Environment. Oxford, England: Oxford University Press.

32. World Bank Water Demand Research Team. 1993. The demand for water in rural areas: Determinants and policy implications. The World Bank Research Observer 8:47-70.

33. Briscoe, J. 1992. Poverty and water supply: How to move forward. Finance and Development 29:16-19.

TABLE 1. Exogenous Variables Used in the Reduced-form Specification

---

Household resources

Household income*
Value of household assets*
Has electricity
Owns refrigerator
Owns television
Owns radio
Owns gas stove
Availability of child care

Prices (minimum, deflated)

Corn*
Cooking oil*
Kerosene
Cerelac
Latundun (banana)
Evaporated milk
Formula
Condensed sweetened milk
Powdered milk
Child care

Household composition

Spouse present*
Mother is head of household
Extended family present*
Mother is absent
Mother's age*
Father's age
Grandmothers present
Grandfathers present
Other adult females present
Household speaks Cebuano*
Household is Catholic*
Sex of the child
Lived in barangay more than 1 year

---

* denotes variables used for baseline instruments.

cont.

Table 1. continued.

Education

Mother's education*
Mother is literate
Father's education

Seasonality

Days of rain in past two weeks*
Total rainfall in past two weeks*
Wet season*
Dry season*

Community characteristics

Community density
Modernization index
Distance to a road
Residence is on an island
Availability of piped water supply
% time that formula is available
% time that other breast milk substitute is available
Travel time to store with breast milk substitute

---

* denotes variables used for baseline instruments.

TABLE 2. 7-day Prevalence of Diarrhea, by Child's Age and Sex

	Child's age (months)					
Category	2	4	6	8	10	12
Overall	7.2	12.7	20.4	25.0	24.4	24.1
Females	7.1	12.3	20.1	24.6	22.9	21.9
Males	7.4	13.1	20.6	25.4	25.7	26.0

TABLE 3. Type of Excreta Disposal Facility Used

Type of excreta disposal facility	Frequency	Percent
Private facilities:		
Cistern toilet (inside)	127	5.4
Water-sealed toilet (inside)	254	10.8
Cistern toilet (outside)	46	2.0
Water-sealed toilet (outside)	655	27.8
Public facilities:		
Water-sealed toilet	194	8.2
Latrine	532	22.6
No facility:		
Open pit	100	4.3
No facility used (field, canal, seashore)	438	18.6
Other	8	0.3
Unknown	1	<0.1
TOTAL	2355	100.0

TABLE 4. Presence of Excreta Around the House

Observation categories	Frequency	Percent
Heavy excreta visible	178	7.6
Some excreta visible	600	25.5
Very little excreta visible	335	14.2
No excreta visible	1219	51.8
No observation made	23	1.0
TOTAL	2355	100.1

TABLE 5. Comparison of Households in Communities with Good and Very Poor Sanitation

	Very poor community sanitation (18%)		Good community sanitation (82%)	
	Mean	(s.d.)	Mean	(s.d.)
Household assets (pesos x 10^{-3})	7.8	(20.1)	13.9	(55.3)
Mother's education (years)	7.6	(3.2)	7.6	(3.3)
Private excreta disposal facility	0.35		0.48	
In-house water connection	0.11		0.08	
Fecal material visible outside house	0.62		0.27	

TABLE 6. Distribution of daily FC doses, adjusted for water boiling, by level of community sanitation.

	Very poor community sanitation		Good community sanitation		Overall	
Daily FC dose	n	(%)	n	(%)	n	(%)
< 1	2179	(92.0)	9029	(93.7)	11208	(93.4)
1 to 10	130	(5.5)	500	(5.2)	630	(5.3)
10 to 100	56	(2.4)	99	(1.0)	155	(1.3)
> 100	3	(0.1)	4	(0.0)	7	(0.1)
TOTAL	2368	(100.0)	9632	(99.9)	12000	(100.1)

Table 7.

In MITABLAN.DOC

Table 8. Risk Differences and Risk Ratios Associated with Water Contamination, by Level of Community Sanitation

	Community Sanitation			
	Good		Very Poor	
Daily FC Dose	Risk Difference[a] (a)	Risk Ratio[b] (b)	Risk Difference[a] (c)	Risk Ratio[b] (d)
1	---- reference level ----		---- reference level ----	
10	0.05 (0.02, 0.08)	1.32 (1.12, 1.53)	-0.01 (-.05, 0.05)	0.97 (0.75, 1.22)
100	0.11 (0.04, 0.18)	1.69 (1.25, 2.11)	-0.01 (-.10, 0.09)	0.94 (0.56, 1.46)
1000	0.17 (0.06, 0.29)	2.12 (1.39, 3.00)	-0.02 (-.13, 0.15)	0.91 (0.39, 1.73)

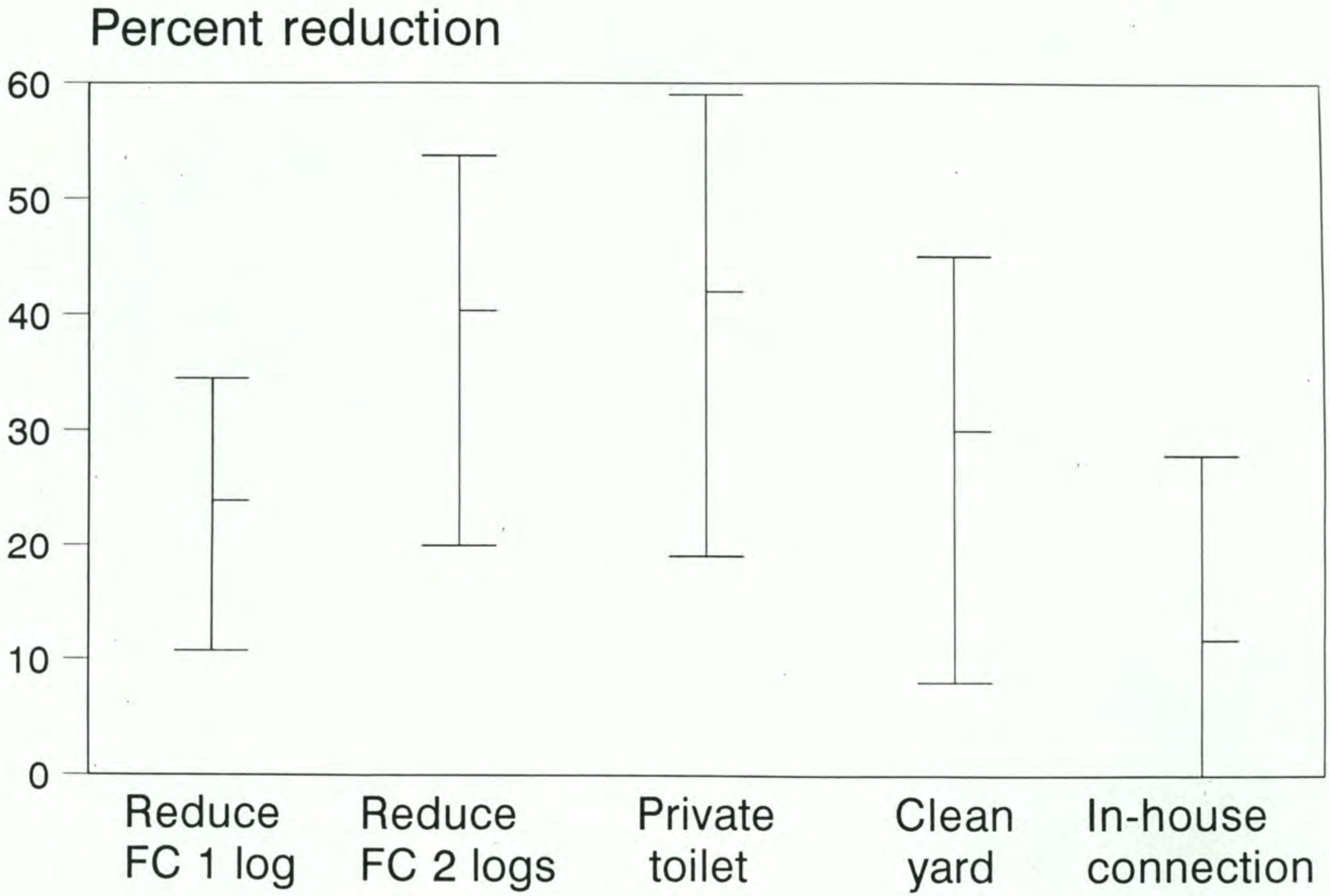
[a] Risk difference =
(expected proportion with diarrhea at specified FC dose)
- (expected proportion with diarrhea for FC dose of 1).

[b] Risk ratio = $\frac{\text{(expected proportion with diarrhea at specified FC dose)}}{\text{(expected proportion with diarrhea for FC dose of 1).}}$

TABLE 9. Comparison of the Effects of Alternative Environmental Interventions on Diarrheal Disease in Children: Mean Reductions in the Predicted Probability of Diarrhea

Intervention	Percent reduction in diarrhea for affected families	(95% CI)
(a) In-house water connections	12%	(-7, 28)
(b) Private excreta disposal facilities	42%	(19, 60)
(c) Removing excreta around the house	30%	(08, 45)
(d) For families in neighborhoods with good community sanitation, reduce water source FC concentration from 100 FC/100ml to:		
10 FC/100 ml	24%	(11, 34)
1 FC/100ml	40%	(20, 54)
(e) For families who use good quality drinking water, improving community sanitation	25%	(16, 33)

Figure 1. Reduction in Diarrheal Prevalence for Alternative Environmental Interventions



Error bars denote 95% confidence intervals

American Journal of Epidemiology
Vol. 133, No. 2
Printed in U.S.A.

Underlying and Proximate Determinants of Child Health: The Cebu Longitudinal Health and Nutrition Study

The Cebu Study Team[1]

A proper understanding of infant health requires the integration of socioeconomic, behavioral, and biomedical models. A methodology is presented for assessing the effects of "underlying" social factors and "proximate" behavioral and biomedical factors on infant morbidity, growth, and mortality. The method is applied to data collected from over 3,000 children in Cebu, Philippines, over the first 2 years of life. Data were collected between 1983 and 1985. A central theme is that mothers recognize certain observable and nonobservable threats to the health of their infants, and that the mothers take measures to reduce the risk from such threats. It is shown that if conventional statistical techniques (which do not take such behaviors into account) are used, the estimates of the effect of the risk factors on health are incorrect. Procedures for obtaining correct estimates are described. The application of the methodology is illustrated by modeling childhood diarrhea, and by showing how maternal education induces behavioral changes, and how these changes, in turn, induce changes in the prevalence of childhood diarrhea. *Am J Epidemiol* 1991;133:185–201.

biological factors; diarrhea; epidemiologic methods; growth; health behavior; models, statistical; socioeconomic factors

The causes for high levels of childhood disease in developing countries have been the subject of numerous investigations by both social and biomedical scientists (1). The focus of the social science literature is on examining the relations between "underlying" socioeconomic variables and health outcomes, with most research focusing on mortality (e.g., references 2–6). Many of the results of these analyses are robust, with increased household income and maternal education, for example, consistently emerging as powerful determinants of health (e.g., references 7, 8). However, this literature usually gives rise to conclusions which are so sweeping (such as "where income and edu-

---

Received for publication November 16, 1988, and in final form February 9, 1990.

[1] The Cebu Study Team members and their affiliations are as follows: Linda S. Adair, John S. Akin, David K. Guilkey, and Barry M. Popkin (the University of North Carolina and the Carolina Population Center, Chapel Hill, NC); John Briscoe (the World Bank, Washington, DC); Robert E. Black (The Johns Hopkins University, Baltimore, MD); and Wilhelm Flieger (the Office of Population Studies of the University of San Carlos, Cebu, Philippines). John Briscoe had principal responsibility for organizing and writing this paper.

Reprint requests to John Briscoe, World Bank, 1818 H Street, N.W., Washington, DC 20433.

This paper is part of a collaborative research project involving the Office of Population Studies, University of San Carlos, Cebu, Philippines, directed by Dr. Wilhelm Flieger; the Nutrition Center of the Philippines, directed by Dr. Florentino S. Solon; and a group from the Carolina Population Center. Barry M. Popkin of the University of North Carolina is the project coordinator. Funding for parts of the project design, data collection, and computerization was provided by the National Institutes of Health (contract nos. R01-HD19983A, R01-HD23137, and R01-HD18880), the Nestles' Coordinating Center for Nutrition Research, Wyeth International, the Ford Foundation, the United States National Academy of Sciences, the Carolina Population Center, the United States Agency for International Development, and the World Bank. David Fugate, Deborah Barnes, Eilene Bisgrove, Frances Dancy, Patricia Dargent, Lionel Deang, Maria E. Fernandez, Gina Gantt, Amanda Lyerly, Christine Moe, James Vanderslice, and Emelita Wong provided invaluable assistance at many stages of the work.

cation are higher, health is better") that they provide little guidance to those concerned with formulating health programs. The mechanisms by which the socioeconomic determinants affect health remain largely an unexplored and unexplained "black box" (1).

The biomedical literature, on the other hand, focuses on the biologic precursors (such as infection and malnutrition) of morbidity and mortality. The virtue of the biomedical literature (namely its reliance on a biologic model and the focused nature of the questions answered) is its curse, too. This is so because: these narrowly focused studies have often ignored the effects of important confounding variables (with the biomedical literature on breast feeding being a good example); the ultimate consequences for mortality in populations at large tend to be neglected; and the fact that people perceive threats to their health and react to these by changing their behavior is often either not recognized or ignored because it is considered analytically intractable. The result is a literature which inevitably leads to policy conclusions favoring strictly medical interventions (1).

Drawing heavily on an analogous situation in the field of fertility research (9), Mosley and Chen (1) and Mosley (10, 11) have argued for the development of a new approach to child health research which incorporates both the social and biomedical approaches into a coherent analytic framework in which the relations between "underlying," "intermediate," and "outcome" variables are investigated. Important steps have been taken in recent years in studies in Malaysia (12–14) and Jordan (15) to conduct empirical research on mortality using this framework. The Cebu study was designed to build on these landmark studies.

This paper describes the methodology used in modeling child health in the Cebu study. Empirical results are given to illustrate the usefulness of the approach. Detailed discussions of the results and their implications have been presented in other papers (16–18).

THE DESIGN OF THE CEBU STUDY

The principal objective of the Cebu study was to correctly estimate the effects of underlying and proximate determinants of child health. Data were collected in the metropolitan area around the city of Cebu in the central Philippines. After a pilot study, a stratified, single-stage sampling procedure was used to select 17 of 158 urban and 16 of 85 rural neighborhoods in the metropolitan Cebu area. Households were surveyed to collect data on all births between May 1, 1983 and April 30, 1984. The sample consisted of 3,080 women (77 percent of whom were urban) having single live births, for whom both baseline pregnancy surveys and birth information are available. Participation rates were high. Over the course of the 2-year period, 311 of the 3,080 women (264 of 2,355 in urban areas) were lost as a result of migration, and 49 of the mothers (39 in urban areas) decided to withdraw from the study.

For each study child, questionnaires were administered in the third trimester of pregnancy, at birth, and at 2-month intervals through the first 2 years of life. Where necessary, the questionnaires were supplemented by observations (e.g., of sanitary conditions) and measurements (e.g., of weight and water quality). Information was collected on "underlying variables" (including family income and assets, education of family members and other socioeconomic variables, prices of foods and other goods and services in the community, and accessibility to health facilities), "intermediate variables" (describing households' consumption choices for health-related goods and services, such as prenatal care and infant feeding patterns, water-use practices, personal hygiene practices, use of preventive health services, maternal smoking and drinking) and "outcome variables" (including gestational age and birth weight, and growth, morbidity, and mortality at each subsequent 2-month interval). Additional details on the survey design and data are available (16–18).

MODELS FOR ASSESSING THE EFFECT OF UNDERLYING AND INTERMEDIATE VARIABLES ON CHILD HEALTH

The mechanisms whereby socioeconomic, behavioral, and biomedical factors affect health can be described in terms of two sets of equations. The first equation describes how the underlying individual, family and community variables determine health-related behaviors; the second equation describes how underlying and intermediate behavioral and biomedical variables affect health outcomes. Following standard economics terminology, these are referred to as "structural equations." For reasons which will become apparent later it is necessary to distinguish between "endogenous variables," whose values are determined by forces operating within the model, and "exogenous variables," whose values, while important to the model, are determined by forces outside the model and are not explained by the model. In the present context (as shown in figure 1 and table 1), variables such as infant feeding patterns, use of medical facilities, type of water supply and sanitation, maternal work status, and health status of the child are treated as endogenous, while variables that are not the result of health-related household decisions (such as maternal education and food prices) are considered exogenous.

The variables entering the models (where the subscript "i" refers to the particular child and the subscript "t" to the time period) are: H_{ti}, the health of the infant; Y_{ti}, endogenous variables measuring the consumption of health-related goods; Z_{ti}, exogenous community and household characteristics; μ_i, an individual-specific disturbance term that does not change through time; and ϵ_{ti}, purely random errors that vary across individuals and through time.

Structural equation 1: Determinants of behavior

The underlying family and community variables (the Z_is) are hypothesized to determine the health-related behaviors (Y_is) as follows:

$$Y_{ti} = \alpha_1 H_{Gt-1,i} + \alpha_2 H_{St-1,i} + \alpha_3 Y_{t-1,i} + \alpha_4 Z_{ti} + \mu_{Yi} + \epsilon_{Yti} \qquad \text{(expression 1)}$$

$$\text{for } i = 1, 2, \cdots, N;$$

$$t = 2, 4, \cdots\ 24 \text{ months}$$

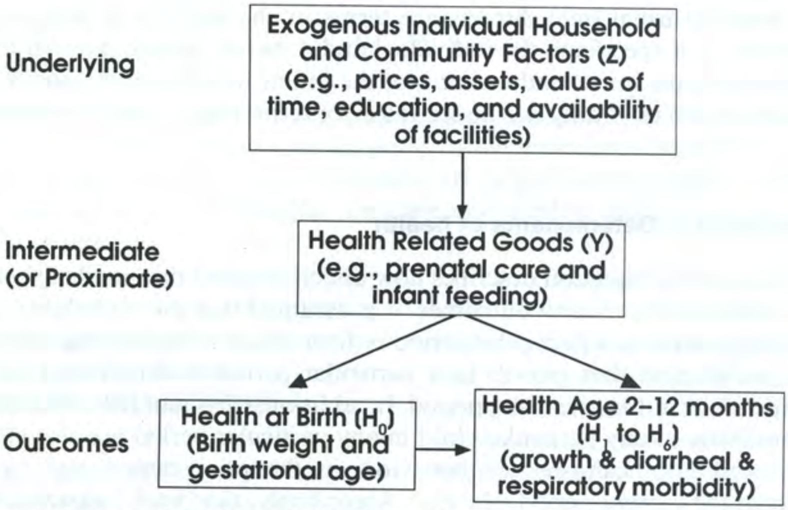


Figure 1. Conceptual framework relating underlying, intermediate and outcome variables.

TABLE 1. Endogenous variables in the diarrhea model for Cebu, Philippines, 1983–1985

I. Health of the infant
  A. At birth (H_0)
    1. Birth weight [continuous, OLS]*
    2. Gestational age [continuous, OLS]
  B. Every 2 months (H_t, $t = 2, 4, \ldots, 12$)
    1. Diarrhea [dichotomous, probit]
    2. Weight [continuous, OLS]
    3. Severe Respiratory Infection [dichotomous, probit]
II. Health-related factors affecting H_t ($t = 2, 4, \ldots, 12$) (Y_t)
  A. Exclusive breast-feeding pattern [dichotomous, probit]
  B. Any breast-feeding pattern [dichotomous, probit]
  C. Infant supplemental food nutrient intake [continuous, OLS]†
  D. Preventive medical care for infant [dichotomous, probit]
  E. Personal hygiene (soap use) [continuous, OLS]
  F. Food processing [dichotomous, probit]
  G. Immunizations (2 DPT‡; 3 DPT; measles) [dichotomous, probit]
  H. Water quality [dichotomous, probit]
  I. Excreta disposal [ordered discrete, ordered probit]

[ ] Indicates type of variable and estimation technique.
* OLS, ordinary least squares regression.
† During the period of exclusive breast feeding, a tobit estimation procedure was used.
‡ DPT, diphtheria-pertussis-tetanus vaccination.

and where the subscripts "G" and "S" refer to growth and sickness (or morbidity), respectively.

Note that the health variables measured before the current time period (the "lagged" health variables) have an effect on current health and non-health-related behaviors. That is, the model takes account of the fact that a mother might, for instance, alter her infant-feeding practices if her child failed to grow in an earlier time period. Note, too, that the model allows for the possibility that children and families have peculiarities which have important effects on how a child is treated, but which cannot be observed. For example, some mothers may know from prior pregnancies that their children are likely to be frail, and may take account of this when deciding whether to seek prenatal care.

The formal procedure for taking this unobserved heterogeneity into account is to have two (rather than the usual one) disturbance terms in the behavioral equations. The first disturbance term, μ, is specific to the individual and does not change through time. In these models, μ represents the initial endowments of the infant which cannot usually be observed by the researcher (such as "frailty"). The second error term, the ϵ, is the conventional random error term.

Structural equation 2: Determinants of health

The second structural equation describes how underlying and intermediate behavioral and biomedical variables affect health outcomes. It is assumed that the probability of being sick (e.g., of having diarrhea) in a particular period is determined in part by the nutritional status in the prior period, and that growth in a particular period is determined in part by the morbidity experience in the preceding period. In addition, it is assumed that both nutritional status and morbidity in any particular child in any particular period are also affected by the health-related and non-health-related behaviors in the prior period and by the family characteristics in the current period (the Z_ts). Accordingly, the "health structural equations" are:

$$\begin{aligned}
H_{Gti} &= \theta_1\, H_{Gt\text{-}1,i} + \theta_2\, H_{St\text{-}1,i} + \theta_3\, Y_{t\text{-}1,i} + \theta_4\, Z_{ti} + \mu_{Gi} + \epsilon_{Gti} \\
H_{Sti} &= \gamma_1\, H_{Gt\text{-}1,i} + \gamma_2\, Y_{t\text{-}1,i} + \gamma_3\, Z_{ti} + \mu_{Si} + \epsilon_{Sti} \qquad \text{(expression 2)} \\
H_{Mti} &= \delta_1\, H_{Gt\text{-}1,i} + \delta_2\, H_{St\text{-}1,i} + \delta_3\, Y_{t\text{-}1,i} + \delta_4\, Z_{ti} + \mu_{Mi} + \epsilon_{Mti} \\
&\text{for } i = 1, 2, \cdots, N; \\
&t = 2, 4, \cdots 24 \text{ months}
\end{aligned}$$

and where the subscripts "*G*," "*S*," and "*M*" refer to growth, sickness, and mortality, respectively.

As a specific example, consider the structural equation for diarrhea. One measure of diarrhea from the available data is the dichotomous variable indicating whether the child had diarrhea in the 7-day period preceding the interview. The endogenous variables include exposure variables (water quality, personal hygiene practices, excreta disposal practices, food hygiene practices, whether exclusively breast-fed, mother's concern with preventive measures as measured by use of a well-baby clinic) and susceptibility variables (nutritional status as measured by weight for age, whether breast-fed at all, and whether vaccinated against measles), while the exogenous variables include exposure to animals in the home, season, sex, household size, community density and child's mobility.

The "reduced form" equations

The two sets of structural equations represent a complete description of how behavior and health are determined. By starting at birth and continuously substituting out for all endogenous right-hand side variables, the so-called "reduced form" equations are obtained.

$$\begin{aligned}
Y_{ti} &= \sum_{r=o}^{t} Z_{t\text{-}r,i}\, \beta_r^Y + \overset{*}{\mu}_{Yi} + \overset{*}{\epsilon}_{Yti} \\
H_{Gti} &= \sum_{r=o}^{t} Z_{t\text{-}r,i}\, \beta_r^G + \overset{*}{\mu}_{Gi} + \overset{*}{\epsilon}_{Gti} \qquad \text{(expression 3)} \\
H_{Sti} &= \sum_{r=o}^{t} Z_{t\text{-}r,i}\, \beta_r^S + \overset{*}{\mu}_{Si} + \overset{*}{\epsilon}_{Sti} \\
H_{Mti} &= \sum_{r=o}^{t} Z_{t\text{-}r,i}\, \beta_r^M + \overset{*}{\mu}_{Mi} + \overset{*}{\epsilon}_{Mti}
\end{aligned}$$

In the reduced form equations, the endogenous variables (the *Y*s) and the outcome variables (the *H*s) can be expressed just in terms of the exogenous variables (the *Z*s). These reduced form equations may be used to examine the full effect of exogenous variables (such as maternal education) on child health. They are also used in estimating the parameters of the structural equations.

PROBLEM 1: ESTIMATING PARAMETERS WHEN SOME VARIABLES ARE ENDOGENOUS

The standard approach would be to estimate the parameters of such models using standard statistical procedures (such as ordinary least squares or logistic regression). This section shows that because people recognize (some) health threats and take measures to reduce their risk from these threats, these standard procedures give the wrong answers. The correct statistical procedures are described.

Statistical procedures

For the reduced form equations (expression 3), it can reasonably be assumed that the "heterogeneity disturbance terms" (the μs) are distributed independently of the exogenous variables (the Zs) and that the error terms (the μs and the ϵs) are normally distributed. Accordingly, standard statistical techniques (such as ordinary least squares or logistic regression) can be used to obtain unbiased estimates of the parameters of the reduced form equations.

It is, however, the structural equations (expressions 1 and 2) that are of most interest to policymakers, because they permit assessment of the effects of different social, economic, and biomedical interventions on behavior and health. These equations present more complex estimation problems. For example, in expression 3 it can be seen that the value of, say, Y_{ti} depends on, and is therefore correlated with, the "heterogeneity disturbance term" (μ_{Yi}). Since the value of μ_{Yi} does not change over time, μ_{Yi} is similarly correlated with Y_{t-1}. In expression 1, therefore, a regressor ($Y_{t-1,i}$) is correlated with a disturbance term (μ_{Yi}). Similarly in expression 2 we also have a regressor ($Y_{t-1,i}$) correlated with a disturbance term (μ_{Gi}).

If the parameters of expressions 1 and 2 are estimated using ordinary least squares, inconsistent estimates will result. This is because, in explaining the dependent variable, as much credit as possible is given to the regressor and as little as possible to the error. When the regressor and error are correlated, some of the effect of the error is wrongly attributed to the regressor (19).

A standard procedure for dealing with this problem is that of "instrumental variables" (20). In estimating the parameters of the structural equation, the value of the problematic regressor (such as Y_{ti}) is not used, but replaced by an instrumental variable. The instrumental variable for Y_{ti} is chosen so that it is correlated with the regressor (Y_{ti}) and is uncorrelated with the disturbance term (μ_{Yi}). In this particular case, the "predicted values" of the Ys obtained from versions of the reduced form equations (expression 3) can be used as instrumental variables for the Ys in the structural equations (expressions 1 and 2). Using these predicted values rather than the actual values of the endogenous variables, standard estimation procedures are used to estimate the parameters of the structural equations. Although not widely known in the epidemiologic literature, such techniques are used routinely by economists (21).

As discussed in more detail elsewhere (16–18), in the Cebu study the models were specified to include all endogenous and exogenous variables, and to allow each of these to vary over time (by including a time interaction term). Only those interactions which were statistically significant were retained. The "random effects" estimation procedure (21) was used to estimate the parameters of the models. It was assumed that the μs and ϵs are normally distributed random variables and that corresponding to the observed dependent variables is a continuous latent variable (the severity of the child's diarrhea in this case). The mother reports that the child has diarrhea if the latent variable is sufficiently large (see reference 22 for a different child health model that also uses latent dependent variables). The actual calculations were done with the HOTZTRAN algorithm (23), using maximum likelihood methods in which the distribution of the disturbance term is taken into account in calculating the standard errors of the coefficients.

Does endogeneity make a difference in practice?

Two examples from the literature. The US Environmental Protection Agency has conducted a large-scale study on the effects of medical care and air pollution on mortality from respiratory disease (24). Whereas prior epidemiologic studies had implicitly assumed that people accept air pollution passively, the study recognizes that "people have an incentive to adapt to environmental conditions (by incurring the expense of seeing a doctor or moving away from a polluted city)" (25, p. 42). In analytic terms, this means that "protective factors" (such as use of medical care) are not exogenous (i.e., determined by forces outside of the model) but endogenous (i.e., determined by the level of air pollution and other factors incorporated into the model). The study (24, 25) shows, first, that the conclusion drawn from a conventional analysis was that the level of medical care had no effect on mortality from respiratory illness, but, second, that when statistical procedures taking account of endogeneity were used, medical care was shown to have a significant protective effect.

The second example deals with the effect of prenatal care on child health. A conventional analysis would treat the quantity of prenatal care as an exogenous variable and examine the relation between the level of this variable and infant health. In fact, however, many mothers seek prenatal care in part because they perceive (for reasons that are valid but which investigators cannot observe) their fetus to be particularly vulnerable. A detailed assessment of the relation between prenatal care and infant mortality in the United States (3, 5) has shown that a conventional analysis (which ignores this behavioral aspect) would conclude that mothers place their children at risk by obtaining prenatal care, but that when statistical procedures take account of this behavioral relation, use of prenatal care is shown to have a strong protective effect.

Later in this paper, the practical consequences of ignoring endogeneity are examined for the Cebu study.

PROBLEM 2: SAMPLE SELECTIVITY IN LONGITUDINAL STUDIES

As children are followed over time, there are inevitably losses to the study from refusal, out-migration, and death. Since children with certain characteristics are more likely to be lost to the sample than other children, the sample has been reduced in a way which is certainly not random. An exogenous factor which affects migration (such as father's occupation) but does not affect child health would tend to emerge from the analysis as a determinant of child health.

The procedure for correcting for this possibility consists of introducing a "correction factor" (technically known as the hazard rate or the inverse of the Mills' ratio (20, 26)) which is equal to zero for those individuals who would, without any doubt, remain in the sample throughout the period, and is relatively large for those who are likely to have been lost from the sample during the period. The statistical procedure involves, first, determining whether there is significant self-selection (if the "correction factor" is significant) and, if so, applying the necessary correction.

AN ILLUSTRATIVE RESULT: THE PATHWAYS THROUGH WHICH MATERNAL EDUCATION AFFECTS BEHAVIOR AND CHILD HEALTH

The Cebu Study Group has already published detailed results from some early analyses (16–18). For the present purposes, some empirical results illustrate how the model may be used to assess the biomedical and socioeconomic determinants of child health. The example chosen is one of major policy interest because of the consistent and strong relation (7, 8) between maternal education and child health and because of the paucity of data delineating the mechanisms by which this effect operates. The example is developed only for the urban sample, only for diarrhea, and only for the first year of life.

Underlying-proximate relations

The parameters of the behavioral structural equations (expression 1) are estimated for each of the health-related behaviors at each particular stage of the child's life. Table 2 shows the effects of maternal education on health-related behaviors during each 2-month period. Table 3 shows the simulated effects on the mean values of the health-related behaviors of increasing the education of each woman in the sample by one year.

Question 1: Is the direction of the effect sensible? Table 2 shows that as maternal education increases, there are increases in food intake, preventive health care, measles

TABLE 2. Effects of maternal education on health-related behavior during the first year of life, urban Cebu, Philippines, 1983–1985

Health-related behavior	Characteristics of dependent variable	Age of infants (months)					
		2	4	6	8	10	12
Feeding practices							
Breast feeding							
Exclusive breast feeding, no exposure to pathogens, 7 days before survey	Binary†	−0.07***	−0.07***	−0.12***	−0.02	N/A	N/A
Any breast feeding, 7 days before survey	Binary†	−0.10***	−0.11***	−0.11***	−0.10***	−0.10***	−0.10***
Food intake							
Total calories	Calories	23.02***	27.93***	23.97***	23.18***	22.96***	24.81***
Health service use							
Preventive health care	Binary‡	0.05***	0.03***	0.02*	0.04***	0.04***	0.04***
Measles immunization	Binary‡	N/A	N/A	N/A	N/A	0.07**	0.06***
Health practices, personal and environmental							
Pathogenicity of food processing	Binary§	0.0004	−0.002	−0.0001	−0.0007	−0.0008	−0.002
Poor type of excreta disposal	Binary‖	−0.09***	−0.09***	−0.09***	−0.09***	−0.09***	−0.09***
Quantity of soap per capita	Grams	1.76***	1.76***	1.76***	1.76***	1.76***	1.76***
Good quality of drinking water source	Binary¶	0.002	0.002	0.002	0.002	0.002	0.002

Entries are the coefficients of mother's formal education in years in creating instrumental variables for the dependent health-related variables. The asterisks indicate the significance level for testing whether the coefficient is zero: * $\alpha = 0.10$, ** $\alpha = 0.05$, *** $\alpha = 0.01$. N/A, not applicable.

† If the infant is breast-fed, the variable is set to 1 and 0 otherwise.

‡ If the infant had a preventive health care visit or had the specified type and dosage of immunization, the variable takes on the value of 1 and 0 otherwise.

§ If food processing is severely pathogenic, the variable is set to 1 and 0 otherwise.

‖ If the household's exreta disposal is poor, the variable is set to 1 and 0 otherwise.

¶ If the source of drinking water is good, the dependent variable takes on the value of 1 and 0 otherwise.

immunization, adequacy of excreta disposal practices, quantity of soap used per capita, and quality of drinking water; and decreases in breast feeding (both exclusive and any) and food contamination risk.

Question 2: Is the effect statistically significant? From table 2 it can also be seen that, for most health-related practices, the effects of maternal education are highly statistically significant. The two exceptions are unhygienic food preparation practices and quality of drinking water. In both cases, the lack of significance is almost certainly because, with the measures employed in these early analyses, there is little variation in these variables in the urban sample.

Question 3: Are the findings of practical significance? From table 3, it can be seen that a one-year increase in the education of each mother would have substantial effects on most of the health-related behaviors. For example, for a 6-month-old child, a one-year increase in maternal education implies a 36 percent reduction in the probability of exclusive breast feeding, a 5 percent reduction in the probability of any breast feeding, a 7 percent increase in caloric intake, a 4 percent increase in the use of preventive health care, a 9 percent reduction in the probability of inadequate excreta disposal practices, and a 2 percent increase in per capita soap use.

Proximate-outcome relations

The second set of structural equations (expression 2) describes the relations between the proximate (behavioral and biomedical) variables and health outcomes. Table 4 presents the proximate-outcome structural equation for diarrhea for the longitudinal model estimated for the full first year of life for the urban population. Table 5 shows the responsiveness of diarrhea to changes in the proximate variables (as measured by the "elasticity," that is, the percent change in diarrhea resulting from a percent change in the proximate variable).

Substantive Question 1: Are the estimates sensible and statistically significant? From table 4, it can be seen that diarrhea is statistically significantly lower for: faster growing infants (with the effect greatest in small infants); infants who are breast-fed (with the protective effect greatest at young ages); infants who are exclusively breast-fed; infants who consume more food; infants whose families use better quality water; infants whose families follow hygienic food preparation practices; and infants whose families have better excreta disposal practices (with the effect greater in the early months of this first year of life). Diarrhea is statistically significantly higher for: male infants in the latter months; crawling infants when there are animals in the house; and children in more densely settled communities.

Substantive Question 2: Are the findings of practical significance? From table 5, it is evident that the level of diarrhea is highly responsive to breast-feeding practices (especially in the early months of life) and to excreta disposal and water supply practices (throughout the first year of life), moderately responsive to caloric intake, especially later in the first year of life, and largely unaffected by preventive health care.

Methodological Question 1: What are the consequences of sample selectivity? Statistical analysis showed that the hazard rate (or inverse of the Mills' ratio) was small, and not significantly different from zero. It was therefore concluded that sample selectivity was not significant (that is, that nonresponse could be viewed as a random event in the sample) and that the Mills' ratios could be excluded in the final specifications of the instrumental variables.

Methodological Question 2: What are the consequences of ignoring endogeneity? The importance of taking account of endogeneity when modeling child health was tested in two ways using the Cebu data set. The first test is a formal statistical test—a chi-square version of the Hausman test (21, 27)—which indicates whether endogeneity was actually present. The critical value for a 1 percent test is 29, while the test statistic was 98: the null hypothesis of no endogeneity is strongly rejected. The results of a second, more intuitive, test (comparing results from two estimation procedures, one taking account

TABLE 3. Simulated effects of one-year increase in maternal education on health-related behavior and diarrhea incidence, at every 2 months during the first year of life, urban Cebu, Philippines, 1983–1985*

	Age of infants (months)											
	2		4		6		8		10		12	
	Mean	% Change	Mean	% Change	Mean	% Change	Mean	% Change	Mean	% Change	Mean	% Change
Feeding practices												
Breast feeding												
Exclusive breast feeding, no exposure to pathogens, 7 days before survey	0.16	−10	0.06	−12	0.0002	−36	N/A		N/A		N/A	
Any breast feeding, 7 days before survey	0.85	−3	0.79	−4	0.74	−5	0.70	−5	0.64	−6	0.57	−7
Food intake												
Total calories	121.70	10	192.60	9	307.70	7	372.00	6	439.40	5	507.50	5
Health service use												
Preventive health care	0.12	9	0.18	5	0.18	4	0.13	7	0.11	8	0.08	7
Measles immunization	N/A		N/A		N/A		N/A		0.01	19	0.01	16
Health practices, personal and environmental												
Pathogenicity of food processing†	0.04	1	0.03	−8	0.03	0	0.03	−2	0.03	−3	0.04	−5
Poor type of excreta disposal†	0.52	−9	0.52	−9	0.52	−9	0.52	−9	0.52	−9	0.52	−9
Quantity of soap per capita†	81.12	2	81.12	2	81.12	2	81.12	2	81.12	2	81.12	2
Good quality of drinking source†	0.99	0	0.99	0	0.99	0	0.99	0	0.99	0	0.99	0

* Simulation is done using the instrumental variable equation. The % change is obtained with this formula: % change = $(M_2 - M_1)/M_1 \times 100$, where M_1 is the simulated mean of the health-related behavior of interest obtained by multiplying the coefficients of the exogenous variables by their corresponding sample means. M_1 is entered as mean in the table. M_2 is the simulated mean when all variables except woman's education are at their sample means and mean maternal education is increased by one year. N/A, not applicable because no infants had this feeding or immunization during this period.

† Current version of these variables are not time-varying yet.

TABLE 4. Longitudinal analysis: Structural equation for diarrhea incidence in week preceding survey, urban Cebu, Philippines, 1983–1985

Explanatory variables	Coefficient	(*t* statistic)
A. ENDOGENOUS		
Susceptibility		
Lagged weight velocity (g/day)	−0.01	(−2.02**)
Lagged weight velocity interacted with weight (g × g/day)	3.30×10^{-6}	(2.90***)
Gestational age (weeks)	0.01	(3.12***)
Gestational age interacted with age (weeks × days)	9.60×10^{-6}	(1.68*)
Susceptibility/exposure		
Feeding practices		
Any breast feeding 7 days before survey (prob)†	−0.68	(−2.51**)
Any breast feeding interacted with age (prob × days)	1.60×10^{-3}	(1.76*)
Exclusive breast feeding with no exposure to pathogens, 7 days before survey (prob)	−1.53	(−5.91***)
Total calories (cal)	-4.40×10^{-4}	(−1.73*)
Exposure		
Health service use		
Preventive health care (prob)	−0.24	(−1.27)
Health practices, personal and environmental		
Good quality water source (prob)	−0.32	(−3.35***)
Soap purchased/capita/week (g)	-3.70×10^{-5}	(−0.06)
Pathogenic food processing (prob)	0.91	(1.85*)
Poor excreta disposal (prob)	0.92	(4.97***)
Poor excreta disposal interacted with age (prob × days)	-1.80×10^{-3}	(−2.45**)
B. EXOGENOUS		
Susceptibility		
Child's age (days)	7.10×10^{-4}	(1.00)
Child's sex (0–1)	−0.02	(−0.24)
Child's sex interacted with age (0–1 × days)	5.60×10^{-4}	(2.09**)
Exposure		
Animals in the house (0–1)	-6.90×10^{-3}	(−0.22)
Animals under the house (0–1)	−0.02	(−0.53)
Baby crawling interacted with animals in the house (0–1)	0.08	(1.93*)
Crowding		
No. of preschoolers (0–6)	−0.03	(−2.20**)
No. of persons/room (0–9.5)	9.90×10^{-3}	(1.05)
Community density (persons/km^2)	6.50×10^{-6}	(7.09***)
Cumulative rainfall in last 2 weeks before survey (mm)	2.05×10^{-4}	(0.45)
Cumulative rainfall interacted with age (mm × days)	1.20×10^{-6}	(0.69)
C. OTHERS		
Constant	−6.05	(−3.63***)
Rho	0.12	(7.22***)

Note: Sample size for this analysis is 11,807. The significance levels for testing whether the coefficient is zero are indicated by: * $\alpha = 0.10$, ** $\alpha = 0.05$, *** $\alpha = 0.01$.

† Prob, the predicted probability of the explanatory variable.

of endogeneity and one ignoring it) are presented in table 6 and summarized in figure 2.

"Column 1" of table 6 presents the results of the analysis which ignores endogeneity. For this analysis, observations are needed on all of the variables (both those which are considered exogenous and endogenous in the analysis using instrumental variables). A total of 6,674 observations are available. "Column 2" of table 6 presents the results of the instrumental variable analysis, for this same sample. A comparison of columns 1 and 2 shows that the analysis which does not account for endogeneity is reasonably specific. In only one case—total calories—does

TABLE 5. Percent change in diarrhea for a 1% increase in explanatory variables during the first year of life, urban Cebu, Philippines, 1983–1985*

	Age of infants (months)					
	2	4	6	8	10	12
Feeding practices						
Exclusive breast feeding, no exposure to pathogens	−0.59	−0.33	−0.05	N/A	N/A	N/A
Any breast feeding	−1.08	−0.91	−0.76	−0.70	−0.66	−0.60
Total calories	−0.12	−0.15	−0.22	−0.24	−0.29	−0.34
Health service use						
Preventive health care	−0.07	−0.09	−0.08	−0.06	−0.05	−0.04
Measles immunization	N/A	N/A	N/A	N/A	−0.02	−0.03
Health practices, personal and environmental						
Pathogenicity of food processing	0.08	0.05	0.04	0.05	0.04	0.06
Poor type of excreta disposal	0.90	0.71	0.54	0.43	0.34	0.24
Quantity of soap per capita	−0.01	−0.01	0	0	0	0
Good quality of drinking water source	−0.55	−0.49	−0.44	−0.43	−0.44	−0.43

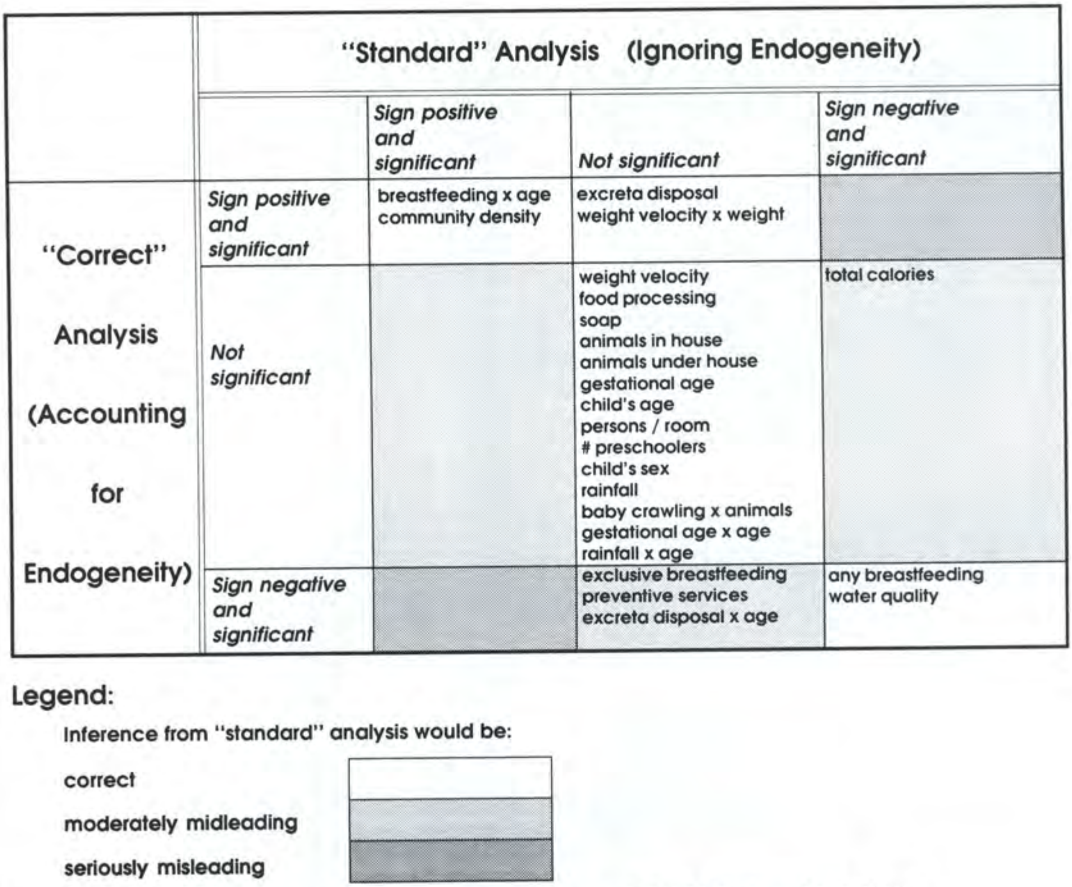
* Entries are computed by the following formula: % change = $(D_2 - D_1)/D_1 \times 100$, where D_1 is the simulated mean diarrhea when the coefficients of the explanatory variables in the diarrhea structural equation are multiplied by their corresponding sample means. D_2 is the simulated mean when the value of the variable of interest is increased by 1.01 of its sample mean and the rest of the explanatory variables are at their means. N/A, not available because no infant had this type of feeding or immunization during this period.

TABLE 6. The effect of ignoring endogeneity: Structural equation for diarrhea incidence in week preceding survey, urban Cebu, Philippines, 1983–1985

Explanatory variables	Column 1		Column 2	
	Estimates when endogeneity is ignored	(t-statistic)	Instrumental variable estimates	(t-statistic)
A. ENDOGENOUS				
Susceptibility				
Lagged weight velocity (g/day)	0.08	(1.39)	−0.01	(−1.20)
Lagged weight velocity interacted with weight (g × g/day)	-2.40×10^{-6}	(−0.24)	4.10×10^{-6}	(2.13**)
Gestational age (weeks)	0.18	(0.74)	0.09	(1.48)
Gestational age interacted with age (weeks × days)	-3.20×10^{-6}	(−0.29)	-1.50×10^{-6}	(−1.41)
Susceptibility/exposure				
Feeding practices				
Any breast feeding 7 days before survey (prob)†	−0.68	(−5.58***)	−1.06	(−2.29**)
Any breast feeding interacted with age (prob × days)	2.00×10^{-3}	(3.93***)	4.30×10^{-3}	(2.20**)
Exclusive breast feeding with no exposure to pathogens, 7 days before survey (prob)	−0.23	(−1.46)	−1.35	(−3.64***)
Total calories (cal)	-2.30×10^{-4}	(−2.94***)	-3.10×10^{-4}	(−0.84)

Exposure				
Health service use				
Preventive health care (prob)	0.02×10^{-3}	(−0.31)	−0.06	(−2.30**)
Health practices, personal & environmental				
Good quality water source (prob)	−0.17	(−1.87*)	−0.22	(−1.70*)
Soap purchased/capita/week (g)	-1.20×10^{-5}	(−0.37)	3.00×10^{-5}	(0.35)
Pathogenic food processing (prob)	0.15	(1.30)	0.87	(1.27)
Poor excreta disposal (prob)	0.09	(1.04)	1.11	(3.68***)
Poor excreta disposal interacted with age (prob × days)	2.70×10^{-3}	(0.07)	-3.60×10^{-3}	(−2.71***)
B. EXOGENOUS				
Susceptibility				
Child's age (days)	2.80×10^{-4}	(0.65)	1.30×10^{-4}	(0.98)
Child's sex (0–1)	−0.05	(−0.47)	−0.07	(−0.70)
Child's sex interacted with age (0–1 × days)	6.50×10^{-4}	(1.40)	7.10×10^{-4}	(1.55)
Exposure				
Animals in the house (0–1)	−0.02	(−0.39)	-6.60×10^{-3}	(−0.12)
Animals under the house (0–1)	−0.01	(0.05)	−0.03	(−0.68)
Baby crawling interacted with animals in the house (0–1)	0.02	(1.04)	−0.01	(0.13)
Crowding				
No. of preschoolers (0–6)	0.03	(0.14)	−0.03	(−1.49)
No. of persons/room (0–9.5)	0.02	(1.15)	0.01	(0.73)
Community density (persons/km^2)	5.80×10^{-6}	(5.13***)	7.60×10^{-6}	(6.26***)
Cumulative rainfall in last two weeks before survey (mm)	−0.07	(−1.20)	-4.47×10^{-4}	(−0.68)
Cumulative rainfall interacted with age (mm × days)	-4.50×10^{-6}	(−1.59)	3.30×10^{-6}	(1.17)
C. OTHERS				
Constant	−1.87	(−1.85**)	−4.51	(−1.88**)
Rho	0.16	(6.27***)	0.15	(5.86***)

Note: Sample size for this analysis is 6,674. The significance levels for testing whether the coefficient is zero are indicated by: * $\alpha = 0.10$, ** $\alpha = 0.05$, *** $\alpha = 0.01$.
† Prob is the predicted probability of the explanatory variable.



"Correct" Analysis (Accounting for Endogeneity)	"Standard" Analysis (Ignoring Endogeneity): *Sign positive and significant*	*Not significant*	*Sign negative and significant*
Sign positive and significant	breastfeeding x age community density	excreta disposal weight velocity x weight	
Not significant		weight velocity food processing soap animals in house animals under house gestational age child's age persons / room # preschoolers child's sex rainfall baby crawling x animals gestational age x age rainfall x age	total calories
Sign negative and significant		exclusive breastfeeding preventive services excreta disposal x age	any breastfeeding water quality

Legend:

Inference from "standard" analysis would be:

correct ☐

moderately midleading ▨

seriously misleading ■

Figure 2. Inferences from the "correct" and "standard" analyses, Cebu, Philippines, 1983–1985.

the simpler model (column 1) suggest a statistically significant relation which is not significant in the "correct" model (column 2). More serious, however, are those behaviors—improved excreta disposal, preventive health care and exclusive breast feeding—which would appear to have no effect if endogeneity is ignored (column 1) but which, in fact, have strong protective effects (as shown in the "correct" analysis in column 2). Furthermore, where the analysis ignoring endogeneity gave statistically significant estimates of the correct sign (any breast feeding, water quality, community density, and breast feeding × age), the parameter values were substantially biased toward the null. In short, if endogeneity is ignored, incorrect conclusions will be drawn on the determinants of health.

Tracing the paths by which education affects health

Interesting and important as the above results are, the major potential contribution of the Cebu Project is the integration of these two levels of analysis into a single, integrated behavioral-cum-biomedical description of child health. This integration can be illustrated by tracing through the pathways by which maternal education affects health-related behavior, and how such behavior, in turn, affects health.

Before tracing this path, the aggregate effect on diarrhea of a one-year increase in maternal education can be calculated using the reduced form (expression 3). The net effect of a one-year increase in maternal education would be to reduce the incidence

TABLE 7. Percent change in diarrhea due to behavioral change induced by a one-year increase in maternal education at every 2 months during the first year of life, urban Cebu, Philippines, 1983–1985*

	Age of infants in months					
	2	4	6	8	10	12
Feeding practices						
Exclusive breast feeding, no exposure to pathogens	+ 4.50	+ 1.73	0.02	N/A	N/A	N/A
Any breast feeding	2.53	2.55	2.03	1.43	0.94	0.46
Total calories	−1.00	−1.18	−1.23	−1.35	−1.30	−1.34
Health service use						
Preventive health care	−0.50	−0.35	−0.21	−0.29	−0.27	−0.16
Measles immunization	N/A	N/A	N/A	N/A	0	0
Health practices, personal and environmental						
Pathogenicity of food processing	0.06	−0.35	−0.01	−0.08	−0.09	−0.23
Poor type of excreta disposal	−6.74	−4.97	−3.76	−2.96	−2.26	−1.54
Quantity of soap per capita	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01
Good quality of drinking water source	0	0	0	0	0	0

Bad effect of education (if independent of prob. of outside employment)

* Entries are computed with the following formula: % change = $(D_3 - D_1) \times 100/D_1$, where D_1 is the simulated mean diarrhea when coefficients of the explanatory variables in the diarrhea structural equation are multiplied by their corresponding sample means. D_3 is the simulated mean diarrhea where the mean of instrumental variable of interest is changed by increasing mean maternal education by one year and the rest of the explanatory variables are at their sample means. N/A, not applicable because no infant had this type of feeding or immunization during this period.

of diarrhea episodes by about 5 percent in each time period. (Over the first year of life, the mean incidence of diarrhea in a 7-day period increased from under 4 percent in the first 2 months to over 15 percent in the final 2 months.)

By combining the simulated effects of education on health (table 3) with the simulated effect of behavioral changes on diarrhea (table 5), the education-behavior-diarrhea pathway can be traced. Table 7 shows that the three major pathways through which maternal education affects health in this population are: a large reduction in diarrhea (about 4 percent) because of improved excreta disposal practices, with the effect being particularly strong in the early months of life; a substantial reduction in diarrhea because of the increase in calories given to the child, with the effect greater toward the end of the first year of life; and a substantial, offsetting, increase in diarrhea because of a reduction in the number of mothers who breast-fed, with reduced exclusive breast feeding most important in the early months, and reduced breast feeding most deleterious early but remaining serious throughout the first year of life.

This information is presented graphically (for 6-month-old urban children) in figure 3. From figure 3, it is evident that some pathways are not important in this population either: because maternal education has little effect on behavior (as is the case of water supply for this urban population); or because the prevalence of the particular behavior is low (only 9 percent of mothers are exclusively breast-feeding their children at this age, for instance); or because changes in the particular behavior have little effect on health in this period (such as changes in the use of soap).

SUMMARY AND CONCLUSIONS

This paper shows that an integrated socioeconomic-biomedical model of child health can be specified and the parameters estimated. The results show that if endogeneity is ignored, incorrect conclusions are

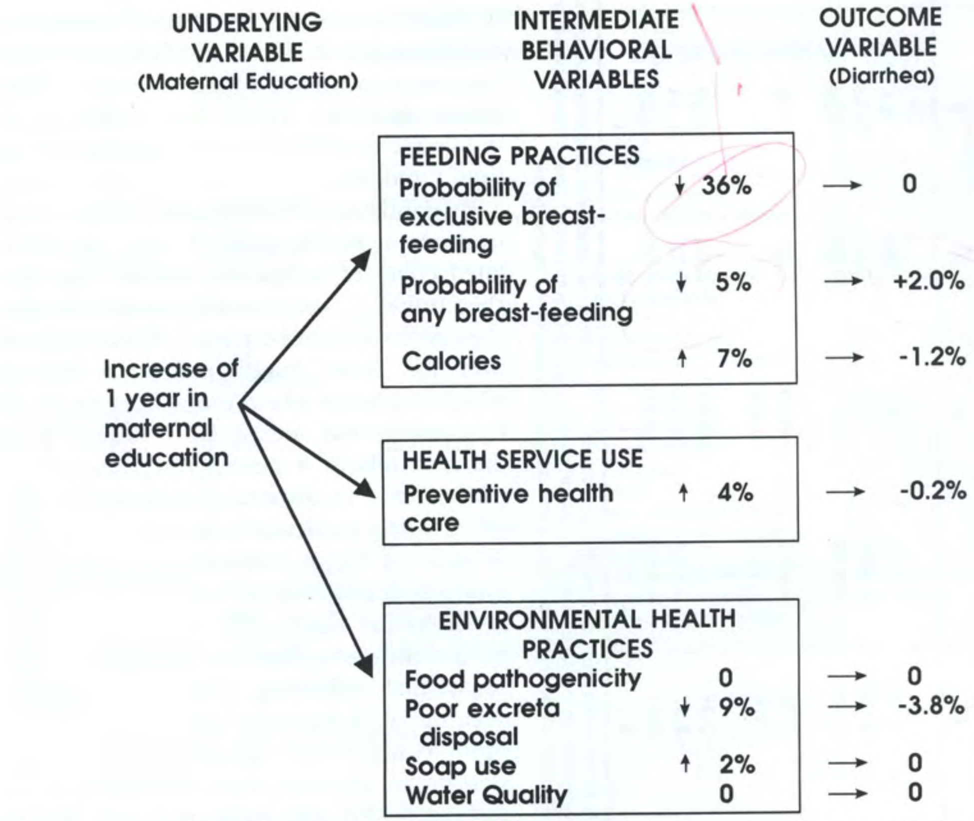


Figure 3. The pathways through which maternal education affects health at age 6 months, urban Cebu, Philippines, 1983–1985.

drawn concerning the effects of several determinants of child health. The analytic approach permits unique and readily understandable disaggregation of the effects of underlying and intermediate determinants of child health.

REFERENCES

1. Mosley WH, Chen LC. An analytic framework for the study of child survival in developing countries. Population and Development Review 1984;10:25–48.
2. Preston SH. Causes and consequences of mortality declines in less developed countries during the twentieth century. In: Easterlin RA, ed. Population and economic change in developing countries. Chicago: University of Chicago Press, 1980.
3. Schultz TP. Studying the impact of household economic and community variables on child mortality. Population and Development Review 1984;10:215–36.
4. Trussell J, Mencken J. Estimation of levels, trends and covariates of infant and child mortality in statistically poor countries. Population and Development Review 1984;10:325–46.
5. Rosensweig MR, Schultz TP. Estimating a household production function: heterogeneity and the demand for health inputs, and their effects on birthweight. Journal of Political Economy 1983;92:723–46.
6. Behm H. Socioeconomic determinants of mortality in Latin America. Proceedings of the meeting on socioeconomic determinants and consequences of mortality, el Colegio de Mexico, Mexico City, June 19–25, 1979. New York: World Health Organization, 1980.
7. Caldwell JC, McDonald P. Influence of maternal education on infant and child mortality. Proceedings of International Population Conference. Liège, Belgium: International Union for the Scientific Study of Population, 1981.
8. Cochrane SH, O'Hara DJ, Leslie J. The effects of education on health. World Bank Staff Working Paper no. 405, Washington, DC: World Bank, 1980.
9. Bongaarts J. A framework for analyzing the proximate determinants of fertility. Population and Development Review 1978;4:105–32.
10. Mosley WH. Biological contamination of the environment by man. In: Biological and social aspects

of mortality and length of life. Preston SH, ed. Liège, Belgium: International Union for Scientific Study of Population, 1980:36–67.

11. Mosley WH. The demographic impact of child survival programs. International Symposium, Center for Public Health Research, Mexico City, July 1986.
12. Davanzo J. A household survey of child mortality determinants in Malaysia. Population and Development Review 1984;10:307–24.
13. Davanzo J, Butz WP, Habicht JP. How biological and behavioral influences on mortality in Malaysia vary during the first year of life. Population Studies 1983;37:381–402.
14. Butz WP, Davanzo J, Habicht JP. Biological and behavioral influences on the mortality of Malaysian infants. Santa Monica, CA: Rand Corporation, 1982.
15. Tekce B, Shorter F. Determinants of child mortality: a study of squatter settlements in Jordan. Population and Development Review 1984;10:257–80.
16. Adair LS, Popkin B. Birth weight, maturity and proportionality in Filipino infants. Hum Biol 1988; 60:319–39.
17. Guilkey DK, Popkin BM, Akin JS, et al. Prenatal care and pregnancy outcomes in the Philippines. Journal of Development Economics 1989;30: 241–72.
18. Wong E, Popkin B, Akin JS, et al. Accessibility, quality of care and prenatal care use in the Philippines. Soc Sci Med 1987;24:927–44.
19. Wonnacott TH, Wonnacott RJ. Introductory statistics for business and economics. New York: John Wiley and Sons, 1972.
20. Judge GG, Griffiths WE, Hill RC, et al. The theory and practice of econometrics. New York: John Wiley and Sons, 1980.
21. Maddala GS. Introduction to econometrics. New York: Macmillan, 1988.
22. Behrman JR, Wolfe BL. How does mothers' schooling affect family health, nutrition, medical care usage, and household sanitation? Journal of Econometrics 1987;36:185–204.
23. Avery R, Hotz J. Hotztran users' manual. Old Greenwich, CT: CERA Economic Consultants, 1985.
24. Crocker TD, Schulze W, Ben-David S, et al. Experiments with economics of air pollution epidemiology: volume 1 of methods development for assessing air pollution control benefits. Washington DC: United States Environmental Protection Agency, 1979. (Report EPA-600/5-79-00 1a).
25. Kneese AV. Measuring the benefits of clean air and water. Washington, DC: Resources for the Future, 1986.
26. Heckman JJ. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement 1976;5:475–92.
27. Hausman JA. Specification tests in econometrics. Econometrica 1978;46:1251–71.

Record Removal Notice

The World Bank Group
Archives
& Records Management

File Title Project Management Records - Synergistic Health Effects from Water Supply and Sanitation Interventions - (RPO # 677-25) - 1v	**Barcode No.** 1572444

Document Date 13 February, 1992	**Document Type** Form	

Correspondents / Participants

Subject / Title
The World Bank Research Program Request for Research Support Budget (RSB) Funding
Synergistic Effects from Water Supply and Sanitation Interventions

Exception(s)
Personal Information

Corporate Administrative Matters

Additional Comments

The item(s) identified above has/have been removed in accordance with The World Bank Policy on Access to Information or other disclosure policies of the World Bank Group.

Withdrawn by	**Date**
Sherrine M. Thompson	November 11, 2017

Archives 01 (March 2017)

Proposal for INUWS Research on:

Synergistic health effects from water supply and sanitation interventions

While there have been numerous attempts to quantify the health impacts from improving water supply or sanitation, few studies have considered factors which may alter the relationship between environmental sanitation and diarrheal disease. For example, improving water quality appears to have a much stronger effect on diarrheal disease in areas with adequate excreta disposal.[1] Similarly, improvements in sanitation may provide larger health benefits to children who are not breast-fed.[2]

The proposed research is an investigation of these two potentially "synergistic" effects. The specific questions to be addressed are:

1. Do simultaneous improvements to both water supply and sanitation result in a larger health impact than the sum of the impacts from each intervention alone? That is, do multiple environmental interventions have a "synergistic" effect in reducing diarrheal disease?

2. Does breast-feeding affect the health impact from improved water supply and sanitation? Do children who are not breast-fed realize greater health benefits?

P. 1st part of this is →
good para 2, but

what do you do about 1 & 2? Pretty micro for policy implications

The presence of such interaction effects would have important implications for planning water supply and sanitation programs, and for identifying high risk children. If improving both water supply and sanitation has a "synergistic" health effect, then a "package" of multiple environmental interventions should be encouraged. A strong interaction between breast-feeding and environmental sanitation may indicate that nonbreast-fed children are at a particularly high risk of diarrhea from environmental contamination, and thus should be a high priority target for water supply and sanitation interventions. → How at such micro level?

The proposed study will utilize existing data from the Cebu Longitudinal Health and Nutrition Survey (CLHNS), a prospective community-based investigation of 3327 children in Cebu, Philippines. A baseline survey and twelve bimonthly interviews were used to collect information on socioeconomic and demographic characteristics of the family, drinking water source and sanitation facilities, hygienic conditions around the house, feeding patterns and food preparation practices, and diarrheal morbidity from birth to two years of age. Data from all surveys has been cleaned and is available for this project.

Estimating interaction effects in health impact assessments is not easy, which may be one reason why they have been studied so little. In multivariate models, interaction terms typically have little power, requiring large sample sizes for stable parameter estimates. Simple stratified analyses,

754 households in
784 valid samples "

---

[1] Esrey S A, Feachem R G, Hughes J M. Interventions for the control of diarrheal diseases among young children: improving water supplies and excreta disposal facilities. Bull WHO. 1985;63:757-772.

[2] Butz W P, Habicht J P, DaVanzo J. Environmental factors in the relationship between breastfeeding and infant mortality: The role of sanitation and water in Malaysia. Am J Epid. 1984;119:516-525.

on the other hand, cannot control for potential confounders. The CLHNS is one of the few longitudinal data sets with sufficiently detailed behavioral and environmental information to explore these important interactions. Within the study population there is wide variation in the socioeconomic characteristics, feeding practices, and sanitary conditions. Furthermore, previous research with this data has shown water quality, sanitation, and breast-feeding to be significantly associated with diarrheal disease.[3] Selected papers from this research are attached.

---

[3] Cebu Study Team. Underlying and proximate determinants of child health: The Cebu longitudinal health and nutrition survey. Am J Epid. 1991;122:185-201.

Selected Prior Publications from the Same Research Effort

All Coliforms Are Not Created Equal: A Comparison of the Effects of Water Source and In-House Contamination on Infantile Diarrheal Disease

James VanDerslice, M.S.E.E.
Carolina Population Center and
Department of Environmental Sciences and Engineering
University of North Carolina at Chapel Hill

John Briscoe, Ph.D.
The World Bank
Washington, D.C.

In press

Submitted ~~to~~ Water Resources Research March 21, 1991

Revised August 30, 1991

NOT FOR CITATION

Short title: All Coliforms Are Not Created Equal

ABSTRACT

Storing drinking water in the home is common in the developing world, even among those served by in-house connections to piped supplies. Several studies have documented increased concentrations of fecal coliforms during household storage. This has led to the belief that in-house water contamination is an important transmission route for enteric pathogens and, moreover, that improving water source quality is not warranted until that quality can be maintained in the home. We contend that in-house contamination does not pose a serious risk of diarrhea because the pathogens contaminating a household's water are those already infecting the household members. A contaminated water source poses much more of a risk since it may introduce new pathogens into the household. The effects of water source and in-house contamination on diarrheal disease are estimated for 2355 Filipino infants. The results confirm our hypothesis: contaminated water sources pose a serious risk of diarrhea while contamination of drinking water in the home does not. Water boiling is shown to eliminate the risk of diarrhea due to water source contamination. The results imply that improvements in water source quality are more important than improving water storage practices.

Not
all right
for
infants

Very weak
indicator
of in-house
contamination
(184 observations
vs. 1600 for
source contamination)
don't adequately capture

Introduction

The past decade has seen a major effort to improve water quality and availability in the developing world. A principal goal of this effort has been to reduce the high levels of waterborne and water-washed diseases, notably diarrheal disease. However, families served by hand pumps, standpipes and even in-house connections from intermittent piped supplies, commonly store drinking water in the home where it may be contaminated by fecal material. This observation has lead to the belief that providing a high quality water supply is not worthwhile if the quality cannot be maintained during household storage [Feachem et al., 1983, p. 211].

Upon further reflection, however, this would appear to be incorrect. Consider a hypothetical person living by himself. He draws water from a well containing 100 fecal coliforms (FC) per 100 ml. He does not practice good hygiene so his stored drinking water contains 1000 fecal coliforms/100 ml. Does the higher level of contamination in his stored water mean that in-house contamination is more important than water source contamination? No. Since our hypothetical person himself is the source of the pathogens contaminating his drinking water during storage, these pathogens can do him no further harm.

Now consider a more realistic case of a family with poor hygienic practices. Pathogens infecting any family member are easily spread from his or her hands (contaminated during defecation) to food, utensils, or stored water. Soon the whole family would be infected by and would excrete a shared set of "internal" pathogens. Contamination of stored water by these "internal" pathogens would pose little additional risk to the family's health since family members would already be infected.

Contamination of the family's water source, however, is a very different matter. The contaminated water source may contain pathogens from other peoples' feces, pathogens that are new to the household environment. It is these "external" pathogens that can cause new infections that put the family at risk of diarrheal disease. Providing a high quality water source eliminates this source of "external" pathogens, reducing the family's risk of diarrhea.

Obviously this is a simplistic description of waterborne pathogen transmission. A contaminated water supply will probably contain both "internal" and "external" pathogens. However, if a contaminated water source is a principal means of introducing new pathogens into the home, then water source contamination will have a significant effect on diarrheal disease. Moreover, if in-house contamination is simply passing on "internal" pathogens already infecting the family, then it will not be significantly associated with diarrhea. The objective of this study is to estimate and compare the effects of water source contamination and in-house contamination on diarrheal disease.

Previous Studies of In-House Water Contamination

In-house water contamination has been recognized as a possible transmission route for enteric pathogens for over 25 years [van Zijl, 1966]. Even so, the existing literature provides little information on the risk it poses for diarrheal disease.

Most studies of in-house contamination have documented changes in water quality during storage by simply comparing the distribution of indicator organisms in water sources to the distribution in the storage containers. These studies have found substantial increases in coliform levels [Rajasekaran et al., 1977; Shiffman et al., 1978; El Attar et al., 1982; Lloyd-Evans et al., 1984; Magnani et al., 1984; Pickering, 1985; Lehmusluoto, 1986; Blum et al., 1990; Pinfold, 1990], little or no change in overall coliform levels [Oluwande, 1980; Esrey et al., 1986; Young and Briscoe, 1986; Sutton and Mubiana, 1989], and, in one case, a large decrease in mean coliform levels [Tompkins et al., 1978]. These aggregate measures are of limited use because they conceal the changes occurring in each household. Household-level changes in quality are observed by collecting paired water samples from the household's water source and storage container. Three studies, in Lesotho [Feachem et al., 1978], Malawi [Lindskog and Lindskog, 1988], and Sri Lanka [Mertens et al., 1990a] have used this method. These studies found considerable variation in the difference between water source and stored-water fecal coliform concentrations. The differences ranged from small decreases to increases of more than 1000 FC/100 ml.

The use of fecal coliforms as an indicator of fecal contamination in tropical waters has been seriously questioned [Hazen, 1988]. Three studies have circumvented this problem by analyzing stored water for the presence of pathogens. Spira et al. [1980] sampled water sources and stored water in Bangladesh for V. cholerae in neighborhoods where cholera cases had occurred. In a similar fashion, Echeverria et al. [1987] sampled for enterotoxigenic Escherichia coli (ETEC) in Thai neighborhoods where ETEC-positive diarrhea cases had come from. Neither researcher found any evidence that drinking water had been contaminated with these pathogens during storage in the home. A study in a rural Egyptian village [Khairy et al., 1982], however, did find Stronglyoides and Ascaris in 10% and 15% of the household storage jars, respectively, while the source water was free from these organisms. While stored water appears to be subject to contamination by parasites, the studies in Thailand and Bangladesh indicate that bacterial pathogen transmission via in-house contamination may not be as common as increases in fecal coliform levels suggest.

To our knowledge there have been no studies of the effect of in-house contamination on diarrheal disease. A cross-sectional study conducted in Nigeria [Huttly et al., 1987] found no association between the presence of a cover on the water storage container and the prevalence of diarrhea in any age group. Four studies of diarrheal disease, conducted in the Philippines [Magnani et al., 1984], Lesthoto [Esrey et al., 1986b], Bangladesh [Henry and Rahim, 1990] and Sri Lanka [Mertens et al., 1990b], used stored water quality as their measure of exposure to waterborne pathogens. None of these studies found a statistically significant relationship with diarrheal disease. These results provide no evidence that in-house contamination or water source contamination are important risk factors. Moreover, even if a significant association had been observed, it would have been impossible to distinguish the separate effects of water source and in-house contamination.

Methods

Study Design

This study uses data from the Cebu Longitudinal Health and Nutrition Survey, a prospective community-based investigation of infant health and nutrition in Cebu, Philippines. The study area consists of Cebu City and surrounding peri-urban areas and has an estimated population of 1 million. Seventeen of the 95 barangays (political districts) were randomly selected and pregnant women residing in these barangays were recruited. Of the 2555 women recruited, 2355 had single live births and agreed to participate in the study.

A baseline survey conducted during the third trimester of pregnancy collected information on the household's income, assets, water source, sanitation facilities and hygienic conditions. Bimonthly interviews, conducted through the first two years of the child's life, documented feeding patterns and food preparation practices, the volume of water consumed by the child, whether the water was boiled, and the specific water source used. Diarrheal morbidity for the previous 7 days and the child's weight and height were also recorded. Additional details on the survey design and content are available elsewhere [Cebu Study Team, 1991a].

Water Sampling

Special efforts were made to characterize exposure to waterborne pathogens from the household's water source and from in-house water contamination. Water sources were sampled between two and five times over the course of a year. Water samples were collected in the same manner that users collected their water. Spigots, pump spouts, and outflow pipes were not sterilized. Open dug wells without pumps were sampled using an aluminum bucket that was sterilized by flaming just before sampling.

Household water collection and storage practices were documented during a special survey of 254 households randomly selected from the larger study population. At each household two water samples were collected, one from the drinking water storage container and the other from the water source that had supplied the water in the storage container. Samples were collected from the storage container in the same manner as the family removed water. These pairs of samples were used to estimate the level of in-house contamination.

one time collection

Fecal coliforms were used as an indicator of fecal contamination. These organisms are continually present in large numbers in the feces of warm-blooded animals [Feachem et al., 1983] so their presence in drinking water indicates that the water has been contaminated with fecal material. While fecal coliforms appear to be more sensitive and specific than total coliforms in tropical waters [Lavoie, 1983], they are not ideal indicators of fecal contamination [Hazen et al., 1988]. They have been isolated in areas thought to be devoid of fecal material [Fujioka, et al. 1988], and may not be as persistent as some enteric pathogens [McFeters et al., 1974]. Nevertheless, the use of fecal coliforms is consistent with bacteriological water quality standards [W.H.O., 1984].

All water samples were transported on ice to the laboratory where they were refrigerated overnight and analyzed the following morning. Membrane filtration [American Public Health Association, 1985] was used to culture fecal coliform colonies using M-FC agar incubated at 44.5°C for 24 hours. Volumes of 1, 10, and 100 ml were filtered from each sample. Dark blue colonies were counted as fecal coliforms.

Of the 1650 water source samples collected, 154 (9%) produced unreliable estimates due to the presence of uncharacteristic colonies or heavy background growth. Only 2% of the 233 stored-water samples produced unreliable FC estimates. These estimates were excluded from the analyses.

When the number of colonies on a filter was in the "countable range" (10 to 100), that count was used to estimate the concentration of fecal coliforms. When more than one filter provided counts in the countable range, or when all filters had low counts (< 10), the total number of colonies counted was divided by the total volume of water filtered. When some of the filters had low counts and the other(s) were too numerous to count (TNTC), a maximum likelihood estimator was used [Haas and Heller, 1988]. Finally when all filters were TNTC, the estimate was set at 200 FC per 1 ml (20,000 FC/100 ml).

Estimating In-House Contamination

In-house water contamination is difficult to measure as the actual number of organisms added to the stored water can not be readily observed. The cumulative effect of in-house contamination will be reflected by an increase in bacterial concentrations. The bacteria observed in the storage container, however, will be a combination of those introduced by contaminated hands or cups, and those originating in the water source. Thus the concentration of FC due to in-house contamination (C_H) is the FC concentration observed in the storage vessel (C_v) minus the concentration of FC originally from the water source (C_s).

$$C_H = C_v - C_s \tag{1}$$

Standard bacteriological methods can not differentiate between fecal coliforms from in-house contamination and those from a contaminated source water. In order to estimate in-house contamination by this method, the concentration of FC in the storage container which came from the water source must be determined. While re-growth of fecal coliforms has been observed in nutrient-rich surface waters [Kinney et al., 1978; Hendricks, 1972; Carillo et al., 1985], re-growth during household storage was assumed to be negligible as virtually all households used groundwater for drinking. Therefore the concentration of water source FC in the storage container at the time of sampling is the concentration observed at the source (C_w) minus the concentration that died during storage (C_d).

$$C_s = C_w - C_d \tag{2}$$

There was no way to reliably estimate the number of FC from the water source which died during storage. However, even with no information on the level of die-off, upper and lower bounds on the level of in-house contamination can be calculated. The smallest value for in-house contamination (C_{Hmin}) occurs when all water source bacteria are assumed to survive (i.e., $C_s = C_w$).

$$C_{Hmin} = C_v - C_w, \text{ for } C_v > C_w, \text{ (net increase)} \tag{3}$$

$$C_{Hmin} = 0, \text{ for } C_v < C_w, \text{ (net decrease)} \tag{4}$$

The largest possible value for in-house contamination (C_{Hmax}) occurs when none of the water source fecal coliforms are assumed to survive (i.e., $C_s = 0$).

$$C_{Hmax} = C_v - 0 \tag{5}$$

A reasonable estimate of in-house contamination would be some point in this interval. Since exponential increases in pathogen dose are related to linear increases in the risk of diarrhea [Akin, 1981], $\log_{10}$ FC concentrations are used to model the effects of water contamination on diarrheal disease. As such, the level of in-house contamination was estimated as the mid-point between the minimum and maximum values measured on a log scale.

$$\log_{10}(C_H) = ½ [\log_{10}(C_{Hmax}) - \log_{10}(C_{Hmin})]. \quad (6)$$

In many cases the minimum and maximum values for $\log_{10}(C_H)$ were almost equal, indicating that the estimate of in-house contamination was not sensitive to the assumed level of water source FC die-off. For example, when the stored water contained ten times as many fecal coliforms as the source water, the difference between $\log_{10}(C_{Hmax})$ and $\log_{10}(C_{Hmin})$ was only 0.05. When the FC concentration in the storage container was only twice that observed at the water source, this difference was 0.3.

When the storage vessel was free of fecal coliforms, in-house contamination (C_H) was set to 0.9 FC/100 ml, the lower limit of detection. When the water source contained no fecal coliforms, the minimum and maximum estimates were equal and C_H was set to the concentration of FC observed in the storage vessel.

Diarrheal Model Specification

This research employs a previously developed longitudinal model of diarrheal disease [Cebu Study Team, 1991a, b]. In this model diarrhea

Δ weight ~~of the infant?~~ of the FC?

(D) results from past growth (G), behavioral factors (Y), and "underlying" socioeconomic and environmental factors (Z),

why longer lag here than in environment?

I don't understand the lag structure

$$D_{t,i} = \beta_1 G_{t-1,i} + \beta_2 Y_{t-1,i} + \beta_3 Z_{t,i} + \mu_{Di} + \epsilon_{Dt,i} \tag{7}$$

for t = 1 to 6 two-month time periods, and i = 1 to N study infants. Furthermore, the behavioral factors (Y) are determined by growth and diarrhea in the previous time period, past behaviors, and underlying socioeconomic and environmental factors.

H_2O boiling
soap use
breastfeeding
⋮

$$Y_{t,i} = \alpha 1\ G_{t-1,i} + \alpha_2 D_{t-1,i} + \alpha_3 Y_{t-1,i} + \alpha_4 Z_{t,i} + \mu_{Yi} + \epsilon_{Yt,i} \tag{8}$$

These equations contain two error terms. The first error term, μ, represents unobserved differences unique to each child or family. These may be the genetic endowment of the child, the parent's perceptions of risk, or other factors. These differences are expected to persist over the course of the study so the same error term is used for all time periods. The unobserved variations may affect each outcome differently. As such the μ's are different for each equation but correlated across equations. The second error term, ϵ, is a purely random disturbance which varies across individuals and with time, and is not correlated across equations.

Description of Variables

Diarrhea. Diarrhea is measured by a binary variable indicating whether the child experienced a diarrheal episode in the 7 days preceding the interview. It can be thought of as arising from a latent continuous measure of diarrheal severity. If the severity is greater than some threshold level, then the diarrheal episode is reported by the mother and the variable takes on the value of one. Otherwise the episode is not observed, and the variable will have the value of zero.

Behavioral factors. Many determinants of diarrheal disease are governed by behavior. 1 Exposure to waterborne pathogens, for example, is in part determined by the 1a choice of water source, 1b amount of water consumed, and 1c household water treatment. The behavioral variables used in the model measure various exposures to pathogens and factors affecting the child's susceptibility to infection.

Exposure to contaminated source water is measured as the $\log_{10}$ daily dose of FC from the water source. The dose was estimated by multiplying the infant's 24-hour total water intake by the expected FC concentration for the water source used by the household for the period two weeks prior to the interview date. In-house contamination is measured by the $\log_{10}$ daily dose of FC added during household storage. It was calculated by multiplying the infant's 24-hour total water intake by the estimated increase in FC concentration due to in-house contamination.

intake × (source × in-house increase coeffect)

Water boiling is expected to reduce the risk of diarrhea due to contaminated water to the same extent that contaminated water increases that risk. Interactions of water boiling with the two water contamination variables are included to model this effect. Since water boiling may also indicate a greater awareness of good hygiene, the main effect is also included. The water boiling variable indicates that the water consumed by the child the day before the interview had been boiled.

$C_H \Rightarrow \uparrow B$

$\uparrow B \Rightarrow \downarrow C_H$

Exposure to fecal contamination around the house is measured by two variables: the lack of toilet or latrine, and the presence of feces in the yard. The presence of feces was assessed through direct observations by trained fieldworkers.

what a job!

Several variables measure hygienic behaviors. The level of water service is used as a proxy for water use as families with an on-site water source are assumed to use more water for bathing, cleaning, and hand washing. Per capita nonlaundry soap usage, estimated from reported household expenditures for soap, is used as a proxy for personal hygiene. Household crowding, measured as the number of family members divided by the number of rooms, is used an indicator of higher person-to-person pathogen transmission. Finally, a variable indicating a high potential for food contamination was constructed from food preparation and storage practices at each longitudinal survey.

Breast-feeding may reduce the child's susceptibility to infection via maternal antibodies and provides nourishment which is free from contamination. Feeding patterns are measured by three dichotomous variables signifying whether the child was exclusively breast-fed, breast-fed and given nonnutritive supplements (such as plain water or juice), or given nutritive foods in addition to breastmilk. The omitted category is not breast-fed.

Use of preventive health care services is expected to improve the child's susceptibility to infection and may indicate that the mother has a greater awareness of her child's health. The variable indicates that some type of preventive health care (e.g., immunizations, well-baby check-up) was used in the two months preceding the interview.

Growth. The child's weight at the previous survey is included as a measure of nutritional status, an indicator of susceptibility to infection. The values are standardized at each cross-section.

Underlying factors. x Several underlying risk factors are thought to have direct effects on diarrheal disease. Age may reflect the immunological development of the child, secular trends in economic factors, and may capture age-related factors not adequately represented by the intermediate behavioral variables. Age-squared is included to capture nonlinearities. The child's sex, another commonly observed risk factor, may act as a proxy for unmeasured differences in immunological development between males and females, or represent differences in child-related behaviors.

Diarrhea has frequently been associated with season or rainfall. This may be due to enhanced survival of bacteria in humid weather, increases in water source contamination after large storms, or changes in food availability and prices during the growing season. The total rainfall in centimeters (cm) over the past two weeks is used to model these effects. Finally, community density is included as high-density areas are characterized by higher levels of environmental contamination.

Estimation Methods

In equation 7, diarrhea in the present time-period is specified as a function of past growth, past behaviors, (i.e. D_{t-1}) current socioeconomic and environmental conditions, and two error terms. This model differs from traditional research in two important ways. First, the model explicitly acknowledges that behaviors are determined in part by the child's health (equation 8). Secondly, the model allows for unobserved differences between children or their families, differences affecting both the family's behaviors and their child's health. These two refinements capture an important aspect of diarrheal disease in children: parents may recognize risks to their children's health and modify their behaviors to reduce those risks [Briscoe et al., 1990].

Failure to account for these effects can lead to biased parameter estimates and spurious results [Cebu Study Team, 1991a, b; Briscoe et al., 1990]. Consider the effect of a behavioral factor on diarrhea. The behavioral structural equation (8) specifies that all behaviors are determined in part by μ_Y, the random error representing unobserved differences between children. Since the unobserved differences which affect behaviors may also affect diarrhea, μ_Y is correlated with μ_D, the equivalent disturbance term in the diarrhea equation. As a result, the behavioral variables in the diarrhea equation are correlated with the error term μ_D. This is a violation of one of the basic assumptions of the ordinary-least-squares (OLS) estimator and if OLS is used the estimated effect of behaviors on diarrhea will be biased (inconsistent). This is because some of the variability in the dependent variable due to random error is mistakenly attributed to the independent variables.

Consistent estimates can be obtained if the behavioral variables are purged of their association with the unobserved factors, μ_D. This can be accomplished by using instrumental variables in place of the behavioral variables. Instrumental variables are variables correlated with the behavioral risk factors, but not correlated with the individual-specific error term.

Suitable instruments can be derived from the behavioral structural equation. If the growth, diarrhea and behavioral variables on the right-hand side are substituted out using their respective structural equations, equation 8 becomes

$$Y_{t,i} = \delta_1 G_{t-2,i} + \delta_2 D_{t-2,i} + \delta_3 Y_{t-2,i} + \delta_4 Z_{t,i} + \delta_5 Z_{t-1,i} + \mu_{Yi} + \epsilon_i \qquad (9).$$

The same variables (at t-2) can be substituted out again. This process is repeated until only the underlying variables and error terms remain.

$$Y_{t,i} = \gamma_1 Z_{t,i} + \gamma_2 Z_{t-1,i} + \ldots + \gamma_s Z_{t-s,i} + \mu_{Yi} + \epsilon_i \qquad (10)$$

This is the reduced-form of the behavior structural equation, specifying any behavior as a function of strictly exogenous, underlying variables.

The reduced-form equation can be used to create instruments for the behavioral variables. Since the underlying variables, Z, are not affected by the family's behaviors, they are not correlated with the individual-specific error term, μ_D. As such equation 10 can be estimated using standard techniques. Predicted values for the behavior variable, $Y_{t,i}$, are then generated and used as the instrumental variable. The predicted values should be well correlated with the actual values and not correlated with the error term μ_D. The same procedure was used to generate predicted values for the growth variable. Using these predicted values in place of the growth and behavioral variables produces consistent estimates of the parameters in the diarrheal equation [Judge et al., 1982].

Since the dependent variable in eqtn 7 i.e. $D_{t,i}$ is binary and the error term is assumed to be normally distributed, a random-effects probit estimator was used. This estimator assumes the same "random-effect" for all observations from a given child [Judge et al., 1982; Avery and Hotz, 1985]. This random effect represents the "unobserved" characteristics of that child (or family) affecting the child's health and the family's behaviors.

A maximum likelihood procedure, found in the HOTZTRAN software [Avery and Hotz, 1985], was used to estimate the parameters. The standard errors may be underestimated because the variation associated with use of instruments is not taken into account. While it is theoretically possible to correct for the use of instruments [Maddala, 1983], it is not feasible given the large number of instruments used.

The coefficients estimated from a probit model can not be interpreted as the marginal effect of an independent variable on the probability of diarrhea. The marginal change in the probability of diarrhea resulting from a unit change in a dependent variable, X_k, was calculated by [Maddala, 1983, p. 23]:

?

$$\partial\, P(D=1)/\partial\, X_k \;=\; \phi(X\,\beta)\,\beta_k \qquad (11).$$

Approximate confidence intervals for the marginal effects were calculated by using the end-points of the 95% confidence interval of the parameter estimate in place of the parameter estimate itself.

$$\text{lower 95\% confidence limit} = \phi(X\,\beta)\ [\beta - 1.96(\text{s.e.}_{\beta})] \qquad (12)$$

$$\text{upper 95\% confidence limit} = \phi(X\,\beta)\ [\beta + 1.96(\text{s.e.}_{\beta})] \qquad (13)$$

Confidence intervals for the marginal effects will include the null value of zero when the confidence interval for the parameter estimate includes zero.

Results

Characteristics of the Study Population

There is wide variation in the demographic characteristics of the study population (Table 1). Education levels are quite high in Cebu; over 90% of the parents have completed primary education and 15% have graduated from high school. Most of the households (70%) are headed by waged or salaried workers and one-fourth are self-employed. Household incomes range from 0 to 12,500 pesos per week with a median of 200 (approx. $ U.S. 10). Total household assets range from 0 to almost 1.5 million pesos with a median of 2,400 (approx. $ U.S. 120).

Environmental sanitation conditions are also quite variable. Over three-quarters of the households use an "adequate" excreta disposal facility (i.e., flush or pour-flush toilet or latrine). However, there is no sewerage in Cebu City and most on-site disposal systems would be considered inadequate. Almost 20% of the families report that they defecate into a canal or on the seashore. Fecal material was observed at one-third of the sample houses.

Water Source Use and Quality

Over 500 water sources are used by the study population. Almost all households use an "improved" water source; 59% are served by boreholes and 30% by the municipal piped supply. The remaining households rely on open dug wells (5%), or dug wells fitted with pumps (5%). The sample population also enjoys a relatively high level of service; 10% have in-house connections and another 48% are within 1 minute of their water source. Only a small proportion (5%) must walk more than 5 minutes to fetch water.

Boreholes and the piped supply generally provide high quality water (Table 2). Over three-fourths of the samples from these sources produced no FC colonies, and another 10% had less than 10 FC/100 ml. Still, over 10% of the boreholes and 10% of the samples from the piped supply were contaminated with more than 100 FC/100 ml. Dug wells had much higher levels of contamination. Those fitted with covers and pumps were grossly contaminated (> 100 FC/100 ml) less often than open dug wells (41% vs. 78%).

Water Collection and Storage Practices

Over 99% of the households report storing drinking water in the home, including many of those with in-house connections. Almost all households have only one storage container, and the stored water is used for several purposes (e.g., drinking, cooking, bathing and cleaning). While many types of containers were used, they can be classified into four categories: small containers (e.g., pitchers and used Clorox bottles), large containers (e.g., 6-gallon gasoline cans and used cooking oil cans), traditional clay jars, and pails. The small containers, large containers, and clay jars were each used by about one-third of the study households (Table 3). Small containers were frequently used for both collection and storage. About half of the earthen jars were subject to contamination from scoops or cups. The remaining jars were fitted with spigots. Water was usually poured from the other types of containers. Most of the containers had covers or caps.

Changes in Water Quality During Storage

When both the water source and stored-water samples were "too numerous to count", it was impossible to determine if the concentration of fecal coliforms had increased, decreased, or remained the same during storage. These samples, as well as those with unreliable counts due to heavy background growth, were not used in the analysis, leaving only 184 of the 233 pairs of water samples collected.

Table 4 presents the distribution of the change in concentration between the source and the storage container and the levels of in-house contamination estimated using equation 6. One-third of the household

$C_H = C_v - C_s =$ $\log_{10} C_H = (\max \log_{10} C_H - \min \log_{10} C_H)$

samples had substantially higher concentrations of fecal coliforms (>100 FC 100 ml) than the respective water source sample. Over 30% of the sample pairs demonstrated no net change (-1 to 1) in fecal coliform concentrations. This may reflect no change in quality, or the combination of high levels of in-house contamination and die-off. Surprisingly, 16% of the sample pairs demonstrate a net decrease in FC concentration, indicating that bacterial die-off can be greater than increases due to in-house contamination.

Covering the storage container appears to have little effect on in-house contamination. The geometric mean (G.M.) of the estimated increase in the concentration fecal coliforms per 100 ml was 1.82 for covered containers and 1.51 for containers that were not covered. Samples taken from containers from which water was scooped demonstrated slightly larger increases (G.M.=1.96) than samples from containers where water was poured or flowed through a spigot (G.M.=1.48). Small storage containers were subject to less in-house contamination (G.M.=1.16) than the large containers (G.M.=1.79) or earthen storage jars (G.M.=1.92).

Water Consumption and Boiling

The proportion of children fed plain water (i.e., not as a part of food or a prepared drink) increased substantially over the first 6 months (Table 5). By the time the children were 8 months old, over 99% had received water in the 24 hours preceding the interview. Mean consumption for those fed any water almost doubled over the first year. About 90% of the mothers reported that they boiled the water given to their 2 month old infants. This proportion dropped to about half by the time the child was 6 months old and remained constant for the rest of the child's first year of life.

Exposure to Water Source and In-House contamination

The distributions of predicted daily FC doses from the water source and from in-house contamination for all 6 time periods combined baseline + up to 5 observations are presented in Table 6. The low doses were censored to 1 fecal coliform per day. Ten percent of the in-house contamination doses and 84% of the water source doses were so censored. ?

Diarrheal Disease in Cebuano Infants

The proportion of children experiencing diarrhea during the week previous to the interview increased dramatically over the first 8 months of the child's life, from just over 7% when the infants were 2 months old to 25% just 6 months later. During this period the prevalence was the same for males and females. From 8 months to a year of age the prevalence of diarrhea prevalence among males continued to increase while the prevalence among females decreased slightly to 22%. At one year of age, female children were experiencing about 20% fewer cases of diarrhea than male children.

Effects of Water Source and In-House Contamination on Diarrheal Disease

The parameter estimates and t-statistics for two models of diarrheal disease are presented in Table 7. In the first model of the effects of water contamination on diarrhea, water source dose is a small but very significant ~~strong~~ risk factor (t=3.0 2.6). Increasing the dose from 1 to 100 FC per day increases the probability of diarrhea 22%, from 0.18 to 0.22. In-house contamination, however, is not associated with diarrhea ✓ and has a point estimate very close to zero. Excreta disposal, food pathogenicity, age,

?,

and community density are all significant risk factors, and exclusive breast-feeding a significant protective factor. The signs of all significant and marginally significant coefficients are as expected.

This simple model ignores the fact that water boiling should have a greater protective effect when water contamination levels are high. This effect was modeled by including interactions of boiling with the two water contamination variables (model #2). In the case of water source contamination, the boiling interaction coefficient is negative indicating that boiling reduces diarrhea more as the level of contamination increases. Water source dose is still statistically significant and the interaction is marginally significant. Neither in-house contamination nor its interaction with boiling are statistically significant. The other parameter estimates do not change appreciably from the first model.

Figure 1 illustrates the importance of the interaction between water source contamination and boiling. Predicted probabilities of diarrhea over a 7-day period were computed for various levels of water source dose when the water was boiled and not boiled. Separate predictions were made using the model including the interaction between boiling and water contamination, and the model containing only their main effects.

When the interaction is not included, the increase in the probability of diarrhea from a ten-fold increase in water source dose is the same whether or not the water is boiled. The model including the interaction term gives much more intuitive results. When water is boiled, water source contamination does not increase the probability of diarrhea. However, source contamination has a considerable effect on diarrhea when water is not boiled.

Marginal increases in the probability of diarrhea resulting from unit increases in the water contamination and water boiling variables are presented in Table 8. Each $\log_{10}$ increase in the water source fecal coliform dose increases the probability of diarrhea by 0.043. The interaction of source contamination with water boiling has exactly the opposite effect, reducing the probability of diarrhea by 0.044 per $\log_{10}$ increase in dose. Thus the effect of the boiling interaction variable is to cancel out the risk due to water source contamination.

where

Discussion

The results from the diarrhea models confirm our hypothesis: in-house water contamination does not pose a serious risk of diarrhea. Its parameter estimate is not significant, or even positive, in either of the models estimated.

The model results, however, indicate that water source contamination poses a significant risk for diarrhea. When water is not boiled, water source contamination substantially increases the probability of diarrhea. A ten-fold increase in the concentration of fecal coliforms would lead to a 17% increase in diarrheal prevalence. Conversely, if families using moderately contaminated dug wells (100 FC/100 ml) were able to use a high quality water source, diarrhea among their children would be reduced by over 30%.

Sanitation is also an important risk factor for diarrhea. The relative importance of water contamination, sanitation, and the level of service, as well as the effects of multiple interventions will be addressed in a forthcoming paper.

In summary, all coliforms are not created equal. There are important differences between in-house water contamination by "internal" pathogens and contamination of one's water source by "external" pathogens. The implications for planning improvements to water supplies are clear. Improving water source quality can have a substantial impact on diarrheal disease. Eliminating in-house contamination may have no impact unless other household transmission routes are eliminated as well. In any case, there is no reason to delay making improvements in water source quality because of contamination occurring in the home.

Notation List

C_d	concentration of fecal coliforms from the water source which died during storage.
C_H	concentration of fecal coliform due to in-house contamination.
C_{Hmax}	maximum value for the concentration of fecal coliform due to in-house contamination.
C_{Hmin}	minimum value for the concentration of fecal coliform due to in-house contamination.
C_s	concentration of water source fecal coliform in the storage container at the time of sampling.
C_v	fecal coliform concentration observed in the storage vessel.
C_w	concentration of fecal coliform observed at the water source.
D	diarrhea.
ETEC	enterotoxigenic *Escherichia coli*.
FC	fecal coliforms.
G	growth.
G.M.	geometric mean.
OLS	ordinary-least-squares.
$s.e._{\beta}$	standard error of the estimate of β.

Notation list cont.

TNTC	too numerous to count.
Y	behavioral factors.
Z	"underlying" socioeconomic and environmental factors.
ϕ	normal density function.
μ_D	individual-specific random error affecting diarrhea, D.
μ_Y	individual-specific random error affecting behaviors, Y.

Acknowledgments

This article is part of a collaborative research project involving the Office of Population Studies, directed by Wilhelm Flieger, and the Water Resources Center, directed by Herman van Engelen, both of the University of San Carlos, Cebu, Philippines; the Nutrition Center of the Philippines, directed by Florentino S. Solon; and a group from the Carolina Population Center, University of North Carolina at Chapel Hill (UNC-CH). Barry Popkin of UNC-CH is the project coordinator. Funding for parts of the project design, data collection, and computerization was provided by the National Institutes of Health (Contract R01-HD19983A, R01-HD23137, and R01-HD18880), the Nestle's Coordinating Center for Nutrition Research, Wyeth International, the Ford Foundation, the U.S. National Academy of Sciences, the Carolina Population Center, the U.S. Agency for International Development, and The World Bank. Funds for data analysis for this study were provided by the National Institutes of Health (R01-HD19983A).

We acknowledge the valuable help of David Guilkey for the development of the statistical models and methods. We also thank David Fugate, Gina Dahiya, Christine Moe, Arlene Quijada, Rexy Ruiz, and Ruben Quijada for their technical assistance.

References

Akin, E. W., A review of infective dose data for enteroviruses and other enteric microorganisms in human subjects, in *Microbiological Health Considerations of Soil Disposal of Domestic Wastewater*, National Center for Groundwater Research, and U.S. Environmental Protection Agency, Norman, Okla., 1981.

American Public Health Association, *Standard Methods for the Examination of Water and Wastewater, 15th ed.*, pp. 814-822, Washington D.C., 1985.

Avery, R. B., and V. J. Hotz, *Hotztran User's Manual*, CERA Economic Consultants, Inc., Old Greenwich, Conn., 1985.

Blum, D., R. N. Emeh, S. R. A. Huttly, O. Dosunmu-Ogunbi, N. Okeke, M. Ajala, J. I. Okoro, C. Akujobi, B. R. Kirkwood, and R. G. Feachem, The Imo State (Nigeria) drinking water supply and sanitation project, 1. Description of the project, evaluation methods, and impact on intervening variables, *Trans. Royal Soc. Trop. Med. Hyg.*, 84, 309-315, 1990.

Briscoe, J., J. Akin, and D. Guilkey, People are not passive acceptors of threats to health: Endogeneity and its consequences, *Int. J. Epid.*, 19, 147-153, 1990.

Carillo, M., E. Estrada, and T. C. Hazen, Survival and enumeration of fecal indicators Bifidobacterium adolescentis and Escherichia coli in a tropical rain forest watershed, *Appl. Envr. Micro.*, 50, 468-476, 1985.

Cebu Study Team, Underlying and proximate determinants of child health: The Cebu longitudinal health and nutrition survey, *Am. J. Epid.*, 133, 185-201, 1991a.

Cebu Study Team, A child health production function estimated from longitudinal data, *J. Dev. Econ.*, in press, 1991b.

Echeverria, P., D. N. Taylor, J. Seriwatnana, U. Leksomboon, W. Chaicumpa, C. Tirapat, and B. Rowe, Potential sources of enterotoxigenic Escherichia coli in homes of children with diarrhoea in Thailand, *Bull. WHO*, 65, 207-215, 1987.

El Attar, L., A. A. Gawad, A. E. M. Khairy, and O. El Sebaie, The sanitary condition of rural drinking water in a Nile Delta village, II. Bacterial contamination of drinking water in a Nile Delta village, *J. Hyg.*, 88, 63-67, 1982.

Esrey, S. A., J. P. Habicht, G. Casella, M. D. Miliotis, A. H. Kidd, J. Collett, V. Qheku, and M. C. Latham, Infection, diarrhea and growth rates of young children following the installation of village water supplies in Lesotho, paper presented at the 22nd Annual Conference of the American Water Resources Association, International Conference on Water and Human Health, Atlanta, Ga., November 9-14, 1986.

Feachem, R. G., E. Bruns, A. Cronin, P. Cross, D. Curtis, M. K. Khan, D. Lamb, and H. Southall, *Water, Health and Development: An Interdisciplinary Evaluation*, Tri-med books, LTD., London, 1978.

Feachem, R. G., D. J. Bradley, H. Garelick, and D. D. Mara, *Sanitation and Disease: Health Aspects of Excreta and Wastewater Management*, John Wiley and Sons, New York, 1983.

Fujioka, R. S., K. Tenno, and S. Kansako, Naturally occurring fecal coliforms and fecal streptococci in Hawaii's freshwater streams, *Tox. Asses.*, 3, 613-630, 1988.

Haas, C. N., and B. Heller, Averaging of TNTC counts, *Appl. Environ. Microbiol.*, 54, 2069-2072, 1988.

Hazen, T. C., Fecal coliforms as indicators in tropical waters: A review, *Tox. Asses.*, 3, 461-477, 1988.

Hendricks, C. W., Enteric bacterial growth rates in river water, *Appl. Micro.*, 24, 168-174, 1972.

Henry, F. J., and Z. Rahim, Transmission of diarrhoea in two crowded areas with different sanitary facilities in Dhaka, Bangladesh, *J. Trop. Med. Hyg.*, 93, 121-126, 1990.

Huttly, S. R. A., D. Blum, B. R. Kirkwood, R. N. Emeh, N. Okeke, M. Ajala, G. S. Smith, D. C. Carson, O. Dosunmu-Ogunbi, and R. G. Feachem, The Imo State (Nigeria) drinking water supply and sanitation project, 2. Impact on dracunculiasis, diarrhoea and nutritional status, *Trans. Royal Soc. Trop. Med. Hyg.*, 84, 316-321, 1990.

Judge, G. G., R. C. Hill, W. Griffiths, H. Lutkepohl, and T. Lee, Introduction to the theory and practice of econometrics, John Wiley and Sons, New York, 1982.

Khairy, A. E. M., O. El Sebaie, A. A. Gawad, and L. El Attar, The sanitary condition of rural drinking water in a Nile Delta village, I. Parasitiological assessment of 'zir' stored and direct tap water, *J. Hyg.*, 68, 57-61, 1982.

Kinney, E. C., Drummond, D. W., and Hanes, N. B., Effects of chlorination on differentiated coliform groups, *J. Water Poll. Control Fed.*, 50, 2307-2312, 1978

Lavoie, M. C., Identification of strains isolated as total and fecal coliforms and comparisons of both groups of indicators of fecal pollution in tropical climates, *Can. J. Micro.*, 29, 689-693, 1983

Lehmusluoto, P. O., *Survey on Contamination of Water Sources and Household Waters as an Integrated Part of Impact Management*, University of Tampere, Tampere, Finland, 1986.

Lindskog, P., and U. Lindskog, Bacterial contamination of water in rural areas. An intervention study from Malawi, *J. Trop. Med. Hyg.*, 91, 1-7, 1988.

Lloyd-Evans, N., H. Pickering, S. G. J. Goh, and M. G. M. Rowland, Food and water hygiene and diarrhoea in young Gambian children: a limited case control study, *Trans. Royal Soc. Trop. Med. Hyg.*, 78, 209-211, 1984.

Magnani, R., S. Tourkin, and M. Hartz, *Evaluation of the Provincial Water Project in the Philippines*, International Statistics Program Center, Bureau of the Census, U.S. Department of Commerce, Washington, D.C., 1984.

McFeters, G. A., G. K. Bissonnette, J. J. Jezeski, C. A. Thomson, and D. G. Stuart, Comparative survival of indicator bacteria and enteric pathogens in well water, *Appl. Micro.*, 27, 823-829, 1974.

Mertens, T. E., M. A. Fernando, T. F. de C. Marshall, B. R. Kirkwood, S. Cairncross, and A. Radalowicz, Determinants of water quality, availability and use in Kurunegala, Sri Lanka, *Trop. Med. Parasit.*, 41, 89-97, 1990a.

Mertens, T. E., M. A. Fernando, S. N. Cousens, B. R. Kirkwood, T. F. de C. Marshall, and R. G. Feachem, Childhood diarrhoea in Sri Lanka: a case-control study of the impact of improved water sources, *Trop. Med. Parasit.*, 41, 98-104, 1990b.

Oluwande, P. A., Public health implications of chemical and bacteriological characteristics of treated water in Western State of Nigeria, *Nig. Med. J.*, 7, 86-90, 1980.

Pickering, H., Social and environmental factors associated with diarrhoea and growth in young children: Child health in urban Africa, *Soc. Sci. Med.*, 21, 121-127, 1985.

Pinfold, J. V., Faecal contamination of water and fingertip-rinses as a method for evaluating the effect of low-cost water supply and sanitation activities on faeco-oral disease transmission, I. A case study in rural northeast Thailand, *Epid. and Inf.*, 2, 363-375, 1990.

Rajasekaran, P., P. R. Dutt, and K. A. Pisharoti, Impact of water supply on the incidence of diarrhoea and sigellosis among children in rural communities in Madurai, *Ind. J. Med. Res.*, 66, 189-199, 1977.

Shiffman, M. A., R. Schneider, J. M. Faigenblum, R. Helms, and A. Turner, Field studies on water, sanitation and health education in relation to health status in Central America, *Prog. Wat. Tech.*, 11, 143-150, 1978.

Spira, W. M., M. U. Khan, Y. A. Saeed and M. A. Sattar, Microbiological surveillance of intra-neighborhood El Tor cholera transmission in rural Bangladesh, *Bull. WHO*, 58, 731-740, 1980.

Sutton, S., and D. Mubiana, Household water quality in rural Zambia, *Waterlines*, 8, 20-21, 1989.

Tompkins, A. M., B. S. Drasar, A. K. Bradley, and W. A. Williamson, Water supply and nutritional status in rural Northern Nigeria, *Trans. Royal Soc. Trop. Med. Hyg.*, 72, 239-243, 1978.

van Zijl, W. J., Studies on diarrhoeal diseases in seven countries by the WHO Diarrhoeal Diseases Advisory Team, *Bull. WHO*, 35, 249-261, 1966.

Young, B., and J. Briscoe, *Water and Health in Rural Malawi*, U.S. Agency for International Development, Washington, D.C., 1986.

World Health Organization, *Guidelines for Drinking Water Quality, Vol. 1, Recommendations*, Geneva, 1984

Author address list

James VanDerslice
Carolina Population Center
University Square, CB# 8120
University of North Carolina
Chapel Hill, N.C. 27516-3997

Phone: (919) 966-7143
Fax: (919) 966-6638

John Briscoe
The World Bank
S11 - 105
1616 H Street, N.W.
Washington D.C. 20433

Figure legends

Figure 1. Effect of water source fecal coliform dose on the predicted probability of diarrhea

TABLE 1. Means and Standard Deviations of Selected Household and Community Factors

Variable	Mean or proportion	S.D.
Child is male	0.53	
Mother's age	25.89	5.84
Father's age	28.71	6.73
Spouse is present	0.94	
Mother's highest grade completed	7.61	3.30
Father's highest grade completed	7.97	3.41
Household has extended family members	0.42	
Household has electricity	0.60	
Household owns radio	0.55	
Household owns television	0.22	
Household owns refrigerator	0.08	
Household income	231.18	362.56
Value of assets (pesos x 10^{-3})	12.99	51.49
Municipal piped supply is available	0.56	
Boreholes are available	1.00	
Dug wells are available	0.24	
Springs are available	<0.01	
Household has good excreta disposal facility	0.77	
Excreta observed around house	0.33	

Table 1. cont.

Distance to nearest road (m)	95.36	214.20
Population density (pop. x 10^{-3} / km^2)	19.32	20.39
Total rainfall in past 2 weeks (cm)	6.31	4.0
Number of days with rain in past 2 weeks	6.32	2.4

TABLE 2. Water Source Fecal Coliform Concentrations, by Type of Water Source

Water source type	# of samples[a]	$\log_{10}$ concentrations: Mean	$\log_{10}$ concentrations: Std. dev.	% of samples with: No colonies	% of samples with: TNTC
Piped supply	111	-2.22	2.90	78	5
Boreholes	403	-1.95	3.00	75	2
Improved dug wells	46	1.95	1.33	9	7
Unimproved dug wells	60	2.55	1.00	3	15

[a] Due to the estimation difficulties arising from the censored observations (i.e., TNTC and zero counts) and the variable number of samples taken from each water source, these statistics are based on one randomly selected sample per water source, except for the piped supply where all samples were used.

TABLE 3. Household Water Storage Practices by Type of Container

Type of container	number & percent	% of containers used for collection	water scooped	covered	Mean # of trips per day
Small containers	84 (33%)	86	1	79	3.1
Large containers	78 (31%)	18	0	67	2.5
Clay jars	85 (33%)	0	44	94	2.6
Pails	7 (3%)	25	100	29	9.0
Overall	254(100%)	45	18	79	3.0

TABLE 4. Difference in Fecal Coliform Concentrations between Source and Storage, and Estimated Levels of In-house Contamination

Change in concentration of fecal coliforms	Observed change #	Observed change %	Estimated in-house contamination #	Estimated in-house contamination %
Net decrease:				
< -10,000 (TNTC)	7	3.8		
-1,000 to -10,000	5	2.7		
-1,000 to -100	10	5.4		
-100 to -10	3	1.6		
-10 to -1	4	2.2		
No change: (-1 to 1)	58	31.5	56	30.4
Net increase:				
1 to 10	19	10.3	26	14.1
10 to 100	19	10.3	38	20.7
100 to 1,000	22	12.0	25	13.6
1,000 to 10,000	19	10.3	21	11.4
> 10,000 (TNTC)	18	9.8	18	9.8
TOTAL	184	100.0	184	100.0

TABLE 5. Water Consumption and Water Boiling in Past 24 Hours, by Child's Age

	Child's age (months)					
Variable	2	4	6	8	10	12
Percent fed plain water	38.1	58.2	86.5	95.1	97.0	97.9
Percent fed any water	75.9	82.9	96.1	99.2	99.3	99.7
Among those fed any water:						
Total amount consumed per day (ml):						
Mean	363	408	425	489	558	647
Std. dev.	416	482	482	491	506	509
Percent that boiled water before serving	86.8	78.0	61.4	52.0	49.8	50.6

TABLE 6. Distributions of Predicted Daily Fecal Coliform Doses

Predicted daily fecal coliform dose	Water source Frequency	Water source Percent	In-house Frequency	In-house Percent
$< 10^{-3}$ 0.9FC/100 ml.	941	7.8	0	0.0
10^{-3} to 10^{-2}	3329	27.7	58	0.5
10^{-2} to 10^{-1}	3813	31.8	272	2.3
10^{-1} to 10^{0}	1971	16.4	837	7.0
10^{0} to 10^{1}	701	5.8	1954	16.3
10^{1} to 10^{2}	601	5.0	1980	16.5
10^{2} to 10^{3}	539	4.5	3307	27.6
10^{3} to 10^{4}	96	0.8	2413	20.1
10^{4} to 10^{5}	8	0.1	932	7.8
10^{5} to 10^{6}	0	0.0	176	1.5
$> 10^{6}$	1	0.0	71	0.6
TOTAL	12000	99.9	12000	100.2

TABLE 7. Parameter Estimates and T-Statistics from Probit Models of Diarrheal Disease

	Main effects		Boiling interaction	
	(Model #1) eqtn 8		(Model #2) eqtn 10	
Variable	β	t	β	t
Intercept	-1.813	-5.2***a	-1.778	-5.0***
Rho	0.121	7.6***	0.121	7.6***
Water contamination:				
Water source $\log_{10}$ FC dose $\log_{10}(C_s)$	0.068	2.6***	0.168	2.6***
In-house $\log_{10}$ FC dose $\log_{10}(C_H)$	-0.002	-0.1	-0.028	-0.9
Water boiling:[b]				
Main effect	-0.094	-0.5	-0.193	-0.9
Interacted with:				
Water source FC dose ↑FC → ↓diarrhea w/ boiling			-0.171	-1.7*
In-house FC dose			0.049	1.0
Poor excreta disposal[b]	0.238	1.8*	0.220	1.6*
Excreta around the house[b]	0.364	2.7***	0.365	2.7***
Water source on-site[b]	-0.157	-1.2	-0.106	-0.8
High food pathogenicity[b]	0.718	1.8*	0.736	1.8*
Soap use (mkg/person day)	-0.011	-0.2	-0.006	-0.1
Household density (persons/room)	-0.024	-0.5	-0.018	-0.3
Preventive health care use[b]	-0.125	-0.6	-0.113	-0.6
Breast-feeding and nutritive supplements[b]	-0.028	-0.2	-0.033	-0.2
Breast-feeding and nonnutritive supplements[b]	-0.320	-1.2	-0.360	-1.3

N=1496
N=184?

4

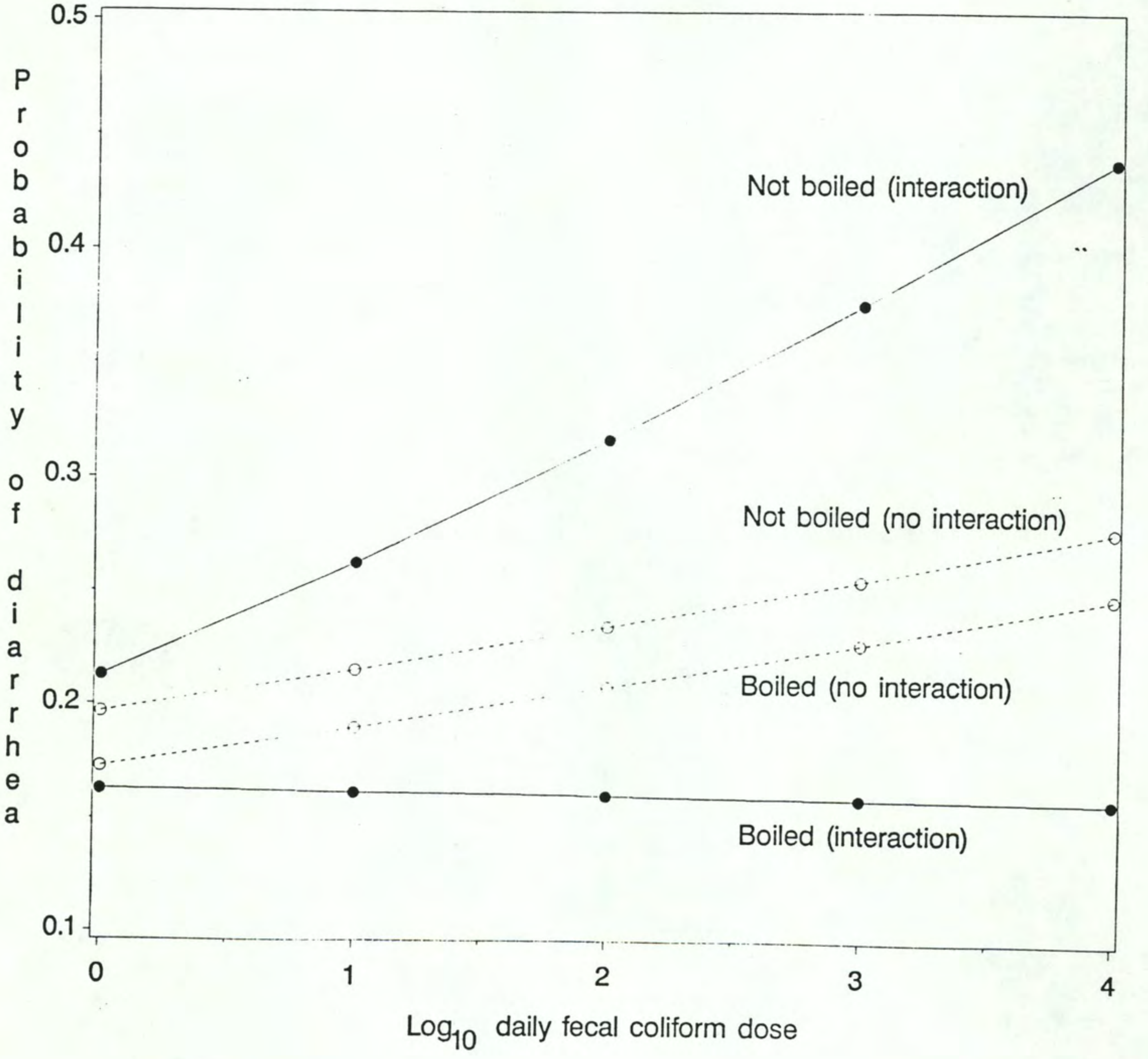
Table 7. cont.

2 Breast-feeding only[b]	-0.615	-2.2**	-0.560	-2.0**
growth Standardized weight (std. dev.)	0.003	0.1	0.003	0.1
Male child	0.060	1.4	0.062	1.4
Child's age (weeks)	0.048	6.6***	0.048	6.5***
Child's age squared (weeks^2)	-0.001	-6.1***	-0.001	-6.0***
Community density (10^3 persons/km^2)	0.004	2.8***	0.004	2.7***
Cumulative rainfall (cm)	0.004	1.2	0.004	1.3

---

[a] The stars denote the level of significance for a two-tailed test: *** = $p<0.01$, ** = $p<0.05$, * = $p<0.10$.

[b] These variables are the predicted probability that the child has the stated characteristic.


Probability of diarrhea
0.5
0.4
0.3
0.2
0.1
Not boiled (interaction)
Not boiled (no interaction)
Boiled (no interaction)
Boiled (interaction)
0
1
2
3
4
Log10 daily fecal coliform dose

Correspondence

F

ALL-IN-1 NOTE

DATE: 09-Apr-1993 05:56pm

TO: Gregory Ingram (GREGORY INGRAM)

FROM: John Briscoe, TWUWS (JOHN BRISCOE)

EXT.: 35557

Research Grant

Greg:

I am sending you copies of two papers produced by myself and Dr. James Vanderslice under the grant provided to us.

I hope that you will be pleased with the outcome of this effort, which I think has produced two papers which are significant contributions both in terms of methodology and policy content.

As you can see from the letter from Dr. Vanderslice, the research specifically addressed the methodological issues brought to our attention in the review process. As you can see, both papers are being submitted to professional journals. I'll keep you apprised on the outcome.

Once again many thanks for your efficient help on this.

John

CC: Louis Pouliquen (LOUIS POULIQUEN)

F

THE WORLD BANK / IFC / MIGA

OFFICE MEMORANDUM

DATE: June 5, 1992

TO: Mr. John Briscoe, INUWS

FROM: Gregory Ingram, Administrator, Research Advisory Staff

EXTENSION: 31052

SUBJECT: Research Project 677-25: "Synergistic Health Effects from Water Supply and Sanitation Intervention"

Thank you for your memorandum of May 5th, the related correspondence regarding the addition of an econometrics adviser to the project and your efforts to respond to the issues I raised in my March 10th memorandum. I regret that you have not found common ground with Dr. Bhargava, and Dr. Atkin does not find it appropriate to provide Dr. Bhargava with the created variables or instruments to allow him to work with you to estimate the continuous dependent variable, as he proposed.

Regardless of your decision with regard to Dr. Bhargava's role in the research, I hope you will continue to work with Dr. Guilkey to alleviate some of the methodological problems. Insofar as you intend to proceed with the approach outlined in the proposal, however, I would expect to see greater acknowledgement of the econometric constraints and their effects on the interpretations of your findings in the output of this study than I find in the other Cebu study papers.

Please let me know what proportion of the $4,000 increment to the RPO you have used or intend to use in addressing these problems. We will reduce the FY93 allocation to your RPO by the difference.

GKI:et

cc: L. Pouliquen, INUDR; C. Carnemark, INUWS
S. Shah, C. Else, V. Mataac, RAD

ALL-IN-1 NOTE

DATE: 03-Jun-1992 03:32pm

TO: Clara Else (CLARA ELSE)

FROM: Gregory Ingram, RAD (GREGORY INGRAM)

EXT.: 31052

SUBJECT: RE: Briscoe research project (RP0677-25); Bhargava's role

I like your draft memo. But first, I discussed the matter with Briscoe about whether the data would be available to Bharghava, as he says. I had lunch with Briscoe today. It is. But not the constructed instruments. There is little hope of progress here. Have Edith set up the memo you have drafted for signature.

CC: Edith Thomas (EDITH THOMAS)
CC: Shekhar Shah (SHEKHAR SHAH)

Record Removal Notice

The World Bank Group
Archives
& Records Management

File Title Project Management Records - Synergistic Health Effects from Water Supply and Sanitation Interventions - (RPO # 677-25) - 1v	**Barcode No.** 1572444

Document Date 02 June, 1992	**Document Type** Memorandum	

Correspondents / Participants
To: Gregory Ingram and Shekhar Shah
From: Clara Else, RAD

Subject / Title
Briscoe research project (ROP677-25); Bhargava's role

Exception(s)
Personal Information

Additional Comments

The item(s) identified above has/have been removed in accordance with The World Bank Policy on Access to Information or other disclosure policies of the World Bank Group.

Withdrawn by	**Date**
Sherrine M. Thompson	November 11, 2017

RECEIVED
MAY 26 1992

THE WORLD BANK/INTERNATIONAL FINANCE CORPORATION

OFFICE MEMORANDUM

DATE: May 5, 1992

TO: Greg Ingram, RAD

FROM: John Briscoe, INUWS

COPIED TO: Jim Vanderslice, Alok Bhargava, David Guilkey

EXTENSION: 3-5557

SUBJECT: Research project on health impacts of water and sanitation

---

When approving this proposal, you suggested that Professor Bhargava might provide some help with some of the estimation issues which arise when using probit estimators with instrumental variables. Allowance was made for this in the revised (approved) budget.

I have spent quite a bit of time trying to facilitate this. The bottom line seems to me to be the following. Using the discrete dependent variable (which is what is used in the epidemiologic literature, has been used throughout the Cebu studies, and has been widely accepted in reputable journals) the Cebu team has done as well as anyone can. Bhargava acknowledges this and that there is little that he (or anyone else) could add to this.

Bhargava is interested in extending the analysis by constructing and using a continuous dependent variable (including severity). I tried to facilitate this link. However, the Cebu team has considered this carefully and does not wish to go down this road. As you can see from the attached, he can (like any other researcher) get the raw Cebu data and work with these as he chooses.

I thus propose that Dr. Vanderslice, with the help from Professor Guilkey indicated in the original proposal, proceed with the work as proposed, drawing attention (as has been done in the other Cebu papers) to some of the unresolved and unresolveable (in the short-term) estimation problems.

Finally, I'd like to add that all involved -- Professor Bhargava, Dr. Vanderslice and others -- have been cooperative and helpful. It's just that there's not enought commong ground on which to proceed.

Attached: Relevant correspondence

THE UNIVERSITY OF NORTH CAROLINA
AT
CHAPEL HILL

Carolina Population Center
FAX: 919/966-6638
TELEX: 6501529429 6501529429MCI UW
CABLE: POPCENTER, Chapel Hill, N.C.
TELEPHONE: 919/966-2157

The University of North Carolina at Chapel Hill
CB# 8120, University Square
Chapel Hill, N.C. 27516-3997

April 23, 1992

Dr. John Briscoe
Chief, Water and Sanitation Policy Unit
The World Bank
S11 - 105
1818 H Street, NW
Washington, D.C. 20433

Dear Dr. Briscoe:

Some questions have been raised regarding model specification for our study of the synergistic health effects from multiple environment interventions. We proposed using a dichotomous dependent variable indicating whether the child experienced a bout of diarrhea in the seven days before each interview. Such dichotomous measures of diarrhea are commonly used in the literature.

We are planning to use a probit estimator to model the risks associated with unimproved water supply and sanitation. Since the choice of water source and excreta disposal method are behavioral, they must be treated as endogenous and instrumental variables will be used in the estimation. As a result, the estimated standard errors of the coefficients from the probit model will be biased, and the t-statistics will be inflated.

I have spoken with Dr. Bhargava about this problem. He suggested that a continuous measure of diarrheal severity be developed in lieu of the dichotomous variable, and expressed his strong interest in working on such a model. However, diarrheal severity has rarely been used in the literature because factors other than exposure are important determinants of the severity of the episode (i.e., the specific pathogen causing the infection, the child's nutritional status, and how the child is treated and fed during the episode).

A few Monte Carlo simulations have been conducted to assess the degree of bias in the estimated standard errors when instrumental variables are used in probit models. The bias appears to be relatively minor, however, the effect may vary from case to case. Furthermore, this limitation in the use of probit estimators has been clearly set out in several papers which have been thoroughly review and accepted by reputable journals[1,2,3].

Since Dr. Bhargava's expertise is not directly applicable to modeling limited dependent variables, it does not seem that he could contribute much to this research project. If you feel that some outside review of these methods is still advisable, Dr. Guilkey has suggested that we contact Dr. Tom Mroz from the Department of Economics at UNC, or Dr. Robin Sickles of Rice University. Both of these researchers have considerable experience estimating structural equations with limited dependent variables.

Sincerely,

James VanDerslice, Ph.D.
Research Associate
Carolina Population Center

---

[1] Cebu Study Team, Underlying and proximate determinants of child health: The Cebu longitudinal health and nutrition survey, Am. J. Epid., 133, 185-201, 1991a.

[2] Cebu Study Team, A child health production function estimated from longitudinal data, J. Dev. Econ., in press, 1992.

[3] VanDerslice, J., and J. Briscoe, All coliforms are not created equal: A comparison of the effects of water source and in-house contamination on infantile diarrheal disease, Water Resources Research, in press, 1992.

chron

The World Bank
INTERNATIONAL BANK FOR RECONSTRUCTION AND DEVELOPMENT
INTERNATIONAL DEVELOPMENT ASSOCIATION

1818 H Street, NW
Washington DC 20433
U.S.A.

Tel: (202) 473-5557
FAX: (202) 477-0164
Cable Address: INDEVAS

FAX

TO: **David Guilkey, Barry Popkin and Jim VanDerslice**

FAX #: 919-966-6638

FROM: **John Briscoe**
Chief, Water and Sanitation Policy Unit

PHONE #: (202) 473-5557

FAX #: (202) 477-0164

Number of pages (including cover sheet): 4

Date sent: April 24, 1992 Time sent: 9:01 AM

David, Barry and Jim:

Re: Bhargava and Jim's project

I've had a couple of conversations with both Jim and Bhragava over the past couple of days. I have two suggestions, and want to run them by you.

Suggestion 1: Estimation issues for the existing models

I will explain to the Research Committee, along the lines of Jim's letter (attached), that we have done as well as can be done on the econometric problems (using instruments in a probit) and that Bhargavan agrees that this is so. We therefore intend to proceed as originally proposed using the "standard Cebu models" for Jim's study and need no further expert econometric input into this.

Suggestion 2: Specifying and estimating new models

Bhargavan's principal interest is in "deepening" the dependent variable, so that we incorporate diarrhea severity. He is also interested in learning more about environmental determinants of diarrhea. What Bhargavan and I discussed was that he work with Jim on a "parallel" analysis, the principal difference being in the specification of the dependent variable (and the corresponding use of his bag of estimation techniques). To do this would require that he be given full access to the data and that he be given assistance in transferring the data to his computer system. This would require a modest addition to the grant to accomodate Bhargavan's additional work -- I believe I could get this from the Research Committee.

There are a couple of perspectives on this. On the one hand, to a disinterested observer it seems an unmitigated good thing to have another first-rate researcher working on the data. On the other hand, I understand that there are some fundamental differences (he talked to me a bit about "health production functions" and his scepticism about them, for instance) and am concerned that there may be a hidden agenda, and that this may be a way for him to get "priveleged access" to the Cebu data set. Here, again, my point of view is that we have done good work of which we are confident, and that we should welcome it when others of his quality use these data. This is, of course, not a new discussion, but one that ties back into the long discussion we had on making data available to others. (As I recall, there were some differences among us regarding what "making data available" meant, and that might have some bearing on this issue.)

Enough for now. Please let me know your thoughts on this and make specific suggestions as to how I should proceed.

Regards to all.

John

April 29, 1992

John:

An additional thought regarding Dr. Bhargava's involvement.

If Dr. Bhargava is provided with just the raw data I can't see him being able to estimate structural equations without a lot of effort over a long period of time, most probably more time and effort than is currently funded. On the other hand, John Akin does not feel that it would be appropriate to provide created variables or instruments, unless his research is a direct part of my research. Do you want me working this closely with Bhargava on these analyses? It seems to me that this would be the only way for him to be able to produce any results in the limited amount of time being funded. I may be wrong; perhaps he could develop his own variables from the raw data much more quickly than we did, but I have my doubts.

I would be happy to work closely with him if you think that it would be worthwhile. A priori I don't think that the measures of diarrheal severity proposed by Bhargava would be very sensitive to the measures of exposure to water supply and sanitation that we'll be using.

I am sorry if this has turned into an administrative hassle for you.

Thanks for the letter re consulting fees.

Jim

Reviews

ALL-IN-1 NOTE

DATE: 18-Mar-1992 06:00pm EST

TO: Gregory Ingram (GREGORY INGRAM)

FROM: John Briscoe, INUWS (JOHN BRISCOE)

EXT.: 35557

SUBJECT: **FYI**

Greg:

I thought you might be interested in this. It was just sent to Raj and so, I suppose, should be considered lightly confidential.

John

A L L - I N - 1 N O T E

DATE: 18-Mar-1992 05:56pm EST

TO: V.N. Rajagopalan (V.N. RAJAGOPALAN)

FROM: John Briscoe, INUWS (JOHN BRISCOE)

EXT.: 35557

SUBJECT: OSP Retreat -- Research

A group comprising Jock Anderson, John Briscoe, Anthony Churchill, Leighton Cumming, Cornelis de Haan, Ian Johnson, Anthony Measham, Mohan Munasinghe and Michel Pommier discussed "research", both RSB and externally-funded. The following is a summary of the group's discussion.

I: OBSERVATIONS:

The Research Portfolio

OSP's portfolio of research activities can be classified (somewhat schematically) into:
(a) Short-term activities: These include documents on "best practice" and syntheses of existing knowledge. There is a clear constituency, and strong support for such products in Operations.
(b) Long-term activities: These are the major research activities, which make major contributions to the Bank's role in the development debate, and through which the Bank builds an important part of its intellectual capital. However, such activities are often seen s staff as motivated by idiosyncratic, individual interests, not focused and esoteric.

The group strongly believed that the latter activities are central to OSP's mission, and concluded that "cultivation of stakeholders" in support of these is an important task for OSP managers.

Research Staff

It was agreed that the most productive researchers in the Bank are often young and new to the Bank. It was also agreed that contracting out of research was not usually satisfactory, because the Bank often lost the human capital built in the process.

It was further agreed that management of research is a major task for many managers and staff in OSP, and that greater emphasis should be given to developing research management skills. This would include greater attention to selecting appropriate staff, and to training and providing research support services for such staff.

The Quality of Research

It was generally agreed that the quality of research in OSP could be considerably improved. There were two perspectives on this.

First, there is a problem of consistency. On the one hand, research funded by the Research Committee is subject to an intense review process. This requires much greater attention to the design of the research, and takes considerable resources and time. However, these investments give high returns, in the assessment of the group, in terms of the quality of the research emanating from this process. On the other hand, there is externally-funded research, which can be obtained on the basis of little more than a statement of intentions. The group felt strongly that there is a need for greater quality control, or, rather, greater uniformity in quality control on externalunded researc

Second, the group noted that the quality of DEC-initiated research was often higher than OSP-initiated group believed that a significant factor in this discrepancy is the lack of a compelling and articulated "development vision" in OSP. Such a vision, the group suggested, was the starting point for the articulation of a focussed, quality research program in OSP.

The Nature of OSP Research

The familiar theme of "too much econome technical research" was addressed briefly in the group's discussion and in the subsequent plenary discussion.

The group saw the dissension around this issue primarily as a consequence of the previous issue, namely lack of a clear articulation of OSP's mission and a definition of a derived, focussed research effort. The group also discussed the principle which should guide the definition of research priorities, namely that of the Bank's comparative advantage.

On the specific issue of research on technical issues, the sense of the group was that the Bank did not generally have a comparative advantage in the development of technologies. There did, however, seem to be two important areas in which more work on technical issues is appropriate. The first is in a more effective process for reviewing technological developments and disseminating these reviews to staff. The second is in a more serious focus on the implications of technological change for economic development in general and for specific sectors. This work could range from a better understanding of how, for instance, Japan and Korea have organized to absorb technology over the s to an assessment of the effects of technological changes in telecommunications on trade and hence on the demand for different quality and forms of transportation.

II: ACTIONS FOR OSP:

Advocacy:

As discussed above, the group felt that high-quality, focussed research on priority issues was a central task for OSP. The group felt that OSP management failed to articulate and advocate a position consistent with this. There were three adverse consequences of this. First, it was the group's perception that when increasing demands were made on OSP, management did not aggressively defend the importance of, and allocations to, research. The second, related point is that the research agenda is increasingly funded by outside sources, a development which may have adverse consequences both for the subject of research and (as discussed above) for the quality of the research. Third and finally, it meant that operations staff saw research by OSP a the mission of the Bank.

Management:

The single most important recommendation from the group was that there is an urgent need for the development and transmission of a strategic vision for OSP. The group believes that once such a vision is articulated, then it will be feasible to develop a focussed research program for OSP.

At the level of the sector departments, the group distinguished between the focus of research and the quality of research. With regard to the focus, the situation generally seemed satisfactory, with considerable involvement of Operations and others in the Bank in the setting of the agenda. Here the main point would be the tightening of the focus and the setting of priorities. With regard to the methodological quality of research, the group felt that atte given to the development of more consistent and appropriate procedures for reviewing and enhancing quality of OSP research, with a particular emphasis on more and more effective peer review of externally-funded research.

CC: Mohan Munasinghe (MOHAN MUNASINGHE)
CC: Jock Anderson (JOCK ANDERSON)
CC: Anthony A. Churchill (ANTHONY A. CHURCHILL)
CC: Leighton H. Cumming (LEIGHTON H. CUMMING)
CC: Cornelis De Haan (CORNELIS DE HAAN)
CC: Ian Johnson (IAN JOHNSON)
CC: Anthony R. Measham (ANTHONY R. MEASHAM)
CC: Michel Pommier (MICHEL POMMIER)

The World Bank
INTERNATIONAL BANK FOR RECONSTRUCTION AND DEVELOPMENT
INTERNATIONAL DEVELOPMENT ASSOCIATION

1818 H Street, N.W.
Washington, D.C. 20433
U.S.A.

(202) 477-1234
Cable Address: INTBAFRAD
Cable Address: INDEVAS

March 13, 1992

Professor Alok Bhargava
Department of Economics
University of Houston
Houston
Texas 77204-5882
(Tel) 713-749-1362

Dear Professor Bhargava:

Attached please find a set of documents pertaining to a small research project which I and Dr. James Vanderslice of the University of North Carolina will be carrying out with funding from the Research Committee at the World Bank.

As you will see, in approving the request, the Research Committee counselled us to seek additional expert input regarding a couple of key estimation questions, and suggested that you would be an excellent person to assist in this. What we did was add a small amoung ($2,000 each) so that we could get input from you and from Professor David Guilkey (who, as you know, is the econometrician with principal responsibility for the Cebu work).

This letter is a formal request for you to provide such an input, for the amount specified above. We would appreciate hearing from you at your earliest convenience whether you are interested in and prepared to do this so that, in case you are not, we can obtain advice elsewhere.

Dr. Vanderslice will have principal responsibility for this work, with the direct help of Professor Guilkey on the econometric aspects. They will be in touch with you directly on this.

I hope that it will be possible for you to work with us on this.

Yours sincerely,

John Briscoe

John Briscoe
Chief
Water and Sanitation Policy Unit

cc. Jim Vanderslice, David Guilkey, Clara Else

ALL-IN-1 NOTE

DATE: 12-Mar-1992 03:01pm

TO: John Briscoe (JOHN BRISCOE)

FROM: Clara Else, RAD (CLARA ELSE)

EXT.: 31049

SUBJECT: RE: Proposal

Sorry for the delay: the memo in this regard was written Tuesday, and I am not sure why it did not get to you yet. The gist of it is that the proposal is funded, but that we would like to have an econometric guru take a crack at the very difficult technical issues with which you are wrestling here. Thus we would be willing to supplement your budget to a negotiable degeree to pay for expert advice and subsequent revision of the analytical framework. Prof. Alok Bhargava is suggested, and would be willing to consider it, or to recommend the right person if you should want to turn to him.

You need merely indicate what you want to do and tell me a (reasonable) amount you would want for this fiscal year, and the account will be opened.

CC: Edith Thomas (EDITH THOMAS)

A L L - I N - 1   N O T E

DATE: 12-Mar-1992 12:45pm

TO: Clara Else (CLARA ELSE)

FROM: Gregory Ingram, RAD (GREGORY INGRAM)

EXT.: 31052

SUBJECT: **Briscoe's response**

This strikes me as the kind of return se were looking for. His budget proposal looks a bit steep, given that it is twice the inputs that we probable had in mind, but in the scheme of things it is a small matter. Unless you have any problems with this that you want to raise, I would suggest proceeding with a decision memo.

Record Removal Notice

The World Bank Group
Archives
& Records Management

File Title Project Management Records - Synergistic Health Effects from Water Supply and Sanitation Interventions - (RPO # 677-25) - 1v	**Barcode No.** 1572444

Document Date 12 March, 1992	**Document Type** Memorandum	

Correspondents / Participants
To: Gregory Ingram
From: John Briscoe, INUWS

Subject / Title
Research Proposal

Exception(s)
Personal Information

Additional Comments

The item(s) identified above has/have been removed in accordance with The World Bank Policy on Access to Information or other disclosure policies of the World Bank Group.

Withdrawn by Sherrine M. Thompson	**Date** November 11, 2017

Archives 01 (March 2017)

ALL-IN-1 NOTE

DATE: 12-Mar-1992 11:03am EST

TO: Clara Else (CLARA ELSE)

FROM: John Briscoe, INUWS (JOHN BRISCOE)

EXT.: 35557

SUBJECT: Proposal

Clara:

I trust you got the material I sent last week. When can I expect to hear back?

Could you also do me a favour, and send me a copy of last year's Annual Report on Research (Room S 11-043)?

Many thanks

John

THE WORLD BANK/IFC/MIGA

OFFICE MEMORANDUM

DATE: March 10, 1992

TO: John Briscoe, INUWS

FROM: Gregory K. Ingram, Administrator, Research Advisory Staff

EXTENSION: 31052

SUBJECT: Research proposal, "Synergistic Health Effects from Water Supply and Sanitation Interventions"

I have reviewed your request for $20,000 to fund research on the above topic and am impressed by the importance of the issues addressed and the richness of the data set which you will use. It appears, however, that there are some technical aspects of the proposal approach that could weaken the validity and persuasiveness of your analysis. Thus I have approved RSB funding of the project contingent upon your obtaining expert advice on the methodological formulation of the study and revising the analytical framework in response to such advice.

The questions that arise pertain to (1) the validity of deriving the instrumental variables from the behavioral structural equations when you have observations from a small number of time periods, (2) the effect of including the lagged dependent variable in estimation of the probit model, and (3) the effect on the significance of the in-house contamination variable of using a single observation of only 184 households for it, relative to the time series observations of more than 600 water sources and of other variables in nearly 1650 households. Given the centrality of the water contamination issue in the study, the validity or generalizability of the very small and insignificant coefficient on in-house contamination would seem to be a critical issue that I would hope to see thoroughly addressed.

I would be willing to supplement the proposal's budget to cover some additional consultant fees incurred in obtaining assistance to address the methodological issues outlined above. Prof. Alok Bhargava of the University of Houston, for instance, has worked extensively in and developed new approaches to modelling random effects probit estimators from panel data covering short time periods. I expect that you would find his work quite germane, and his advice most helpful in dealing with the technical issues in this project.

Please call Clara Else (31049) to let us know if such a supplement to the budget would be useful and to indicate your final fiscal year allocation of the project's budget. Since FY91, RPOs have been treated in the same manner as other World Bank accounts. Funds not committed and disbursed by the fiscal year deadlines will be lost to both your project and to the Research Support Budget.

Upon receipt of a memorandum indicating your agreement to the above terms, your decision concerning consultant supplementary funding, and your fiscal year allocation of the budget, an account will be opened and funds transferred to initiate the research project.

cc: V. Rajagopalan, OSPVP; L. Pouliquen, INUDR; C. Carnemark, INUWS;
S. Shah, C. Else, V. Mataac, E. Thomas, RAD;
Research Committee Members

3/10/92

Greg, Shekhar:

Re: Briscoe - Synergistic Health Effects from Water Supply and Sanitation Interventions

The proposed study would examine the synergy between water source improvements (i.e. a policy variable) and household sanitation and breastfeeding in reducing infant and childhood diarrhea, with an emphasis on the interactions between maternal behavioral responses and the relationships estimated.

Upon initial reading there appeared to be econometric problems with the methodology set out in the attachment to the proposal, so I took it to Martin Ravallion at Ravi's urging. Ravallion confirmed my concerns (and added some of his own). Ravallion also indicated that the exposition of the paper was too muddled to discern the approach used in some cases.

OK/

Briscoe makes no claims to econometric training, the main consultant (who drafted the problematic paper with Briscoe) has only an M.S.E.E. and what they are attempting here is both complex and important. Thus I suggest that we ask them to consult with Prof. Alok Bhargava (U. Houston), who was highly recommended by Ravallion, to work over and suggest improvements in their methodology before they proceed. The costs of this consultation could be added to the budget. (Please note that we are already getting the benefit of a very large NIH grant on causes of infant and child mortality here.)

This is all I really had to say - the rest is just background to my arriving at this conclusion

The problem that I have with this approach is my inability to clearly convey the econometric weaknesses. My central concern is the validity or generalizability of the very small and insignificant coefficient on in-house contamination. In-house contamination is the only variable for which there is only one observation per household; the others were measured in a baseline interview during the mother's pregnancy, and at bimonthly interviews until the child was two years old. Thus not only do they have no time series for in-house contamination, they also are working with 184 observations on that variable while measurements across 1650 households have generated time-series data on the others. Thus it seems possible that the relative insignificance of the variable would be as much a product of the method as of the relationship.

Larger and significant coefficients are estimated on several other variables associated with household sanitation that would not be expected to be as direct a contributor to infant and child diarrhea as water contamination. This also suggests that the effects of high levels of in-house contamination are not being properly picked up here.

VanDerslice and Briscoe also use instrumental variables derived by substituting prior-term estimates for the endogenous variables "until only the underlying variables and error terms remain" (p. 18). Ravallion has confirmed that the limited number of periods in their time series (t = 1 - 6) also undermines the appropriateness of the approach. Ravallion also pointed out that the t stats are inaccurate as they do not account for the increased uncertainty introduced through the use of instrumental rather than the behavioral variables.

Ravallion indicated that including the lagged dependent variable in the equation introduces "massive bias" when there are relatively few time periods over which the regression is run.

Finally, rather than measuring diarrhea incidence as a binary variable, Ravallion has suggested that they could count numbers of incidents to get some more information from the survey data.

Clara

Bhargava, Alok, Dept. Econ, U. Houston,
Houston TX 77204-5882;
(713) 749-1362 (o); (713) 664-8115 (h)

→ lagged dependent variable (rho?)
massive bias if few

Remove i.d. lagged dependent var. effect

dichotomous variable ≠ feasible
count incidence over time

Longitudinal model of diarrheal disease in infants and children

Use dichotomous variable for diarrhea;
n (10) behavioral variables (most related to sanitation), plus 4 "environmental" variables, growth and 2 error terms.

Uses instrumental variables derived from the behavioral structural equation by substituting the structural equations for each of the endogenous variables until the 1st time period when they are f^s only of exogenous variables (i.e. don't have prior period observations); yields predicted values for the behavioral variable used as the instrumental variable for behavioral and growth variables.

Used a random-effects probit estimator and a maximum likelihood procedure to estimate the parameters

Models interaction of boiling water with other variables and relative relationship w/ diarrheal disease

Includes the lagged dependent variable in the regression

3/11/92 Alok Bhargava called: explained situation. He may be willing: asked me to send material. Told him I'd ask Briscoe first

Martin
Ravallion

36859

Call to confirm 11AM

S-12-013

```
                      OPTION 4.  DISPLAY RECORDS FOUND
**  BIBLIO DATABASE INFORMATION **

  1.  AUTHOR                      :  alok bhargava             Found: 6

-----------------------------------------------------------------------------

    Bhargava, Alok
       Estimating short and long run income elasticities of foods and nutrients
    for rural South India.  JOURNAL OF THE ROYAL STATISTICAL SOCIETY. SERIES A,
    GENERAL (U.K.) 154, pt. 1:157-74, 1991.
    INDIA - NUTRITION
    INDIA - FOOD CONSUMPTION
    INDIA - RURAL INCOME
    ELASTICITY (ECONOMICS)
    INDIA - HOUSEHOLD SURVEYS
       Location:  Joint Library.
                                                                    [171041]

       1 OF 6
Enter number of record, HELP, RETURN to continue, or EXIT:

2
    Bhargava, Alok
       Testing covariance restrictions in systems of simultaneous equations with
    vector autoregressive errors.  INTERNATIONAL ECONOMIC REVIEW (U.S.)
    30:357-72, May 1989.
    COVARIANCE
    VECTOR AUTOREGRESSION
       Location:  Joint Library.
                                                                    [136294]

       2 OF 6
Enter number of record, HELP, RETURN to continue, or EXIT:

    Bhargava, Alok
       On the theory of testing for unit roots in observed time series.  REVIEW
    OF ECONOMIC STUDIES (U.K.) 53:369-84, July 1986. The unit root null
    hypotheses are tested using U.S. data on the velocity of money and the
    Michigan PSID.
    UNIT ROOTS
    TIME SERIES ANALYSIS
    TEST STATISTICS
    UNITED STATES - MONEY. CIRCULATION
    UNITED STATES - INCOME DISTRIBUTION
       Location:  Joint Library.
                                                                    [80007]

       3 OF 6
Enter number of record, HELP, RETURN to continue, or EXIT:

    Bhargava, Alok and Sargan, J. D.
       Estimating dynamic random effects models from panel data covering short
    time periods.  ECONOMETRICA (U.S.) 51:1635-59, November 1983.
    ECONOMETRICS
       Location:  Joint Library.
                                                                    [19432]
```

```
         4 OF 6
Enter number of record, HELP, RETURN to continue, or EXIT:

    Sargan, J.D. and Bhargava, Alok
       Maximum likelihood estimation of regression models with first order
    moving average errors when the root lies on the unit circle.  ECONOMETRICA
    (U.S.) 51:799-820, May 1983.
    REGRESSION ANALYSIS
    MAXIMUM LIKELIHOOD ESTIMATION
    ECONOMETRIC MODELS
       Location:  Joint Library.
                                                                      [8301]

         5 OF 6
Enter number of record, HELP, RETURN to continue, or EXIT:

    Sargan, J.D. and Bhargava, Alok
       Testing residuals from least squares regression for being generated by
    the Gaussian random walk.  ECONOMETRICA (U.S.) 51:153-74, Jan. 1983.
    RANDOM WALKS (MATHEMATICS)
    LEAST SQUARES
       Location:  Joint Library.
                                                                      [3140]

         6 OF 6
DISPLAY OF INFORMATION IS COMPLETE
Enter number of record, HELP, RETURN to continue, or EXIT:
exit


                  SEARCH THE BIBLIO  DATABASE

     OPTIONS                                     FIELDS (6-12)
---------------------------------------------------------------------------
 1.  ERASE all searches                    6.  SEARCH WORD
 2.  COMBINE searches                      7.  ARTICLE TITLE
 3.  LIST possible search terms            8.  AUTHOR
 4.  DISPLAY records ONE at a time         9.  SUBJECT
 5.  DISPLAY records TWENTY at a time     10.  PUBLICATION DATE (Ex: 1990@)
```

Hand deliver to
Clara Else
5-3

RECEIVED
MAR 06 1992
RESEARCH ADVISORY STAFF

A L L - I N - 1   N O T E

DATE: 06-Mar-1992 02:04pm

TO: Clara Else (CLARA ELSE)

FROM: John Briscoe, INUWS (JOHN BRISCOE)

EXT.: 35557

SUBJECT: Research proposal on water and sanitation

Clara:

I am sending down responses to questions 3 through 6.

We discussed questions 1 and 2. Regarding question 3, we would like $9,000 in FY 92 and the rest in FY 93.

John

3) The 184 observations of in-house contamination are used to estimate parameters for the reduced-form model. These parameter estimates are used to generate predicted in-house contamination levels for all children at all time periods. These predicted values are used in the diarrhea model. While the small number of actual observations increases the uncertainty associated with the predicted values, the values themselves are consistent (unbiased). As a result, the parameter estimate associated with in-house contamination in the diarrhea model will also be consistent. Since this parameter estimate is essentially zero, there is no evidence that in-house water contamination is associated with diarrheal disease.

Two alternate measures of in-house contamination were also used in the diarrhea model: a binary variable indicating that the in-house water contamination FC dose was greater than 1000/day, and a variable indicating that the method of in-house water storage was inadequate (i.e., container was not covered or a scoop was used to remove water). Neither of these variables were associated with diarrhea. Thus in-house water contamination does not appear to be an important risk factor for diarrhea.

4)In order to create an instrument for an endogenous variable, all endogenous variables in its structural equation are substituted out, replaced by its exogenous and endogenous determinants. This substitution is repeated until the first time period when the endogenous variables are functions of only exogenous variables. Thus the reduced-form equation includes all lagged values of the exogenous variables. None of the endogenous variables remain. However, due to the large number of exogenous variables it is not practical to include all lagged values when estimating the instrumental models. As such only the exogenous variables from the same time period, and selected exogenous variables from the preceding time period are included in the model specification.

5) First, exposure to pathogens from the water source can not be assumed to be exogenous. While the actions of the household members may not directly affect the quality of water from any given water source, the quality of water does vary significantly between water sources. As a result, the child's exposure to pathogens from a water source will depend on which source the family collects water from. Furthermore, household water treatment (boiling) may also be used to reduce the level of contamination from either the water source or in-house water contamination.

Secondly, while it is possible that concerned parents may have improved their sanitary practices and thus reduced the level of in-house water contamination, over 35% of the households we sampled had in-house contamination levels greater than 100 FC/100ml. As a result over 30% of the study children were estimated to consume more than 1,000 FC per day due to in-house water contamination. Thus in-house water contamination posed higher levels of exposure than contaminated water sources. If in-house water contamination was an important transmission route, its association with diarrhea should have been observable.

6) The parameter estimate for water source FC dose in the main effects model is 0.068 (with a t-statistic of 2.6). However, this is the average effect for those who boil and do not boil their childrens' drinking water. The full effect of contaminated drinking water is given in the second model where water boiling has been interacted with the FC dose. In this model the parameter estimate for log FC dose is 0.17, indicating that each order of magnitude decrease in the concentration of fecal coliform would lead to a 20% decrease in the probability of diarrhea.

A L L - I N - 1   N O T E

DATE: 05-Mar-1992 02:38pm

TO: John Briscoe (JOHN BRISCOE)

FROM: Clara Else, RAD (CLARA ELSE)

EXT.: 31049

SUBJECT: <u>Research proposal on water supply and sanitation</u>

I have just a few questions that I would like to briefly discuss with you on the proposal:

1) What is the related work of Guilkey and Popkin funded by NIH, and how does it differ from this?

2) You indicate that no regional consultation took place (Part II, A) but that the results of the work have been discussed with EA1IN. If you have not done so, it would be a good idea to send a copy of the proposal to the EA1 director or at least country officer.

3) It appears that the stored water sampling was not repeated "between two and five times over the course of a year" as were the water <u>source</u> samplings (p.7), but starts with only the 233 (p.9) winnowed down to only 184 total observations (p.22). If so, are 184 observations sufficient to generate stable parameter estimates of the interaction terms given the degrees of freedom that are lost in these tests? The t stat on the in-house FC dose interaction term is 1.0: if the in-house contamination interaction is specifically what you are exploring in this study, it would be nice to feel that you are able to capture it well. Could you explain a bit more?

4) I am relatedly confused by the procedure you use to derive the instruments from the behavioral structural equations. You iteratively substitute the structural equations from prior periods for most of the RHS variables. What number of iterations would be expected to reduce the coefficients on the behavioral variables to zero? Is this possible if you have only a one-time measurement of in-house contamination, or only 2 others as suggested on page 7? Or do you omit this from the behavioral vector (and wouldn't that give you an omitted variable bias -- see my question below)?

5) The level of in-house contamination is endogenous to the household; the water source contamination is not. Thus one would expect severe or repeated bouts of diarrhea caused by in-house contamination to induce the parent to ask around about causes (if they were not previously known), and to adopt better in-house sanitation to reduce that source of contamination and diarrhea. This would suggest that there are upper bounds on the levels of in-house-induced diarrhea that would be captured in the study; would the measures capture its full potential as a cause of infant diarrhea relative to the fully exogenous source? One

might also see significant improvements in sanitary practices that induce a subsequent decrease in diarrhea, where the length of the lag from contamination to diarrhea to behavioral response to drop in diarrhea would be difficult to capture. Could this partially account for the very small point estimate shown in table 7? I do not see how the model, as described in the "Coliforms" paper, can capture a possible upper bound or the lagged response or, if there is only one-time measure of in-house contamination, of the overall endogeneity. Could you explain?

6) A quick one: you say that "water source dose is a very strong risk factor (t = 3.0)" in the first model (p. 24). Table 7 suggests that it is a very small (0.068) but significant factor (t = 2.6). Am I looking at the right thing?

7) What is the FY92/FY93 distribution of the budget that you would like?

I doubt that we will need this in writing: please call when you have a minute. We will then give you a fast response on your request.

CC: Edith Thomas (EDITH THOMAS)

ALL-IN-1 NOTE

DATE: 28-Feb-1992 05:14pm

TO: Clara Else (CLARA ELSE)

FROM: Gregory Ingram, RAD (GREGORY INGRAM)

EXT.: 31052

SUBJECT: **briscoe**

ALL-IN-1 NOTE

DATE: 28-Feb-1992 04:03pm EST

TO: Gregory Ingram (GREGORY INGRAM)

FROM: John Briscoe, INUWS (JOHN BRISCOE)

EXT.: 35557

SUBJECT: **Proposal on Water and Health in Cebu**

Greg:

I expect in others an efficiency that I am never capable of. A couple of weeks ago I sent you a proposal for funding for about $20,000 but haven't heard back yet. Can you let me know when I might hear?

Many thanks

John

Proposal

Record Removal Notice

The World Bank Group
Archives
& Records Management

File Title Project Management Records - Synergistic Health Effects from Water Supply and Sanitation Interventions - (RPO # 677-25) - 1v	**Barcode No.** 1572444

Document Date 13 February, 1992	**Document Type** Form	

Correspondents / Participants

Subject / Title
The World Bank Research Program Request for Research Support Budget (RSB) Funding
Synergistic Health Effects from Water Supply and Sanitation Interventions

Exception(s)
Personal Information

Corporate Administrative Matters

Additional Comments

The item(s) identified above has/have been removed in accordance with The World Bank Policy on Access to Information or other disclosure policies of the World Bank Group.

Withdrawn by	**Date**
Sherrine M. Thompson	November 11, 2017

Proposal for INUWS Research on:

Synergistic health effects from water supply and sanitation interventions

While there have been numerous attempts to quantify the health impacts from improving water supply or sanitation, few studies have considered factors which may alter the relationship between environmental sanitation and diarrheal disease. For example, improving water quality appears to have a much stronger effect on diarrheal disease in areas with adequate excreta disposal.[1] Similarly, improvements in sanitation may provide larger health benefits to children who are not breast-fed.[2]

The proposed research is an investigation of these two potentially "synergistic" effects. The specific questions to be addressed are:

1. Do simultaneous improvements to both water supply and sanitation result in a larger health impact than the sum of the impacts from each intervention alone? That is, do multiple environmental interventions have a "synergistic" effect in reducing diarrheal disease?

2. Does breast-feeding affect the health impact from improved water supply and sanitation? Do children who are not breast-fed realize greater health benefits?

The presence of such interaction effects would have important implications for planning water supply and sanitation programs, and for identifying high risk children. If improving both water supply and sanitation has a "synergistic" health effect, then a "package" of multiple environmental interventions should be encouraged. A strong interaction between breast-feeding and environmental sanitation may indicate that nonbreast-fed children are at a particularly high risk of diarrhea from environmental contamination, and thus should be a high priority target for water supply and sanitation interventions.

The proposed study will utilize existing data from the Cebu Longitudinal Health and Nutrition Survey (CLHNS), a prospective community-based investigation of 3327 children in Cebu, Philippines. A baseline survey and twelve bimonthly interviews were used to collect information on socioeconomic and demographic characteristics of the family, drinking water source and sanitation facilities, hygienic conditions around the house, feeding patterns and food preparation practices, and diarrheal morbidity from birth to two years of age. Data from all surveys has been cleaned and is available for this project.

Estimating interaction effects in health impact assessments is not easy, which may be one reason why they have been studied so little. In multivariate models, interaction terms typically have little power, requiring large sample sizes for stable parameter estimates. Simple stratified analyses,

---

[1] Esrey S A, Feachem R G, Hughes J M. Interventions for the control of diarrheal diseases among young children: improving water supplies and excreta disposal facilities. Bull WHO. 1985;63:757-772.

[2] Butz W P, Habicht J P, DaVanzo J. Environmental factors in the relationship between breastfeeding and infant mortality: The role of sanitation and water in Malaysia. Am J Epid. 1984;119:516-525.

on the other hand, cannot control for potential confounders. The CLHNS is one of the few longitudinal data sets with sufficiently detailed behavioral and environmental information to explore these important interactions. Within the study population there is wide variation in the socioeconomic characteristics, feeding practices, and sanitary conditions. Furthermore, previous research with this data has shown water quality, sanitation, and breast-feeding to be significantly associated with diarrheal disease.[3] Selected papers from this research are attached.

---

[3] Cebu Study Team. Underlying and proximate determinants of child health: The Cebu longitudinal health and nutrition survey. Am J Epid. 1991;122:185-201.

Selected Prior Publications from the Same Research Effort

American Journal of Epidemiology
Copyright © 1991 by The Johns Hopkins University School of Hygiene and Public Health
All rights reserved

Vol. 133, No. 2
Printed in U.S.A.

Underlying and Proximate Determinants of Child Health: The Cebu Longitudinal Health and Nutrition Study

The Cebu Study Team[1]

A proper understanding of infant health requires the integration of socioeconomic, behavioral, and biomedical models. A methodology is presented for assessing the effects of "underlying" social factors and "proximate" behavioral and biomedical factors on infant morbidity, growth, and mortality. The method is applied to data collected from over 3,000 children in Cebu, Philippines, over the first 2 years of life. Data were collected between 1983 and 1985. A central theme is that mothers recognize certain observable and nonobservable threats to the health of their infants, and that the mothers take measures to reduce the risk from such threats. It is shown that if conventional statistical techniques (which do not take such behaviors into account) are used, the estimates of the effect of the risk factors on health are incorrect. Procedures for obtaining correct estimates are described. The application of the methodology is illustrated by modeling childhood diarrhea, and by showing how maternal education induces behavioral changes, and how these changes, in turn, induce changes in the prevalence of childhood diarrhea. *Am J Epidemiol* 1991;133:185–201.

biological factors; diarrhea; epidemiologic methods; growth; health behavior; models, statistical; socioeconomic factors

The causes for high levels of childhood disease in developing countries have been the subject of numerous investigations by both social and biomedical scientists (1). The focus of the social science literature is on examining the relations between "underlying" socioeconomic variables and health outcomes, with most research focusing on mortality (e.g., references 2–6). Many of the results of these analyses are robust, with increased household income and maternal education, for example, consistently emerging as powerful determinants of health (e.g., references 7, 8). However, this literature usually gives rise to conclusions which are so sweeping (such as "where income and edu-

---

Received for publication November 16, 1988, and in final form February 9, 1990.

[1] The Cebu Study Team members and their affiliations are as follows: Linda S. Adair, John S. Akin, David K. Guilkey, and Barry M. Popkin (the University of North Carolina and the Carolina Population Center, Chapel Hill, NC); John Briscoe (the World Bank, Washington, DC); Robert E. Black (The Johns Hopkins University, Baltimore, MD); and Wilhelm Flieger (the Office of Population Studies of the University of San Carlos, Cebu, Philippines). John Briscoe had principal responsibility for organizing and writing this paper.

Reprint requests to John Briscoe, World Bank, 1818 H Street, N.W., Washington, DC 20433.

This paper is part of a collaborative research project involving the Office of Population Studies, University of San Carlos, Cebu, Philippines, directed by Dr. Wilhelm Flieger; the Nutrition Center of the Philippines, directed by Dr. Florentino S. Solon; and a group from the Carolina Population Center. Barry M. Popkin of the University of North Carolina is the project coordinator. Funding for parts of the project design, data collection, and computerization was provided by the National Institutes of Health (contract nos. R01-HD19983A, R01-HD23137, and R01-HD18880), the Nestles' Coordinating Center for Nutrition Research, Wyeth International, the Ford Foundation, the United States National Academy of Sciences, the Carolina Population Center, the United States Agency for International Development, and the World Bank. David Fugate, Deborah Barnes, Eilene Bisgrove, Frances Dancy, Patricia Dargent, Lionel Deang, Maria E. Fernandez, Gina Gantt, Amanda Lyerly, Christine Moe, James Vanderslice, and Emelita Wong provided invaluable assistance at many stages of the work.

cation are higher, health is better") that they provide little guidance to those concerned with formulating health programs. The mechanisms by which the socioeconomic determinants affect health remain largely an unexplored and unexplained "black box" (1).

The biomedical literature, on the other hand, focuses on the biologic precursors (such as infection and malnutrition) of morbidity and mortality. The virtue of the biomedical literature (namely its reliance on a biologic model and the focused nature of the questions answered) is its curse, too. This is so because: these narrowly focused studies have often ignored the effects of important confounding variables (with the biomedical literature on breast feeding being a good example); the ultimate consequences for mortality in populations at large tend to be neglected; and the fact that people perceive threats to their health and react to these by changing their behavior is often either not recognized or ignored because it is considered analytically intractable. The result is a literature which inevitably leads to policy conclusions favoring strictly medical interventions (1).

Drawing heavily on an analogous situation in the field of fertility research (9), Mosley and Chen (1) and Mosley (10, 11) have argued for the development of a new approach to child health research which incorporates both the social and biomedical approaches into a coherent analytic framework in which the relations between "underlying," "intermediate," and "outcome" variables are investigated. Important steps have been taken in recent years in studies in Malaysia (12–14) and Jordan (15) to conduct empirical research on mortality using this framework. The Cebu study was designed to build on these landmark studies.

This paper describes the methodology used in modeling child health in the Cebu study. Empirical results are given to illustrate the usefulness of the approach. Detailed discussions of the results and their implications have been presented in other papers (16–18).

THE DESIGN OF THE CEBU STUDY

The principal objective of the Cebu study was to correctly estimate the effects of underlying and proximate determinants of child health. Data were collected in the metropolitan area around the city of Cebu in the central Philippines. After a pilot study, a stratified, single-stage sampling procedure was used to select 17 of 158 urban and 16 of 85 rural neighborhoods in the metropolitan Cebu area. Households were surveyed to collect data on all births between May 1, 1983 and April 30, 1984. The sample consisted of 3,080 women (77 percent of whom were urban) having single live births, for whom both baseline pregnancy surveys and birth information are available. Participation rates were high. Over the course of the 2-year period, 311 of the 3,080 women (264 of 2,355 in urban areas) were lost as a result of migration, and 49 of the mothers (39 in urban areas) decided to withdraw from the study.

For each study child, questionnaires were administered in the third trimester of pregnancy, at birth, and at 2-month intervals through the first 2 years of life. Where necessary, the questionnaires were supplemented by observations (e.g., of sanitary conditions) and measurements (e.g., of weight and water quality). Information was collected on "underlying variables" (including family income and assets, education of family members and other socioeconomic variables, prices of foods and other goods and services in the community, and accessibility to health facilities), "intermediate variables" (describing households' consumption choices for health-related goods and services, such as prenatal care and infant feeding patterns, water-use practices, personal hygiene practices, use of preventive health services, maternal smoking and drinking) and "outcome variables" (including gestational age and birth weight, and growth, morbidity, and mortality at each subsequent 2-month interval). Additional details on the survey design and data are available (16–18).

MODELS FOR ASSESSING THE EFFECT OF UNDERLYING AND INTERMEDIATE VARIABLES ON CHILD HEALTH

The mechanisms whereby socioeconomic, behavioral, and biomedical factors affect health can be described in terms of two sets of equations. The first equation describes how the underlying individual, family and community variables determine health-related behaviors; the second equation describes how underlying and intermediate behavioral and biomedical variables affect health outcomes. Following standard economics terminology, these are referred to as "structural equations." For reasons which will become apparent later it is necessary to distinguish between "endogenous variables," whose values are determined by forces operating within the model, and "exogenous variables," whose values, while important to the model, are determined by forces outside the model and are not explained by the model. In the present context (as shown in figure 1 and table 1), variables such as infant feeding patterns, use of medical facilities, type of water supply and sanitation, maternal work status, and health status of the child are treated as endogenous, while variables that are not the result of health-related household decisions (such as maternal education and food prices) are considered exogenous.

The variables entering the models (where the subscript "i" refers to the particular child and the subscript "t" to the time period) are: H_{ti}, the health of the infant; Y_{ti}, endogenous variables measuring the consumption of health-related goods; Z_{ti}, exogenous community and household characteristics; μ_i, an individual-specific disturbance term that does not change through time; and ϵ_{ti}, purely random errors that vary across individuals and through time.

Structural equation 1: Determinants of behavior

The underlying family and community variables (the Z_ts) are hypothesized to determine the health-related behaviors (Y_ts) as follows:

$$Y_{ti} = \alpha_1\, H_{Gt\text{-}1,i} + \alpha_2\, H_{St\text{-}1,i} + \alpha_3\, Y_{t\text{-}1,i} + \alpha_4\, Z_{ti} + \mu_{Yi} + \epsilon_{Yti} \qquad \text{(expression 1)}$$

$$\text{for } i = 1, 2, \cdots, N;$$

$$t = 2, 4, \cdots\ 24 \text{ months}$$

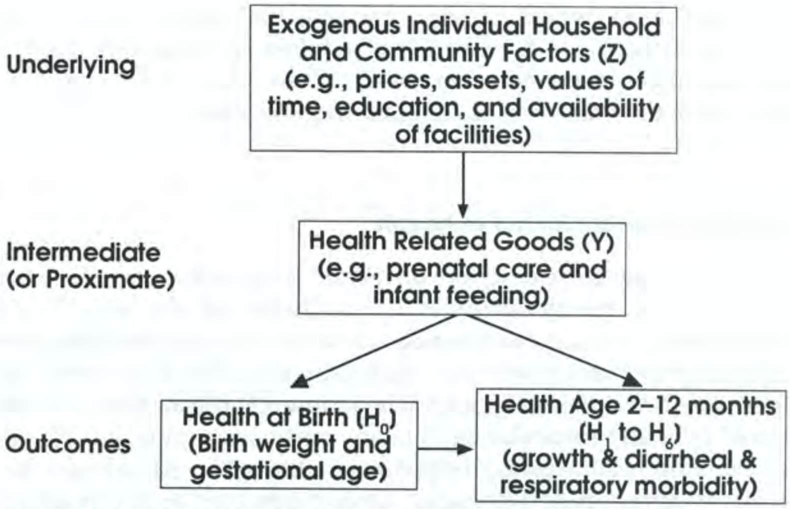


Figure 1. Conceptual framework relating underlying, intermediate and outcome variables.

TABLE 1. Endogenous variables in the diarrhea model for Cebu, Philippines, 1983–1985

I. Health of the infant
  A. At birth (H_0)
    1. Birth weight [continuous, OLS]*
    2. Gestational age [continuous, OLS]
  B. Every 2 months (H_t, $t = 2, 4, \ldots, 12$)
    1. Diarrhea [dichotomous, probit]
    2. Weight [continuous, OLS]
    3. Severe Respiratory Infection [dichotomous, probit]
II. Health-related factors affecting H_t ($t = 2, 4, \ldots, 12$) (Y_t)
  A. Exclusive breast-feeding pattern [dichotomous, probit]
  B. Any breast-feeding pattern [dichotomous, probit]
  C. Infant supplemental food nutrient intake [continuous, OLS]†
  D. Preventive medical care for infant [dichotomous, probit]
  E. Personal hygiene (soap use) [continuous, OLS]
  F. Food processing [dichotomous, probit]
  G. Immunizations (2 DPT‡; 3 DPT; measles) [dichotomous, probit]
  H. Water quality [dichotomous, probit]
  I. Excreta disposal [ordered discrete, ordered probit]

[ ] Indicates type of variable and estimation technique.
* OLS, ordinary least squares regression.
† During the period of exclusive breast feeding, a tobit estimation procedure was used.
‡ DPT, diphtheria-pertussis-tetanus vaccination.

and where the subscripts "G" and "S" refer to growth and sickness (or morbidity), respectively.

Note that the health variables measured before the current time period (the "lagged" health variables) have an effect on current health and non-health-related behaviors. That is, the model takes account of the fact that a mother might, for instance, alter her infant-feeding practices if her child failed to grow in an earlier time period. Note, too, that the model allows for the possibility that children and families have peculiarities which have important effects on how a child is treated, but which cannot be observed. For example, some mothers may know from prior pregnancies that their children are likely to be frail, and may take account of this when deciding whether to seek prenatal care.

The formal procedure for taking this unobserved heterogeneity into account is to have two (rather than the usual one) disturbance terms in the behavioral equations. The first disturbance term, μ, is specific to the individual and does not change through time. In these models, μ represents the initial endowments of the infant which cannot usually be observed by the researcher (such as "frailty"). The second error term, the ϵ, is the conventional random error term.

Structural equation 2: Determinants of health

The second structural equation describes how underlying and intermediate behavioral and biomedical variables affect health outcomes. It is assumed that the probability of being sick (e.g., of having diarrhea) in a particular period is determined in part by the nutritional status in the prior period, and that growth in a particular period is determined in part by the morbidity experience in the preceding period. In addition, it is assumed that both nutritional status and morbidity in any particular child in any particular period are also affected by the health-related and non-health-related behaviors in the prior period and by the family characteristics in the current period (the Z_ts). Accordingly, the "health structural equations" are:

$$
\begin{aligned}
H_{Gti} &= \theta_1 H_{Gt\text{-}1,i} + \theta_2 H_{St\text{-}1,i} + \theta_3 Y_{t\text{-}1,i} + \theta_4 Z_{ti} + \mu_{Gi} + \epsilon_{Gti} \\
H_{Sti} &= \gamma_1 H_{Gt\text{-}1,i} + \gamma_2 Y_{t\text{-}1,i} + \gamma_3 Z_{ti} + \mu_{Si} + \epsilon_{Sti} \qquad \text{(expression 2)} \\
H_{Mti} &= \delta_1 H_{Gt\text{-}1,i} + \delta_2 H_{St\text{-}1,i} + \delta_3 Y_{t\text{-}1,i} + \delta_4 Z_{ti} + \mu_{Mi} + \epsilon_{Mti} \\
&\text{for } i = 1, 2, \cdots, N; \\
&t = 2, 4, \cdots \text{ 24 months}
\end{aligned}
$$

and where the subscripts "*G*," "*S*," and "*M*" refer to growth, sickness, and mortality, respectively.

As a specific example, consider the structural equation for diarrhea. One measure of diarrhea from the available data is the dichotomous variable indicating whether the child had diarrhea in the 7-day period preceding the interview. The endogenous variables include exposure variables (water quality, personal hygiene practices, excreta disposal practices, food hygiene practices, whether exclusively breast-fed, mother's concern with preventive measures as measured by use of a well-baby clinic) and susceptibility variables (nutritional status as measured by weight for age, whether breast-fed at all, and whether vaccinated against measles), while the exogenous variables include exposure to animals in the home, season, sex, household size, community density and child's mobility.

The "reduced form" equations

The two sets of structural equations represent a complete description of how behavior and health are determined. By starting at birth and continuously substituting out for all endogenous right-hand side variables, the so-called "reduced form" equations are obtained.

$$
\begin{aligned}
Y_{ti} &= \sum_{r=o}^{t} Z_{t\text{-}r,i}\, \beta_r^Y + \overset{*}{\mu}_{Yi} + \overset{*}{\epsilon}_{Yti} \\
H_{Gti} &= \sum_{r=o}^{t} Z_{t\text{-}r,i}\, \beta_r^G + \overset{*}{\mu}_{Gi} + \overset{*}{\epsilon}_{Gti} \qquad \text{(expression 3)} \\
H_{Sti} &= \sum_{r=o}^{t} Z_{t\text{-}r,i}\, \beta_r^S + \overset{*}{\mu}_{Si} + \overset{*}{\epsilon}_{Sti} \\
H_{Mti} &= \sum_{r=o}^{t} Z_{t\text{-}r,i}\, \beta_r^M + \overset{*}{\mu}_{Mi} + \overset{*}{\epsilon}_{Mti}
\end{aligned}
$$

In the reduced form equations, the endogenous variables (the *Y*s) and the outcome variables (the *H*s) can be expressed just in terms of the exogenous variables (the *Z*s). These reduced form equations may be used to examine the full effect of exogenous variables (such as maternal education) on child health. They are also used in estimating the parameters of the structural equations.

PROBLEM 1: ESTIMATING PARAMETERS WHEN SOME VARIABLES ARE ENDOGENOUS

The standard approach would be to estimate the parameters of such models using standard statistical procedures (such as ordinary least squares or logistic regression). This section shows that because people recognize (some) health threats and take measures to reduce their risk from these threats, these standard procedures give the wrong answers. The correct statistical procedures are described.

Statistical procedures

For the reduced form equations (expression 3), it can reasonably be assumed that the "heterogeneity disturbance terms" (the μs) are distributed independently of the exogenous variables (the Zs) and that the error terms (the μs and the ϵs) are normally distributed. Accordingly, standard statistical techniques (such as ordinary least squares or logistic regression) can be used to obtain unbiased estimates of the parameters of the reduced form equations.

It is, however, the structural equations (expressions 1 and 2) that are of most interest to policymakers, because they permit assessment of the effects of different social, economic, and biomedical interventions on behavior and health. These equations present more complex estimation problems. For example, in expression 3 it can be seen that the value of, say, Y_{ti} depends on, and is therefore correlated with, the "heterogeneity disturbance term" (μ_{Yi}). Since the value of μ_{Yi} does not change over time, μ_{Yi} is similarly correlated with $Y_{t\text{-}1}$. In expression 1, therefore, a regressor ($Y_{t\text{-}1,i}$) is correlated with a disturbance term (μ_{Yi}). Similarly in expression 2 we also have a regressor ($Y_{t\text{-}1,i}$) correlated with a disturbance term (μ_{Gi}).

If the parameters of expressions 1 and 2 are estimated using ordinary least squares, inconsistent estimates will result. This is because, in explaining the dependent variable, as much credit as possible is given to the regressor and as little as possible to the error. When the regressor and error are correlated, some of the effect of the error is wrongly attributed to the regressor (19).

A standard procedure for dealing with this problem is that of "instrumental variables" (20). In estimating the parameters of the structural equation, the value of the problematic regressor (such as Y_{ti}) is not used, but replaced by an instrumental variable. The instrumental variable for Y_{ti} is chosen so that it is correlated with the regressor (Y_{ti}) and is uncorrelated with the disturbance term (μ_{Yi}). In this particular case, the "predicted values" of the Ys obtained from versions of the reduced form equations (expression 3) can be used as instrumental variables for the Ys in the structural equations (expressions 1 and 2). Using these predicted values rather than the actual values of the endogenous variables, standard estimation procedures are used to estimate the parameters of the structural equations. Although not widely known in the epidemiologic literature, such techniques are used routinely by economists (21).

As discussed in more detail elsewhere (16–18), in the Cebu study the models were specified to include all endogenous and exogenous variables, and to allow each of these to vary over time (by including a time interaction term). Only those interactions which were statistically significant were retained. The "random effects" estimation procedure (21) was used to estimate the parameters of the models. It was assumed that the μs and ϵs are normally distributed random variables and that corresponding to the observed dependent variables is a continuous latent variable (the severity of the child's diarrhea in this case). The mother reports that the child has diarrhea if the latent variable is sufficiently large (see reference 22 for a different child health model that also uses latent dependent variables). The actual calculations were done with the HOTZTRAN algorithm (23), using maximum likelihood methods in which the distribution of the disturbance term is taken into account in calculating the standard errors of the coefficients.

Does endogeneity make a difference in practice?

Two examples from the literature. The US Environmental Protection Agency has conducted a large-scale study on the effects of medical care and air pollution on mortality from respiratory disease (24). Whereas prior epidemiologic studies had implicitly assumed that people accept air pollution passively, the study recognizes that "people have an incentive to adapt to environmental conditions (by incurring the expense of seeing a doctor or moving away from a polluted city)" (25, p. 42). In analytic terms, this means that "protective factors" (such as use of medical care) are not exogenous (i.e., determined by forces outside of the model) but endogenous (i.e., determined by the level of air pollution and other factors incorporated into the model). The study (24, 25) shows, first, that the conclusion drawn from a conventional analysis was that the level of medical care had no effect on mortality from respiratory illness, but, second, that when statistical procedures taking account of endogeneity were used, medical care was shown to have a significant protective effect.

The second example deals with the effect of prenatal care on child health. A conventional analysis would treat the quantity of prenatal care as an exogenous variable and examine the relation between the level of this variable and infant health. In fact, however, many mothers seek prenatal care in part because they perceive (for reasons that are valid but which investigators cannot observe) their fetus to be particularly vulnerable. A detailed assessment of the relation between prenatal care and infant mortality in the United States (3, 5) has shown that a conventional analysis (which ignores this behavioral aspect) would conclude that mothers place their children at risk by obtaining prenatal care, but that when statistical procedures take account of this behavioral relation, use of prenatal care is shown to have a strong protective effect.

Later in this paper, the practical consequences of ignoring endogeneity are examined for the Cebu study.

PROBLEM 2: SAMPLE SELECTIVITY IN LONGITUDINAL STUDIES

As children are followed over time, there are inevitably losses to the study from refusal, out-migration, and death. Since children with certain characteristics are more likely to be lost to the sample than other children, the sample has been reduced in a way which is certainly not random. An exogenous factor which affects migration (such as father's occupation) but does not affect child health would tend to emerge from the analysis as a determinant of child health.

The procedure for correcting for this possibility consists of introducing a "correction factor" (technically known as the hazard rate or the inverse of the Mills' ratio (20, 26)) which is equal to zero for those individuals who would, without any doubt, remain in the sample throughout the period, and is relatively large for those who are likely to have been lost from the sample during the period. The statistical procedure involves, first, determining whether there is significant self-selection (if the "correction factor" is significant) and, if so, applying the necessary correction.

AN ILLUSTRATIVE RESULT: THE PATHWAYS THROUGH WHICH MATERNAL EDUCATION AFFECTS BEHAVIOR AND CHILD HEALTH

The Cebu Study Group has already published detailed results from some early analyses (16–18). For the present purposes, some empirical results illustrate how the model may be used to assess the biomedical and socioeconomic determinants of child health. The example chosen is one of major policy interest because of the consistent and strong relation (7, 8) between maternal education and child health and because of the paucity of data delineating the mechanisms by which this effect operates. The example is developed only for the urban sample, only for diarrhea, and only for the first year of life.

Underlying-proximate relations

The parameters of the behavioral structural equations (expression 1) are estimated for each of the health-related behaviors at each particular stage of the child's life. Table 2 shows the effects of maternal education on health-related behaviors during each 2-month period. Table 3 shows the simulated effects on the mean values of the health-related behaviors of increasing the education of each woman in the sample by one year.

Question 1: Is the direction of the effect sensible? Table 2 shows that as maternal education increases, there are increases in food intake, preventive health care, measles

TABLE 2. Effects of maternal education on health-related behavior during the first year of life, urban Cebu, Philippines, 1983–1985

Health-related behavior	Characteristics of dependent variable	Age of infants (months)					
		2	4	6	8	10	12
Feeding practices							
Breast feeding							
Exclusive breast feeding, no exposure to pathogens, 7 days before survey	Binary†	−0.07***	−0.07***	−0.12***	−0.02	N/A	N/A
Any breast feeding, 7 days before survey	Binary†	−0.10***	−0.11***	−0.11***	−0.10***	−0.10***	−0.10***
Food intake							
Total calories	Calories	23.02***	27.93***	23.97***	23.18***	22.96***	24.81***
Health service use							
Preventive health care	Binary‡	0.05***	0.03***	0.02*	0.04***	0.04***	0.04***
Measles immunization	Binary‡	N/A	N/A	N/A	N/A	0.07**	0.06***
Health practices, personal and environmental							
Pathogenicity of food processing	Binary§	0.0004	−0.002	−0.0001	−0.0007	−0.0008	−0.002
Poor type of excreta disposal	Binary‖	−0.09***	−0.09***	−0.09***	−0.09***	−0.09***	−0.09***
Quantity of soap per capita	Grams	1.76***	1.76***	1.76***	1.76***	1.76***	1.76***
Good quality of drinking water source	Binary¶	0.002	0.002	0.002	0.002	0.002	0.002

Entries are the coefficients of mother's formal education in years in creating instrumental variables for the dependent health-related variables. The asterisks indicate the significance level for testing whether the coefficient is zero: * $\alpha = 0.10$, ** $\alpha = 0.05$, *** $\alpha = 0.01$. N/A, not applicable.

† If the infant is breast-fed, the variable is set to 1 and 0 otherwise.

‡ If the infant had a preventive health care visit or had the specified type and dosage of immunization, the variable takes on the value of 1 and 0 otherwise.

§ If food processing is severely pathogenic, the variable is set to 1 and 0 otherwise.

‖ If the household's exreta disposal is poor, the variable is set to 1 and 0 otherwise.

¶ If the source of drinking water is good, the dependent variable takes on the value of 1 and 0 otherwise.

immunization, adequacy of excreta disposal practices, quantity of soap used per capita, and quality of drinking water; and decreases in breast feeding (both exclusive and any) and food contamination risk.

Question 2: Is the effect statistically significant? From table 2 it can also be seen that, for most health-related practices, the effects of maternal education are highly statistically significant. The two exceptions are unhygienic food preparation practices and quality of drinking water. In both cases, the lack of significance is almost certainly because, with the measures employed in these early analyses, there is little variation in these variables in the urban sample.

Question 3: Are the findings of practical significance? From table 3, it can be seen that a one-year increase in the education of each mother would have substantial effects on most of the health-related behaviors. For example, for a 6-month-old child, a one-year increase in maternal education implies a 36 percent reduction in the probability of exclusive breast feeding, a 5 percent reduction in the probability of any breast feeding, a 7 percent increase in caloric intake, a 4 percent increase in the use of preventive health care, a 9 percent reduction in the probability of inadequate excreta disposal practices, and a 2 percent increase in per capita soap use.

Proximate-outcome relations

The second set of structural equations (expression 2) describes the relations between the proximate (behavioral and biomedical) variables and health outcomes. Table 4 presents the proximate-outcome structural equation for diarrhea for the longitudinal model estimated for the full first year of life for the urban population. Table 5 shows the responsiveness of diarrhea to changes in the proximate variables (as measured by the "elasticity," that is, the percent change in diarrhea resulting from a percent change in the proximate variable).

Substantive Question 1: Are the estimates sensible and statistically significant? From table 4, it can be seen that diarrhea is statistically significantly lower for: faster growing infants (with the effect greatest in small infants); infants who are breast-fed (with the protective effect greatest at young ages); infants who are exclusively breast-fed; infants who consume more food; infants whose families use better quality water; infants whose families follow hygienic food preparation practices; and infants whose families have better excreta disposal practices (with the effect greater in the early months of this first year of life). Diarrhea is statistically significantly higher for: male infants in the latter months; crawling infants when there are animals in the house; and children in more densely settled communities.

Substantive Question 2: Are the findings of practical significance? From table 5, it is evident that the level of diarrhea is highly responsive to breast-feeding practices (especially in the early months of life) and to excreta disposal and water supply practices (throughout the first year of life), moderately responsive to caloric intake, especially later in the first year of life, and largely unaffected by preventive health care.

Methodological Question 1: What are the consequences of sample selectivity? Statistical analysis showed that the hazard rate (or inverse of the Mills' ratio) was small, and not significantly different from zero. It was therefore concluded that sample selectivity was not significant (that is, that nonresponse could be viewed as a random event in the sample) and that the Mills' ratios could be excluded in the final specifications of the instrumental variables.

Methodological Question 2: What are the consequences of ignoring endogeneity? The importance of taking account of endogeneity when modeling child health was tested in two ways using the Cebu data set. The first test is a formal statistical test—a chi-square version of the Hausman test (21, 27)—which indicates whether endogeneity was actually present. The critical value for a 1 percent test is 29, while the test statistic was 98: the null hypothesis of no endogeneity is strongly rejected. The results of a second, more intuitive, test (comparing results from two estimation procedures, one taking account

TABLE 3. Simulated effects of one-year increase in maternal education on health-related behavior and diarrhea incidence, at every 2 months during the first year of life, urban Cebu, Philippines, 1983–1985*

	Age of infants (months)											
	2		4		6		8		10		12	
	Mean	% Change	Mean	% Change	Mean	% Change	Mean	% Change	Mean	% Change	Mean	% Change
Feeding practices												
Breast feeding												
Exclusive breast feeding, no exposure to pathogens, 7 days before survey	0.16	−10	0.06	−12	0.0002	−36	N/A		N/A		N/A	
Any breast feeding, 7 days before survey	0.85	−3	0.79	−4	0.74	−5	0.70	−5	0.64	−6	0.57	−7
Food intake												
Total calories	121.70	10	192.60	9	307.70	7	372.00	6	439.40	5	507.50	5
Health service use												
Preventive health care	0.12	9	0.18	5	0.18	4	0.13	7	0.11	8	0.08	7
Measles immunization	N/A		N/A		N/A		N/A		0.01	19	0.01	16
Health practices, personal and environmental												
Pathogenicity of food processing†	0.04	1	0.03	−8	0.03	0	0.03	−2	0.03	−3	0.04	−5
Poor type of excreta disposal†	0.52	−9	0.52	−9	0.52	−9	0.52	−9	0.52	−9	0.52	−9
Quantity of soap per capita†	81.12	2	81.12	2	81.12	2	81.12	2	81.12	2	81.12	2
Good quality of drinking source†	0.99	0	0.99	0	0.99	0	0.99	0	0.99	0	0.99	0

* Simulation is done using the instrumental variable equation. The % change is obtained with this formula: % change = $(M_2 - M_1)/M_1 \times 100$, where M_1 is the simulated mean of the health-related behavior of interest obtained by multiplying the coefficients of the exogenous variables by their corresponding sample means. M_1 is entered as mean in the table. M_2 is the simulated mean when all variables except woman's education are at their sample means and mean maternal education is increased by one year. N/A, not applicable because no infants had this feeding or immunization during this period.

† Current version of these variables are not time-varying yet.

TABLE 4. Longitudinal analysis: Structural equation for diarrhea incidence in week preceding survey, urban Cebu, Philippines, 1983–1985

Explanatory variables	Coefficient	(*t* statistic)
A. ENDOGENOUS		
Susceptibility		
Lagged weight velocity (g/day)	−0.01	(−2.02**)
Lagged weight velocity interacted with weight (g × g/day)	3.30×10^{-6}	(2.90***)
Gestational age (weeks)	0.01	(3.12***)
Gestational age interacted with age (weeks × days)	9.60×10^{-6}	(1.68*)
Susceptibility/exposure		
Feeding practices		
Any breast feeding 7 days before survey (prob)†	−0.68	(−2.51**)
Any breast feeding interacted with age (prob × days)	1.60×10^{-3}	(1.76*)
Exclusive breast feeding with no exposure to pathogens, 7 days before survey (prob)	−1.53	(−5.91***)
Total calories (cal)	-4.40×10^{-4}	(−1.73*)
Exposure		
Health service use		
Preventive health care (prob)	−0.24	(−1.27)
Health practices, personal and environmental		
Good quality water source (prob)	−0.32	(−3.35***)
Soap purchased/capita/week (g)	-3.70×10^{-5}	(−0.06)
Pathogenic food processing (prob)	0.91	(1.85*)
Poor excreta disposal (prob)	0.92	(4.97***)
Poor excreta disposal interacted with age (prob × days)	-1.80×10^{-3}	(−2.45**)
B. EXOGENOUS		
Susceptibility		
Child's age (days)	7.10×10^{-4}	(1.00)
Child's sex (0–1)	−0.02	(−0.24)
Child's sex interacted with age (0–1 × days)	5.60×10^{-4}	(2.09**)
Exposure		
Animals in the house (0–1)	-6.90×10^{-3}	(−0.22)
Animals under the house (0–1)	−0.02	(−0.53)
Baby crawling interacted with animals in the house (0–1)	0.08	(1.93*)
Crowding		
No. of preschoolers (0–6)	−0.03	(−2.20**)
No. of persons/room (0–9.5)	9.90×10^{-3}	(1.05)
Community density (persons/km^2)	6.50×10^{-6}	(7.09***)
Cumulative rainfall in last 2 weeks before survey (mm)	2.05×10^{-4}	(0.45)
Cumulative rainfall interacted with age (mm × days)	1.20×10^{-6}	(0.69)
C. OTHERS		
Constant	−6.05	(−3.63***)
Rho	0.12	(7.22***)

Note: Sample size for this analysis is 11,807. The significance levels for testing whether the coefficient is zero are indicated by: * $\alpha = 0.10$, ** $\alpha = 0.05$, *** $\alpha = 0.01$.

† Prob, the predicted probability of the explanatory variable.

of endogeneity and one ignoring it) are presented in table 6 and summarized in figure 2.

"Column 1" of table 6 presents the results of the analysis which ignores endogeneity. For this analysis, observations are needed on all of the variables (both those which are considered exogenous and endogenous in the analysis using instrumental variables). A total of 6,674 observations are available. "Column 2" of table 6 presents the results of the instrumental variable analysis, for this same sample. A comparison of columns 1 and 2 shows that the analysis which does not account for endogeneity is reasonably specific. In only one case—total calories—does

TABLE 5. Percent change in diarrhea for a 1% increase in explanatory variables during the first year of life, urban Cebu, Philippines, 1983–1985*

	Age of infants (months)					
	2	4	6	8	10	12
Feeding practices						
Exclusive breast feeding, no exposure to pathogens	−0.59	−0.33	−0.05	N/A	N/A	N/A
Any breast feeding	−1.08	−0.91	−0.76	−0.70	−0.66	−0.60
Total calories	−0.12	−0.15	−0.22	−0.24	−0.29	−0.34
Health service use						
Preventive health care	−0.07	−0.09	−0.08	−0.06	−0.05	−0.04
Measles immunization	N/A	N/A	N/A	N/A	−0.02	−0.03
Health practices, personal and environmental						
Pathogenicity of food processing	0.08	0.05	0.04	0.05	0.04	0.06
Poor type of excreta disposal	0.90	0.71	0.54	0.43	0.34	0.24
Quantity of soap per capita	−0.01	−0.01	0	0	0	0
Good quality of drinking water source	−0.55	−0.49	−0.44	−0.43	−0.44	−0.43

* Entries are computed by the following formula: % change = $(D_2 - D_1)/D_1 \times 100$, where D_1 is the simulated mean diarrhea when the coefficients of the explanatory variables in the diarrhea structural equation are multiplied by their corresponding sample means. D_2 is the simulated mean when the value of the variable of interest is increased by 1.01 of its sample mean and the rest of the explanatory variables are at their means. N/A, not available because no infant had this type of feeding or immunization during this period.

TABLE 6. The effect of ignoring endogeneity: Structural equation for diarrhea incidence in week preceding survey, urban Cebu, Philippines, 1983–1985

Explanatory variables	Column 1		Column 2	
	Estimates when endogeneity is ignored	(t-statistic)	Instrumental variable estimates	(t-statistic)
A. ENDOGENOUS				
Susceptibility				
Lagged weight velocity (g/day)	0.08	(1.39)	−0.01	(−1.20)
Lagged weight velocity interacted with weight (g × g/day)	-2.40×10^{-6}	(−0.24)	4.10×10^{-6}	(2.13**)
Gestational age (weeks)	0.18	(0.74)	0.09	(1.48)
Gestational age interacted with age (weeks × days)	-3.20×10^{-6}	(−0.29)	-1.50×10^{-6}	(−1.41)
Susceptibility/exposure				
Feeding practices				
Any breast feeding 7 days before survey (prob)†	−0.68	(−5.58***)	−1.06	(−2.29**)
Any breast feeding interacted with age (prob × days)	2.00×10^{-3}	(3.93***)	4.30×10^{-3}	(2.20**)
Exclusive breast feeding with no exposure to pathogens, 7 days before survey (prob)	−0.23	(−1.46)	−1.35	(−3.64***)
Total calories (cal)	-2.30×10^{-4}	(−2.94***)	-3.10×10^{-4}	(−0.84)

Exposure				
Health service use				
Preventive health care (prob)	0.02×10^{-3}	(−0.31)	−0.06	(−2.30**)
Health practices, personal & environmental				
Good quality water source (prob)	−0.17	(−1.87*)	−0.22	(−1.70*)
Soap purchased/capita/week (g)	-1.20×10^{-5}	(−0.37)	3.00×10^{-5}	(0.35)
Pathogenic food processing (prob)	0.15	(1.30)	0.87	(1.27)
Poor excreta disposal (prob)	0.09	(1.04)	1.11	(3.68***)
Poor excreta disposal interacted with age (prob × days)	2.70×10^{-3}	(0.07)	-3.60×10^{-3}	(−2.71***)
B. EXOGENOUS				
Susceptibility				
Child's age (days)	2.80×10^{-4}	(0.65)	1.30×10^{-4}	(0.98)
Child's sex (0–1)	−0.05	(−0.47)	−0.07	(−0.70)
Child's sex interacted with age (0–1 × days)	6.50×10^{-4}	(1.40)	7.10×10^{-4}	(1.55)
Exposure				
Animals in the house (0–1)	−0.02	(−0.39)	-6.60×10^{-3}	(−0.12)
Animals under the house (0–1)	−0.01	(0.05)	−0.03	(−0.68)
Baby crawling interacted with animals in the house (0–1)	0.02	(1.04)	−0.01	(0.13)
Crowding				
No. of preschoolers (0–6)	0.03	(0.14)	−0.03	(−1.49)
No. of persons/room (0–9.5)	0.02	(1.15)	0.01	(0.73)
Community density (persons/km^2)	5.80×10^{-6}	(5.13***)	7.60×10^{-6}	(6.26***)
Cumulative rainfall in last two weeks before survey (mm)	−0.07	(−1.20)	-4.47×10^{-4}	(−0.68)
Cumulative rainfall interacted with age (mm × days)	-4.50×10^{-6}	(−1.59)	3.30×10^{-6}	(1.17)
C. OTHERS				
Constant	−1.87	(−1.85**)	−4.51	(−1.88**)
Rho	0.16	(6.27***)	0.15	(5.86***)

Note: Sample size for this analysis is 6,674. The significance levels for testing whether the coefficient is zero are indicated by: * $\alpha = 0.10$, ** $\alpha = 0.05$, *** $\alpha = 0.01$.
† Prob is the predicted probability of the explanatory variable.

		"Standard" Analysis (Ignoring Endogeneity)		
		Sign positive and significant	*Not significant*	*Sign negative and significant*
"Correct" Analysis (Accounting for Endogeneity)	*Sign positive and significant*	breastfeeding x age community density	excreta disposal weight velocity x weight	▓
	Not significant	▒	weight velocity food processing soap animals in house animals under house gestational age child's age persons / room # preschoolers child's sex rainfall baby crawling x animals gestational age x age rainfall x age	▒ total calories
	Sign negative and significant	▓	▒ exclusive breastfeeding preventive services excreta disposal x age	any breastfeeding water quality

Legend:

Inference from "standard" analysis would be:

correct	☐
moderately midleading	▒
seriously misleading	▓

Figure 2. Inferences from the "correct" and "standard" analyses, Cebu, Philippines, 1983–1985.

the simpler model (column 1) suggest a statistically significant relation which is not significant in the "correct" model (column 2). More serious, however, are those behaviors—improved excreta disposal, preventive health care and exclusive breast feeding—which would appear to have no effect if endogeneity is ignored (column 1) but which, in fact, have strong protective effects (as shown in the "correct" analysis in column 2). Furthermore, where the analysis ignoring endogeneity gave statistically significant estimates of the correct sign (any breast feeding, water quality, community density, and breast feeding × age), the parameter values were substantially biased toward the null. In short, if endogeneity is ignored, incorrect conclusions will be drawn on the determinants of health.

Tracing the paths by which education affects health

Interesting and important as the above results are, the major potential contribution of the Cebu Project is the integration of these two levels of analysis into a single, integrated behavioral-cum-biomedical description of child health. This integration can be illustrated by tracing through the pathways by which maternal education affects health-related behavior, and how such behavior, in turn, affects health.

Before tracing this path, the aggregate effect on diarrhea of a one-year increase in maternal education can be calculated using the reduced form (expression 3). The net effect of a one-year increase in maternal education would be to reduce the incidence

TABLE 7. Percent change in diarrhea due to behavioral change induced by a one-year increase in maternal education at every 2 months during the first year of life, urban Cebu, Philippines, 1983–1985*

	Age of infants in months					
	2	4	6	8	10	12
Feeding practices						
Exclusive breast feeding, no exposure to pathogens	4.50	1.73	0.02	N/A	N/A	N/A
Any breast feeding	2.53	2.55	2.03	1.43	0.94	0.46
Total calories	−1.00	−1.18	−1.23	−1.35	−1.30	−1.34
Health service use						
Preventive health care	−0.50	−0.35	−0.21	−0.29	−0.27	−0.16
Measles immunization	N/A	N/A	N/A	N/A	0	0
Health practices, personal and environmental						
Pathogenicity of food processing	0.06	−0.35	−0.01	−0.08	−0.09	−0.23
Poor type of excreta disposal	−6.74	−4.97	−3.76	−2.96	−2.26	−1.54
Quantity of soap per capita	−0.01	−0.01	−0.01	−0.01	−0.01	−0.01
Good quality of drinking water source	0	0	0	0	0	0

* Entries are computed with the following formula: % change = $(D_3 - D_1) \times 100/D_1$, where D_1 is the simulated mean diarrhea when coefficients of the explanatory variables in the diarrhea structural equation are multiplied by their corresponding sample means. D_3 is the simulated mean diarrhea where the mean of instrumental variable of interest is changed by increasing mean maternal education by one year and the rest of the explanatory variables are at their sample means. N/A, not applicable because no infant had this type of feeding or immunization during this period.

of diarrhea episodes by about 5 percent in each time period. (Over the first year of life, the mean incidence of diarrhea in a 7-day period increased from under 4 percent in the first 2 months to over 15 percent in the final 2 months.)

By combining the simulated effects of education on health (table 3) with the simulated effect of behavioral changes on diarrhea (table 5), the education-behavior-diarrhea pathway can be traced. Table 7 shows that the three major pathways through which maternal education affects health in this population are: a large reduction in diarrhea (about 4 percent) because of improved excreta disposal practices, with the effect being particularly strong in the early months of life; a substantial reduction in diarrhea because of the increase in calories given to the child, with the effect greater toward the end of the first year of life; and a substantial, offsetting, increase in diarrhea because of a reduction in the number of mothers who breast-fed, with reduced exclusive breast feeding most important in the early months, and reduced breast feeding most deleterious early but remaining serious throughout the first year of life.

This information is presented graphically (for 6-month-old urban children) in figure 3. From figure 3, it is evident that some pathways are not important in this population either: because maternal education has little effect on behavior (as is the case of water supply for this urban population); or because the prevalence of the particular behavior is low (only 9 percent of mothers are exclusively breast-feeding their children at this age, for instance); or because changes in the particular behavior have little effect on health in this period (such as changes in the use of soap).

SUMMARY AND CONCLUSIONS

This paper shows that an integrated socioeconomic-biomedical model of child health can be specified and the parameters estimated. The results show that if endogeneity is ignored, incorrect conclusions are

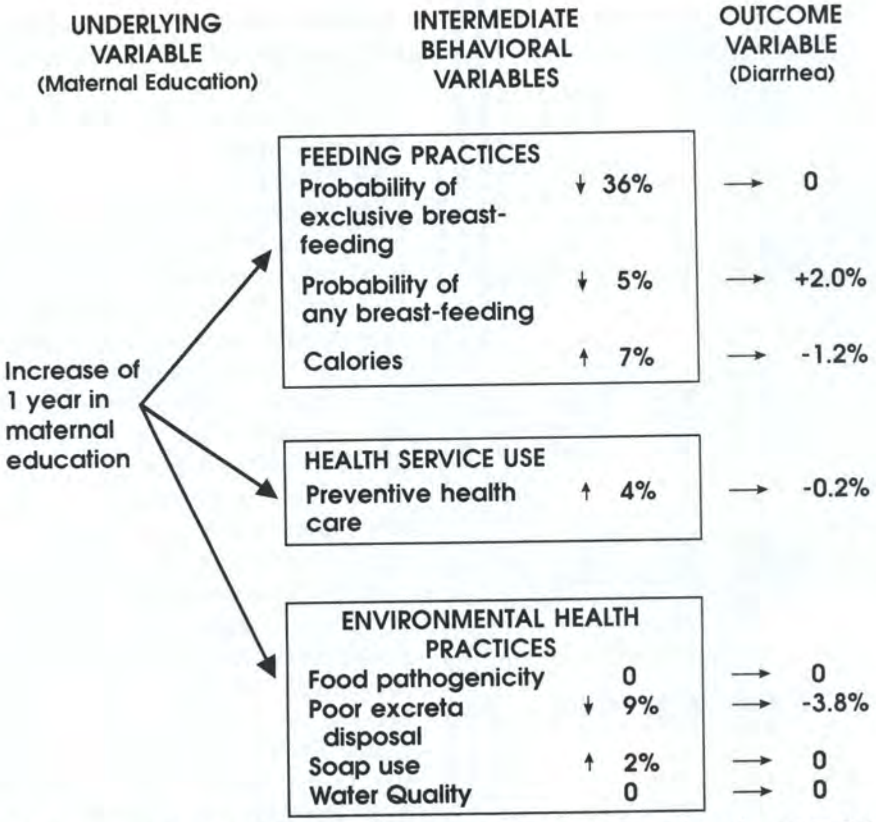


Figure 3. The pathways through which maternal education affects health at age 6 months, urban Cebu, Philippines, 1983–1985.

drawn concerning the effects of several determinants of child health. The analytic approach permits unique and readily understandable disaggregation of the effects of underlying and intermediate determinants of child health.

REFERENCES

1. Mosley WH, Chen LC. An analytic framework for the study of child survival in developing countries. Population and Development Review 1984;10: 25–48.
2. Preston SH. Causes and consequences of mortality declines in less developed countries during the twentieth century. In: Easterlin RA, ed. Population and economic change in developing countries. Chicago: University of Chicago Press, 1980.
3. Schultz TP. Studying the impact of household economic and community variables on child mortality. Population and Development Review 1984; 10:215–36.
4. Trussell J, Mencken J. Estimation of levels, trends and covariates of infant and child mortality in statistically poor countries. Population and Development Review 1984;10:325–46.
5. Rosensweig MR, Schultz TP. Estimating a household production function: heterogeneity and the demand for health inputs, and their effects on birthweight. Journal of Political Economy 1983; 92:723–46.
6. Behm H. Socioeconomic determinants of mortality in Latin America. Proceedings of the meeting on socioeconomic determinants and consequences of mortality, el Colegio de Mexico, Mexico City, June 19–25, 1979. New York: World Health Organization, 1980.
7. Caldwell JC, McDonald P. Influence of maternal education on infant and child mortality. Proceedings of International Population Conference. Liège, Belgium: International Union for the Scientific Study of Population, 1981.
8. Cochrane SH, O'Hara DJ, Leslie J. The effects of education on health. World Bank Staff Working Paper no. 405, Washington, DC: World Bank, 1980.
9. Bongaarts J. A framework for analyzing the proximate determinants of fertility. Population and Development Review 1978;4:105–32.
10. Mosley WH. Biological contamination of the environment by man. In: Biological and social aspects

of mortality and length of life. Preston SH, ed. Liège, Belgium: International Union for Scientific Study of Population, 1980:36–67.

11. Mosley WH. The demographic impact of child survival programs. International Symposium, Center for Public Health Research, Mexico City, July 1986.
12. Davanzo J. A household survey of child mortality determinants in Malaysia. Population and Development Review 1984;10:307–24.
13. Davanzo J, Butz WP, Habicht JP. How biological and behavioral influences on mortality in Malaysia vary during the first year of life. Population Studies 1983;37:381–402.
14. Butz WP, Davanzo J, Habicht JP. Biological and behavioral influences on the mortality of Malaysian infants. Santa Monica, CA: Rand Corporation, 1982.
15. Tekce B, Shorter F. Determinants of child mortality: a study of squatter settlements in Jordan. Population and Development Review 1984;10:257–80.
16. Adair LS, Popkin B. Birth weight, maturity and proportionality in Filipino infants. Hum Biol 1988; 60:319–39.
17. Guilkey DK, Popkin BM, Akin JS, et al. Prenatal care and pregnancy outcomes in the Philippines. Journal of Development Economics 1989;30: 241–72.
18. Wong E, Popkin B, Akin JS, et al. Accessibility, quality of care and prenatal care use in the Philippines. Soc Sci Med 1987;24:927–44.
19. Wonnacott TH, Wonnacott RJ. Introductory statistics for business and economics. New York: John Wiley and Sons, 1972.
20. Judge GG, Griffiths WE, Hill RC, et al. The theory and practice of econometrics. New York: John Wiley and Sons, 1980.
21. Maddala GS. Introduction to econometrics. New York: Macmillan, 1988.
22. Behrman JR, Wolfe BL. How does mothers' schooling affect family health, nutrition, medical care usage, and household sanitation? Journal of Econometrics 1987;36:185–204.
23. Avery R, Hotz J. Hotztran users' manual. Old Greenwich, CT: CERA Economic Consultants, 1985.
24. Crocker TD, Schulze W, Ben-David S, et al. Experiments with economics of air pollution epidemiology: volume 1 of methods development for assessing air pollution control benefits. Washington DC: United States Environmental Protection Agency, 1979. (Report EPA-600/5-79-00 1a).
25. Kneese AV. Measuring the benefits of clean air and water. Washington, DC: Resources for the Future, 1986.
26. Heckman JJ. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement 1976;5:475–92.
27. Hausman JA. Specification tests in econometrics. Econometrica 1978;46:1251–71.

All Coliforms Are Not Created Equal: A Comparison of the Effects of Water Source and In-House Contamination on Infantile Diarrheal Disease

James VanDerslice, M.S.E.E.
Carolina Population Center and
Department of Environmental Sciences and Engineering
University of North Carolina at Chapel Hill

John Briscoe, Ph.D.
The World Bank
Washington, D.C.

In press
Sub~~mitted~~ ~~to~~ Water Resources Research March 21, 1991

Revised August 30, 1991

NOT FOR CITATION

Short title: All Coliforms Are Not Created Equal

ABSTRACT

Storing drinking water in the home is common in the developing world, even among those served by in-house connections to piped supplies. Several studies have documented increased concentrations of fecal coliforms during household storage. This has led to the belief that in-house water contamination is an important transmission route for enteric pathogens and, moreover, that improving water source quality is not warranted until that quality can be maintained in the home. We contend that in-house contamination does not pose a serious risk of diarrhea because the pathogens contaminating a household's water are those already infecting the household members. A contaminated water source poses much more of a risk since it may introduce new pathogens into the household. The effects of water source and in-house contamination on diarrheal disease are estimated for 2355 Filipino infants. The results confirm our hypothesis: contaminated water sources pose a serious risk of diarrhea while contamination of drinking water in the home does not. Water boiling is shown to eliminate the risk of diarrhea due to water source contamination. The results imply that improvements in water source quality are more important than improving water storage practices.

Introduction

The past decade has seen a major effort to improve water quality and availability in the developing world. A principal goal of this effort has been to reduce the high levels of waterborne and water-washed diseases, notably diarrheal disease. However, families served by hand pumps, standpipes and even in-house connections from intermittent piped supplies, commonly store drinking water in the home where it may be contaminated by fecal material. This observation has lead to the belief that providing a high quality water supply is not worthwhile if the quality cannot be maintained during household storage [Feachem et al., 1983, p. 211].

Upon further reflection, however, this would appear to be incorrect. Consider a hypothetical person living by himself. He draws water from a well containing 100 fecal coliforms (FC) per 100 ml. He does not practice good hygiene so his stored drinking water contains 1000 fecal coliforms/100 ml. Does the higher level of contamination in his stored water mean that in-house contamination is more important than water source contamination? No. Since our hypothetical person himself is the source of the pathogens contaminating his drinking water during storage, these pathogens can do him no further harm.

Now consider a more realistic case of a family with poor hygienic practices. Pathogens infecting any family member are easily spread from his or her hands (contaminated during defecation) to food, utensils, or stored water. Soon the whole family would be infected by and would excrete a shared set of "internal" pathogens. Contamination of stored water by these "internal" pathogens would pose little additional risk to the family's health since family members would already be infected.

Contamination of the family's water source, however, is a very different matter. The contaminated water source may contain pathogens from other peoples' feces, pathogens that are new to the household environment. It is these "external" pathogens that can cause new infections that put the family at risk of diarrheal disease. Providing a high quality water source eliminates this source of "external" pathogens, reducing the family's risk of diarrhea.

Obviously this is a simplistic description of waterborne pathogen transmission. A contaminated water supply will probably contain both "internal" and "external" pathogens. However, if a contaminated water source is a principal means of introducing new pathogens into the home, then water source contamination will have a significant effect on diarrheal disease. Moreover, if in-house contamination is simply passing on "internal" pathogens already infecting the family, then it will not be significantly associated with diarrhea. The objective of this study is to estimate and compare the effects of water source contamination and in-house contamination on diarrheal disease.

Previous Studies of In-House Water Contamination

In-house water contamination has been recognized as a possible transmission route for enteric pathogens for over 25 years [van Zijl, 1966]. Even so, the existing literature provides little information on the risk it poses for diarrheal disease.

Most studies of in-house contamination have documented changes in water quality during storage by simply comparing the distribution of indicator organisms in water sources to the distribution in the storage containers. These studies have found substantial increases in coliform levels [Rajasekaran et al., 1977; Shiffman et al., 1978; El Attar et al., 1982; Lloyd-Evans et al., 1984; Magnani et al., 1984; Pickering, 1985; Lehmusluoto, 1986; Blum et al., 1990; Pinfold, 1990], little or no change in overall coliform levels [Oluwande, 1980; Esrey et al., 1986; Young and Briscoe, 1986; Sutton and Mubiana, 1989], and, in one case, a large decrease in mean coliform levels [Tompkins et al., 1978]. These aggregate measures are of limited use because they conceal the changes occurring in each household. Household-level changes in quality are observed by collecting paired water samples from the household's water source and storage container. Three studies, in Lesotho [Feachem et al., 1978], Malawi [Lindskog and Lindskog, 1988], and Sri Lanka [Mertens et al., 1990a] have used this method. These studies found considerable variation in the difference between water source and stored-water fecal coliform concentrations. The differences ranged from small decreases to increases of more than 1000 FC/100 ml.

The use of fecal coliforms as an indicator of fecal contamination in tropical waters has been seriously questioned [Hazen, 1988]. Three studies have circumvented this problem by analyzing stored water for the presence of pathogens. Spira et al. [1980] sampled water sources and stored water in Bangladesh for V. cholerae in neighborhoods where cholera cases had occurred. In a similar fashion, Echeverria et al. [1987] sampled for enterotoxigenic Escherichia coli (ETEC) in Thai neighborhoods where ETEC-positive diarrhea cases had come from. Neither researcher found any evidence that drinking water had been contaminated with these pathogens during storage in the home. A study in a rural Egyptian village [Khairy et al., 1982], however, did find Stronglyoides and Ascaris in 10% and 15% of the household storage jars, respectively, while the source water was free from these organisms. While stored water appears to be subject to contamination by parasites, the studies in Thailand and Bangladesh indicate that bacterial pathogen transmission via in-house contamination may not be as common as increases in fecal coliform levels suggest.

To our knowledge there have been no studies of the effect of in-house contamination on diarrheal disease. A cross-sectional study conducted in Nigeria [Huttly et al., 1987] found no association between the presence of a cover on the water storage container and the prevalence of diarrhea in any age group. Four studies of diarrheal disease, conducted in the Philippines [Magnani et al., 1984], Lesthoto [Esrey et al., 1986b], Bangladesh [Henry and Rahim, 1990] and Sri Lanka [Mertens et al., 1990b], used stored water quality as their measure of exposure to waterborne pathogens. None of these studies found a statistically significant relationship with diarrheal disease. These results provide no evidence that in-house contamination or water source contamination are important risk factors. Moreover, even if a significant association had been observed, it would have been impossible to distinguish the separate effects of water source and in-house contamination.

Methods

Study Design

This study uses data from the Cebu Longitudinal Health and Nutrition Survey, a prospective community-based investigation of infant health and nutrition in Cebu, Philippines. The study area consists of Cebu City and surrounding peri-urban areas and has an estimated population of 1 million. Seventeen of the 95 barangays (political districts) were randomly selected and pregnant women residing in these barangays were recruited. Of the 2555 women recruited, 2355 had single live births and agreed to participate in the study.

A baseline survey conducted during the third trimester of pregnancy collected information on the household's income, assets, water source, sanitation facilities and hygienic conditions. Bimonthly interviews, conducted through the first two years of the child's life, documented feeding patterns and food preparation practices, the volume of water consumed by the child, whether the water was boiled, and the specific water source used. Diarrheal morbidity for the previous 7 days and the child's weight and height were also recorded. Additional details on the survey design and content are available elsewhere [Cebu Study Team, 1991a].

Water Sampling

Special efforts were made to characterize exposure to waterborne pathogens from the household's water source and from in-house water contamination. Water sources were sampled between two and five times over the course of a year. Water samples were collected in the same manner that users collected their water. Spigots, pump spouts, and outflow pipes were not sterilized. Open dug wells without pumps were sampled using an aluminum bucket that was sterilized by flaming just before sampling.

Household water collection and storage practices were documented during a special survey of 254 households randomly selected from the larger study population. At each household two water samples were collected, one from the drinking water storage container and the other from the water source that had supplied the water in the storage container. Samples were collected from the storage container in the same manner as the family removed water. These pairs of samples were used to estimate the level of in-house contamination.

Fecal coliforms were used as an indicator of fecal contamination. These organisms are continually present in large numbers in the feces of warm-blooded animals [Feachem et al., 1983] so their presence in drinking water indicates that the water has been contaminated with fecal material. While fecal coliforms appear to be more sensitive and specific than total coliforms in tropical waters [Lavoie, 1983], they are not ideal indicators of fecal contamination [Hazen et al., 1988]. They have been isolated in areas thought to be devoid of fecal material [Fujioka, et al. 1988], and may not be as persistent as some enteric pathogens [McFeters et al., 1974]. Nevertheless, the use of fecal coliforms is consistent with bacteriological water quality standards [W.H.O., 1984].

All water samples were transported on ice to the laboratory where they were refrigerated overnight and analyzed the following morning. Membrane filtration [American Public Health Association, 1985] was used to culture fecal coliform colonies using M-FC agar incubated at 44.5°C for 24 hours. Volumes of 1, 10, and 100 ml were filtered from each sample. Dark blue colonies were counted as fecal coliforms.

Of the 1650 water source samples collected, 154 (9%) produced unreliable estimates due to the presence of uncharacteristic colonies or heavy background growth. Only 2% of the 233 stored-water samples produced unreliable FC estimates. These estimates were excluded from the analyses.

When the number of colonies on a filter was in the "countable range" (10 to 100), that count was used to estimate the concentration of fecal coliforms. When more than one filter provided counts in the countable range, or when all filters had low counts (< 10), the total number of colonies counted was divided by the total volume of water filtered. When some of the filters had low counts and the other(s) were too numerous to count (TNTC), a maximum likelihood estimator was used [Haas and Heller, 1988]. Finally when all filters were TNTC, the estimate was set at 200 FC per 1 ml (20,000 FC/100 ml).

Estimating In-House Contamination

In-house water contamination is difficult to measure as the actual number of organisms added to the stored water can not be readily observed. The cumulative effect of in-house contamination will be reflected by an increase in bacterial concentrations. The bacteria observed in the storage container, however, will be a combination of those introduced by contaminated hands or cups, and those originating in the water source. Thus the concentration of FC due to in-house contamination (C_H) is the FC concentration observed in the storage vessel (C_V) minus the concentration of FC originally from the water source (C_s).

$$C_H = C_V - C_s \tag{1}$$

Standard bacteriological methods can not differentiate between fecal coliforms from in-house contamination and those from a contaminated source water. In order to estimate in-house contamination by this method, the concentration of FC in the storage container which came from the water source must be determined. While re-growth of fecal coliforms has been observed in nutrient-rich surface waters [Kinney et al., 1978; Hendricks, 1972; Carillo et al., 1985], re-growth during household storage was assumed to be negligible as virtually all households used groundwater for drinking. Therefore the concentration of water source FC in the storage container at the time of sampling is the concentration observed at the source (C_w) minus the concentration that died during storage (C_d).

$$C_s = C_w - C_d \tag{2}$$

There was no way to reliably estimate the number of FC from the water source which died during storage. However, even with no information on the level of die-off, upper and lower bounds on the level of in-house contamination can be calculated. The smallest value for in-house contamination (C_{Hmin}) occurs when all water source bacteria are assumed to survive (i.e., $C_s = C_w$).

$$C_{Hmin} = C_v - C_w, \text{ for } C_v > C_w, \text{ (net increase)} \tag{3}$$

$$C_{Hmin} = 0, \text{ for } C_v < C_w, \text{ (net decrease)} \tag{4}$$

The largest possible value for in-house contamination (C_{Hmax}) occurs when none of the water source fecal coliforms are assumed to survive (i.e., $C_s = 0$).

$$C_{Hmax} = C_v - 0 \tag{5}$$

A reasonable estimate of in-house contamination would be some point in this interval. Since exponential increases in pathogen dose are related to linear increases in the risk of diarrhea [Akin, 1981], $\log_{10}$ FC concentrations are used to model the effects of water contamination on diarrheal disease. As such, the level of in-house contamination was estimated as the mid-point between the minimum and maximum values measured on a log scale.

$$\log_{10}(C_H) = \frac{1}{2}\,[\log_{10}(C_{Hmax}) - \log_{10}(C_{Hmin})]. \tag{6}$$

In many cases the minimum and maximum values for $\log_{10}(C_H)$ were almost equal, indicating that the estimate of in-house contamination was not sensitive to the assumed level of water source FC die-off. For example, when the stored water contained ten times as many fecal coliforms as the source water, the difference between $\log_{10}(C_{Hmax})$ and $\log_{10}(C_{Hmin})$ was only 0.05. When the FC concentration in the storage container was only twice that observed at the water source, this difference was 0.3.

When the storage vessel was free of fecal coliforms, in-house contamination (C_H) was set to 0.9 FC/100 ml, the lower limit of detection. When the water source contained no fecal coliforms, the minimum and maximum estimates were equal and C_H was set to the concentration of FC observed in the storage vessel.

Diarrheal Model Specification

This research employs a previously developed longitudinal model of diarrheal disease [Cebu Study Team, 1991a, b]. In this model diarrhea

(D) results from past growth (G), behavioral factors (Y), and "underlying" socioeconomic and environmental factors (Z),

$$D_{t,i} = \beta_1 G_{t-1,i} + \beta_2 Y_{t-1,i} + \beta_3 Z_{t,i} + \mu_{Di} + \epsilon_{Dt,i} \qquad (7)$$

for t = 1 to 6 two-month time periods, and i = 1 to N study infants. Furthermore, the behavioral factors are determined by growth and diarrhea in the previous time period, past behaviors, and underlying socioeconomic and environmental factors.

$$Y_{t,i} = \alpha 1 G_{t-1,i} + \alpha_2 D_{t-1,i} + \alpha_3 Y_{t-1,i} + \alpha_4 Z_{t,i} + \mu_{Yi} + \epsilon_{Yt,i} \qquad (8)$$

These equations contain two error terms. The first error term, μ, represents unobserved differences unique to each child or family. These may be the genetic endowment of the child, the parent's perceptions of risk, or other factors. These differences are expected to persist over the course of the study so the same error term is used for all time periods. The unobserved variations may affect each outcome differently. As such the μ's are different for each equation but correlated across equations. The second error term, ϵ, is a purely random disturbance which varies across individuals and with time, and is not correlated across equations.

Description of Variables

Diarrhea. Diarrhea is measured by a binary variable indicating whether the child experienced a diarrheal episode in the 7 days preceding the interview. It can be thought of as arising from a latent continuous measure of diarrheal severity. If the severity is greater than some threshold level, then the diarrheal episode is reported by the mother and the variable takes on the value of one. Otherwise the episode is not observed, and the variable will have the value of zero.

Behavioral factors. Many determinants of diarrheal disease are governed by behavior. Exposure to waterborne pathogens, for example, is in part determined by the choice of water source, amount of water consumed, and household water treatment. The behavioral variables used in the model measure various exposures to pathogens and factors affecting the child's susceptibility to infection.

Exposure to contaminated source water is measured as the $\log_{10}$ daily dose of FC from the water source. The dose was estimated by multiplying the infant's 24-hour total water intake by the expected FC concentration for the water source used by the household for the period two weeks prior to the interview date. In-house contamination is measured by the $\log_{10}$ daily dose of FC added during household storage. It was calculated by multiplying the infant's 24-hour total water intake by the estimated increase in FC concentration due to in-house contamination.

Water boiling is expected to reduce the risk of diarrhea due to contaminated water to the same extent that contaminated water increases that risk. Interactions of water boiling with the two water contamination variables are included to model this effect. Since water boiling may also indicate a greater awareness of good hygiene, the main effect is also included. The water boiling variable indicates that the water consumed by the child the day before the interview had been boiled.

Exposure to fecal contamination around the house is measured by two variables: the lack of toilet or latrine, and the presence of feces in the yard. The presence of feces was assessed through direct observations by trained fieldworkers.

Several variables measure hygienic behaviors. The level of water service is used as a proxy for water use as families with an on-site water source are assumed to use more water for bathing, cleaning, and hand washing. Per capita nonlaundry soap usage, estimated from reported household expenditures for soap, is used as a proxy for personal hygiene. Household crowding, measured as the number of family members divided by the number of rooms, is used an indicator of higher person-to-person pathogen transmission. Finally, a variable indicating a high potential for food contamination was constructed from food preparation and storage practices at each longitudinal survey.

Breast-feeding may reduce the child's susceptibility to infection via maternal antibodies and provides nourishment which is free from contamination. Feeding patterns are measured by three dichotomous variables signifying whether the child was exclusively breast-fed, breast-fed and given nonnutritive supplements (such as plain water or juice), or given nutritive foods in addition to breastmilk. The omitted category is not breast-fed.

Use of preventive health care services is expected to improve the child's susceptibility to infection and may indicate that the mother has a greater awareness of her child's health. The variable indicates that some type of preventive health care (e.g., immunizations, well-baby check-up) was used in the two months preceding the interview.

Growth. The child's weight at the previous survey is included as a measure of nutritional status, an indicator of susceptibility to infection. The values are standardized at each cross-section.

Underlying factors. Several underlying risk factors are thought to have direct effects on diarrheal disease. Age may reflect the immunological development of the child, secular trends in economic factors, and may capture age-related factors not adequately represented by the intermediate behavioral variables. Age-squared is included to capture nonlinearities. The child's sex, another commonly observed risk factor, may act as a proxy for unmeasured differences in immunological development between males and females, or represent differences in child-related behaviors.

Diarrhea has frequently been associated with season or rainfall. This may be due to enhanced survival of bacteria in humid weather, increases in water source contamination after large storms, or changes in food availability and prices during the growing season. The total rainfall in centimeters (cm) over the past two weeks is used to model these effects. Finally, community density is included as high-density areas are characterized by higher levels of environmental contamination.

Estimation Methods

In equation 7, diarrhea in the present time-period is specified as a function of past growth, past behaviors, current socioeconomic and environmental conditions, and two error terms. This model differs from traditional research in two important ways. First, the model explicitly acknowledges that behaviors are determined in part by the child's health (equation 8). Secondly, the model allows for unobserved differences between children or their families, differences affecting both the family's behaviors and their child's health. These two refinements capture an important aspect of diarrheal disease in children: parents may recognize risks to their children's health and modify their behaviors to reduce those risks [Briscoe et al., 1990].

Failure to account for these effects can lead to biased parameter estimates and spurious results [Cebu Study Team, 1991a, b; Briscoe et al., 1990]. Consider the effect of a behavioral factor on diarrhea. The behavioral structural equation (8) specifies that all behaviors are determined in part by μ_Y, the random error representing unobserved differences between children. Since the unobserved differences which affect behaviors may also affect diarrhea, μ_Y is correlated with μ_D, the equivalent disturbance term in the diarrhea equation. As a result, the behavioral variables in the diarrhea equation are correlated with the error term μ_D. This is a violation of one of the basic assumptions of the ordinary-least-squares (OLS) estimator and if OLS is used the estimated effect of behaviors on diarrhea will be biased (inconsistent). This is because some of the variability in the dependent variable due to random error is mistakenly attributed to the independent variables.

Consistent estimates can be obtained if the behavioral variables are purged of their association with the unobserved factors, μ_D. This can be accomplished by using instrumental variables in place of the behavioral variables. Instrumental variables are variables correlated with the behavioral risk factors, but not correlated with the individual-specific error term.

Suitable instruments can be derived from the behavioral structural equation. If the growth, diarrhea and behavioral variables on the right-hand side are substituted out using their respective structural equations, equation 8 becomes

$$Y_{t,i} = \delta_1 G_{t-2,i} + \delta_2 D_{t-2,i} + \delta_3 Y_{t-2,i} + \delta_4 Z_{t,i} + \delta_5 Z_{t-1,i} + \mu_{Yi} + \epsilon_i \qquad (9).$$

The same variables (at t-2) can be substituted out again. This process is repeated until only the underlying variables and error terms remain.

$$Y_{t,i} = \gamma_1 Z_{t,i} + \gamma_2 Z_{t-1,i} + \cdots + \gamma_s Z_{t-s,i} + \mu_{Yi} + \epsilon_i \qquad (10)$$

This is the reduced-form of the behavior structural equation, specifying any behavior as a function of strictly exogenous, underlying variables.

The reduced-form equation can be used to create instruments for the behavioral variables. Since the underlying variables, Z, are not affected by the family's behaviors, they are not correlated with the individual-specific error term, μ_D. As such equation 10 can be estimated using standard techniques. Predicted values for the behavior variable, $Y_{t,i}$, are then generated and used as the instrumental variable. The predicted values should be well correlated with the actual values and not correlated with the error term μ_D. The same procedure was used to generate predicted values for the growth variable. Using these predicted values in place of the growth and behavioral variables produces consistent estimates of the parameters in the diarrheal equation [Judge et al., 1982].

Since the dependent variable is binary and the error term is assumed to be normally distributed, a random-effects probit estimator was used. This estimator assumes the same "random-effect" for all observations from a given child [Judge et al., 1982; Avery and Hotz, 1985]. This random effect represents the "unobserved" characteristics of that child (or family) affecting the child's health and the family's behaviors.

A maximum likelihood procedure, found in the HOTZTRAN software [Avery and Hotz, 1985], was used to estimate the parameters. The standard errors may be underestimated because the variation associated with use of instruments is not taken into account. While it is theoretically possible to correct for the use of instruments [Maddala, 1983], it is not feasible given the large number of instruments used.

The coefficients estimated from a probit model can not be interpreted as the marginal effect of an independent variable on the probability of diarrhea. The marginal change in the probability of diarrhea resulting from a unit change in a dependent variable, X_k, was calculated by [Maddala, 1983, p. 23]:

$$\partial P(D=1)/\partial X_k = \phi(X\,\beta)\,\beta_k \qquad (11).$$

Approximate confidence intervals for the marginal effects were calculated by using the end-points of the 95% confidence interval of the parameter estimate in place of the parameter estimate itself.

$$\text{lower 95\% confidence limit} = \phi(X\,\beta)\ [\beta - 1.96(\text{s.e.}_{\beta})] \tag{12}$$

$$\text{upper 95\% confidence limit} = \phi(X\,\beta)\ [\beta + 1.96(\text{s.e.}_{\beta})] \tag{13}$$

Confidence intervals for the marginal effects will include the null value of zero when the confidence interval for the parameter estimate includes zero.

Results

Characteristics of the Study Population

There is wide variation in the demographic characteristics of the study population (Table 1). Education levels are quite high in Cebu; over 90% of the parents have completed primary education and 15% have graduated from high school. Most of the households (70%) are headed by waged or salaried workers and one-fourth are self-employed. Household incomes range from 0 to 12,500 pesos per week with a median of 200 (approx. $ U.S. 10). Total household assets range from 0 to almost 1.5 million pesos with a median of 2,400 (approx. $ U.S. 120).

Environmental sanitation conditions are also quite variable. Over three-quarters of the households use an "adequate" excreta disposal facility (i.e., flush or pour-flush toilet or latrine). However, there is no sewerage in Cebu City and most on-site disposal systems would be considered inadequate. Almost 20% of the families report that they defecate into a canal or on the seashore. Fecal material was observed at one-third of the sample houses.

Water Source Use and Quality

Over 500 water sources are used by the study population. Almost all households use an "improved" water source; 59% are served by boreholes and 30% by the municipal piped supply. The remaining households rely on open dug wells (5%), or dug wells fitted with pumps (5%). The sample population also enjoys a relatively high level of service; 10% have in-house connections and another 48% are within 1 minute of their water source. Only a small proportion (5%) must walk more than 5 minutes to fetch water.

Boreholes and the piped supply generally provide high quality water (Table 2). Over three-fourths of the samples from these sources produced no FC colonies, and another 10% had less than 10 FC/100 ml. Still, over 10% of the boreholes and 10% of the samples from the piped supply were contaminated with more than 100 FC/100 ml. Dug wells had much higher levels of contamination. Those fitted with covers and pumps were grossly contaminated (> 100 FC/100 ml) less often than open dug wells (41% vs. 78%).

Water Collection and Storage Practices

Over 99% of the households report storing drinking water in the home, including many of those with in-house connections. Almost all households have only one storage container, and the stored water is used for several purposes (e.g., drinking, cooking, bathing and cleaning). While many types of containers were used, they can be classified into four categories: small containers (e.g., pitchers and used Clorox bottles), large containers (e.g., 6-gallon gasoline cans and used cooking oil cans), traditional clay jars, and pails. The small containers, large containers, and clay jars were each used by about one-third of the study households (Table 3). Small containers were frequently used for both collection and storage. About half of the earthen jars were subject to contamination from scoops or cups. The remaining jars were fitted with spigots. Water was usually poured from the other types of containers. Most of the containers had covers or caps.

Changes in Water Quality During Storage

When both the water source and stored-water samples were "too numerous to count", it was impossible to determine if the concentration of fecal coliforms had increased, decreased, or remained the same during storage. These samples, as well as those with unreliable counts due to heavy background growth, were not used in the analysis, leaving only 184 of the 233 pairs of water samples collected.

Table 4 presents the distribution of the change in concentration between the source and the storage container and the levels of in-house contamination estimated using equation 6. One-third of the household samples had substantially higher concentrations of fecal coliforms (>100 FC 100 ml) than the respective water source sample. Over 30% of the sample pairs demonstrated no net change (-1 to 1) in fecal coliform concentrations. This may reflect no change in quality, or the combination of high levels of in-house contamination and die-off. Surprisingly, 16% of the sample pairs demonstrate a net decrease in FC concentration, indicating that bacterial die-off can be greater than increases due to in-house contamination.

Covering the storage container appears to have little effect on in-house contamination. The geometric mean (G.M.) of the estimated increase in the concentration fecal coliforms per 100 ml was 1.82 for covered containers and 1.51 for containers that were not covered. Samples taken from containers from which water was scooped demonstrated slightly larger increases (G.M.=1.96) than samples from containers where water was poured or flowed through a spigot (G.M.=1.48). Small storage containers were subject to less in-house contamination (G.M.=1.16) than the large containers (G.M.=1.79) or earthen storage jars (G.M.=1.92).

Water Consumption and Boiling

The proportion of children fed plain water (i.e., not as a part of food or a prepared drink) increased substantially over the first 6 months (Table 5). By the time the children were 8 months old, over 99% had received water in the 24 hours preceding the interview. Mean consumption for those fed any water almost doubled over the first year. About 90% of the mothers reported that they boiled the water given to their 2 month old infants. This proportion dropped to about half by the time the child was 6 months old and remained constant for the rest of the child's first year of life.

Exposure to Water Source and In-House contamination

The distributions of predicted daily FC doses from the water source and from in-house contamination for all 6 time periods combined are presented in Table 6. The low doses were censored to 1 fecal coliform per day. Ten percent of the in-house contamination doses and 84% of the water source doses were so censored.

Diarrheal Disease in Cebuano Infants

The proportion of children experiencing diarrhea during the week previous to the interview increased dramatically over the first 8 months of the child's life, from just over 7% when the infants were 2 months old to 25% just 6 months later. During this period the prevalence was the same for males and females. From 8 months to a year of age the prevalence of diarrhea prevalence among males continued to increase while the prevalence among females decreased slightly to 22%. At one year of age, female children were experiencing about 20% fewer cases of diarrhea than male children.

Effects of Water Source and In-House Contamination on Diarrheal Disease

The parameter estimates and t-statistics for two models of diarrheal disease are presented in Table 7. In the first model of the effects of water contamination on diarrhea water source dose is a very strong risk factor (t=3.0). Increasing the dose from 1 to 100 FC per day increases the probability of diarrhea 22%, from 0.18 to 0.22. In-house contamination, however, is not associated with diarrhea and has a point estimate very close to zero. Excreta disposal, food pathogenicity, age, and community density are all significant risk factors, and exclusive breast-feeding a significant protective factor. The signs of all significant and marginally significant coefficients are as expected.

This simple model ignores the fact that water boiling should have a greater protective effect when water contamination levels are high. This effect was modeled by including interactions of boiling with the two water contamination variables (model #2). In the case of water source contamination, the boiling interaction coefficient is negative indicating that boiling reduces diarrhea more as the level of contamination increases. Water source dose is still statistically significant and the interaction is marginally significant. Neither in-house contamination nor its interaction with boiling are statistically significant. The other parameter estimates do not change appreciably from the first model.

Figure 1 illustrates the importance of the interaction between water source contamination and boiling. Predicted probabilities of diarrhea over a 7-day period were computed for various levels of water source dose when the water was boiled and not boiled. Separate predictions were made using the model including the interaction between boiling and water contamination, and the model containing only their main effects.

When the interaction is not included, the increase in the probability of diarrhea from a ten-fold increase in water source dose is the same whether or not the water is boiled. The model including the interaction term gives much more intuitive results. When water is boiled, water source contamination does not increase the probability of diarrhea. However, source contamination has a considerable effect on diarrhea when water is not boiled.

Marginal increases in the probability of diarrhea resulting from unit increases in the water contamination and water boiling variables are presented in Table 8. Each $\log_{10}$ increase in the water source fecal coliform dose increases the probability of diarrhea by 0.043. The interaction of source contamination with water boiling has exactly the opposite effect, reducing the probability of diarrhea by 0.044 per $\log_{10}$ increase in dose. Thus the effect of the boiling interaction variable is to cancel out the risk due to water source contamination.

Discussion

The results from the diarrhea models confirm our hypothesis: in-house water contamination does not pose a serious risk of diarrhea. Its parameter estimate is not significant, or even positive, in either of the models estimated.

The model results, however, indicate that water source contamination poses a significant risk for diarrhea. When water is not boiled, water source contamination substantially increases the probability of diarrhea. A ten-fold increase in the concentration of fecal coliforms would lead to a 17% increase in diarrheal prevalence. Conversely, if families using moderately contaminated dug wells (100 FC/100 ml) were able to use a high quality water source, diarrhea among their children would be reduced by over 30%.

Sanitation is also an important risk factor for diarrhea. The relative importance of water contamination, sanitation, and the level of service, as well as the effects of multiple interventions will be addressed in a forthcoming paper.

In summary, all coliforms are not created equal. There are important differences between in-house water contamination by "internal" pathogens and contamination of one's water source by "external" pathogens. The implications for planning improvements to water supplies are clear. Improving water source quality can have a substantial impact on diarrheal disease. Eliminating in-house contamination may have no impact unless other household transmission routes are eliminated as well. In any case, there is no reason to delay making improvements in water source quality because of contamination occurring in the home.

Notation List

C_d	concentration of fecal coliforms from the water source which died during storage.
C_H	concentration of fecal coliform due to in-house contamination.
C_{Hmax}	maximum value for the concentration of fecal coliform due to in-house contamination.
C_{Hmin}	minimum value for the concentration of fecal coliform due to in-house contamination.
C_s	concentration of water source fecal coliform in the storage container at the time of sampling.
C_v	fecal coliform concentration observed in the storage vessel.
C_w	concentration of fecal coliform observed at the water source.
D	diarrhea.
ETEC	enterotoxigenic *Escherichia coli*.
FC	fecal coliforms.
G	growth.
G.M.	geometric mean.
OLS	ordinary-least-squares.
$s.e._{\beta}$	standard error of the estimate of β.

Notation list cont.

TNTC	too numerous to count.
Y	behavioral factors.
Z	"underlying" socioeconomic and environmental factors.
ϕ	normal density function.
μ_D	individual-specific random error affecting diarrhea, D.
μ_Y	individual-specific random error affecting behaviors, Y.

Acknowledgments

This article is part of a collaborative research project involving the Office of Population Studies, directed by Wilhelm Flieger, and the Water Resources Center, directed by Herman van Engelen, both of the University of San Carlos, Cebu, Philippines; the Nutrition Center of the Philippines, directed by Florentino S. Solon; and a group from the Carolina Population Center, University of North Carolina at Chapel Hill (UNC-CH). Barry Popkin of UNC-CH is the project coordinator. Funding for parts of the project design, data collection, and computerization was provided by the National Institutes of Health (Contract R01-HD19983A, R01-HD23137, and R01-HD18880), the Nestle's Coordinating Center for Nutrition Research, Wyeth International, the Ford Foundation, the U.S. National Academy of Sciences, the Carolina Population Center, the U.S. Agency for International Development, and The World Bank. Funds for data analysis for this study were provided by the National Institutes of Health (R01-HD19983A).

We acknowledge the valuable help of David Guilkey for the development of the statistical models and methods. We also thank David Fugate, Gina Dahiya, Christine Moe, Arlene Quijada, Rexy Ruiz, and Ruben Quijada for their technical assistance.

References

Akin, E. W., A review of infective dose data for enteroviruses and other enteric microorganisms in human subjects, in *Microbiological Health Considerations of Soil Disposal of Domestic Wastewater*, National Center for Groundwater Research, and U.S. Environmental Protection Agency, Norman, Okla., 1981.

American Public Health Association, *Standard Methods for the Examination of Water and Wastewater, 15th ed.*, pp. 814-822, Washington D.C., 1985.

Avery, R. B., and V. J. Hotz, *Hotztran User's Manual*, CERA Economic Consultants, Inc., Old Greenwich, Conn., 1985.

Blum, D., R. N. Emeh, S. R. A. Huttly, O. Dosunmu-Ogunbi, N. Okeke, M. Ajala, J. I. Okoro, C. Akujobi, B. R. Kirkwood, and R. G. Feachem, The Imo State (Nigeria) drinking water supply and sanitation project, 1. Description of the project, evaluation methods, and impact on intervening variables, *Trans. Royal Soc. Trop. Med. Hyg.*, 84, 309-315, 1990.

Briscoe, J., J. Akin, and D. Guilkey, People are not passive acceptors of threats to health: Endogeneity and its consequences, *Int. J. Epid.*, 19, 147-153, 1990.

Carillo, M., E. Estrada, and T. C. Hazen, Survival and enumeration of fecal indicators Bifidobacterium adolescentis and Escherichia coli in a tropical rain forest watershed, *Appl. Envr. Micro.*, 50, 468-476, 1985.

Cebu Study Team, Underlying and proximate determinants of child health: The Cebu longitudinal health and nutrition survey, *Am. J. Epid.*, 133, 185-201, 1991a.

Cebu Study Team, A child health production function estimated from longitudinal data, *J. Dev. Econ.*, in press, 1991b.

Echeverria, P., D. N. Taylor, J. Seriwatnana, U. Leksomboon, W. Chaicumpa, C. Tirapat, and B. Rowe, Potential sources of enterotoxigenic Escherichia coli in homes of children with diarrhoea in Thailand, *Bull. WHO*, 65, 207-215, 1987.

El Attar, L., A. A. Gawad, A. E. M. Khairy, and O. El Sebaie, The sanitary condition of rural drinking water in a Nile Delta village, II. Bacterial contamination of drinking water in a Nile Delta village, *J. Hyg.*, 88, 63-67, 1982.

Esrey, S. A., J. P. Habicht, G. Casella, M. D. Miliotis, A. H. Kidd, J. Collett, V. Qheku, and M. C. Latham, Infection, diarrhea and growth rates of young children following the installation of village water supplies in Lesotho, paper presented at the 22nd Annual Conference of the American Water Resources Association, International Conference on Water and Human Health, Atlanta, Ga., November 9-14, 1986.

Feachem, R. G., E. Bruns, A. Cronin, P. Cross, D. Curtis, M. K. Khan, D. Lamb, and H. Southall, *Water, Health and Development: An Interdisciplinary Evaluation*, Tri-med books, LTD., London, 1978.

Feachem, R. G., D. J. Bradley, H. Garelick, and D. D. Mara, *Sanitation and Disease: Health Aspects of Excreta and Wastewater Management*, John Wiley and Sons, New York, 1983.

Fujioka, R. S., K. Tenno, and S. Kansako, Naturally occurring fecal coliforms and fecal streptococci in Hawaii's freshwater streams, *Tox. Asses.*, 3, 613-630, 1988.

Haas, C. N., and B. Heller, Averaging of TNTC counts, *Appl. Environ. Microbiol.*, 54, 2069-2072, 1988.

Hazen, T. C., Fecal coliforms as indicators in tropical waters: A review, *Tox. Asses.*, 3, 461-477, 1988.

Hendricks, C. W., Enteric bacterial growth rates in river water, *Appl. Micro.*, 24, 168-174, 1972.

Henry, F. J., and Z. Rahim, Transmission of diarrhoea in two crowded areas with different sanitary facilities in Dhaka, Bangladesh, *J. Trop. Med. Hyg.*, 93, 121-126, 1990.

Huttly, S. R. A., D. Blum, B. R. Kirkwood, R. N. Emeh, N. Okeke, M. Ajala, G. S. Smith, D. C. Carson, O. Dosunmu-Ogunbi, and R. G. Feachem, The Imo State (Nigeria) drinking water supply and sanitation project, 2. Impact on dracunculiasis, diarrhoea and nutritional status, *Trans. Royal Soc. Trop. Med. Hyg.*, 84, 316-321, 1990.

Judge, G. G., R. C. Hill, W. Griffiths, H. Lutkepohl, and T. Lee, Introduction to the theory and practice of econometrics, John Wiley and Sons, New York, 1982.

Khairy, A. E. M., O. El Sebaie, A. A. Gawad, and L. El Attar, The sanitary condition of rural drinking water in a Nile Delta village, I. Parasitiological assessment of 'zir' stored and direct tap water, *J. Hyg.*, 68, 57-61, 1982.

Kinney, E. C., Drummond, D. W., and Hanes, N. B., Effects of chlorination on differentiated coliform groups, *J. Water Poll. Control Fed.*, 50, 2307-2312, 1978

Lavoie, M. C., Identification of strains isolated as total and fecal coliforms and comparisons of both groups of indicators of fecal pollution in tropical climates, *Can. J. Micro.*, 29, 689-693, 1983

Lehmusluoto, P. O., *Survey on Contamination of Water Sources and Household Waters as an Integrated Part of Impact Management*, University of Tampere, Tampere, Finland, 1986.

Lindskog, P., and U. Lindskog, Bacterial contamination of water in rural areas. An intervention study from Malawi, *J. Trop. Med. Hyg.*, 91, 1-7, 1988.

Lloyd-Evans, N., H. Pickering, S. G. J. Goh, and M. G. M. Rowland, Food - and water hygiene and diarrhoea in young Gambian children: a limited case control study, *Trans. Royal Soc. Trop. Med. Hyg.*, 78, 209-211, 1984.

Magnani, R., S. Tourkin, and M. Hartz, *Evaluation of the Provincial Water Project in the Philippines*, International Statistics Program Center, Bureau of the Census, U.S. Department of Commerce, Washington, D.C., 1984.

McFeters, G. A., G. K. Bissonnette, J. J. Jezeski, C. A. Thomson, and D. G. Stuart, Comparative survival of indicator bacteria and enteric pathogens in well water, *Appl. Micro.*, 27, 823-829, 1974.

Mertens, T. E., M. A. Fernando, T. F. de C. Marshall, B. R. Kirkwood, S. Cairncross, and A. Radalowicz, Determinants of water quality, availability and use in Kurunegala, Sri Lanka, *Trop. Med. Parasit.*, 41, 89-97, 1990a.

Mertens, T. E., M. A. Fernando, S. N. Cousens, B. R. Kirkwood, T. F. de C. Marshall, and R. G. Feachem, Childhood diarrhoea in Sri Lanka: a case-control study of the impact of improved water sources, *Trop. Med. Parasit.*, 41, 98-104, 1990b.

Oluwande, P. A., Public health implications of chemical and bacteriological characteristics of treated water in Western State of Nigeria, *Nig. Med. J.*, 7, 86-90, 1980.

Pickering, H., Social and environmental factors associated with diarrhoea and growth in young children: Child health in urban Africa, *Soc. Sci. Med.*, 21, 121-127, 1985.

Pinfold, J. V., Faecal contamination of water and fingertip-rinses as a method for evaluating the effect of low-cost water supply and sanitation activities on faeco-oral disease transmission, I. A case study in rural northeast Thailand, *Epid. and Inf.*, 2, 363-375, 1990.

Rajasekaran, P., P. R. Dutt, and K. A. Pisharoti, Impact of water supply on the incidence of diarrhoea and sigellosis among children in rural communities in Madurai, *Ind. J. Med. Res.*, 66, 189-199, 1977.

Shiffman, M. A., R. Schneider, J. M. Faigenblum, R. Helms, and A. Turner, Field studies on water, sanitation and health education in relation to health status in Central America, *Prog. Wat. Tech.*, 11, 143-150, 1978.

Spira, W. M., M. U. Khan, Y. A. Saeed and M. A. Sattar, Microbiological surveillance of intra-neighborhood El Tor cholera transmission in rural Bangladesh, *Bull. WHO*, 58, 731-740, 1980.

Sutton, S., and D. Mubiana, Household water quality in rural Zambia, *Waterlines*, 8, 20-21, 1989.

Tompkins, A. M., B. S. Drasar, A. K. Bradley, and W. A. Williamson, Water supply and nutritional status in rural Northern Nigeria, *Trans. Royal Soc. Trop. Med. Hyg.*, 72, 239-243, 1978.

van Zijl, W. J., Studies on diarrhoeal diseases in seven countries by the WHO Diarrhoeal Diseases Advisory Team, *Bull. WHO*, 35, 249-261, 1966.

Young, B., and J. Briscoe, *Water and Health in Rural Malawi*, U.S. Agency for International Development, Washington, D.C., 1986.

World Health Organization, *Guidelines for Drinking Water Quality, Vol. 1, Recommendations*, Geneva, 1984

Author address list

James VanDerslice
Carolina Population Center
University Square, CB# 8120
University of North Carolina
Chapel Hill, N.C. 27516-3997

Phone: (919) 966-7143
Fax: (919) 966-6638

John Briscoe
The World Bank
S11 - 105
1616 H Street, N.W.
Washington D.C. 20433

Figure legends

Figure 1. Effect of water source fecal coliform dose on the predicted probability of diarrhea

TABLE 1. Means and Standard Deviations of Selected Household and Community Factors

Variable	Mean or proportion	S.D.
Child is male	0.53	
Mother's age	25.89	5.84
Father's age	28.71	6.73
Spouse is present	0.94	
Mother's highest grade completed	7.61	3.30
Father's highest grade completed	7.97	3.41
Household has extended family members	0.42	
Household has electricity	0.60	
Household owns radio	0.55	
Household owns television	0.22	
Household owns refrigerator	0.08	
Household income	231.18	362.56
Value of assets (pesos x 10^{-3})	12.99	51.49
Municipal piped supply is available	0.56	
Boreholes are available	1.00	
Dug wells are available	0.24	
Springs are available	<0.01	
Household has good excreta disposal facility	0.77	
Excreta observed around house	0.33	

Table 1. cont.

Distance to nearest road (m)	95.36	214.20
Population density (pop. x 10^{-3} / km^2)	19.32	20.39
Total rainfall in past 2 weeks (cm)	6.31	4.0
Number of days with rain in past 2 weeks	6.32	2.4

TABLE 2. Water Source Fecal Coliform Concentrations, by Type of Water Source

	# of	$\log_{10}$ concentrations		% of samples with:	
Water source type	samples[a]	Mean	Std. dev.	No colonies	TNTC
Piped supply	111	-2.22	2.90	78	5
Boreholes	403	-1.95	3.00	75	2
Improved dug wells	46	1.95	1.33	9	7
Unimproved dug wells	60	2.55	1.00	3	15

[a] Due to the estimation difficulties arising from the censored observations (i.e., TNTC and zero counts) and the variable number of samples taken from each water source, these statistics are based on one randomly selected sample per water source, except for the piped supply where all samples were used.

TABLE 3. Household Water Storage Practices by Type of Container

Type of container	number & percent	% of containers used for collection	% of containers water scooped	% of containers covered	Mean # of trips per day
Small containers	84 (33%)	86	1	79	3.1
Large containers	78 (31%)	18	0	67	2.5
Clay jars	85 (33%)	0	44	94	2.6
Pails	7 (3%)	25	100	29	9.0
Overall	254(100%)	45	18	79	3.0

TABLE 4. Difference in Fecal Coliform Concentrations between Source and Storage, and Estimated Levels of In-house Contamination

Change in concentration of fecal coliforms	Observed change #	Observed change %	Estimated in-house contamination #	Estimated in-house contamination %
Net decrease:				
< -10,000 (TNTC)	7	3.8		
-1,000 to -10,000	5	2.7		
-1,000 to -100	10	5.4		
-100 to -10	3	1.6		
-10 to -1	4	2.2		
No change: (-1 to 1)	58	31.5	56	30.4
Net increase:				
1 to 10	19	10.3	26	14.1
10 to 100	19	10.3	38	20.7
100 to 1,000	22	12.0	25	13.6
1,000 to 10,000	19	10.3	21	11.4
> 10,000 (TNTC)	18	9.8	18	9.8
TOTAL	184	100.0	184	100.0

TABLE 5. Water Consumption and Water Boiling in Past 24 Hours, by Child's Age

	Child's age (months)					
Variable	2	4	6	8	10	12
Percent fed plain water	38.1	58.2	86.5	95.1	97.0	97.9
Percent fed any water	75.9	82.9	96.1	99.2	99.3	99.7
Among those fed any water:						
Total amount consumed per day (ml):						
Mean	363	408	425	489	558	647
Std. dev.	416	482	482	491	506	509
Percent that boiled water before serving	86.8	78.0	61.4	52.0	49.8	50.6

TABLE 6. Distributions of Predicted Daily Fecal Coliform Doses

Predicted daily	Water source		In-house	
fecal coliform dose	Frequency	Percent	Frequency	Percent
$< 10^{-3}$	941	7.8	0	0.0
10^{-3} to 10^{-2}	3329	27.7	58	0.5
10^{-2} to 10^{-1}	3813	31.8	272	2.3
10^{-1} to 10^{0}	1971	16.4	837	7.0
10^{0} to 10^{1}	701	5.8	1954	16.3
10^{1} to 10^{2}	601	5.0	1980	16.5
10^{2} to 10^{3}	539	4.5	3307	27.6
10^{3} to 10^{4}	96	0.8	2413	20.1
10^{4} to 10^{5}	8	0.1	932	7.8
10^{5} to 10^{6}	0	0.0	176	1.5
$> 10^{6}$	1	0.0	71	0.6
TOTAL	12000	99.9	12000	100.2

TABLE 7. Parameter Estimates and T-Statistics from Probit Models of Diarrheal Disease

	Main effects (Model #1)		Boiling interaction (Model #2)	
Variable	β	t	β	t
Intercept	-1.813	-5.2***[a]	-1.778	-5.0***
Rho	0.121	7.6***	0.121	7.6***
Water contamination:				
Water source $\log_{10}$ FC dose	0.068	2.6***	0.168	2.6***
In-house $\log_{10}$ FC dose	-0.002	-0.1	-0.028	-0.9
Water boiling:[b]				
Main effect	-0.094	-0.5	-0.193	-0.9
Interacted with:				
Water source FC dose			-0.171	-1.7*
In-house FC dose			0.049	1.0
Poor excreta disposal[b]	0.238	1.8*	0.220	1.6*
Excreta around the house[b]	0.364	2.7***	0.365	2.7***
Water source on-site[b]	-0.157	-1.2	-0.106	-0.8
High food pathogenicity[b]	0.718	1.8*	0.736	1.8*
Soap use (mkg/person day)	-0.011	-0.2	-0.006	-0.1
Household density (persons/room)	-0.024	-0.5	-0.018	-0.3
Preventive health care use[b]	-0.125	-0.6	-0.113	-0.6
Breast-feeding and nutritive supplements[b]	-0.028	-0.2	-0.033	-0.2
Breast-feeding and nonnutritive supplements[b]	-0.320	-1.2	-0.360	-1.3

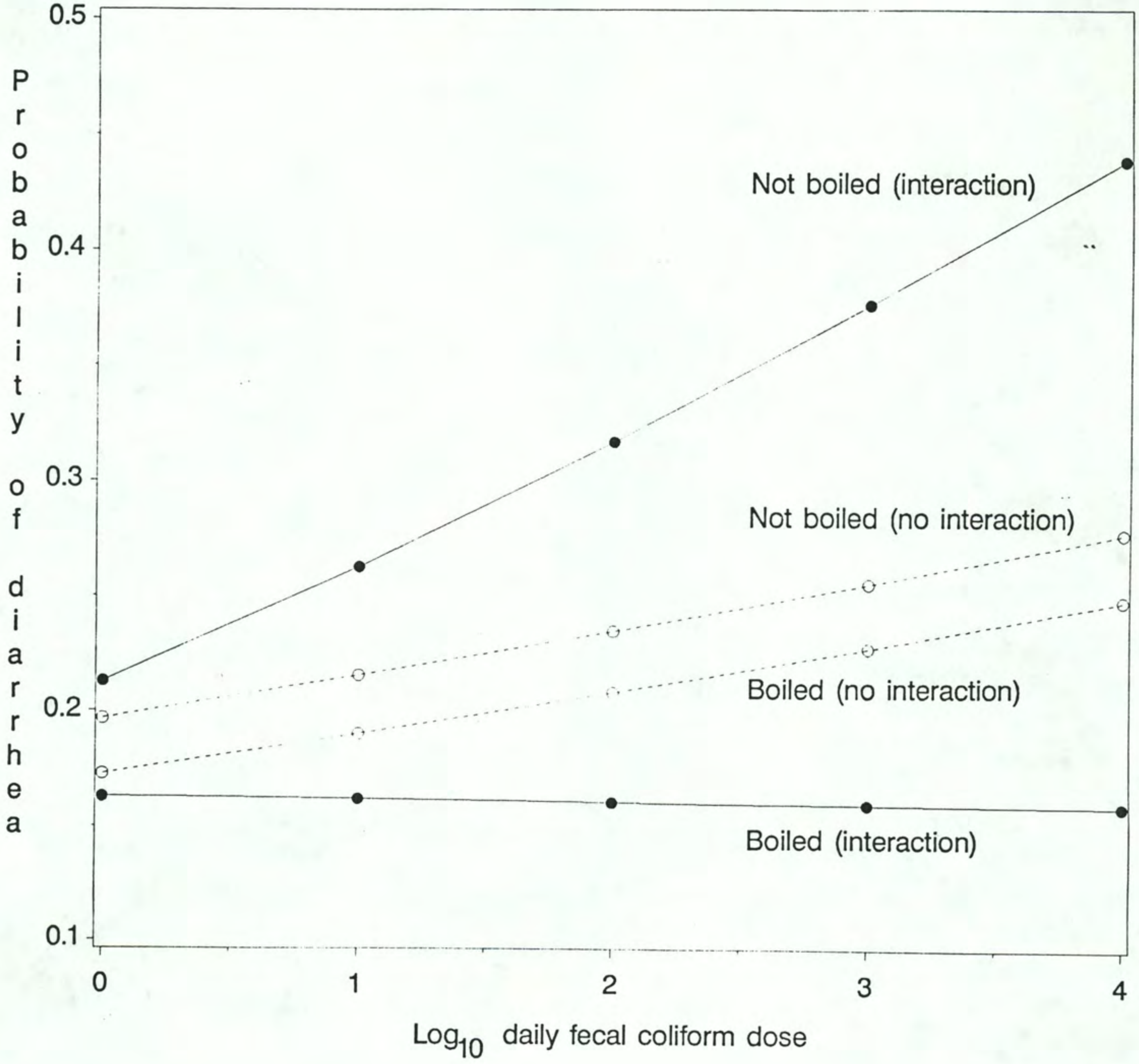
Table 7. cont.

Breast-feeding only[b]	-0.615	-2.2**	-0.560	-2.0**
Standardized weight (std. dev.)	0.003	0.1	0.003	0.1
Male child	0.060	1.4	0.062	1.4
Child's age (weeks)	0.048	6.6***	0.048	6.5***
Child's age squared (weeks^2)	-0.001	-6.1***	-0.001	-6.0***
Community density (10^3 persons/km^2)	0.004	2.8***	0.004	2.7***
Cumulative rainfall (cm)	0.004	1.2	0.004	1.3

---

[a] The stars denote the level of significance for a two-tailed test: *** = $p<0.01$, ** = $p<0.05$, * = $p<0.10$.

[b] These variables are the predicted probability that the child has the stated characteristic.


Probability of diarrhea
0.5
0.4
0.3
0.2
0.1
Not boiled (interaction)
Not boiled (no interaction)
Boiled (no interaction)
Boiled (interaction)
0
1
2
3
4
Log10 daily fecal coliform dose