Socioeconomic Determinants of Disease Transmission in Cambodia

Anil B. Deolalikar and Ramanan Laxminarayan

July 2000 • Discussion Paper 00–32

RESOURCES
FOR THE FUTURE

Resources for the Future
1616 P Street, NW
Washington, D.C. 20036
Telephone: 202–328–5000
Fax: 202–939–3460
Internet: http://www.rff.org

Socioeconomic Determinants of Disease Transmission in Cambodia

Anil B. Deolalikar and Ramanan Laxminarayan

Abstract

The process of acquiring an infection has two components: first, exposure through proximity to another infected individual, and second, transmission of the disease. Earlier studies of the socioeconomic factors that affect the probability of acquiring an illness assume uniform exposure to infected individuals and may therefore result in biased estimates. This paper develops an empirical model, consistent with epidemiological models of spread of infections, to estimate the impact of socioeconomic variables on the extent of disease transmission within villages in Cambodia. Data from the 1997 *Cambodia Socioeconomic Survey* are used in this analysis.

Key Words: disease transmission, sanitation, infection, public health

JEL Classification Numbers: C1, I1, O1

Contents

Socioeconomic Determinants of Disease Transmission in Cambodia

Anil B. Deolalikar and Ramanan Laxminarayan*

1. INTRODUCTION

Developing countries bear a disproportionately large burden of infectious diseases in the world. In 1996, they accounted for over 90% of all cases of infectious diseases reported worldwide (World Resources Institute 1998). In the same year, these diseases were the leading cause of mortality in Asia, accounting for roughly 50% of all deaths. Although the broad impact of infectious disease on economic development and prosperity in developing countries has been widely recognized, and the mechanisms by which *epidemic* transmission of diseases occurs are well understood, there are many gaps in our understanding of the *socioeconomic* determinants of disease transmission.

A large body of economics literature focuses on the determinants of individual and household morbidity (e.g., Behrman and Deolalikar 1988; Strauss and Thomas 1995). However, there are hardly any studies in the social sciences that have looked at the socioeconomic and demographic factors that promote or inhibit the spread of diseases among individuals within a household or a community. One study that used data from Jakarta, Indonesia, found an association between socioeconomic, behavioral, and community-level variables and the incidence of diarrhea; it concluded that the lack of defensive behavior such as boiling water and washing hands frequently were significant predictors of diarrhea (Alberini et al. 1995).

In general, most previous studies have examined the impact of a given set of socioeconomic factors on the probability of an individual falling ill. However, there is some evidence that even in communities that are both culturally and economically homogenous, the burden of disease falls disproportionately on a few households—an empirical fact that is not predicted by just water or sanitation variables alone. For instance, evidence from Zimbabwe has shown that a small fraction of residents in villages make a large number of visits to natural water bodies and so are heavily exposed to schistosome infection (Chandiwana and Woolhouse 1991; Woolhouse 1994).

* Anil B. Deolalikar (anil@u.washington.edu) and Ramanan Laxminarayan (ramanan@rff.org) are from the University of Washington and Resources for the Future, respectively.

From the perspective of infectious-disease epidemiology, two pieces of information are needed in order to determine the risk of an individual being infected with a communicable disease, such as pneumonia. First, how many other individuals who are in proximity to this individual suffer from pneumonia? Second, how strong are the transmission mechanisms that pass on the disease from the infected individuals to uninfected individuals? Current studies seek to identify a broad set of socioeconomic risk factors for infectious disease, and implicitly assume uniform exposure to disease. However, these studies may fail to include information on prevailing levels of infection, and are, therefore, susceptible to omitted variable bias.

Most common infectious disease pathogens are either water-borne or food-borne. Other pathogens are transmitted by physical or sexual contact. Finally, pathogens of diseases such as malaria are transmitted through insect vectors. Regardless of the mechanism of disease transmission, it is likely that geographical proximity to an infected individual significantly increases the risk of acquiring infection. Consequently, there is a need for research that characterizes the socioeconomic factors that affect this likelihood of disease transmission between members of the same household or community. From the policy standpoint, we can frame the question as one that contemplates the relative importance of individual, household, and community determinants of the transmission process.

Given the overwhelmingly large burden of infectious disease in Cambodia, the formulation of successful public health policies focusing on prevention requires an understanding of the socioeconomic determinants of disease and illness transmission within households and communities. Our hypothesis in this paper is that the likelihood of an individual being ill during a reference period is positively associated with the number of individuals in his/her household and village that are ill, even after controlling for the individual’s characteristics and socioeconomic status (e.g., age, sex, income, and education). In addition, the rate at which other individuals’ illnesses influence the likelihood of a given individual being ill (vis à vis, the rate of disease or illness transmission) depends on socioeconomic and demographic characteristics of the individual and his/her family as well as on the health infrastructure in a community. For example, better-educated individuals are more likely to wash their hands before cooking or eating, and this in turn is expected to lower the probability of catching an infection from others in the same household or community. Likewise, in communities that have primary health clinics and community health workers, individuals are more likely to practice good preventative health and hygiene practices, which would reduce the likelihood of disease and illness transmission.

In this paper, we use 1997 household survey data from Cambodia to estimate the magnitude and determinants of intra-household and intra-community disease transmission. Cambodia is an ideal country in which to study this problem because of its high levels of morbidity and the highly communicable nature of this morbidity. More specifically, we estimate a reduced-form model of individual morbidity that is conditional on the number of other ill individuals in the household or the community. Since the morbidity of others is a potentially endogenous variable, we use instrumental-variable estimation methods to estimate the conditional demand equations. Within this model, we also explicitly allow socioeconomic and demographic characteristics of individuals, as well as community health infrastructure, to influence the rate at which other individuals' illnesses affect the morbidity of a given individual. Section 2 sets out the framework for this analysis. Section 3 contains a description of the data set and Section 4 provides a discussion of our estimates. Section 5 concludes the paper.

2. ANALYTICAL CONSIDERATIONS

Since the theory of health-care demand care is well-developed,[1] there is little need here to develop an elaborate model of individual health determination. If it is assumed that individuals maximize a utility function using health status and other consumption as its argument, subject to a budget constraint and a health production function that includes medical care as inputs, the resulting reduced-form demand function for health care and the derived demand function for health status will include, as their arguments, health care prices, household income, and socioeconomic and demographic characteristics of the individual and the household.

This model, however, does not explicitly recognize the communicable nature of most infections. In particular, the likelihood of an individual being ill is a direct function of his or her exposure to infections—in other words, the number of other individuals in the household or community who are ill. This model of disease transmission is based on the Kermack-McKendrick model (Kermack and McKendrick 1927) where

$$\dot{I} = \beta I(1 - I) - g(w)I \tag{1}$$

where I represents the fraction of the population that is infected; β is the transmission coefficient of illness and a function of the virulence of the disease, and of hygienic measures undertaken by

[1] See Grossman (1972). Behrman and Deolalikar (1988) and Strauss and Thomas (1995) discuss generic models of health determination for less-developed countries.

the individual and/or the household and w is the rate of recovery from illness that is, in turn, a function of medical care. The function, $g(w)$, represents the health production function. The conditional probability of acquiring an infection (or illness) can be written as

$$P(Ill) = f(I, w) \tag{2}$$

Combining equation (2) with the usual reduced-form derived demand function for health status, we can write the probability of an individual acquiring an infection, conditional on the number of other individuals ill, as

$$P(Ill_{ihc}) = \frac{e^{aX_{ihc}+bY_{hc}+cZ_c+dN_{ihc}+\mu_{ihc}}}{1+e^{aX_{ihc}+bY_{hc}+cZ_c+dN_{ihc}+\mu_{ihc}}} \tag{3}$$

and

$$N_{ihc} = \sum_i Ill_{-i,hc}$$

where

Ill = a reported illness episode in a two-week reference period,

X = vector of individual characteristics influencing illness (e.g., age, sex, education),

Y = vector of household characteristics (e.g., income, age of household head),

Z = vector of community characteristics (e.g., presence of clinic, village population),

N = number of individuals other than the reference individual who reported to be ill in the village during the two-week reference period,

μ = error term, and

i, *h*, and *c* index individuals, households, and communities, respectively.

The coefficient *d* in equation (3) indicates the extent to which infections and illnesses are communicable. *A priori*, one would expect *d* to be positive. We further assume that *d* is not a constant coefficient, but instead varies systematically with certain individual, household, and community characteristics (such as age, education, and availability of health facilities), such that

$$d_{ihc} = d_0 X_{ihc} + d_1 Y_{hc} + d_2 Z_c \tag{4}$$

Substituting the value of d from relation (4) into relation (3) gives us:

$$P(Ill_{ihc}) = \frac{e^{aX_{ihc}+bY_{hc}+cZ_c+d_0N_{ihc}+d_1X_{ihc}N_{ihc}+d_2Y_{hc}N_{ihc}+d_3Z_cN_{ihc}+\mu_{ihc}}}{1+e^{aX_{ihc}+bY_{hc}+cZ_c+d_0N_{ihc}+d_1X_{ihc}N_{ihc}+d_2Y_{hc}N_{ihc}+d_3Z_cN_{ihc}+\mu_{ihc}}}, \tag{5}$$

which is the estimating equation in this paper.

The only problem with estimating equation (5) is that the number of other individuals who are ill, *N*, is a potentially endogenous variable. In that case, estimates of (5) that do not treat *N* as an endogenous variable will likely be biased. To tackle this simultaneity problem, we use instrumental-variable estimates, with the mean characteristics of all individuals other than the reference individual (i.e., $\Sigma_i X_{-i,hc}$) serving as identifying instruments for *N*.

3. DATA AND EMPIRICAL MODEL

This paper uses data from the 1997 *Cambodia Socioeconomic Survey* (CSES), which is the first nationally-representative survey of households in that country. The CSES was administered in June 1997 to 6,010 randomly selected households, in a stratified sample of 474 randomly selected villages, in 20 of Cambodia's provinces.[2] In addition to the household survey, a village-level survey was conducted simultaneously. The household survey obtained information on the demographic characteristics of household members, current and previous school enrollments, employment and earnings, morbidity and health care utilization, housing characteristics, household consumption expenditures, and the ownership of durables. The village questionnaire collected information on land use and access to community and social services (e.g., roads, electricity, markets, schools, health facilities).

The analysis in this paper is based on 29,345 individuals whose survey data on any of the dependent or independent variables were not missing. The dependent variable is the occurrence of an illness episode during the two weeks preceding the survey, as reported by survey respondents. While a more objective measure of illness or infection would have been desirable, the only morbidity ever collected in the vast majority of household health surveys in developing countries is via respondent recall. Since there are often major differences between child and adult morbidity, we estimate separate equations for children (ages 16 years and below) and adults. But we also present estimates for a pooled sample of adults and children.

The independent variables include individual-specific variables, such as age, age squared, sex, and completed schooling years; and household-level variables, such as log consumption expenditure per capita (as a proxy for permanent income), lack of a toilet in the house, and the

[2] Excluded provinces were Mondul Kiri (included in the sampling frame, but not represented in the randomly selected sample), Preah Vihear, and Oddar Meanchey.

sex and completed schooling years of the household head. Since school enrollment and attendance decisions for children are likely to be endogenous with respect to morbidity, the completed schooling variable is excluded from the illness equation for children. In the pooled adult-child equation, the completed schooling variable is included in the form of an interaction with an adult dummy variable (taking on the value of one if an individual is greater than 16 years of age and zero otherwise). This effectively restricts completed schooling to have an effect on only adult—not child—morbidity.

Three community-level variables are also included as independent variables in the model. These are urban/rural residence, the availability of a doctor in the village, and the availability of safe drinking water in the village. While there are many other community-health infrastructure variables in the data set, most of them are highly collinear with the doctor and drinking water availability variables.

In addition, as indicated in equation (5), the model includes the number of individuals in the village *other than the reference individual* who reported being ill during the same two-week reference period, as well as the interactions of this variable with most of the individual, household, and community variables listed above. This set of variables effectively measures the magnitude and determinants of the infection transmission effect. As noted earlier, the number of other ill individuals is a potentially endogenous variable. In particular, if there are unobserved factors or endowments in the community that result in poor health among all individuals (i.e., the reference individual as well as all other individuals residing in the community), then standard estimates of the coefficient on the “other ill individuals” variable (vis à vis, the disease transmission effect) will be biased upwards.

We tackle this problem by using an instrumental variable for the “other ill individuals” variable (and the interaction of this variable with other characteristics). The mean characteristics of all individuals in the village other than the reference individual (e.g., mean age, mean completed schooling years, and mean sex) serve as identifying instruments for the “other ill individuals” variable and its interactions.

As reported illness is a dichotomous variable, the model in equation (5) is estimated by the maximum-likelihood logit method. The reported standard errors are corrected for a general, unknown form of heteroscedasticity using the Huber-White method (White 1978).

Table 1 reports the percentage of individuals reporting an illness episode in the four weeks preceding the survey, disaggregated by per capita expenditure quintile (using a proxy for socioeconomic status) and rural/urban residence. At about 15% of the population, the proportion of respondents reporting an illness episode in the four-week reference period is comparable to other countries. If annualized, it suggests an illness prevalence rate of about 1.8 episodes per person per year.[3]

Table 1: Proportion of individuals reporting an illness episode in the four weeks prior to the survey (by per capita expenditure quintile and rural/urban residence, Cambodia, 1997)

Per capita expenditure quintile	*% individuals reporting an illness*	
	Rural	***Urban***
Poorest	0.095	0.123
Second	0.115	0.150
Third	0.159	0.124
Fourth	0.186	0.139
Richest	0.197	0.121
ALL	0.150	0.132

*Source:*CSES, 1997.

Not surprisingly, there is a positive association between the reported incidence of illness and per capita expenditure among rural residents, perhaps reflecting the greater propensity to report illnesses among better-off individuals.[4] Interestingly, the poorer quintiles in the urban areas have higher reported morbidity than their counterparts in the rural areas; however, the opposite is the case with the richer quintiles.

Table 2, which displays the reported incidence of illnesses by age and sex, shows that illness episodes are concentrated among children ages 0-5 and persons ages 46 and over. These results are a common feature of most developing-country populations, and reflect the high likelihood of morbidity associated with childhood and old-age diseases. Overall, women are at

[3]For example, Berman et al. (1987) reports, on the basis of various sample surveys, an annual prevalence rate of 2.5 illness episodes per person per year in Indonesia.

[4]This has been observed elsewhere. See Chernichovsky and Meesook (1986) and Deolalikar (1997) for Indonesia and Deolalikar and Vashishtha (1992) for India.

greater risk of illness than men across almost all age groups (with the exception of those under five years old.)

Table 2: Proportion of individuals reporting an illness episode in the four weeks prior to the survey (by per capita expenditure quintile and rural/urban residence, Cambodia, 1997)

Age group (years)	*Proportion reporting an illness*		
	Males	*Females*	*All persons*
0-5	0.203	0.188	0.195
6-15	0.079	0.093	0.086
16-25	0.077	0.087	0.083
26-45	0.149	0.166	0.158
46-60	0.230	0.243	0.237
61and over	0.315	0.332	0.325
All persons	0.139	0.154	0.147

*Source:*CSES, 1997.

Data from Cambodia's health management information system indicate that infectious diseases were the leading cause of morbidity in the country, as is the case in most other developing countries (see Table 3) (Kingdom of Cambodia 1997). In 1996, approximately 75% of all reported illnesses among Cambodian children under the age of 4 were symptomatically classifiable as communicable diseases. The corresponding estimates for children ages 5-14 and adults were 70% and 56%, respectively. The most commonly reported outpatient condition for all age groups was acute respiratory illness. These patterns were similar among a cohort of hospital inpatients—roughly 70% of inpatients had been admitted to hospitals for infectious diseases (Table 4).

Table 3: Number of Medical Outpatient Consultations at Hospitals (by Type of Illness and Age Group, Cambodia, 1996)

Health Problem	*0-4 years*		*5-14 years*		*15 years and over*		*All ages*	
	Number	*%*	*Number*	*%*	*Number*	*%*	*Number*	*%*
Malaria	10,606	4.4	17,062	6.8	58,967	6.9	86,635	6.5
Other fevers	44,910	18.8	48,113	19.3	116,018	13.6	209,041	15.6
Diarrhea	35,602	14.9	29,405	11.8	66,409	7.8	131,416	9.8
Acute respiratory infections	76,746	32.1	69,837	28.0	172,104	20.2	318,687	23.8
Cough more than 21 days	1,922	0.8	2,712	1.1	17,730	2.1	22,364	1.7
Measles	632	0.3	462	0.2	779	0.1	1,873	0.1
Malnutrition	3,905	1.6					3,905	0.3
Skin infection	7,281	3.1	8,322	3.3	18,094	2.1	33,697	2.5
Gynecological infections					53,461	6.3	53,461	4.0
STD & infections (males)					9,939	1.2	9,939	0.7
Other diseases	57,109	23.9	73,184	29.4	339,630	39.8	469,923	35.0
All diseases and problems	238,713	100.0	249,097	100.0	853,131	100.0	1,340,941	100.0

*Source:*Ministry of Health, *National Health Statistics 1996*, Phnom Penh, 1997. PPNR

4. ESTIMATES

Table 5 shows the maximum likelihood logit estimates of the illness equation for children, adults, and the pooled sample, while Table 6 reports the elasticity of the probability of illness (evaluated at sample means) with respect to the independent variables. Since the results with respect to disease transmission are of the greatest interest in this paper, we discuss these results first.

Table 4: Frequency and Age Distribution of Main Health Problems Seen Among Hospital Inpatients (Cambodia, 1996)

Health Problem	*0-4 years*		*5-14 years*		*15 years and over*		*All ages*	
	Number	*%*	*Number*	*%*	*Number*	*%*	*Number*	*%*
Malaria	3,013	13.0	4,012	19.0	14,751	12.0	21,776	13.0
Diarrhea	4,077	17.5	1,742	8.3	3,505	2.8	9,324	5.6
Dysentry	941	4.0	702	3.3	1,440	1.2	3,083	1.8
Acute respiratory infections	7,204	31.0	3,918	18.6	11,623	9.4	22,745	13.6
Dengue/Dengue Hemorrhagic Fever	777	3.3	603	2.9	53	0.0	1,433	0.9
Meningitis	408	1.8	233	1.1	217	0.2	858	0.5
Polio	3	0.0	7	0.0	4	0.0	14	0.0
Measles	55	0.2	39	0.2	4	0.0	98	0.1
Neonatal tetanus	7	0.0					7	0.0
Tetanus	24	0.1	29	0.1	104	0.1	157	0.1
Tuberculosis	13	0.1	79	0.4	14,158	11.5	14,250	8.5
Gynecological-obstetric					19,107	15.5	19,107	11.4
Mine accidents	22	0.1	237	1.1	2,159	1.8	2,418	1.4
Road accidents	199	0.9	1,053	5.0	5,970	4.8	7,222	4.3
Other diseases	6,494	27.9	8,444	40.0	50,012	40.6	64,950	38.8
All diseases and problems	23,237	100.0	21,098	100.0	123,107	100.0	167,442	100.0

Source:Ministry of Health, *National Health Statistics 1996*, Phnom Penh, 1997.

Table 5: Regression estimates for maximum likelihood logit estimation including other ill individuals. (Dependent variable: Probability of illness)*

	Children	*t-stat*	*Adults*	*t-stat*	*All*	*t-stat*
Urban	0.117	1.31	**0.325**	4.68	**0.282**	5.17
Doctor in village?	**0.540**	1.92	**0.442**	2.33	**0.500**	3.24
Age	**-0.160**	-1.76	**0.132**	5.26	**0.025**	1.92
Age squared	0.003	0.45	**-0.001**	-3.57	0.000	0.55
Female	-0.133	-0.59	**0.285**	1.76	0.145	1.14
Log Expenditure	**0.489**	3.51	**0.377**	3.65	**0.410**	4.96
Household head female	0.005	0.05	**0.295**	4.51	**0.168**	3.20
Years of adult education			**-0.037**	-1.76	**-0.053**	-2.82
Years of HH head education	-0.005	-0.49	0.012	1.32	0.011	1.63
No toilet	**0.279**	2.99	0.069	0.99	**0.171**	3.07
Access to safe water	-0.094	-1.16	0.030	0.49	-0.014	-0.28
Transmission Effects						
Other ill in village	**0.332**	2.42	**0.546**	4.62	**0.431**	5.07
Other ill * Age	0.002	0.20	**-0.006**	-2.68	**-0.004**	-3.24
Other ill * Age square	0.000	-0.26	**5.9E-05**	2.21	**0.000**	2.38
Other ill * Female	0.016	0.79	-0.02356	-1.51	-0.007	-0.61
Other ill * Log Expenditure	**-0.022**	-1.92	**-0.0235**	-2.51	**-0.023**	-3.18
Other ill * Adult education			0.00137	0.64	0.003	1.59
Other ill * Doctor	**-0.117**	-2.75	**-0.08659**	-3.01	**-0.097**	-4.17
_cons	**-7.572**	-4.78	**-10.8365**	-8.57	**-8.561**	-9.06
N	474		474		474	

* Coefficients that are significant at 90% confidence level are in bold.

The empirical findings suggest a high degree of infection transmission among village residents, with the elasticity of the probability of an individual being ill with respect to the number of other ill individuals in the village being nearly unity (0.94). Surprisingly, the elasticity of the transmission effect is lower for children than for adults (0.41 versus 0.74), suggesting that children are less susceptible to catching infections from others in the community.

The empirical results indicate that, among adults, the susceptibility to catching infections from others is strongly related to age, but in a nonlinear manner. Younger adults are less likely to get infected by others, but susceptibility increases with age. Age does not appear to be a significant determinant of illness transmission among children under the age of 16 years. Among adults, there is some evidence of females being at slightly lower risk of disease transmission than males, although this effect barely achieves statistical significance.

An individual's economic status (as represented by his household consumption expenditure per capita), however, has a strong effect on disease transmission, with better-off individuals having a significantly lower risk than poorer individuals of catching infections from others in the community. Probably the single most important reason why this might occur is that richer individuals are more likely to reside in less-crowded surroundings, which would significantly lower their risk of acquiring infections from neighbors (including other family members).

Schooling, however, does not have a significant effect on disease transmission. We find this result surprising, as better-schooled individuals are generally thought to engage in defensive health and hygiene practices, such as engaging in safe sex and washing one's hands before eating and cooking, that better protect them from the infections of others in the community. However, the number of years of adult education is negatively correlated with the probability of illness, as indicated by our elasticity estimates in Table 6.

Finally, the presence of a doctor in the village is observed to have the single largest effect on disease transmission. The results suggest that having a doctor in the village lowers the probability of catching an infection from others by as much as 10% in both children and in adults. This result is in line with our expectations, as the presence of a doctor in a community is indicative of a higher level of health-care infrastructure in that community. Individuals residing in such communities are much more likely to have access to preventative health services (such as vaccinations) and knowledge of defensive health and hygiene practices that would reduce their

Table 6: Elasticity of probability of illness with respect to socioeconomic variables

Dependent Variable	*Total Elasticities*		
	Children	***Adults***	***All***
Urban	0.011	0.043	0.049
Doctor in village?	-0.035	-0.031	-0.047
Age	-0.540	1.539	-0.240
Female	0.007	0.020	0.032
Log Expenditure	1.392	1.034	1.707
Female head of household	0.000	0.039	0.026
Adult education (years)		-0.007	-0.040
Household head education (years)	-0.010	0.028	0.034
No toilet	0.106	0.032	0.110
Access to safe water	-0.013	0.006	-0.003
Fraction of other ill individuals in village	0.407	0.741	0.935

risk of acquiring communicable illnesses. It is interesting to note that the presence of a clinic in the village was not a strong predictor of illness, which indicates that having a clinic in the village does not necessarily mean that a doctor is available. In our sample, only 14 out of a total of 474 villages surveyed reported having access to both a doctor *and* a *khum* clinic. Seventy-five villages had a doctor, but no clinic, while 61 villages had a clinic that was not staffed with a doctor.

It should be noted that the discussion above refers only to the effect of variables such as age, socioeconomic status, and the presence of a doctor on the magnitude of infection transmission—not on the overall probability of an individual being ill. These variables have two separate effects on the likelihood of an individual being ill: a *ceteris paribus* direct effect (controlling for infection transmission across individuals) and a *mutatis mutandis* indirect effect via illness transmission. As an example, consider the effect of the presence of a doctor in a village on the probability of an individual being ill. The results in Table 5 indicate that having a doctor in the village raises the probability of an individual being ill *controlling for infection transmission across individuals*. However, there is a negative association between the presence of a doctor and the probability of an individual being ill when transmission effects are taken into account. This is because an individual is more likely to be ill when many others in his/her village are also ill, and the presence of a doctor significantly reduces the transmission of infections from one person to another in a community.

Table 6 suggests that the net effect of household consumption expenditure per capita on morbidity is positive and large, with elasticity in excess of one. This suggests that even though the transmission of infections is lowered among affluent individuals, the direct effect of economic status on the reporting of morbidity is positive and outweighs the transmission effect.

There are four other significant findings. First, toilets in the dwelling are associated with a significantly reduced prevalence of illness, especially among children. This is not surprising, given the prevalence of water-borne and water contamination-related diseases in Cambodia.

Second, sex—both of the individual as well as of the household head—has significant effects on the likelihood of an illness. Controlling for other factors, females are more likely than males to suffer from an illness episode. Individuals residing in female-headed households are 18% more likely to fall ill than those residing in male-headed households. However, it is unclear from the results whether females are genuinely more prone to illness than males or whether they are simply more likely to report illness episodes.

Third, urban residence is associated positively with the probability of an illness. Again, it is not clear whether this reflects a greater risk of morbidity in urban relative to rural areas, or whether it reflects better reporting of illness episodes in the urban areas.

Finally, the net effect of age on morbidity is different for adults than for children. Among children, age is associated with a significantly reduced probability of illness, but among adults, the probability of an illness sharply increases with age.

Comparison with naïve estimation

In order to compare the utility of including other illnesses as an explanatory variable for illness in a transmission consistent (TC) regression, with a naïve estimation that assumes uniform exposure, we ran regressions that used all the explanatory variables as in the main estimation, but excluded the other ill variables. Results from these regressions are presented in Tables 7 and 8. Since household factors such as access to safe water or toilets, and education were negatively correlated with the likelihood of exposure, it was our hypothesis that the impact of these factors on disease risk would be elevated. This hypothesis was validated in the regression results. For example, the elasticities of the probability of illness with respect to access to safe water and access to toilets were –3.6 and –12.8 respectively, for the pooled naïve regression. In comparison, the estimates for these elasticities from the TC regression (described earlier) were –0.3 and –11. The effect of the expenditure variable on the probability of illness was much greater in the naïve regression than in the TC regression because of the potential positive correlation between expenditure and exposure to other ill individuals.

Table 7: Regression estimates for maximum likelihood logit estimation without including other ill individuals. (Dependent variable: Probability of illness)*

	Children	*t-stat*	*Adults*	*t-stat*	*All*	*t-stat*
Urban	-0.084	-1.12	0.057	0.93	0.019	0.41
Doctor in village?	**-0.482**	-4.44	**-0.417**	-5.85	**-0.431**	-7.30
Age	**-0.148**	-5.54	**0.060**	8.04	-0.021	-5.24
Age squared	0.001	0.69	**0.000**	-4.19	**0.001**	10.18
Female	0.040	0.62	0.020	0.35	**0.071**	1.74
Log Expenditure	**0.417**	7.70	**0.355**	8.37	**0.383**	11.52
Household head female	-0.125	-1.30	**0.151**	2.41	0.011	0.21
Years of adult education			**-0.036**	-3.60	**-0.022**	-2.92
Years of HH head education	**-0.026**	-2.60	-0.002	-0.25	**-0.013**	-2.00
No toilet	**0.291**	3.23	0.095	1.39	**0.199**	3.68
Access to safe water	**-0.194**	-2.52	**-0.107**	-1.85	**-0.146**	-3.18
_cons	-5.664	-9.23	-7.324	-14.75	-6.124	-16.28
N	474		474		474	

[*]Coefficients that are significant at 90% confidence level are in bold.

Table 8: Elasticity of probability of illness with respect to socioeconomic variables

Dependent variable	*Total Elasticities*		
	Children	*Adults*	*All*
Urban	-0.016	0.009	0.003
Doctor in village?	-0.052	-0.040	-0.043
Age	-1.130	1.640	-0.412
Female	0.020	0.008	0.031
Log Expenditure	4.518	2.994	3.567
Female head of household	-0.018	0.024	0.002
Adult education (years)		-0.103	-0.039
Household head education (years)	-0.099	-0.006	-0.040
No toilet	0.224	0.054	0.128
Access to safe water	-0.054	-0.025	-0.036

However, the availability of a doctor in the village was unlikely to be correlated with exposure to disease and we did not expect any bias in the naïve regression. Accordingly, the elasticity of illness for all individuals with respect to availability of a doctor was –4.3 in the naïve regression and –4.7 in the TC regression.

5. CONCLUSIONS

The empirical approach followed in this analysis demonstrates that it is possible to estimate the impact of socioeconomic variables on the transmission of infectious diseases. The results of this analysis permit both a qualitative as well as a quantitative comparison of the effect of different socioeconomic and environmental factors on the likelihood of disease acquisition, given that a person is in close proximity to other infected individuals. For instance, income and the availability of a doctor exercise a protective effect with respect to disease transmission. Similarly, access to safe water and the availability of a toilet both reduce the probability of falling ill when the effect of being exposed to other ill individuals is controlled for.

The analysis presented also demonstrates the potential pitfalls of excluding disease exposure from socioeconomic studies of disease. Omitting disease exposure automatically implies that all individuals face the same level of exposure to disease. Since exposure also tends to be correlated with other disease determinants like lack of access to water and toilets, omitting exposure biases the estimates of the effect of these variables on disease risk upwards. The policy implications of this bias are important. Although factors that reduce disease transmission are important in reducing the burden of infectious disease, treatment is also an important factor in keeping the population of infected individuals low.

The results of this analysis highlight the public externality associated with infectious diseases and point to the need for timely medical treatment to reduce the likelihood that the disease will be transmitted to other individuals. Ignoring disease exposure could lead to a greater weight being placed on factors like water and sanitation alone, and diminishes the importance of public health treatment interventions—such as medical camps and mobile clinics—that may help reduce disease incidence.

It is prudent to repeat that this analysis is subject to an important caveat. Our data is limited to self-reported illness variables, which we have used as a proxy to measure the incidence of infectious diseases. Although the data in Tables 3 and 4 indicate that infectious diseases are a significant source of morbidity in Cambodia, our results are biased to the extent that systematic measurement errors are induced by the use of this proxy for the incidence of infectious diseases.

REFERENCES

Alberini, A., G. S. Eskeland, et al. (1995). Determinants of Diarrheal Disease in Jakarta. Washington DC, Resources for the Future.

Behrman, J. R. and A. B. Deolalikar (1988). Health and Nutrition. Handbook of Development Economics. H. Chenery and T. N. Srinivasan. North Holland, Elsevier Science Publishers BV.

Berman, P., B. A. Osmond, et al. (1987). "Treatment Use and Expenditure on Curative Care in Rural Indonesia." Health Policy and Planning **2**(4): 289-300.

Chandiwana, S. K. and M. E. J. Woolhouse (1991). "Heterogeneities in human water contact behaviour and the epidemiology of Schistosoma haematobium." Parasitology **103**: 363-70.

Chernichovsky, D. and O. A. Meesook (1986). "Utilization of Health Services in Indonesia." Social Science and Medicine **23**(6): 611-620.

Deolalikar, A. B. (1997). The Demand for Health Services in Indonesia: The Role of Prices, Service Quality, and Reporting of Illnesses. Handbook of Applied Economic Statistics. A. Ullah and D. E. Giles. New York, Marcel Dekker Inc.

Deolalikar, A. B. and P. Vashishta (1992). The Utilization of Government and Private Health Services in India, Options Project, AID Contract no. DPE-3035-C-00-0053-00.

Grossman, M. (1972). "On the Concept of Health Capital and Demand for Health." Journal of Political Economy **80**(2).

Kermack, W. O. and A. G. McKendrick (1927). "A Contribution to the Mathematical Theory of Epidemics." Proceedings of the Royal Society **A115**: 700-21.

Kingdom of Cambodia (1997). National Health Statistics Report 1996. Phnom Penh, Planning and Statistics Unit, Ministry of Health, Kingdom of Cambodia.

Strauss, J. and D. Thomas (1995). Human Resources: Empirical Modeling of Household and Family Decisions. Handbook of Development Economics. J. R. Behrman and T. N. Srinivasan. Amsterdam, North-Holland Publishing Co. **3**.

White, E. B. (1978). "A Heteroskedasticity Consistent Covarience Matrix and a Direct Test of Heteroskedasticity." Econometrica: 817-838.

Woolhouse, M. E. J. (1994). "Immunoepidemiology of human schistsomes: Taking the theory into the field." Parasitology Today **10**: 196-202.

World Resources Institute (1998). World Resources: A Guide to the Global Environment. New York, Oxford University Press.