INTRODUCTION

This past decade has seen a major effort to improve water supply and sanitation 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 in children. Providing a source of high-quality drinking water has traditionally been an important strategy in this effort.

However, families provided high-quality water through hand pumps or standposts commonly store drinking water in the home where it may be contaminated by fecal material. Large increases in indicator organism levels during household storage have been observed in several settings. Furthermore, many studies of the health impacts from improving water source quality have found little or no association with diarrheal disease. In-house contamination of drinking water has often been cited as a likely reason (Ryder et al., 1985; Rahman et al., 1985; Esrey and Habichi, 1986; Lindskog et al., 1987; Huttly et al., 1990). These observations have led to the general 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).

This view, however, may be mistaken, for several reasons. First, it is quite possible that family members would develop some level of immunity to pathogens commonly encountered in the household environment. As such, ingesting pathogens via stored water may be inefficient relative to other household transmission routes, such as person-to-person contact or food contamination. Finally, although fecal coliforms are used as indicators of fecal contamination for water sources and stored drinking water, the pathogens contaminating a family's water source are likely to be quite different than those contaminating its drinking water in the home. As a result, improving water source quality could reduce the risk of diarrhea regardless of the level of in-house water contamination. These arguments are discussed in turn below.

The "Immunity" Argument

Consider a hypothetical person living by himself. He does not practice good hygiene so his drinking water, which was free from fecal coliforms (FC) when it was collected, now contains 1000 FC/100 mL. Does this in-house water contamination increase his risk of diarrhea? Perhaps not. The pathogens contaminating his drinking water during storage may be easily spread from their hands, contaminated during defecation (Han and Hlaing, 1989), to other family members by direct contact, or through contamination of food, utensils or drinking water. Soon other family members may be infected by, and excrete, these same "internal" pathogens.
Such intrahousehold transmission of diarrheal pathogens appears to be quite common in developing countries [de Zoya and Feachem, 1985; Riley et al., 1987].

Subsequent exposure to these internal pathogens would not increase the family's risk of diarrhea if they were already infected by, or immune to, these particular organisms. Previous exposure to a particular enteric pathogen does appear to reduce the risk of diarrhea during subsequent infections [Black et al., 1989, 1982c, 1981b; Ryder et al., 1985; Bishop et al., 1983; Welsh and May, 1979], and such asymptomatic infections are common in developing countries [Black et al., 1989, 1982a, c: Baltazar and Solon, 1989; Saniel et al., 1985; Feachem, 1982]. Thus contamination of stored drinking water by internal pathogens may not increase the family's risk of diarrheal disease.

The "Efficiency" Argument

Even when family members have not developed immunity to a particular internal pathogen, transmission via stored drinking water may be much less efficient than other household transmission routes. Hands, for example, appear to be an important carrier of internal pathogens and several studies have shown hand washing to reduce the spread of diarrheal disease in the home [Khan, 1982; Stanion and Clemens, 1987, Alam et al., 1989]. Contaminated foods are a particularly important transmission route for bacterial pathogens as bacteria may multiply in foods stored without refrigeration [Holmberg et al., 1984; Black et al., 1982b, 1989].

The number of internal pathogens ingested via drinking water, however, may be quite small relative to these other household transmission routes. For instance, contaminated hands may introduce pathogens into stored drinking water, or may pass pathogens directly to other members of the family. Direct transmission via contaminated hands or utensils would be much more efficient than indirect transmission through drinking water. Furthermore, unlike food, drinking water is not an environment conducive to the survival or multiplication of enteric pathogens [Feachem et al., 1983]. Therefore in the presence of other, more efficient household transmission routes, contamination of drinking water during storage may not significantly increase the family's risk of diarrhea [Briscoe, 1984].

The "Different Pathogens" Argument

While fecal coliforms are used as an indicator of fecal contamination for water sources and stored drinking water, the pathogens associated with these indicators are likely to be quite different. Pathogens contaminating drinking water during storage are most likely internal pathogens, originating from the family members' own feces. In contrast, a contaminated water source may well contain pathogens from other people's feces, pathogens which are new to the household environment. It is these "external" pathogens that may initiate new infections which put the family at risk. In fact, a contaminated water source may be a family's only exposure to a particular external pathogen. Providing a high-quality water supply would eliminate this source of external pathogens and reduce the family's risk of diarrhea, regardless of the risk posed by in-house water contamination.

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 water contamination is not an important transmission route, or if the family has developed some level of immunity to these internal pathogens, then in-house water contamination will not be significantly associated with diarrhea. The objective of this study is to estimate and compare the effects of water source and in-house water 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; Melbak et al., 1989; Morin et al., 1990; 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 [Lindsaykog and Lindsaykog, 1988], and Sri Lanka [Mertens et al., 1990*] 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, Echeverrra et al. [1987] sampled for enterotoxigenic Escherichia coli (ETEC) in Thai neighborhoods where ETEC-positive diarrhea cases lived. 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, isolated Strongyloides 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.

There is very little research on the health effects of in-house water contamination. Two studies which investigated the relationship between drinking water storage and diarrheal disease found conflicting results. A cross-sectional study conducted in Nigeria [Hutton et al., 1987] found no association between open drinking water storage containers and the prevalence of diarrhea for any age group. However, in a peri-urban area of Lima, Peru, use of a storage container that was not fitted with a faucet was significantly associated with a higher incidence of diarrhea in children under 3 years of age, even though 98% of the families reported boiling their drinking water [Yegar et al., 1991].

A few studies have used the quality of water in the storage container as a measure of exposure to waterborne pathogens [Mognani et al., 1984; Esrey et al., 1986; Henry and Rahim, 1990; Merrens et al., 1990a]. While none of these studies found any significant relationship with diarrheal disease, contaminated water in the storage container was significantly associated with poor growth in children [Mognani et al., 1984] and with the presence of enteric pathogens in children's stools [Esrey et al., 1986]. However, stored water quality is a poor measure of in-house water contamination since it also reflects contamination of the water source. As such, it is not clear whether the observed effects were due to contamination during storage or contamination of the source.

**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 studying 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 hygiene conditions. Bimonthly interviews, conducted through the first 2 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, 1991].

**Water Sampling**

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 a measure 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 [Haas et al., 1988]. They have been isolated in areas thought to be devoid of fecal material [Fujimori 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 [World Health Organization, 1984]. More importantly, concern about in-house water contamination has been precipitated by the observed increases in fecal coliform levels, not increases in the concentration of pathogens.

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. Blanks were used to detect contamination occurring in the laboratory. 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" (10-100), 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 cannot 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, are a mixture of those introduced by contaminated hands or cups during storage, and those which originally came from the water source. Thus the concentration of FC due to in-house contamination (C_H) is the FC...
concentration observed in the storage vessel \((C_H)\) minus the concentration of FC in the container which originated from the water source \((C_0)\):

\[
C_H = C_0 - C_d
\]  

(1)

Standard bacteriological methods can not differentiate between fecal coliforms from in-house contamination and those from a contaminated water source. 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 regrowth of fecal coliforms has been observed in nutrient-rich surface waters [Hendricks, 1972; Kinney et al., 1978; Carillo et al., 1985], regrowth 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 \((C_s)\) minus the concentration that died during storage \((C_d)\):

\[
C_i = C_s - C_d
\]  

(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 \((C_{H\text{min}})\) occurs when all water source bacteria are assumed to survive (i.e., \(C_i = C_s\)):

\[
C_{H\text{min}} = C_s - C_w, \quad C_i > C_w, \quad \text{(net increase)}
\]  

(3)

When the FC concentration in the container is less than that found at the water source, then the minimum value for in-house contamination was set to 0.9 FC/100 mL, the lower limit of detection:

\[
C_{H\text{min}} = 0.9, \quad C_i < C_w, \quad \text{(net decrease)}
\]  

(4)

The largest possible value for in-house contamination \((C_{H\text{max}})\) occurs when none of the water source fecal coliforms are assumed to survive (i.e., \(C_i = 0\)). In this case all the fecal coliforms in the storage container are assumed to be due to in-house contamination:

\[
C_{H\text{max}} = C_0 - 0
\]  

(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, in-house contamination was estimated as the midpoint between the minimum and maximum values measured on a log scale:

\[
\log_{10} \left( C_H \right) = \frac{1}{2} \left[ \log_{10} \left( C_{H\text{max}} \right) - \log_{10} \left( C_{H\text{min}} \right) \right]
\]  

(6)

In many cases the minimum and maximum values for \(\log_{10} \left( C_H \right)\) 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 \(\log_{10} \left( C_{H\text{max}} \right)\) and \(\log_{10} \left( C_{H\text{min}} \right)\) 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, 1991, 1992]. In this model diarrhea \((D)\) results from past growth \((G)\), behavioral factors \((Y)\), and "underlying" socioeconomic and environmental factors \((Z)\):

\[
D_{i,t} = \beta_1 G_{i,t-1} + \beta_2 Y_{i,t-1} + \beta_3 Z_{i,t} + \mu_i + \varepsilon_{i,t}
\]  

(7)

for \(t = 1-6\) 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:

\[
Y_{i,t} = \alpha_1 G_{i,t-1} + \alpha_2 Y_{i,t-1} + \alpha_3 Z_{i,t} + \mu_y + \eta_{i,t}
\]  

(8)

These equations contain two error terms. The first error term, \(\mu\), 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 \(\mu\) are different for each equation but correlated across equations. The second error term, \(\varepsilon\), 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. While recall data is always subject to error, periods of up to 2 weeks provide morbidity information with adequate accuracy [Black, 1984].

Behavorial 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 2 weeks prior to
the interview date. In-house water contamination is mea-
sured by the logarithm of the 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 concen-
tration due to in-house contamination.

Water boiling is expected to reduce the risk of diarrhea
due to contaminated water to the same extent that contam-
inated 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 fieldwork-
ers.

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 nonlaun-
dry soap usage, estimated from reported household expendi-
tures 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 con-
tamination was constructed from food preparation and stor-
age practices at each longitudinal survey.

Breast feeding may reduce the child's susceptibility to
infection via maternal antibodies [Welsh and May, 1979] and
provides nourishment which is free from contamination.
Feeding patterns are measured by three dichotomous vari-
ables 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
breast milk. 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 2 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 be-
havioral variables. Age squared is included to capture non-
linearities. 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 2 weeks is used to model these effects.
Finally, community density is included as high-density areas
are characterized by higher levels of environmental contam-
ination.

Estimation Methods

In (7), diarrhea in the present time period is specified as a
function of past growth, past behaviors, current socioeco-

conic and environmental conditions, and two error terms.
This model differs from traditional research in two important
ways. First, the model explicitly acknowledges that behav-
iors are determined in part by the child's health (equation
(8)). Second, 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, 1990].

Failure to account for these effects can lead to biased
parameter estimates and spurious results [Cebu Study Team, 1991, 1992; 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 \( \mu_y \), the random error representing unobserved
differences between children. Since the unobserved differ-
ences which affect behaviors may also affect diarrhea, \( \mu_y \) is
 correlates with \( \mu_D \), the equivalent disturbance term in the
diarrhea equation. As a result, the behavioral variables in the
diarrhea equation are correlated with the error term \( \mu_D \).
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 the unobserved heterogeneity
\( \mu_D \) 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, \( \mu_D \). 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, (8) becomes

\[
Y_{it} = \delta_1 G_{i,t-2} + \delta_2 D_{i,t-2} + \delta_3 Y_{i,t-2} +
\delta_4 Z_{i,t-1} + \mu_y + \epsilon_i
\]  

(9)
The same variables \((i = 2)\) can be substituted out again.
This process is repeated until only the underlying variables
and error terms remain:

\[
Y_{i,t} = \gamma_1 Z_{i,t-1} + \gamma_2 Z_{i,t-2} + \cdots + \gamma_r Z_{i,t-r} + \mu_y + \epsilon_i
\]  

(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, $\mu_i$. As such, (10) can be estimated using standard techniques. Predicted values for the behavior variable, $Y_{i,k}$, 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 $\mu_i$.

For example, predicted in-house contamination levels were derived as a function of socioeconomic characteristics of the household, education of the mother and father, availability of different types of water sources, community density and level of modernization, prices, and rainfall. The same procedure was used to generate predicted values for the other behavioral determinants and the growth variable. Using these predicted values in place of the actual values 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 HOTZ-TRANS 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]

$$\partial P(D = 1)/\partial X_k = \phi(X_k)\beta_k$$  \hspace{1cm} (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:

lower 95% confidence limit $= \phi(X_k)[\beta_k - 1.96(s.e.\beta)]$  \hspace{1cm} (12)

upper 95% confidence limit $= \phi(X_k)[\beta_k + 1.96(s.e.\beta)]$  \hspace{1cm} (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/week with a median of 200 (approximately U.S. $10). Total household assets range from 0 to almost 1.5 million pesos with a median of 2400 (approximately 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 min of their water source. Only a small proportion (5%) must walk more than 5 min 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

### Table 1. Means and Standard Deviations of Selected Household and Community Factors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean or Proportion</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child is male</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Mother's age</td>
<td>25.89</td>
<td>3.84</td>
</tr>
<tr>
<td>Father's age</td>
<td>28.71</td>
<td>6.73</td>
</tr>
<tr>
<td>Spouse is present</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>Mother's highest grade completed</td>
<td>7.61</td>
<td>3.30</td>
</tr>
<tr>
<td>Father's highest grade completed</td>
<td>7.97</td>
<td>3.41</td>
</tr>
<tr>
<td>Household has extended family members</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Household has electricity</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Household owns radio</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Household owns television</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>Household owns refrigerator</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Household income</td>
<td>231.18</td>
<td>362.56</td>
</tr>
<tr>
<td>Value of assets (pesos x 10⁻³)</td>
<td>12.99</td>
<td>39.49</td>
</tr>
<tr>
<td>Municipal piped supply is available</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>Boreholes are available</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Dug wells are available</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Springs are available</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>Household has good excreta disposal facility</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Excreta observed around house</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Distance to nearest road (m)</td>
<td>95.26</td>
<td>214.20</td>
</tr>
<tr>
<td>Population density</td>
<td>19.32</td>
<td>30.39</td>
</tr>
<tr>
<td>Total rainfall in past 2 weeks (cm)</td>
<td>6.31</td>
<td>4.0</td>
</tr>
<tr>
<td>Number of days with rain in past 2 weeks</td>
<td>6.32</td>
<td>2.4</td>
</tr>
</tbody>
</table>
TABLE 2. Water Source Fecal Coliform Concentrations by Type of Water Source

<table>
<thead>
<tr>
<th>Water Source Type</th>
<th>No. of Samples</th>
<th>Log Concentrations</th>
<th>Percentage of Samples With</th>
<th>No Colonies</th>
<th>TNTC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piped supply</td>
<td>111</td>
<td>-2.22</td>
<td>2.90</td>
<td>78</td>
<td>5</td>
</tr>
<tr>
<td>Boreholes</td>
<td>403</td>
<td>-1.95</td>
<td>3.00</td>
<td>75</td>
<td>2</td>
</tr>
<tr>
<td>Improved dug wells</td>
<td>46</td>
<td>1.95</td>
<td>1.33</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Unimproved dug wells</td>
<td>66</td>
<td>2.55</td>
<td>1.00</td>
<td>3</td>
<td>15</td>
</tr>
</tbody>
</table>

*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.

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 (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 demonstrated a net decrease in FC concentration, indicating that bacterial die-off can be greater than increases due to in-house contamination. The estimated levels of in-house contamination calculated using (6) are presented in the second set of columns of Table 4. Just over 30% of the sample households were estimated as having very little in-house contamination, while 21% were thought to have considerable contamination during storage (i.e., >1000 FC/100 mL).

Covering the storage container appears to have little effect on in-house contamination. The geometric mean (GM) 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 (GM = 1.96) than samples from containers where water was poured or flowed through a spigot (GM = 1.48). Small storage containers were subject to less in-house contamination (GM = 1.16) than the large containers (GM = 1.79) or earthen storage jars (GM = 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

TABLE 3. Household Water Storage Practices by Type of Container

<table>
<thead>
<tr>
<th>Type of Container</th>
<th>No. and Percent</th>
<th>Used for Collection</th>
<th>Cover</th>
<th>Mean No. of Trips per Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small containers</td>
<td>84 (33%)</td>
<td>86</td>
<td>1</td>
<td>79</td>
</tr>
<tr>
<td>Large containers</td>
<td>78 (31%)</td>
<td>18</td>
<td>0</td>
<td>67</td>
</tr>
<tr>
<td>Clay jars</td>
<td>85 (33%)</td>
<td>0</td>
<td>44</td>
<td>94</td>
</tr>
<tr>
<td>Pails</td>
<td>7 (3%)</td>
<td>25</td>
<td>100</td>
<td>29</td>
</tr>
<tr>
<td>Overall</td>
<td>254 (100%)</td>
<td>45</td>
<td>18</td>
<td>79</td>
</tr>
</tbody>
</table>

FC/100 mL) less often than open dug wells (41% versus 78%).
TABLE 4. Difference in Fecal Coliform Concentrations Between Source and Storage and Estimated Levels of In-House Contamination

<table>
<thead>
<tr>
<th>Change in Concentration of Fecal Coliforms</th>
<th>Observed Change</th>
<th>Estimated In-House Contamination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net decrease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; -10,000 (TNTC)</td>
<td>7</td>
<td>3.8</td>
</tr>
<tr>
<td>-1,000 to -10,000</td>
<td>5</td>
<td>2.7</td>
</tr>
<tr>
<td>-100 to -100</td>
<td>10</td>
<td>5.4</td>
</tr>
<tr>
<td>-10 to 1</td>
<td>3</td>
<td>1.6</td>
</tr>
<tr>
<td>No change</td>
<td>4</td>
<td>2.2</td>
</tr>
<tr>
<td>Net increase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 to 10</td>
<td>19</td>
<td>10.3</td>
</tr>
<tr>
<td>10 to 100</td>
<td>19</td>
<td>10.3</td>
</tr>
<tr>
<td>100 to 1,000</td>
<td>22</td>
<td>12.0</td>
</tr>
<tr>
<td>1,000 to 10,000</td>
<td>19</td>
<td>10.3</td>
</tr>
<tr>
<td>&gt; 10,000 (TNTC)</td>
<td>18</td>
<td>9.8</td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>100.0</td>
</tr>
</tbody>
</table>

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 six time periods combined are presented in Table 6. The low doses were censored to 1 fecal coliform/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 1 year of age, female children were experiencing about 20% fewer cases of diarrhea than male children.

TABLE 5. Distributions of Predicted Daily Fecal Coliform Doses

<table>
<thead>
<tr>
<th>Predicted Daily Fecal Coliform Dose</th>
<th>Water Source Frequency</th>
<th>Percent</th>
<th>In-House Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 10^-1</td>
<td>941</td>
<td>7.8</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>10^-1 to 10^-2</td>
<td>3329</td>
<td>27.7</td>
<td>58</td>
<td>0.5</td>
</tr>
<tr>
<td>10^-2 to 10^-1</td>
<td>3813</td>
<td>31.8</td>
<td>272</td>
<td>2.3</td>
</tr>
<tr>
<td>10^-3 to 10^-9</td>
<td>971</td>
<td>16.4</td>
<td>837</td>
<td>7.0</td>
</tr>
<tr>
<td>10^-4 to 10^-1</td>
<td>701</td>
<td>5.8</td>
<td>1954</td>
<td>16.3</td>
</tr>
<tr>
<td>10^-5 to 10^-4</td>
<td>601</td>
<td>5.0</td>
<td>1980</td>
<td>16.5</td>
</tr>
<tr>
<td>10^-6 to 10^-5</td>
<td>139</td>
<td>4.5</td>
<td>3307</td>
<td>27.6</td>
</tr>
<tr>
<td>10^-7 to 10^-6</td>
<td>96</td>
<td>0.8</td>
<td>2413</td>
<td>20.1</td>
</tr>
<tr>
<td>10^-8 to 10^-7</td>
<td>9</td>
<td>0.1</td>
<td>932</td>
<td>7.8</td>
</tr>
<tr>
<td>10^-9 to 10^-8</td>
<td>0</td>
<td>0.0</td>
<td>176</td>
<td>1.5</td>
</tr>
<tr>
<td>&gt; 10^-9</td>
<td>1</td>
<td>0.0</td>
<td>71</td>
<td>0.6</td>
</tr>
<tr>
<td>Total</td>
<td>12,000</td>
<td>99.9</td>
<td>12,000</td>
<td>100.2</td>
</tr>
</tbody>
</table>

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, water source dose is a very strong risk factor (t = 2.6). Increasing the dose from 1 to 100 FC/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...
<table>
<thead>
<tr>
<th>Variable</th>
<th>Main Effects, Model 1</th>
<th>Boiling Interaction, Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$</td>
<td>t</td>
</tr>
<tr>
<td>Intercept</td>
<td>-1.813</td>
<td>-5.2*</td>
</tr>
<tr>
<td>Rho</td>
<td>0.121</td>
<td>7.6*</td>
</tr>
<tr>
<td>Water contamination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water source $\log_{10}$ FC dose</td>
<td>0.068</td>
<td>2.6*</td>
</tr>
<tr>
<td>In-house $\log_{10}$ FC dose</td>
<td>-0.002</td>
<td>-0.1</td>
</tr>
<tr>
<td>Water boiling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect</td>
<td>-0.094</td>
<td>-0.5</td>
</tr>
<tr>
<td>Interacted with</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water source FC dose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-house FC dose</td>
<td>0.236</td>
<td>1.8b</td>
</tr>
<tr>
<td>Excreta around the house $^c$</td>
<td>0.364</td>
<td>2.7a</td>
</tr>
<tr>
<td>Water source on-site $^c$</td>
<td>-0.157</td>
<td>-1.2</td>
</tr>
<tr>
<td>High food pathogenicity $^c$</td>
<td>0.718</td>
<td>1.8b</td>
</tr>
<tr>
<td>Soap use (mg/person/day)</td>
<td>-0.011</td>
<td>-0.2</td>
</tr>
<tr>
<td>Household density (persons/room)</td>
<td>-0.024</td>
<td>-0.5</td>
</tr>
<tr>
<td>Preventive health care use $^c$</td>
<td>-0.125</td>
<td>-0.6</td>
</tr>
<tr>
<td>Breast feeding and nutritive supplements $^d$</td>
<td>-0.028</td>
<td>-0.2</td>
</tr>
<tr>
<td>Breast feeding and nonnutritive supplements $^d$</td>
<td>-0.320</td>
<td>-1.2</td>
</tr>
<tr>
<td>Breast feeding only $^e$</td>
<td>-0.615</td>
<td>-2.2*</td>
</tr>
<tr>
<td>Standardized weight (S.D.)</td>
<td>0.003</td>
<td>0.1</td>
</tr>
<tr>
<td>Male child</td>
<td>0.060</td>
<td>1.4</td>
</tr>
<tr>
<td>Child's age (weeks)</td>
<td>0.048</td>
<td>6.6a</td>
</tr>
<tr>
<td>Child's age squared (weeks$^2$)</td>
<td>-0.001</td>
<td>-6.1c</td>
</tr>
<tr>
<td>Community density (10$^9$ persons/km$^2$)</td>
<td>0.004</td>
<td>2.8a</td>
</tr>
<tr>
<td>Cumulative rainfall (cm)</td>
<td>0.004</td>
<td>1.2</td>
</tr>
</tbody>
</table>

*Here $p < 0.01$ (two-tailed test).

$^b$Here $p < 0.10$ (two-tailed test).

$^c$These variables are the predicted probability that the child has the stated characteristic.

$^d$Here $p < 0.05$ (two-tailed test).

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. While there is considerable uncertainty associated with the predicted in-house contamination levels due to the relatively small number of matched source and household water samples, the predicted values are consistent (unbiased). As a result, the parameter estimate for in-house water contamination in the diarrhea model is also consistent. The uncertainty increases the estimated standard error of this parameter estimate, reducing its significance. However, since the parameter estimate for in-house contami-
greater than 1000/day, and a variable indicating that in-house coliforms would lead to a 17% increase in diarrheal prevalence. A tenfold increase in the concentration of fecal coliforms was associated with diarrhea (results not presented). Neither of these variables were associated with diarrhea (results not presented).

Water source contamination, however, poses a significant risk for diarrhea. When water is not boiled, contamination of the water source substantially increases the probability of diarrhea. A tenfold 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 in this population. The relative importance of water contamination, sanitation, and the level of water service as well as the effects of multiple interventions will be addressed in a forthcoming paper.

It is clear that improving water source quality can substantially reduce diarrheal disease in spite of contamination occurring in the home. Three plausible explanations for this result were presented in the introduction.

The Efficiency Argument

The lack of association between in-house water contamination and diarrheal disease may indicate that transmission of internal pathogens via drinking water is much less efficient than transmission via contaminated hands or foods. While hand washing was not observed in this study, poor food hygiene was a marginally significant risk factor for diarrhea.

In this situation, measures to protect drinking water quality in the home without improvements in household hygiene may do little to reduce pathogen transmission between family members. Improvements in hygiene and hand washing, however, may significantly reduce all household transmission routes, including transmission via in-house water contamination. Thus changes in drinking water quality during household storage may be more an indicator of household hygiene than a risk factor for diarrheal disease [Pinfold, 1990].

The Different Pathogens Argument

In this study, a water source containing 100 fecal coliforms per 100 mL poses a substantial risk of diarrhea, while an equivalent increase in FC levels during household storage presents little or no risk. This implies that the pathogens associated with these indicators are quite different. The strong relationship between water source contamination and diarrheal disease suggests that water sources are a principal means by which new external pathogens are introduced into the household environment.

The lack of association between increases in fecal coliform levels during storage and diarrhea implies that internal pathogens contaminating drinking water pose little risk, or that the fecal coliforms are indigenous to the household environment (e.g., soil) and do not indicate fecal contamination. This explanation is consistent with two stud-
ies which found no evidence of in-house water contamination by bacterial pathogens [Spira et al., 1980; Eschereria et al., 1987], but contrary to one study which documented in-house water contamination by enteric parasites [Khairy et al., 1982].

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 water 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

$C_d$: 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.

$C_c$: 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.

$GM$: geometric mean.

OLS: ordinary least squares.

$s.e.$: standard error of the estimate of $\beta$.

$TNTC$: too numerous to count.

$Y$: behavioral factors.

$Z$: underlying socioeconomic and environmental factors.

$\phi$: normal density function.

$\mu_D$: individual-specific random error affecting diarrhea.

$\mu_Y$: individual-specific random error affecting behaviors, $Y$.


Cebu Study Team. A child health production function estimated from longitudinal data. J. Dev. Econ. 34, 133–151. 1992.


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.


J. VanDerslice, Carolina Population Center, University Square, CB 8120, University of North Carolina, Chapel Hill, NC 27516.

(Received April 4, 1991; revised January 22, 1992; accepted February 27, 1992.)