Does Piped Water Improve Household Welfare? Evidence from Vietnam

Nguyen Viet, Cuong

8 February 2011

Does Piped Water Improve Household Welfare? Evidence from Vietnam

Nguyen Viet Cuong
Vu Thieu

ABSTRACT

Clean water is essential for human survival. Yet, many people do not have access to clean water in Vietnam. Only around 23% of the population had access to piped water in 2006. This study measures the effect of piped water on household welfare using difference-in-differences estimators and panel data from the Vietnam Household Living Standard Surveys. Findings show that the effect of piped water on household income and labor supply is positive but small and not statistically significant. The effect of piped water on sickness of household members is negative but not statistically significant.

This study is funded by the Economy and Environment Program for Southeast Asia (EEPSEA). We would like to thank Prof. Dale Whittington, Prof. Vic Adamowicz, and Dr. Herminia Francisco for their helpful guidance, suggestions, and comments. We would also like to thank Mrs. Catherine Ndiaye for logistic supports.

Comments should be sent to: Nguyen Viet Cuong
Contact address: Indochina Research & Consulting, Suite 1701, C'land Tower 156 Xa Dan 2, Dong Da, Hanoi, Vietnam.

Tel: + 84 - 4 – 3573 9672.
Fax: + 84 - 4 – 3573 9670
Email: c_nguyenviet@yahoo.com

1.0 INTRODUCTION

Clean water is essential for human living. Yet, many poor people are denied access to sufficient and clean water. Using living standard measurement surveys in the 1990s, Komives et al. (2003) found that the use of unimproved water was prevalent in low-income countries. According to UNICEF (2010), “over 884 million people still use unsafe drinking water sources.” Lack of clean water causes diseases and sickness. The World Health Organization (2004) mentioned that contaminated water resulted in thousands of deaths every day, mostly in children under five years old in developing countries. UNDP (2006) claimed that unsafe water and shortage of basic sanitation caused 80% of diseases in the developing countries.

There are many empirical studies on the linkage between water and welfare.[1] Collin et al. (1981) found that epidemics tend to be associated with poor quality water in France. Dasgupta (2004) concluded that increased availability of water reduces the incidence of water-related diseases in India. Galiani et al. (2005) found that the privatization of water services could reduce child mortality in Argentina. In Brazil, Macinko et al. (2005) found that access to clean water was negatively associated with infant mortality. Recently, Gamper-Rabindran et al. (2010) showed that the provision of piped water reduces the infant mortality rate in Brazil. Meanwhile, Jalan and Ravallion (2003) found that piped water reduced the diarrhea of children in rural India.

Although there are numerous studies on the effect of piped water on health, only a few studies investigate the effect of piped water on other household welfare indicators such as labor supply and income (Waddington et al., 2009). In the long-run, piped water can result in an increase in income through several channels. Unclean water can cause health problems, thereby reducing working efforts and income. Without access to piped water, households have to use other water sources, some of which can be very far from their home and require purification before using. Thus, having piped water can save time for people and allow for more productive activities and income increase. An exceptional

[1] For review of studies on the effect of water on health, see Fewtrell and Kaufmann (2005), Günther and Fink (2010), Waddington et al. (2009), and Fewtrell et al. (2009).

study, which examines the effect of water on labor and income, was done by Devoto et al. (2011) in urban Morocco. They found that water connection could improve the households' leisure and social activities, but not their income and labor supply.

This study aims to measure the impact of piped water on several household welfare indicators including income, labor supply, and health in Vietnam using recent household surveys and difference-in-differences estimators. Vietnam is an interesting case to study the effect of piped water.

First, Vietnam has a lower proportion of people with access to improved water. According to the Vietnam Household Living Standard Survey in 2006, around 26% of the population and only 4.3% of the poor had access to piped water.

Around 80% of cases of infectious diseases are related to unclean water (Xuan-Long, 2010). Nearly one million diarrheal cases are reported every year (MOH, 2008). Hence, provision of improved water supply is an important policy of Vietnam. Since 1998, the government has launched the national program on clean water and sanitation for the rural areas to improve the people's health and to reduce poverty. Thus, it is of interest to policy makers and researchers to examine the effect of improved water supply on the welfare of households.

Second, there are no quantitative and evidence-based studies that measure the effect of clean water on the welfare of households in Vietnam. Several studies have focused on the quality of drinking water (Hoang,1990; Le et al.,1993 ; Nguyen et al.,1994; Le and Munekage, 2004; and Agusa et al.,2006), specifically on its chemical aspects. Other studies mention the adverse effects of unclean water on health but without quantitative evidences (World Bank, 2000 and 2004; Xuan-Long, 2010; Suc Khoe Newspaper, 2010).

Third, the Vietnam Household Living Standard Surveys (VHLSS) 2002, 2004, and 2006 can be used to estimate the effect of piped water beyond water-related diseases using the difference-in-differences with matching estimator. The VHLSSs contain data on water use of households and on household welfare indicators including sickness, labor supply, and income. The VHLSSs also contain panel data for the difference-in-differences estimator. A difficulty in measuring the effect of piped water on household welfare is endogeneity of the piped water. The difference-in-differences with matching estimator can solve the endogeneity bias provided that this bias is caused by time-invariant unobserved variables.[2]

The paper is structured into six sections. The second section introduces data sources used in this study. The third section presents the access to piped water in Vietnam. Next, the fourth and fifth sections present the methodology and empirical findings of the impact of piped water on health, labor supply, and income of households. The sixth section gives the conclusion.

2.0 DATA SOURCES

The study relies on the VHLSSs conducted by the General Statistics Office of Vietnam (GSO) with technical supports from the World Bank (WB) in 2002, 2004, and 2006. The surveys contain detailed information on households including basic demography, employment and labor force participation, education, health, income, expenditure, housing, fixed assets and durable goods, participation of households in poverty alleviation programs, and access to different water sources. The 2004 and 2006 VHLSSs contain some information on the sickness of individuals. However, there are no data on diseases, weight, and heights of individuals.

The sample size of households in the VHLSSs were 74,341 in 2002; 45,943 in 2004; and 45,945 households in 2006. It is very useful that the data from the VHLSSs in 2004 and in 2006 came from a panel of 21695 households. The panel data were representative at the urban/rural and regional level. In addition, there are also a panel data from the three VHLSSs in 2002, 2004 and 2006. More specifically, 10,365 households were covered by the three VHLSSs.

[2] Randomization design and instrumental variables are two ideal methods to deal with endogeneity. However, randomization design of piped water is hard to implemented. Finding a good instrument for piped water is also challenging. Using a weak or invalid instrument can lead to a large estimation bias.

3.0 ACCESS TO PIPED WATER IN VIETNAM

Although Vietnam has achieved remarkable successes in poverty reduction, the current poverty rate remains rather high. In rural areas, the poverty rate was 19.7% in 2006.[3] The poor are often characterized by poor living conditions, including reliable water shortage. Limited access to improved water and sanitation is also often mentioned in most Poverty Participatory Assessment reports (MONRE, 2007). Compared to other Southeast Asian countries, Vietnam has a lower proportion of people with access to improved water. In 2000, around 56% of the Vietnamese had access to improved water supply, which was low compared to Indonesia (76%), Thailand (80%), Philippines (87%), and Lao (90%) (WHO, 2000).[4]

It is almost impossible to measure the quality of water supply throughout a country. In studies in other countries, water supply is often considered to be improved as piped water is introduced (Fewtrell et al., 2009). In this study, we use the VHLSSs to examine the impact of piped water on household welfare in Vietnam. Information on whether piped water is really clean and potable for each household is not available in our data sets.

It should be noted that in this study households with access to piped water are defined as those using piped water for drinking and cooking. Other water sources that are used for drinking are deep drill wells, reinforce-concrete wells, bottled water, ponds, and rivers. Some households can use a purification system before using water. Boiling water before drinking is popular in Vietnam, and it can be an effective way to reduce water-related diseases. According to the 2006 VHLSS, around 86% of households always boil their drinking water.

Figure 1 shows that access to piped water has been increasing in Vietnam. The proportion of households using piped water increased from 17.5% in 2002 to 23.2% in 2006. The proportion of urban households with piped water increased from 54% to 62

---

[3] This estimation is based on Vietnam Household Living Standard Survey in 2006. The poor are classified based on the expenditure poverty line constructed by the World Bank and the General Statistics Office of Vietnam. The poverty line in 2006 was 2560 thousands VND.

[4] In WHO (2000), improved water includes water from household connection, public standpipe, borehole, protected dug well, protected spring, and rainwater collection.

percent during the same period. In rural areas, the proportion of households with piped water also increased from 5.9% to 8.4%.

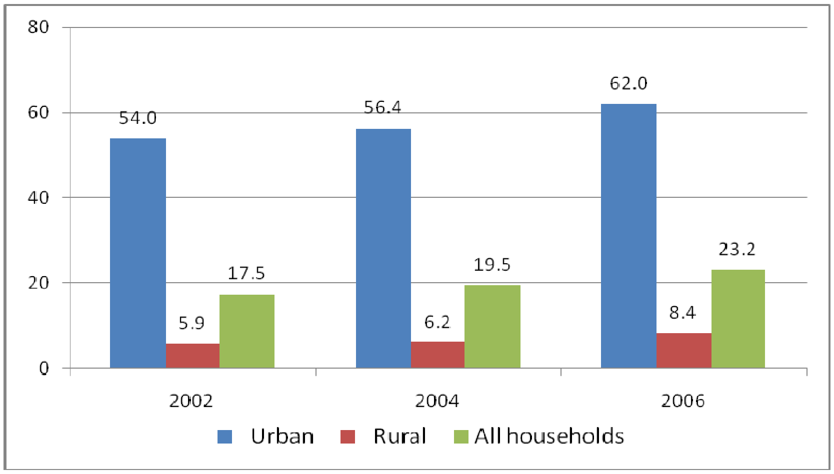


Figure 1. Access to piped water during 2002 to 2006

Source: Estimation from VHLSSs 2002, 2004, and 2006

There is a large disparity in access to piped water across geographic areas and different population groups. The proportion of households with piped water was 62% for urban areas and only 8.4% for rural areas in 2006. Households in delta regions are more likely to have access to piped water than households in mountainous areas (Table 1). Disadvantaged households, including poor and ethnic minority households, have much lower access to tap and clean water than better-off households and Kinh/Chinese households.

Table 1. Percentage of households with access to piped water during 2002 to 2006 by household characteristics

Households	2002	2004	2006
Region			
Red River Delta	17.5	17.9	23.7
North East	10.2	12.2	15.9
North West	11.1	8.1	12.2
North Central Coast	10.7	11.6	13.2
South Central Coast	14.1	17.5	21.0
Central Highlands	10.7	12.3	12.0
South East	31.6	37.5	40.8
Mekong River Delta	20.0	20.6	24.8
Ethnicity			
Kinh, Chinese	19.1	21.3	25.4

Ethnic minorities	3.9	3.7	5.4
Poverty			
Non-Poor	22.5	22.9	26.2
Poor	3.1	3.1	4.2
Expenditure quintiles			
Poorest	3.2	3.0	4.3
Near poorest	3.9	6.5	8.7
Middle	7.6	10.0	15.6
Near richest	17.6	20.2	27.7
Richest	48.5	51.5	52.7
All households	17.5	19.5	23.2

Source: Estimation from VHLSSs 2002, 2004, and 2006

Figures 2 to 5 present the geographical pattern of drinking water in Vietnam. Figures 2 and 3 are based on VHLSS data, which do not allow for estimates at the district level. Figures 4 and 5 use data from the 2006 Agricultural Census, which allows for estimates at the provincial and district levels. The figures show strong spatial differences in water quality in Vietnam. The use of tap and clean water is much more prevalent in delta regions such as Red River Delta, South East, and Mekong River Delta than in mountainous regions such as the North West and North East.

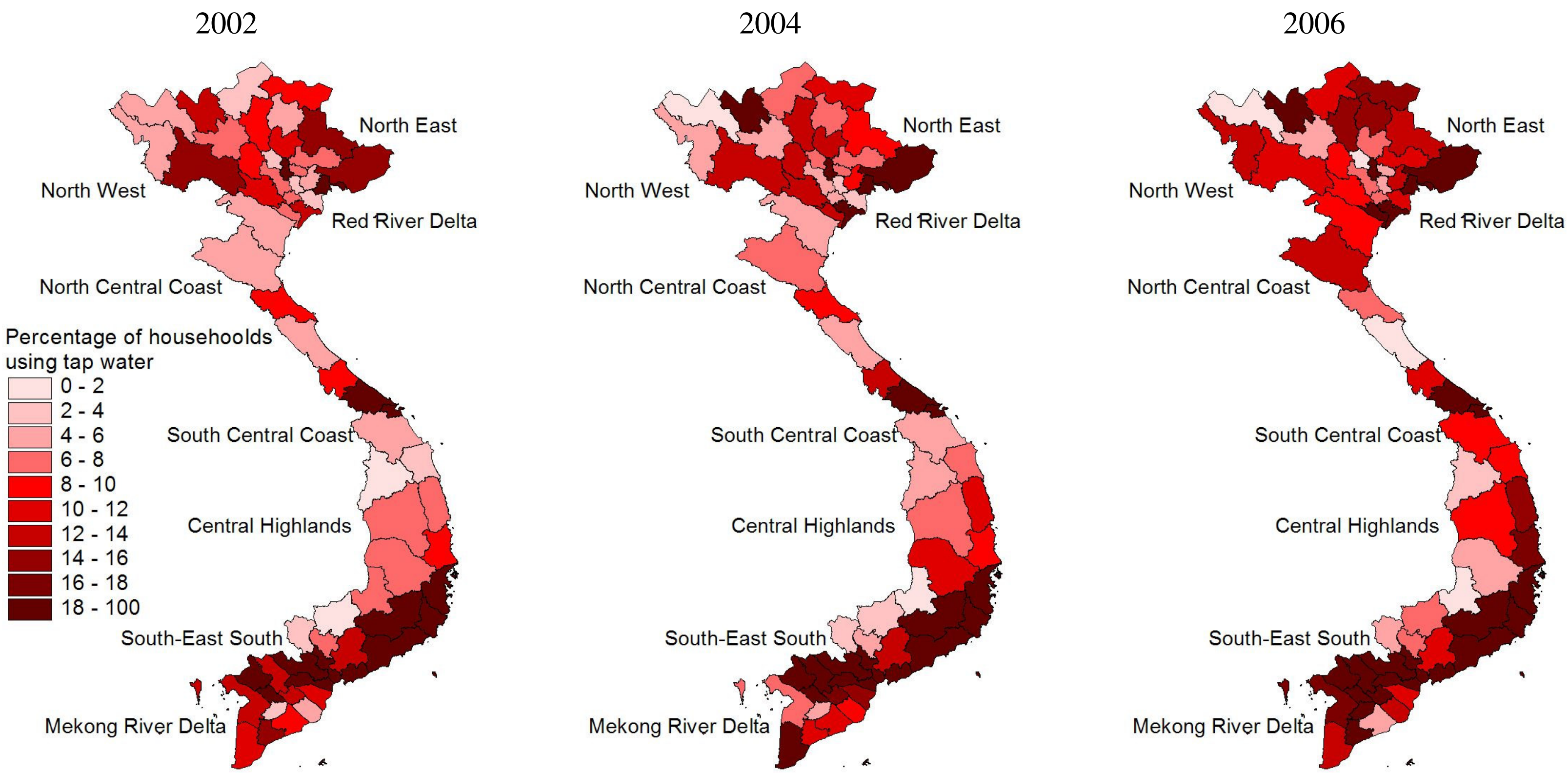


Figure 2. Percentage of households using drinking tap water (all the country)

Source: Estimation from VHLSSs 2002, 2004, and 2006

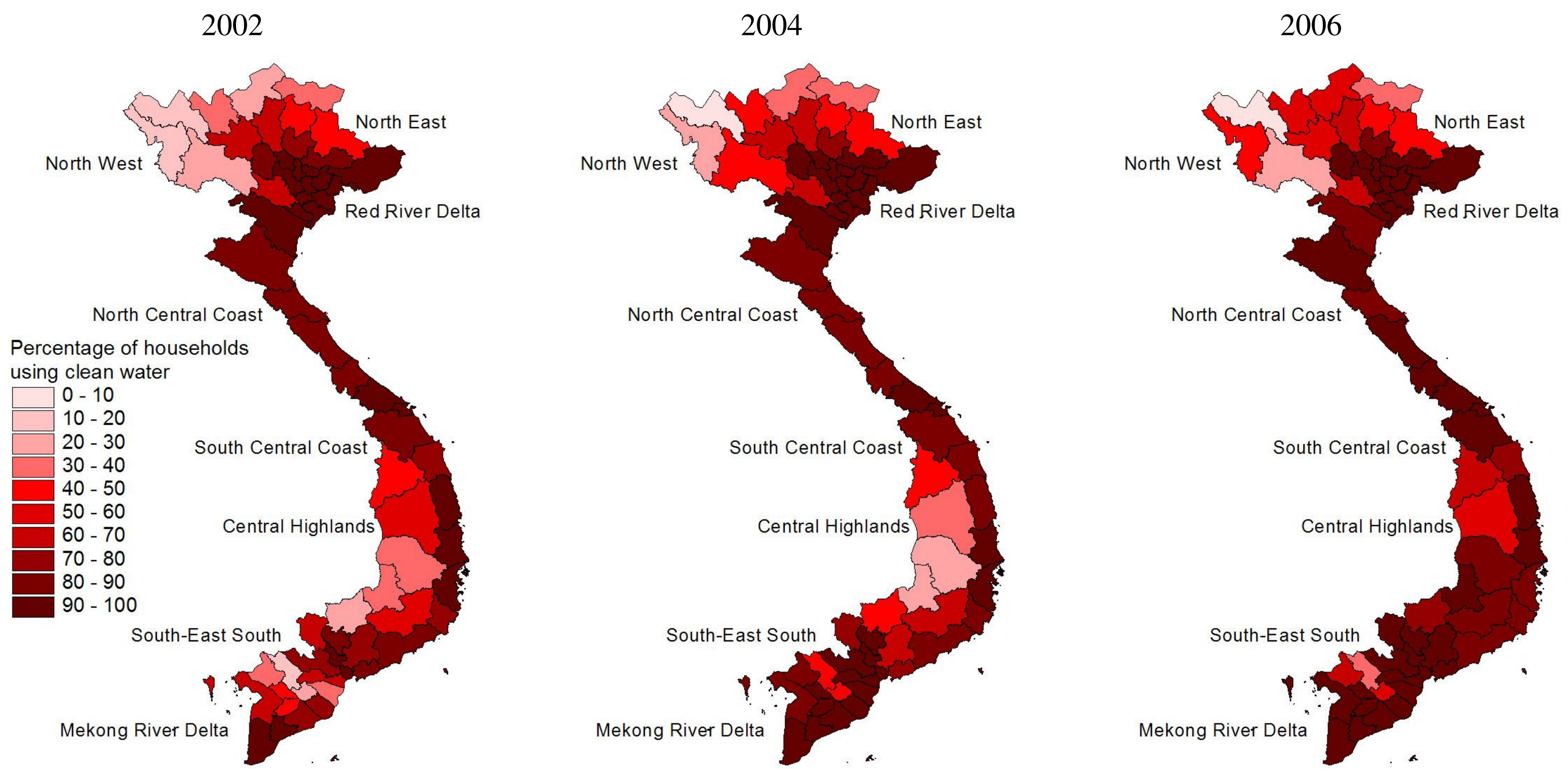


Figure 3. Percentage of households using clean water (including tap water) in Vietnam

Source: Estimation from VHLSSs 2002, 2004, and 200

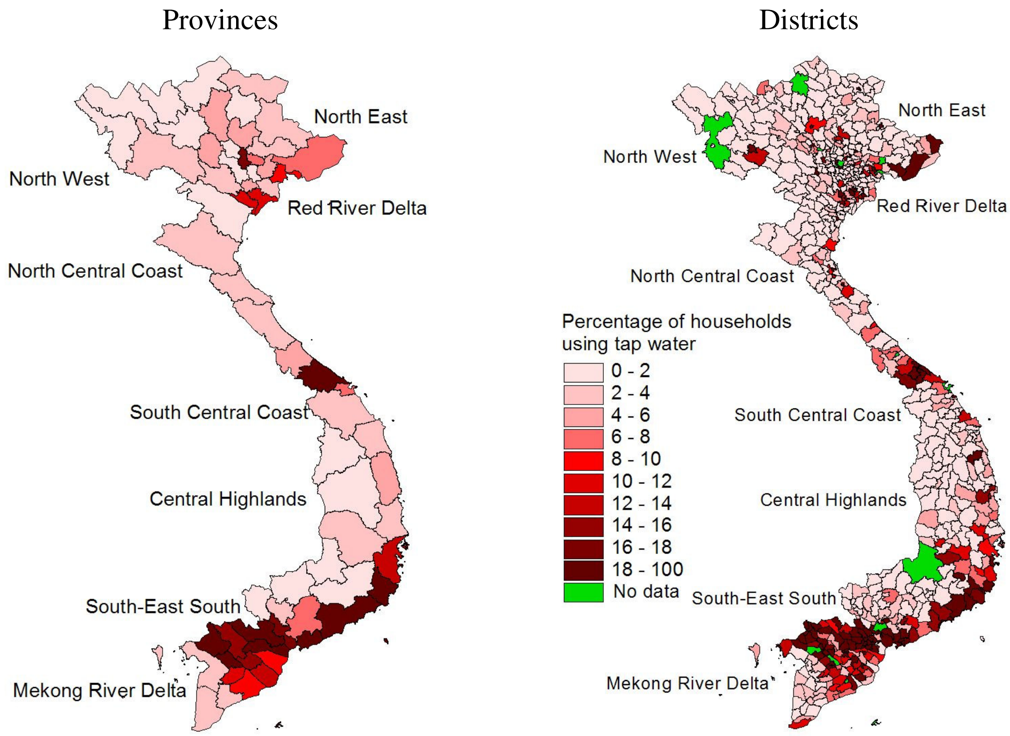


Figure 4. Percentage of households using tap water in rural Vietnam in 2006

Source: Estimation from the Rural Agriculture and Fishery Census, 2006

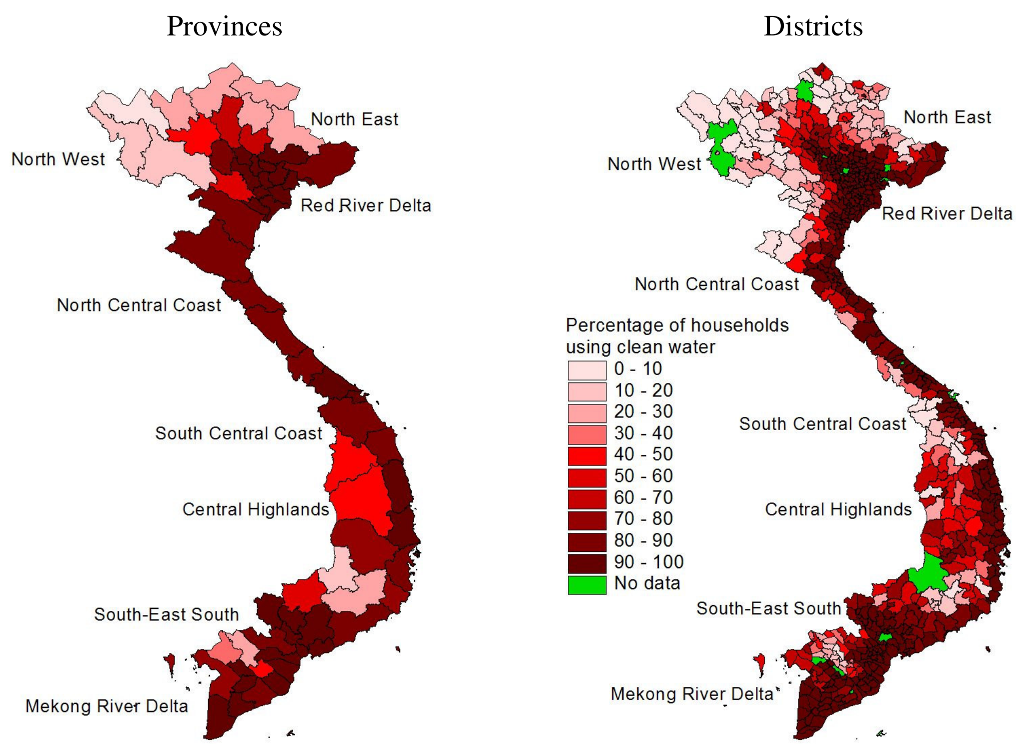


Figure 5. Percentage of households using clean water (including tap water) in rural Vietnam in 2006

Source: Estimation from the Rural Agriculture and Fishery Census, 2006

4.0 IMPACT ESTIMATION METHOD

Parameters of impact

Denote D as a binary variable indicating access to piped water, i.e., D equals 1 if a household is using piped water, and D equals 0 otherwise. Further, let Y denote the observed value of household welfare (so-called outcomes). This variable can have two potential values depending on the value of D, i.e. $Y = Y_1$ for $D = 1$, and $Y = Y_0$ for $D = 0$.

The most popular parameter of the program impact is Average Treatment Effect on the Treated (ATT) (Heckman et al., 1999), which is the expected impact of access to piped water on outcomes of households using piped water:[5]

$$ATT = E(Y_1 - Y_0 | D = 1) = E(Y_1 | D = 1) - E(Y_0 | D = 1) \quad (1)$$

More generally, ATT can vary across a vector of the observed variables X:

$$ATT_{(X)} = E(Y_1 - Y_0 | X, D = 1) = E(Y_1 \mid X, D = 1) - E(Y_0 \mid X, D = 1) \quad (2)$$

Estimation of ATT is not straightforward, since $E(Y_0 \mid D = 1)$ is not observed and cannot be estimated directly. $E(Y_0 \mid D = 1)$ is called counterfactual, which would have been the expected outcome of households if they had not used piped water.

Difference-in-differences with propensity score matching

Since panel data on households with piped water and without piped water are available, we can estimate the impact of piped water on household welfare using a method of difference-in-differences with matching. The basic idea of the matching method is to find a control group who are households without piped water that have similar distribution of

[5] There are other parameters such as average treatment effect (ATE), local average treatment effect, marginal treatment effect, or even effect of "non-treatment or non-treated," which measure what impact the program would have on the non-participants if they had participated in the program, etc.

the variables X as the treatment group who are households with piped water.[6] The matching method can be combined with difference-in-differences method so that access to piped water can be based on unobserved variables. However, these unobserved variables are assumed to be time-invariant.

More specifically, let Y_0^B denote the outcome before access to piped water. After access to piped water, the potential outcomes with and without piped water are denoted by Y_1^A and Y_0^A, respectively. In this study, 2002 (or 2004) and 2006 are the years before and after access to piped water, respectively. The $ATT_{(X)}$ after access to piped water is defined as:

$$ATT_{(X)} = E(Y_1^A | X, D=1) - E(Y_0^A | X, D=1) \tag{3}$$

The difference-in-differences with matching relies on an assumption that conditional on X, the difference in the expectation of outcomes between households with piped water and households without piped water is unchanged before and after access to piped water, i.e.:

$$E(Y_0^B \mid X, D=1) - E(Y_0^B \mid X, D=0) = E(Y_0^A \mid X, D=1) - E(Y_0^A \mid X, D=0). \tag{4}$$

Under this assumption, the conditional parameter $ATT_{(X)}$ can be identified by the matching method, since:

$$\begin{aligned} ATT_{(X)} = \; & E(Y_1^A|X, D=1) - E(Y_0^A|X, D=1) - \left[E(Y_0^A|X,D=0) - E(Y_0^B|X,D=0)\right] \\ & + \left[E(Y_0^A|X,D=1) - E(Y_0^B|X,D=1)\right] \\ = \; & \left[E(Y_1^A|X, D=1) - E(Y_0^A|X,D=0)\right] - \left[E(Y_0^B|X,D=1) - E(Y_0^B|X,D=0)\right] \end{aligned} \tag{5}$$

The unconditional parameter is also identified, since:

$$ATT = \int_{X|D=1} ATT_{(X)} dF(X|D=1). \tag{6}$$

To estimate the impact of clean water, households without piped water are matched with households with piped water based on their variables X. The matched households without piped water form a control (comparison) group. However, to find a

[6] There are many literatures on matching methods of impact evaluation. Important contributions come from Rubin (1977, 1979, 1980), Rosenbaum and Rubin (1983), Heckman et al. (1997), Dehejia, and (1998), and Smith and Todd (2005).

control group who has similar variables *X*, we require a so-called common support assumption:

$$0 < P(D = 1 \mid X) < 1 \quad (7)$$

This assumption means that there are households without piped water who have the *X* variables similar to households with piped water.

The remaining problem is how to match households not having piped water with a households having piped water. There is no problem if there is a single conditioning variable *X*. However, *Xs* are often a vector of variables, and finding "close" non-users to match with a non-user is not straightforward.

A widely-used way to find the matched sample is the propensity score matching.[7] Since a paper by Rosenbaum and Rubin (1983), the matching is often conducted based on the probability of being assigned into a program, which is called the propensity score. In this paper, the matching based on the propensity score was used. Depending on the number of households without piped water that are matched with households with piped water, we can have different matching estimators. We used nearest-neighbors, kernel, and local linear regression matching schemes to examine the sensitivity of the impact estimates. The matching estimator is based on equation (5). It is equal to the difference in differences in outcomes between households with piped water and matched households without piped water before and after access to piped water. The formulas of the estimators are presented in Appendix. The standard errors are calculated using bootstrap techniques.

In this study, we used a treatment group (households with piped water) and a control group (households without piped water). We used the feature of the panel data from VHLSSs 2002, 2004, and 2006 to define the control and treatment groups before and after the use of piped water.

Of the 10,365 households covered by the three VHLSs, 373 households did not have piped water in 2002, but had piped water in 2004 and 2006. These households made up the treatment group. The 7,960 households that did not have piped water during 2002,

[7] Other matching methods can be subclassification (see, e.g., Cochran and Chambers, 1965; Cochran, 1968) and covariate matching (Rubin, 1979, 1980).

2004, and 2006 made up the control group. Other households were dropped. The years 2002 and 2006 refer to the years before and after access to piped water, respectively. Designating a treatment and a control group allows for a longer time for the use of piped water. It is expected that the effect of piped water on labor supply and income requires a longer time to be detected. However, the 2002 to 2006 panel data do not have information on sickness of individuals, since the 2002 VHLSS does not contain this information.

Thus, the second way to construct the treatment and control group is to use the panel data of VHLSSs 2004 and 2006, which contain information on individual sickness in both years. A total of 21,695 households provided the panel data. The treatment group consisted of 1,242 households that had piped water in 2006 but not in 2004, and the control groups consisted of 16,763 households that did not have piped water in 2004 and 2006. Other households not belonging to the treatment and control groups were dropped. This construction of treatment and control groups allowed for more observations, but the time lag of using piped water was short.

The first step in the difference-in-differences with propensity score matching is to estimate propensity scores, which is the probability of having piped water. The control variables are household characteristics in base line years (i.e., 2002 for the first way and 2004 for the second way of constructing treatment and control groups). These variables include household demographic variables, household heads' characteristics, education, land and housing, and regional and urban variables. The probit models to estimate the propensity score are reported in Appendix Table A.1. The predicted propensity scores are graphed in Appendix Figure A.1 and A.2.[8] These show that many households in the treatment and control groups have a similar propensity score (large common support).

[8] We also conduct balancing tests of equality of covariate variables between the treatment group (households with piped water) and the matched control group (households with piped water). For most variables, we cannot reject the hypothesis on the equality of variable means between the treatment and matched control groups. The balancing tests are presented in Tables in Appendix.

Difference-in-differences regression

In addition to the matching, we also ran a standard difference-in-differences regression to examine the impact estimate of piped water to different estimation strategies. The difference-in-differences model is as follows:

$$\ln(Y_{it}) = \beta_0 + T_{it}\beta_1 + D_{it}\beta_2 + D_{it}T_{it}\beta_3 + X_{it}\beta_4 + \varepsilon_{it}\,, \qquad (8)$$

where Y_{it} is the outcome of household *i* in year *t*; T_i is the time dummy, which is equal to 1 for the year 2006 and 0 for the year 2002 (or 2004 if the panel data 2004-2006 are used); D_{it} is the dummy variable of piped water and X_{it} are observed characteristics for household *i* in year *t*. The effect of piped water (ATT) is measured by β_3 (the coefficient of interaction between *D* and *X*).

5.0 ESTIMATION RESULTS

As mentioned above, we used different matching estimators including nearest neighbors, kernel, and local linear regression matching. All the estimators produce very similar results. For interpretation, we used estimates from kernel matching with bandwidth of 0.01.[9]

Estimates of the impact of piped water using the 2002 to 2006 panel data are presented in Table 2. [10] For comparison and correction of inflation, income is adjusted to the price of January 2006. The difference in per capita income between the treatment and control group in the years 2006 and 2002 is 1,512.7 and 829.8 VND thousand, respectively. (ATT) is measured by the 2006 difference minus the 2002 difference. The estimates of ATT for household income and working hours are positive but not statistically significant.

[9] In this paper, estimates from other matching estimators are not presented, but they can be provided on request.

[10] Standard errors are computed using bootstrap techniques. Abadie and Imbens (2008) show that bootstrap produces invalid standard errors for the nearest neighbor matching estimator. However, there is no evidence on the validity of bootstrap in estimating standard errors for other matching estimators.

Table 2. Impact estimation using panel data from 2002 to 2006: kernel matching

Outcomes	In 2006			In 2002			ATT
	Treatment group	Control group	Difference 2006	Treatment group	Control group	Difference 2002	
Per capita annual income (VND thousand)	10391.3****	8878.6****	1512.7***	6782.4****	5952.6****	829.8*	682.9
	[650.8]	[568.2]	[685.1]	[524.1]	[306.7]	[489.2]	[659.3]
Annual working hours per working-age person	1465.7****	1402.6****	63.1	1539.3****	1510.2****	29.1	34.0
	[37.0]	[35.9]	[44.2]	[43.6]	[38.5]	[47.6]	[53.3]

Note: Definition of outcomes:

- 'Per capita annual income of a household' is equal to the total annual income of the household divided by the household size.
- 'Annual working hours per working-age person' is equal to the total annual working hours of household members divided by the number of household members from 15 to 60 years old.

Standard errors in brackets. Standard errors are calculated using bootstrap with 500 replications.
* significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Estimation from panel data of VHLSSs 2002-2006

Table 3 presents the impact estimates from the matching method using the 2004 to 2006 panel data. Similarly, the impact of piped water on income and labor supply is not statistically significant. The effect of piped water on sickness outcome is negative, but very small and not statistically significant.

Table 3. Impact estimation using panel data 2004 to 2006: kernel matching

Outcomes	In 2004			In 2002			ATT
	Treatment group	Control group	Difference 2004	Treatment group	Control group	Difference 2002	
Per capita annual income (VND thousand)	10631.0***	9125.1***	1505.8***	8920.0***	7595.3***	1324.6***	181.2
	[549.1]	[350.5]	[463.2]	[447.8]	[309.9]	[365.0]	[373.9]
Annual working hours per working-age person	1449.1***	1384.8***	64.2**	1489.1***	1451.9***	37.2	27.0
	[28.7]	[19.3]	[28.1]	[29.3]	[20.0]	[29.6]	[29.7]
Proportion of members sick during the past four weeks	0.2056***	0.2031***	0.0026	0.1407***	0.1262***	0.0145	-0.0120
	[0.0099]	[0.0067]	[0.0113]	[0.0102]	[0.0062]	[0.0112]	[0.0151]
Proportion of members sick during the past 12 months	0.3448***	0.3441***	0.0007	0.3288***	0.3141***	0.0147	-0.0140
	[0.0179]	[0.0099]	[0.0194]	[0.0166]	[0.0110]	[0.0178]	[0.0220]
The annual number of sickness days in bed per person	1.7833***	2.2655***	-0.4823*	1.6585***	1.9290***	-0.2705	-0.2118
	[0.2250]	[0.1744]	[0.2701]	[0.2072]	[0.1565]	[0.2538]	[0.3401]

Note: Definition of outcomes:

- 'Per capita annual income of a household' is equal to the total annual income of the household divided by the household size.
- 'Annual working hours per working-age person' is equal to the total annual working hours of household members divided by the number of household members from 15 to 60 years old.
- 'Proportion of members sick during the past four weeks' is the number of household members reporting 'sickness' during the past four weeks divided by the household size.
- 'Proportion of members sick during the past 12 months' is the number of household members reporting 'sickness' during the past 12 months divided by the household size.
- 'The annual number of sickness days in bed per person' is the number of sickness days in bed of all the household members divided by the household size

Standard errors in brackets. Standard errors are calculated using bootstrap with 500 replications.
* significant at 10%; ** significant at 5%; *** significant at 1%.
Source: Estimation from panel data of VHLSSs 2002-2006

The difference-in-differences regressions are also used to estimate the impact of piped water. The regressions are presented in Appendix Tables A.2 to A.4. Similar to the matching method, the impact of piped water on income, working hours, and sickness has the expected sign but not statistically significant.

Finally, it should be noted that the difference-in-differences estimators can still be biased if endogeneity is caused not only by time-invariant but also by time-variant unobserved variables. It is expected that the endogeneity bias caused by time-variant variables is relatively small. In addition, when we examined the sensitivity of the impact estimate of piped water, we found that adding more control variables tended to reduce the impact estimate. Controlling for time-invariant unobserved variables also reduces the impact estimate significantly. It implies that omitted variables may overestimate the impact of piped water. Thus, the unbiased estimate might be even smaller than the difference-in-differences estimates in this study.

6.0 CONCLUSION

Although clean water is essential for health and human development, many poor people still do not have access to clean water in Vietnam. Only around 23% of the households used tap water for drinking in 2006. Other households have to use water from wells, and some households still use drinking water from rivers, pond, and simple wells without any purification. Unclean water can cause diseases, health problems, and low working stamina (?), hence resulting to low income and consumption, thereby poverty. Poverty means low income, poor living conditions, and limited access to clean water.

In Vietnam, there is a strong spatial difference in water quality. The use of tap water and clean water is much more prevalent in delta regions such as the Red River Delta, South East, and Mekong River Delta than in mountainous regions such as the north, west, and northeast. Households with tap water for drinking was 62% for urban areas and 8.4% for rural areas in 2006. Disadvantaged households including poor and ethnic minority households have much lower access to tap and clean water than better-off households and Kinh/Chinese households.

This study aims to measure the effect of piped water on household welfare indicators including income, working effort, and sickness. Our findings are similar to the results of Devoto et al. (2011) that the effect of piped water on household income and labor supply is positive but not statistically significant. The effect of piped water on the sickness of household members is negligible and not statistically significant.

There are two possible explanations for the small effects. First, the difference in the quality of piped water and non-piped water may be small. The VHLSSs does not contain information on the cleanliness of the piped water. However, the media sometimes report that piped water is unclean (e.g., Xuan-Danh, 2011). Households using non-piped water may have also purified and boiled the water first to reduce the incidence of waterborne diseases. According to the 2006 VHLSS, around 86% of households always boiled their drinking water. Thus, piped water did not increase income and labor supply through improved health and time saving.

Second, the time duration in this study might be not long enough - four years between 2002 and 2006 and two years between 2004 and 2006 to see a significant effect of piped water on health and labor supply. In addition, data on labor supply measured by the annual working hours and income might have had large measurement errors that caused large standard errors.

Finally, despite the lack of significant effect of piped water on sickness, labor supply, and income, households still care greatly about the aesthetic and life-style benefits of piped water. They tend to prefer piped water over water from other sources. In addition, a better design for impact evaluation can give more informative results. When measuring the effect of water quality on household welfare, one should use a continuous indicator of water quality such as pollution or arsenic measures, which allow for more variation in water quality. Direct outcomes of piped water such as waterborne diseases should be used to detect the effect of piped water on health. The impact estimation will be more accurate if a randomization design or instrumental variables regressions with valid instruments are used. However, these issues are beyond the scope of this study but certainly important for future studies.

References

Abadie, A., and Imbens, G. W. (2008), "On the failure of the bootstrap for matching estimators," *Econometrica* 76 (6), 1537-1557.

Agusa Tetsuro, Takashi Kunito, Junko Fujihara, Reiji Kubota, Tu Binh Minh, Pham Thi Kim Trang, Hisato Iwata, Annamalai Subramanian, Pham Hung Viet and Shinsuke Tanabe (2006), "Contamination by arsenic and other trace elements in tube-well water and its risk assessment to humans in Hanoi, Vietnam," *Environmental Pollution* 139 (1), 95-106.

Cochran, W. G. (1968), "The effectiveness of adjustment by subclassification in removing bias in observational studies," *Biometrics* 24, 295-313.

Cochran, W. G. and Chambers, S. P. (1965), "The Planning of observational studies of human population," *Journal of the Royal Statistical Society* 128(2), 234-266.

Collin, J.F., Milet, J.J., Morlot, M. and Foliguet, J.M. (1981), "Eau d'adduction et astroentérites en Meurthe-et-Moselle, » *J. Franc. Hydrologie* 12, 155–174.

Dasgupta, P. (2004), "Valuing health damages from water pollution in urban Delhi, India: a health production function approach," *Environment and Development Economics* 9, 83–106.

Dehejia, Rajeev and Sadek Wahba (1998), "Propensity score matching methods for non-experimental causal studies," NBER Working Paper 6829. National Bureau of Economic Research, 1050 Massachusetts Ave., Cambridge, The United States.

Devoto, F., Duflo, E., Dupas, P., Parienté, W. and Pons, V. (2011), "Happiness on tap: piped water adoption in urban Morocco", Unpublished manuscript. The United States. Place where written or their organization?

Fewtrell, L., Kaufmann, R.B., Kay, D., Enanoria, W., Haller, L. and Colford Jr., J.M. (2009), "Water, sanitation, and hygiene interventions to reduce diarrhea in less developed countries: a systematic review and meta-analysis." *Lancet Infectious Diseases*, 5: 42-52.

Fewtrell, L. and R. Kaufmann (2005), "Water sanitation and hygiene interventions to reduce diarrhea in developed countries: A Systematic review and meta-analysis. *Lancet Infectious Diseases, 5*., 42-52

Foster, J., Greer, J. and Thorbecke, E. (1984), "A class of decomposable poverty measures," *Econometrica* 52, 761-765.

Galiani, S. Paul, G. and Ernesto, S. (2005), "Water for life: the impact of the privatization of water services on child mortality," *Journal of Political Economy*, 113(1), 83-120.

Gamper-Rabindran, Shanti, Shakeeb Khan and Timmins, C. (2010), "The impact of piped water provision on infant mortality in Brazil: A quantile panel data approach," *Journal of Development Economics 92*, 188-200

Gerking, S. and Stanley, L.R. (1986), "An economic analysis of air pollution and health: the case of St. Louis," *The Review of Economics and Statistics* 68, 115–121.

Günther, I. and Fink, G. (2010), "Water, sanitation and children's health. Evidence from 172 DHS surveys," World Bank Policy Research Working Paper, No. 5275. Place of publication: World Bank.

Heckman, J., Lalonde, R. and Smith, J. (1999), "The Economics and econometrics of active labor market programs" Chapter 31, pages 1865-2097In Ashenfelter, A. and D. Card (Eds.) *Handbook of Labor Economics, Volume 3*, Amsterdam: Elsevier Science.

Heckman, J., Hidehiko Ichimura and Todd, P. (1997), "Matching as an econometric evaluation estimators: evidence from evaluating a job training programme," *Review of Economic Studies* 64(4), 605- 654.

Hoang Trong Quang (1990), "Vietnam - Country status reports on water quality monitoring," In Water Quality Monitoring in the Asian [sic] and Pacific Region. ESCAP Bangkok. Water Resources Series No. 67, pages 193-199.

Jalan, J., and Ravallion, M. (2003), "Does piped water reduce diarrhea for children in rural India?," *Journal of Econometrics* 112(1), 153-173.

Komives K., Whittington, D. and Xun Wu (2003), "Infrastructure coverage and the poor: a global perspective", Chapter 5 in Brook, P.J. and Irwin, T.C. (Eds.) "Infrastructure for Poor People." Washington, D.C.:World Bank.

Le The Thu, Nguyen Van Ba and Nguyen Van Hiep (1993), "Quality of water sources in six provinces in South Vietnam," Research report, Ho Chi Minh City, Vietnam.

Le Tuan Xuan and Yukihiro Munekage (2004), "Residues of selected antibiotics in water and mud from shrimp ponds in mangrove areas in Viet Nam," *Marine Pollution Bulletin 49(*11-12), 922-92

Macinko, J., Frederico C. G. and Maria de Fátima Marinho de Souza (2005), "Evaluation of the impact of the Family Health Program on infant mortality in Brazil, 1990–2002," *Journal of Epidemiol Community Health,* 60(1), 3-4**.**

Ministry of Natural Resources and Environment of Vietnam (MONRE) (2007), "Lắng nghe tiếng nói của người nghèo." Hanoi, Vietnam: MONRE and UNDP.

MOH (2008), "Report on infectious diseases in Vietnam," Vietnam: Ministry of Health.

Nguyen Tat Ha, Nguyen Hung Long, Bui Van Truong, Hoang Thi Nghia, and Phung Thanh Van (1994), “Chemical analysis of drinking water sources in Hai Hung, Hanoi, and Vinh Phu provinces, Hanoi: Ministry of Health.

Rosenbaum, P. R. and Rubin, R. (1983), “The central role of the propensity score in observational studies for causal effects,” *Biometrika* 70(1), 41-55.

Rubin, D. B. (1977), “Assignment to a treatment group on the basis of a covariate,” *Journal of Educational Statistics* 2(1), 1-26.

Rubin, D. B. (1979), “Using multivariate sampling and regression adjustment to control bias in observational studies,” *Journal of the American Statistical Association*, 74, 318–328.

Rubin, D. B. (1980), “Bias reduction using mahalanobis-metric matching,” *Biometrics* 36(2), 293–298.

Smith, J. and Todd, P. (2005), “Does matching overcome LaLonde’s critique of nonexperimental estimators?,” *Journal of Econometrics* 125(1–2), 305–353.

Suc Khoe Newspaper (2010), "Nhiều bệnh truyền nhiễm có tỷ lệ mắc cao do vệ sinh kém", 24/3/2010, http://suckhoedoisong.vn.

UNDP (2006), Human Development Report 2006. New York, The United States: United Nations Development Programme.

UNICEF (2010), “Water, Sanitation and Hygiene”, http://www.unicef.org/wash/

Waddington. H., Birte Snilstveit, White, H. and Fewtrell, L. (2009), “Water, sanitation and hygiene interventions to combat childhood diarrhoea in developing countries,” Synthetic Review 001. India: International Initiative for Impact Evaluation.

World Bank (2000), *Vietnam Development Report 2000: Attacking Poverty*, Hanoi, Vietnam: World Bank.

World Bank (2004), *Vietnam Development Report 2004: Poverty*, Hanoi, Vietnam: World Bank.

World Health Organization (2000), “Global water supply and sanitation assessment 2000 report.” Geneva, Switzerland: WHO.

World Health Organization (2004), “Meeting the MDG Drinking Water and Sanitation: A Mid-Term Assessment of Progress,” Geneva, Switzerland: WHO.

Xuan-Danh (2011), “Nước máy lại nhiễm bẩn”, Laborer Newspaper, on 6/1/2011. http://www.baomoi.com/Home/PhapLuat/nld.com.vn/Nuoc-may-lai-nhiem-ban/5501597.epi.

Xuan-Long (2010), “80% trường hợp bệnh tật do nguồn nước ô nhiễm", 24/03/2010, Youth Newspaper, http://tuoitre.vn.

Appendices

Propensity score matching estimators

The control group is constructed by matching each participant (i.e., households with piped water) *i* in the treatment group with one or more non-participants (i.e., households without piped water) *j* whose propensity scores are closest to the propensity score of the participant *i*. For a participant *i*, denote n_{ic} as the number of non-participants *j* who are matched with this participant, and *w(i,j)* the weight attached to the outcome of each non-participant. These weights are defined non-negative and sum up to 1, i.e.:

$$\sum_{j=1}^{n_{ic}} w(i, j) = 1 . \tag{A.1}$$

The estimator of ATT is expressed as follows:

$$A\hat{T}T = \frac{1}{n_1}\left\{\sum_{i=1}^{n_1}\left[Y_{1i}^A - \sum_{j=1}^{n_{ic}} w(i, j)Y_{0j}^A\right] - \sum_{i=1}^{n_1}\left[Y_{0i}^B - \sum_{j=1}^{n_{ic}} w(i, j)Y_{0j}^B\right]\right\} \tag{A.2}$$

where n_1 is the number of the participants in the data sample. Y_{1i}^A and Y_{0j}^A are the observed outcomes of participant *i* and matched non-participant *j* after the access to piped water (in 2006), respectively. Y_{0i}^B and Y_{0j}^B are the observed outcomes of participant *i* and non-participant *j* before the access of piped water (in 2002 or 2004), respectively.

If each participant is matched with the one non-participant with the minimum value of *d(i,j)* (where *d(i,j)* is the distance between the propensity scores of participant *i* and that of non-participant *j*), the weight *w(i,j)* equals 1 for all pairs of matches. This is called one nearest neighbor matching. When more than one non-participants are matched with each participant (or vice-versa), we need some ways to define the weights attached to each non-participant.

A number of methods use equal weights for all matches. N-nearest neighbors matching involves matching each participant with n non-participants who have the closest propensity scores. Each matched non-participant will receive weight

$w(i,j) = 1/n$.[11] However, it could be reasonable to assign different weights to different non-participants depending on metric distances between their covariates and the covariates of the matched participant (e.g., Heckman et al., 1997; Smith and Todd, 2005). The kernel matching method matches a participant with one or many non-participants depending on a kernel function G and a selected bandwidth h. The weight is defined as:

$$w(i,j) = \frac{G\left[\frac{d(i,j)}{h}\right]}{\sum_{k=1}^{n_2} G\left[\frac{d(i,k)}{h}\right]}, \quad \text{(A.3)}$$

where n_2 is the number of the non-participants in the data sample. In this paper, we use kernel with bandwidth of 0.01, and the kernel function is the Epanechnikov one:

$$\begin{aligned} G(z) &= 0.75(1-z^2) \text{ if } -1 \le z \le 1 \\ &= 0 \text{ if } |z| > 1, \end{aligned} \quad \text{(A.4)}$$

[11] Caliper matching (e.g., Dehejia and Wahba, 1998; Smith and Todd, 2005) uses equal weights for matched subjects whose distance *d(i,j)* is smaller than a specific value, say 0.05 or 0.1. This criterion aims to ensure the quality of matching. Stratification (interval) matching divides the range of estimated distances into several strata (blocks) of equal ranges. Within each stratum, a participant is matched with all non-participants with equal weights (e.g., Dehejia and Wahba, 1998; Smith and Todd, 2005).

Appendix Table A.1. Probit regressions of access to piped water in 2002
(dependent variable is 1 if households having piped water, and 0 otherwise)

Explanatory variables	Panel data 2002-2006		Panel data 2004-2006	
	Coefficient	Std. Err.	Coefficient	Std. Err.
Head's age	0.0021	0.0031	0.0001	0.0017
Head's gender (male = 1; female =0)	-0.0858	0.0698	-0.1353***	0.0401
Head without education degree	Omitted			
Head with primary education degree	0.0156	0.0756	0.0425	0.0439
Head with lower secondary degree	-0.0783	0.0872	0.0070	0.0508
Head with upper secondary degree	0.0960	0.1131	0.1808***	0.0696
Head with technical degree	-0.2190	0.1918	0.1744**	0.0833
Head with post secondary degree	-0.5971***	0.2522	-0.0061	0.1308
Ethnic minorities (yes = 1)	0.0972	0.1109	0.0214	0.0571
Household size	-0.0241	0.0195	-0.0057	0.0106
Proportion of members over 60	-0.4516	0.1793	0.0429	0.0888
Proportion of members under 16	0.0988	0.1607	-0.0579	0.0904
Proportion of female members	-0.0327	0.1412	-0.1252	0.0831
Proportion of members having technical degree	0.5819	0.3580	0.0930	0.1540
Proportion of members having post-secondary degree	0.9349**	0.4354	0.1055	0.2221
Solid house	Omitted			
Semi-solid house	-0.2175***	0.0801	-0.1607***	0.0440
Temporary house	-0.2979***	0.1011	-0.2631***	0.0582
Having flush toilet	Omitted			
Having toilet	-0.4405***	0.0765	-0.3270***	0.0417
Not having toilet	-0.2694***	0.0974	-0.1483***	0.0569
Annual crop land (1000 m^2)	-0.0078	0.0052	-0.0056	0.0055
Perennial crop land (1000 m^2)	0.0028	0.0048	-0.0247*	0.0139
Urban (yes=1)	0.9011***	0.0654	0.7197***	0.0375
Red River Delta	Omitted			
North East	-0.4534***	0.1190	-0.0813	0.0616
North West	-0.7635***	0.2280	-0.1584	0.1009
North Central Coast	-0.2773**	0.1086	-0.1728***	0.0646
South Central Coast	-0.1023	0.1077	-0.0713	0.0655
Central Highlands	-0.7658***	0.1691	-0.1861**	0.0807
South East	-0.0766	0.1006	0.1477**	0.0582
Mekong River Delta	0.3109***	0.0895	0.4572***	0.0520
Constant	-1.1829***	0.2225	-1.2011***	0.1295
Observations	8333		18005	
Pseudo R2	0.16		0.114	

Note: For definition of dependent variables, see the note in Table 3.
Robust standard errors (corrected for sampling weight and cluster correlation).
* significant at 10%; ** significant at 5%; *** significant at 1%
Source: Estimation VHLSSs 2002, 2004 and 2006

Appendix Table A.2. Difference-in-differences regressions: panel data 2002-2006

Explanatory variables	Dependent variables			
	Log of per capita income	Per capita income	Log of working hours per people above 14	Working hours per people above 14
Drinking piped water (yes = 1) * time dummy (2006 = 1)	-0.0843	683.66	0.0425	58.14
	[0.0807]	[1,027.18]	[0.0424]	[54.30]
Drinking piped water (yes = 1)	0.1020**	507.29	0.012	8.21
	[0.0417]	[378.66]	[0.0326]	[44.33]
Time dummy (2006 = 1)	0.2313***	1,247.01***	-0.0750***	-115.70***
	[0.0128]	[98.60]	[0.0130]	[15.44]
Head's age	0.0055**	43.52***	0.0123***	11.86***
	[0.0027]	[16.72]	[0.0031]	[2.80]
Head's age squared	-0.0001**	-0.43***	-0.0001***	-0.13***
	[0.0000]	[0.16]	[0.0000]	[0.03]
Gender of head (male=1)	-0.0135	-91.66	0.0443***	62.90***
	[0.0156]	[146.85]	[0.0154]	[17.42]
Head without education degree	Omitted			
Head with primary education degree	0.1101***	715.69***	-0.0168	-16.4
	[0.0141]	[112.81]	[0.0140]	[15.52]
Head with lower secondary degree	0.1768***	897.25***	-0.0321*	-26.41
	[0.0172]	[117.68]	[0.0170]	[19.91]
Head with upper secondary degree	0.2687***	1,503.02***	-0.0465**	-37.17
	[0.0226]	[280.31]	[0.0204]	[26.55]
Head with technical degree	0.1311***	440.73	-0.1639***	-165.02***
	[0.0321]	[270.70]	[0.0320]	[38.38]
Head with post secondary degree	0.1460***	149.17	-0.1860***	-197.42***
	[0.0480]	[685.99]	[0.0517]	[51.81]
Ethnic minorities (yes=1)	-0.2329***	-1,048.19***	-0.0025	-17.49
	[0.0238]	[122.10]	[0.0244]	[31.22]
Household size	-0.0699***	-661.75***	-0.0423***	-93.87***
	[0.0118]	[133.38]	[0.0111]	[11.53]
Household size squared	0.0013	25.18**	0.0015*	4.05***
	[0.0010]	[11.88]	[0.0008]	[0.88]
Proportion of members under 16	-0.2647***	-1,274.74***	-0.7268***	-736.60***
	[0.0344]	[293.71]	[0.0423]	[37.11]
Proportion of members over 60	-0.5140***	-2,111.15***	0.5769***	849.07***
	[0.0308]	[318.65]	[0.0279]	[36.60]
Proportion of female members	-0.1720***	-1,262.15***	-0.0391	-61.53*
	[0.0283]	[249.93]	[0.0287]	[33.38]
Proportion of members with technical degree	0.7636***	4,971.49***	0.3261***	351.63***
	[0.0595]	[600.62]	[0.0568]	[70.20]
Proportion of members having post-secondary degree	0.9678***	8,930.23***	0.0674	153.62
	[0.1030]	[1,487.84]	[0.1103]	[105.68]
Solid house	Omitted			

Explanatory variables	Dependent variables			
	Log of per capita income	Per capita income	Log of working hours per people above 14	Working hours per people above 14
Semi-solid house	-0.1304***	-603.27***	0.022	18.67
	[0.0200]	[194.51]	[0.0166]	[20.57]
Temporary house	-0.3162***	-1,417.12***	0.0038	-6.2
	[0.0240]	[215.07]	[0.0223]	[26.59]
Having flush toilet	Omitted			
Having toilet	-0.2781***	-2,623.67***	-0.019	-29.3
	[0.0208]	[272.41]	[0.0181]	[23.08]
Not having toilet	-0.3661***	-2,862.15***	-0.0414*	-44.64
	[0.0275]	[272.71]	[0.0246]	[29.99]
Annual crop land (1000m2)	0.0127***	92.83***	-0.0024**	-2.72**
	[0.0013]	[12.67]	[0.0010]	[1.18]
Perennial crop land (1000m2)	0.0122***	156.04***	-0.0009	-0.63
	[0.0014]	[42.44]	[0.0009]	[1.11]
Urban (yes=1)	0.1450***	756.15***	0.0574**	61.13**
	[0.0275]	[220.93]	[0.0235]	[28.61]
Red River Delta	Omitted			
North East	0.0811***	254.38	0.0864***	98.23**
	[0.0293]	[195.09]	[0.0298]	[40.39]
North West	-0.0728	-343.73	0.0384	32.00
	[0.0450]	[270.00]	[0.0407]	[57.52]
North Central Coast	-0.1042***	-438.14**	-0.0406	-65.91*
	[0.0301]	[202.36]	[0.0291]	[35.19]
South Central Coast	0.0919***	6.62	0.0111	-7.89
	[0.0322]	[205.87]	[0.0309]	[39.01]
Central Highlands	0.027	23.36	-0.0124	-37.37
	[0.0374]	[372.73]	[0.0366]	[48.97]
South East	0.3878***	2,012.31***	0.0254	22.57
	[0.0363]	[292.29]	[0.0340]	[45.17]
Mekong River Delta	0.3640***	1,917.09***	-0.1953***	-230.97***
	[0.0286]	[231.88]	[0.0315]	[36.82]
Constant	8.7512***	8,642.54***	7.0585***	1,491.27***
	[0.0871]	[653.98]	[0.0850]	[86.15]
Observations	16666	16666	16666	16666
R-squared	0.41	0.23	0.24	0.25

Note: For definition of dependent variables, see the note in Table 3.
Robust standard errors in brackets (corrected for sampling weight and cluster correlation).
* significant at 10%; ** significant at 5%; *** significant at 1%
Source: Estimation from panel data of VHLSSs 2002-2006

Appendix Table A.3. Difference-in-differences regressions: panel data 2004-2006

Explanatory variables	Dependent variables			
	Log of per capita income	Per capita income	Log of working hours per people above 14	Working hours per people above 14
Drinking piped water (yes = 1) * time dummy (2006 = 1)	0.0073	650.81*	0.0507	39.65
	[0.0202]	[349.12]	[0.0419]	[27.50]
Drinking piped water (yes = 1)	0.0794***	754.56**	-0.0139	17.26
	[0.0212]	[303.55]	[0.0371]	[25.25]
Time dummy (2006 = 1)	0.0461***	51.83	-0.0872***	-43.30***
	[0.0067]	[101.28]	[0.0116]	[8.80]
Head's age	0.0011	15.92	0.0492***	9.57***
	[0.0021]	[23.31]	[0.0063]	[2.22]
Head's age squared	0.0000	-0.10	-0.0005***	-0.11***
	[0.0000]	[0.22]	[0.0001]	[0.02]
Gender of head (male=1)	-0.0079	32.55	0.0704***	70.57***
	[0.0102]	[110.05]	[0.0186]	[11.51]
Head without education degree	Omitted			
Head with primary education degree	0.1224***	730.61***	-0.0173	-5.02
	[0.0099]	[93.38]	[0.0196]	[11.60]
Head with lower secondary degree	0.1908***	1,104.94***	-0.0616***	-12.17
	[0.0123]	[122.36]	[0.0233]	[14.55]
Head with upper secondary degree	0.2730***	1,939.18***	-0.0831***	-45.27**
	[0.0172]	[253.30]	[0.0274]	[19.38]
Head with technical degree	0.1745***	877.95***	-0.1957***	-154.63***
	[0.0206]	[279.32]	[0.0347]	[23.30]
Head with post secondary degree	0.2046***	583.53	-0.3372***	-216.78***
	[0.0306]	[656.82]	[0.0583]	[34.12]
Ethnic minorities (yes=1)	-0.2464***	-882.00***	-0.0081	-22.97
	[0.0212]	[141.26]	[0.0234]	[24.52]
Household size	-0.0779***	-889.50***	0.2790***	-89.08***
	[0.0074]	[87.04]	[0.0306]	[9.16]
Household size squared	0.0021***	41.95***	-0.0228***	3.42***
	[0.0006]	[7.46]	[0.0028]	[0.74]
Proportion of members under 16	-0.2884***	-1,795.04***	-1.8472***	-798.25***
	[0.0228]	[241.64]	[0.0845]	[27.18]
Proportion of members over 60	-0.4868***	-2,306.11***	0.3946***	863.14***
	[0.0221]	[229.99]	[0.0429]	[27.18]
Proportion of female members	-0.1651***	-1,061.52***	-0.2722***	-74.92***
	[0.0191]	[202.70]	[0.0522]	[22.62]
Proportion of members with technical degree	0.6855***	4,919.42***	0.3133***	285.99***
	[0.0382]	[492.17]	[0.0900]	[47.27]
Proportion of members having post-secondary degree	1.0167***	11,255.05***	0.4813***	291.96***
	[0.0573]	[1,169.36]	[0.1224]	[69.69]
Solid house	Omitted			
Semi-solid house	-0.1796***	-1,574.00***	0.0191	12.10

Explanatory variables	Dependent variables			
	Log of per capita income	Per capita income	Log of working hours per people above 14	Working hours per people above 14
	[0.0123]	[173.67]	[0.0215]	[12.85]
Temporary house	-0.3948***	-2,673.96***	-0.0335	-19.57
	[0.0160]	[191.24]	[0.0287]	[17.57]
Having flush toilet	Omitted			
Having toilet	-0.2677***	-2,620.95***	0.0129	-85.45***
	[0.0129]	[161.40]	[0.0231]	[14.01]
Not having toilet	-0.3956***	-3,197.44***	-0.0075	-89.16***
	[0.0179]	[168.81]	[0.0314]	[18.93]
Annual crop land (1000 m^2)	0.0120***	130.65***	-0.0012	-3.49***
	[0.0008]	[23.69]	[0.0009]	[0.80]
Perennial crop land (1000 m^2)	0.0078***	101.27***	0.0009	-0.31
	[0.0017]	[35.24]	[0.0006]	[0.92]
Urban (yes=1)	0.1058***	954.64***	0.0333	79.37***
	[0.0182]	[220.39]	[0.0289]	[21.08]
Red River Delta	Omitted			
North East	0.1200***	722.98***	0.1351***	111.28***
	[0.0227]	[173.30]	[0.0316]	[27.32]
North West	0.0431	294.19	0.0476	38.13
	[0.0354]	[218.00]	[0.0483]	[49.36]
North Central Coast	-0.1113***	-399.93***	-0.0375	-50.89**
	[0.0216]	[149.40]	[0.0360]	[23.85]
South Central Coast	0.0991***	460.87**	0.0345	-61.98**
	[0.0232]	[185.25]	[0.0381]	[26.75]
Central Highlands	0.2182***	1,345.06***	0.0434	34.85
	[0.0266]	[266.91]	[0.0371]	[30.57]
South East	0.5096***	4,001.19***	0.0440	127.84***
	[0.0258]	[336.79]	[0.0373]	[28.07]
Mekong River Delta	0.3904***	2,593.26***	-0.1014***	-127.82***
	[0.0201]	[201.26]	[0.0342]	[24.87]
Constant	9.0582***	11,102.29***	5.4607***	1,516.24***
	[0.0628]	[676.99]	[0.1739]	[70.57]
Observations	36010	36010	36010	36010
R-squared	0.43	0.27	0.33	0.26

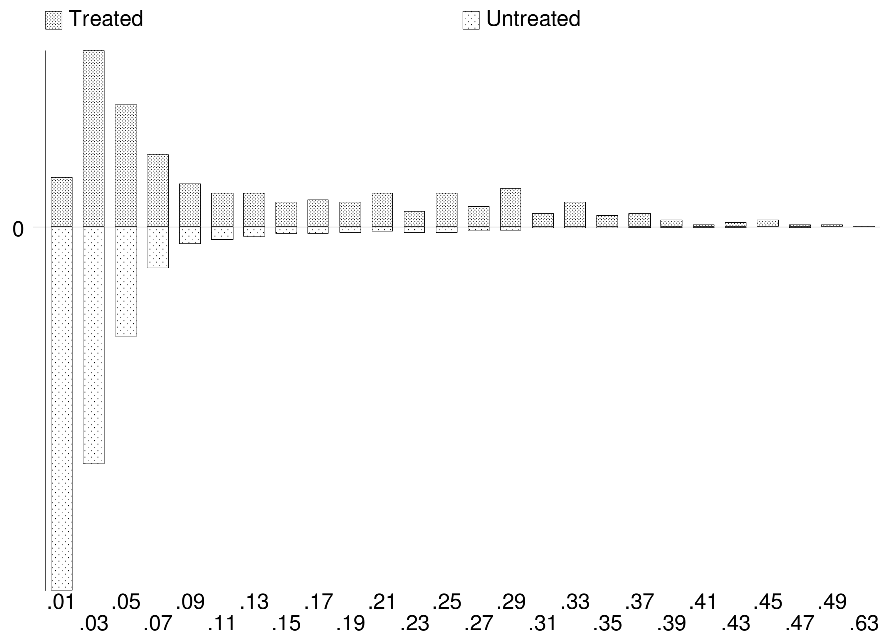
Note: For definition of dependent variables, see the note in Table 3.
Robust standard errors in brackets (corrected for sampling weight and cluster correlation)
* significant at 10%; ** significant at 5%; *** significant at 1%
Source: Estimation from panel data of VHLSSs 2002-2006

Appendix Table A.4. Difference-in-differences regressions: panel data 2004-2006

Explanatory variables	Dependent variables: Proportion of members sick during the past four weeks	Proportion of members sick during the past 12 months	The annual number of sick-days per person	The annual number of sick-days in bed per person
Drinking piped water (yes = 1) * time dummy (2006 = 1)	-0.0068	-0.0202	0.2240	0.0128
	[0.0139]	[0.0208]	[0.4771]	[0.2898]
Drinking piped water (yes = 1)	0.0142	0.0169	0.0097	-0.2886
	[0.0105]	[0.0175]	[0.3661]	[0.2312]
Time dummy (2006 = 1)	0.0756***	0.0356***	-0.1310	0.1021
	[0.0039]	[0.0057]	[0.1601]	[0.1265]
Head’s age	-0.0001	0.0002	-0.0776	-0.3731***
	[0.0009]	[0.0012]	[0.0556]	[0.0727]
Head’s age squared	0.0000	0.0000	0.0012**	0.0039***
	[0.0000]	[0.0000]	[0.0006]	[0.0007]
Gender of head (male=1)	-0.0140***	0.0122**	-0.2881	-0.4142**
	[0.0043]	[0.0060]	[0.2145]	[0.1802]
Head without education degree	Omitted			
Head with primary education degree	-0.0030	-0.0109*	0.4143*	0.1015
	[0.0049]	[0.0061]	[0.2423]	[0.1849]
Head with lower secondary degree	-0.0131**	-0.0134*	0.0081	0.2311
	[0.0057]	[0.0072]	[0.2423]	[0.1941]
Head with upper secondary degree	-0.0277***	-0.0131	-0.6813***	0.2047
	[0.0070]	[0.0101]	[0.2557]	[0.2372]
Head with technical degree	-0.0039	-0.0049	-0.1022	0.1758
	[0.0085]	[0.0113]	[0.3714]	[0.3219]
Head with post secondary degree	-0.0190	0.0091	-0.0069	0.7128*
	[0.0121]	[0.0195]	[0.5248]	[0.3958]
Ethnic minorities (yes=1)	-0.0065	-0.0103	-0.6557*	-0.3148**
	[0.0081]	[0.0119]	[0.3465]	[0.1538]
Household size	-0.0413***	-0.0425***	-1.3249***	-0.3463**
	[0.0042]	[0.0049]	[0.1899]	[0.1459]
Household size squared	0.0025***	0.0020***	0.0786***	0.0287**
	[0.0004]	[0.0004]	[0.0154]	[0.0121]
Proportion of members under 16	0.1516***	0.0349**	6.3067***	3.7236***
	[0.0120]	[0.0138]	[0.8581]	[0.5520]
Proportion of members over 60	0.0437***	0.0511***	-1.7523***	-1.3225***
	[0.0091]	[0.0122]	[0.4267]	[0.2935]
Proportion of female members	0.0343***	0.0315***	0.0165	-0.3621
	[0.0093]	[0.0115]	[0.4921]	[0.4749]
Proportion of members with technical degree	0.0098	0.0461**	0.8532	-0.6030
	[0.0176]	[0.0225]	[1.0090]	[0.6169]
Proportion of members having post-secondary degree	-0.0492**	-0.0092	-1.9984*	-0.7421
	[0.0223]	[0.0328]	[1.0658]	[1.3534]
Solid house	Omitted			
Semi-solid house	0.0070	-0.0062	0.4366**	0.5552***

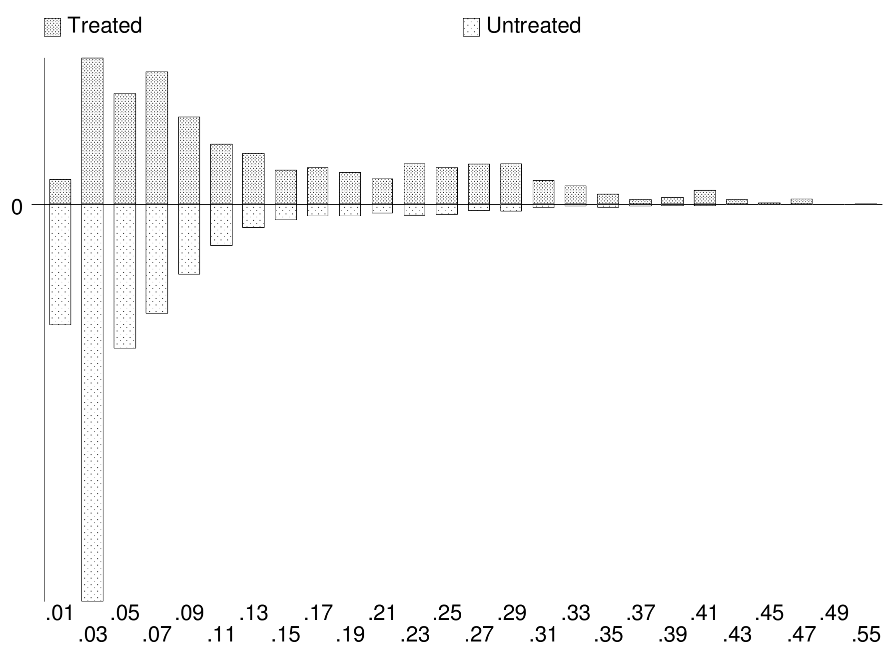
Explanatory variables	Dependent variables			
	Proportion of members sick during the past four weeks	Proportion of members sick during the past 12 months	The annual number of sick-days per person	The annual number of sick-days in bed per person
	[0.0045]	[0.0071]	[0.1871]	[0.1280]
Temporary house	0.0214***	-0.0318***	1.5341***	0.9846***
	[0.0066]	[0.0092]	[0.3359]	[0.2075]
Having flush toilet	Omitted			
Having toilet	0.0177***	0.0159*	0.7904***	0.1077
	[0.0048]	[0.0082]	[0.2098]	[0.1577]
Not having toilet	0.0181**	0.0422***	1.7934***	0.1526
	[0.0076]	[0.0116]	[0.3598]	[0.2662]
Annual crop land (1000 m^2)	-0.0003	0.0007	-0.0156**	-0.0148***
	[0.0003]	[0.0005]	[0.0079]	[0.0050]
Perennial crop land (1000 m^2)	-0.0001	0.0001	0.0030	-0.0031
	[0.0002]	[0.0004]	[0.0103]	[0.0042]
Urban (yes=1)	0.0026	0.0028	-0.4730	-0.1795
	[0.0087]	[0.0138]	[0.2886]	[0.1834]
Red River Delta	Omitted			
North East	0.0150	-0.0416***	-0.6566*	-0.5161**
	[0.0096]	[0.0131]	[0.3608]	[0.2212]
North West	0.0217	-0.0521***	0.5104	0.0969
	[0.0145]	[0.0185]	[0.5630]	[0.3321]
North Central Coast	-0.0052	-0.0391***	0.1969	-0.1509
	[0.0090]	[0.0138]	[0.4673]	[0.2546]
South Central Coast	-0.0031	0.0099	-0.6802	0.0701
	[0.0102]	[0.0178]	[0.4711]	[0.3830]
Central Highlands	0.0682***	0.0960***	1.0577**	0.3411
	[0.0127]	[0.0185]	[0.4499]	[0.3101]
South East	0.0699***	0.1389***	-0.7969*	-0.5756**
	[0.0114]	[0.0196]	[0.4704]	[0.2786]
Mekong River Delta	0.0333***	0.1049***	-0.9255**	-0.5824**
	[0.0089]	[0.0154]	[0.3951]	[0.2663]
Constant	0.1658***	0.3429***	8.8418***	10.9909***
	[0.0270]	[0.0359]	[1.5008]	[1.9117]
Observations	36010	36010	36010	36010
R-squared	0.10	0.06	0.06	0.04

Note: For definition of dependent variables, see the note in Table 3.
Robust standard errors in brackets (corrected for sampling weight and cluster correlation)
* significant at 10%; ** significant at 5%; *** significant at 1%
Source: Estimation from panel data of VHLSSs 2002-2006



Appendix Figure A. 1. Propensity scores: panel data 2002-2006

Source: Estimation from panel data of VHLSSs 2002-2006



Appendix Figure A.2. Propensity scores: panel data 2004-2006

Source: Estimation from panel data of VHLSSs 2004-2006

Appendix Table A.5. Description of covariate variables in 2002 for matching

Variables	Meaning	Type	Mean	Std. Dev.
ethnic02	Ethnic minorities (yes = 1)	Binary	0.132	0.338
hhsize02	Household size	Discrete	4.547	1.793
pelderly02	Proportion of members over 60	Continuous	0.114	0.238
pchild02	Proportion of members under 16	Continuous	0.305	0.219
rtechnical02	Proportion of members having technical diploma	Continuous	0.026	0.099
rposecond02	Proportion of members having post-secondary diploma	Continuous	0.013	0.069
age02	Age of household head	Discrete	47.248	14.171
headedu1	Head without education degree	Binary	0.334	0.472
headedu2	Head with primary education degree	Binary	0.256	0.437
headedu3	Head with lower secondary degree	Binary	0.268	0.443
headedu4	Head with upper secondary degree	Binary	0.076	0.265
headedu5	Head with technical degree	Binary	0.047	0.212
headedu6	Head with post secondary degree	Binary	0.019	0.135
anualand02	Annual crop land (10000m2)	Continuous	0.346	0.608
pereland02	Perennial crop land (10000m2)	Continuous	0.128	0.487
aqualand02	Water surface (10000m2)	Continuous	0.025	0.207
region1	Red River Delta	Binary	0.235	0.424
region2	North East	Binary	0.123	0.329
region3	North West	Binary	0.031	0.173
region4	North Central Coast	Binary	0.154	0.361
region5	South Central Coast	Binary	0.090	0.286
region6	Central Highlands	Binary	0.059	0.235
region7	South East	Binary	0.121	0.326
region8	Mekong River Delta	Binary	0.188	0.391
urban02	Urban (yes =1)	Binary	0.130	0.336

Source: Estimation from the 2002 VHLSS

Appendix Table A.6. Balancing test: one nearest neighbor matching

Variable	Sample	Treated	Control	%bias	bias	t	p>t
ethnic02	Unmatched	0.072	0.174	-31.2		-5.1	0.000
	Matched	0.072	0.091	-5.8	81.5	-0.9	0.362
hhsize02	Unmatched	4.335	4.610	-15.4		-2.8	0.005
	Matched	4.335	4.405	-3.9	74.7	-0.6	0.580
pelderly02	Unmatched	0.083	0.113	-14.2		-2.5	0.014
	Matched	0.083	0.085	-0.7	95.0	-0.1	0.916
pchild02	Unmatched	0.287	0.313	-12.0		-2.3	0.025
	Matched	0.287	0.282	2.4	79.6	0.3	0.738
rtechnical02	Unmatched	0.048	0.024	19.9		4.7	0.000
	Matched	0.048	0.038	8.0	60.0	0.9	0.356
rposecond02	Unmatched	0.031	0.011	22.0		5.5	0.000
	Matched	0.031	0.024	7.5	65.7	0.9	0.362
age02	Unmatched	47.15	46.96	1.3		0.3	0.803
	Matched	47.15	48.13	-7.0	-422.9	-1.0	0.345
headedu1	Unmatched	0.316	0.349	-6.9		-1.3	0.199
	Matched	0.316	0.332	-3.4	50.4	-0.5	0.648
headedu2	Unmatched	0.265	0.258	1.8		0.3	0.739
	Matched	0.265	0.255	2.4	-38.9	0.3	0.745
headedu3	Unmatched	0.204	0.262	-13.9		-2.5	0.012
	Matched	0.204	0.223	-4.4	68.0	-0.6	0.542
headedu4	Unmatched	0.102	0.070	11.2		2.3	0.021
	Matched	0.102	0.118	-5.7	49.0	-0.7	0.494
headedu5	Unmatched	0.078	0.044	14.1		3.1	0.002
	Matched	0.078	0.056	9.0	36.5	1.1	0.254
headedu6	Unmatched	0.035	0.017	11.4		2.6	0.010
	Matched	0.035	0.016	11.8	-4.2	1.6	0.113
anualand02	Unmatched	0.212	0.386	-27.3		-5.1	0.000
	Matched	0.212	0.301	-13.9	49.1	-2.0	0.047
pereland02	Unmatched	0.119	0.143	-3.5		-0.9	0.366
	Matched	0.119	0.111	1.2	64.3	0.2	0.860
aqualand02	Unmatched	0.002	0.027	-16.1		-2.2	0.028
	Matched	0.002	0.002	0.1	99.5	0.1	0.905
region1	Unmatched	0.214	0.211	0.9		0.2	0.871
	Matched	0.214	0.198	3.9	-356.8	0.5	0.597
region2	Unmatched	0.080	0.155	-23.4		-3.9	0.000
	Matched	0.080	0.086	-1.7	92.8	-0.3	0.796
region3	Unmatched	0.011	0.051	-23.6		-3.5	0.000
	Matched	0.011	0.013	-1.6	93.4	-0.3	0.744
region4	Unmatched	0.080	0.128	-15.5		-2.7	0.007
	Matched	0.080	0.059	7.0	54.7	1.1	0.262
region5	Unmatched	0.123	0.095	9.0		1.8	0.074
	Matched	0.123	0.118	1.7	80.8	0.2	0.827
region6	Unmatched	0.024	0.069	-21.4		-3.4	0.001
	Matched	0.024	0.029	-2.6	88.1	-0.4	0.659
region7	Unmatched	0.166	0.109	16.6		3.4	0.001
	Matched	0.166	0.164	0.8	95.3	0.1	0.923
region8	Unmatched	0.300	0.181	28.2		5.8	0.000
	Matched	0.300	0.332	-7.6	73.0	-0.9	0.357
urban02	Unmatched	0.469	0.110	86.1		20.9	0.000
	Matched	0.469	0.469	0.0	100.0	0.0	1.000

Source: Estimation from the 2002 VHLSS

Appendix Table A.7. Balancing test: five nearest neighbor matching

Variable	Sample	Treated	Control	%bias	bias	t	p>t
ethnic02	Unmatched	0.072	0.174	-31.2		-5.1	0.000
	Matched	0.072	0.078	-1.7	94.7	-0.4	0.671
hhsize02	Unmatched	4.335	4.610	-15.4		-2.8	0.005
	Matched	4.335	4.373	-2.1	86.4	-0.5	0.651
pelderly02	Unmatched	0.083	0.113	-14.2		-2.5	0.014
	Matched	0.083	0.095	-5.5	61.3	-1.3	0.213
pchild02	Unmatched	0.287	0.313	-12.0		-2.3	0.025
	Matched	0.287	0.278	4.0	66.7	0.9	0.397
rtechnical02	Unmatched	0.048	0.024	19.9		4.7	0.000
	Matched	0.048	0.044	3.2	84.0	0.6	0.551
rposecond02	Unmatched	0.031	0.011	22.0		5.5	0.000
	Matched	0.031	0.029	1.5	93.0	0.3	0.786
age02	Unmatched	47.15	46.96	1.3		0.3	0.803
	Matched	47.15	47.65	-3.6	-166.0	-0.7	0.456
headedu1	Unmatched	0.316	0.349	-6.9		-1.3	0.199
	Matched	0.316	0.327	-2.3	66.9	-0.5	0.631
headedu2	Unmatched	0.265	0.258	1.8		0.3	0.739
	Matched	0.265	0.252	3.0	-73.6	0.6	0.522
headedu3	Unmatched	0.204	0.262	-13.9		-2.5	0.012
	Matched	0.204	0.216	-2.8	79.9	-0.6	0.545
headedu4	Unmatched	0.102	0.070	11.2		2.3	0.021
	Matched	0.102	0.111	-3.3	71.1	-0.6	0.537
headedu5	Unmatched	0.078	0.044	14.1		3.1	0.002
	Matched	0.078	0.066	4.9	65.1	1.0	0.339
headedu6	Unmatched	0.035	0.017	11.4		2.6	0.010
	Matched	0.035	0.028	4.1	64.3	0.8	0.442
anualand02	Unmatched	0.212	0.386	-27.3		-5.1	0.000
	Matched	0.212	0.251	-6.2	77.5	-1.4	0.150
pereland02	Unmatched	0.119	0.143	-3.5		-0.9	0.366
	Matched	0.119	0.125	-0.8	76.8	-0.2	0.869
aqualand02	Unmatched	0.002	0.027	-16.1		-2.2	0.028
	Matched	0.002	0.002	-0.1	99.5	-0.2	0.880
region1	Unmatched	0.214	0.211	0.9		0.2	0.871
	Matched	0.214	0.190	5.9	-585.3	1.3	0.209
region2	Unmatched	0.080	0.155	-23.4		-3.9	0.000
	Matched	0.080	0.075	1.7	92.8	0.4	0.675
region3	Unmatched	0.011	0.051	-23.6		-3.5	0.000
	Matched	0.011	0.020	-5.3	77.6	-1.6	0.120
region4	Unmatched	0.080	0.128	-15.5		-2.7	0.007
	Matched	0.080	0.059	7.0	54.7	1.8	0.078
region5	Unmatched	0.123	0.095	9.0		1.8	0.074
	Matched	0.123	0.122	0.3	96.2	0.1	0.946
region6	Unmatched	0.024	0.069	-21.4		-3.4	0.001
	Matched	0.024	0.033	-4.1	80.9	-1.1	0.280
region7	Unmatched	0.166	0.109	16.6		3.4	0.001
	Matched	0.166	0.189	-6.7	59.6	-1.3	0.207
region8	Unmatched	0.300	0.181	28.2		5.8	0.000
	Matched	0.300	0.312	-2.7	90.6	-0.5	0.609
urban02	Unmatched	0.469	0.110	86.1		20.9	0.000
	Matched	0.469	0.479	-2.4	97.2	-0.4	0.670

Source: Estimation from the 2002 VHLSS

Appendix Table A.88. Balancing test: kernel matching with bandwidth of 0.01

Variable	Sample	Treated	Control	%bias	bias	t	p>t
ethnic02	Unmatched	0.072	0.174	-31.2		-5.1	0.000
	Matched	0.072	0.079	-2.2	93.1	-1.2	0.228
hhsize02	Unmatched	4.335	4.610	-15.4		-2.8	0.005
	Matched	4.335	4.362	-1.5	90.4	-0.7	0.482
pelderly02	Unmatched	0.083	0.113	-14.2		-2.5	0.014
	Matched	0.083	0.089	-2.9	79.5	-1.5	0.146
pchild02	Unmatched	0.287	0.313	-12.0		-2.3	0.025
	Matched	0.287	0.286	0.5	95.8	0.2	0.818
rtechnical02	Unmatched	0.048	0.024	19.9		4.7	0.000
	Matched	0.048	0.045	2.0	89.9	0.8	0.418
rposecond02	Unmatched	0.031	0.011	22.0		5.5	0.000
	Matched	0.031	0.032	-0.9	95.7	-0.4	0.721
age02	Unmatched	47.15	46.96	1.3		0.3	0.803
	Matched	47.15	47.44	-2.1	-54.3	-1.0	0.341
headedu1	Unmatched	0.316	0.349	-6.9		-1.3	0.199
	Matched	0.316	0.328	-2.5	63.8	-1.2	0.251
headedu2	Unmatched	0.265	0.258	1.8		0.3	0.739
	Matched	0.265	0.253	2.8	-58.9	1.3	0.202
headedu3	Unmatched	0.204	0.262	-13.9		-2.5	0.012
	Matched	0.204	0.209	-1.3	90.8	-0.6	0.545
headedu4	Unmatched	0.102	0.070	11.2		2.3	0.021
	Matched	0.102	0.103	-0.3	97.0	-0.1	0.886
headedu5	Unmatched	0.078	0.044	14.1		3.1	0.002
	Matched	0.078	0.071	2.6	81.4	1.1	0.275
headedu6	Unmatched	0.035	0.017	11.4		2.6	0.010
	Matched	0.035	0.035	-0.3	97.3	-0.1	0.904
anualand02	Unmatched	0.212	0.386	-27.3		-5.1	0.000
	Matched	0.212	0.244	-5.0	81.7	-2.6	0.009
pereland02	Unmatched	0.119	0.143	-3.5		-0.9	0.366
	Matched	0.119	0.116	0.4	87.5	0.2	0.840
aqualand02	Unmatched	0.002	0.027	-16.1		-2.2	0.028
	Matched	0.002	0.004	-1.0	93.5	-1.6	0.104
region1	Unmatched	0.214	0.211	0.9		0.2	0.871
	Matched	0.214	0.187	6.7	-677.5	3.1	0.002
region2	Unmatched	0.080	0.155	-23.4		-3.9	0.000
	Matched	0.080	0.077	0.9	96.0	0.5	0.610
region3	Unmatched	0.011	0.051	-23.6		-3.5	0.000
	Matched	0.011	0.017	-3.7	84.4	-2.5	0.013
region4	Unmatched	0.080	0.128	-15.5		-2.7	0.007
	Matched	0.080	0.068	4.0	74.0	2.1	0.032
region5	Unmatched	0.123	0.095	9.0		1.8	0.074
	Matched	0.123	0.133	-3.1	65.2	-1.3	0.184
region6	Unmatched	0.024	0.069	-21.4		-3.4	0.001
	Matched	0.024	0.034	-4.9	77.3	-2.8	0.006
region7	Unmatched	0.166	0.109	16.6		3.4	0.001
	Matched	0.166	0.184	-5.2	68.9	-2.1	0.033
region8	Unmatched	0.300	0.181	28.2		5.8	0.000
	Matched	0.300	0.299	0.3	98.9	0.1	0.894
urban02	Unmatched	0.469	0.110	86.1		20.9	0.000
	Matched	0.469	0.481	-2.8	96.8	-1.1	0.293

Source: Estimation from the 2002 VHLSS

Appendix Table A.9. Balancing test: local linear regression matching with bandwidth of 0.01

Variable	Sample	Treated	Control	%bias	bias	t	p>t
ethnic02	Unmatched	0.072	0.174	-31.2		-5.1	0.000
	Matched	0.072	0.091	-5.8	81.5	-0.9	0.362
hhsize02	Unmatched	4.335	4.610	-15.4		-2.8	0.005
	Matched	4.335	4.405	-3.9	74.7	-0.6	0.580
pelderly02	Unmatched	0.083	0.113	-14.2		-2.5	0.014
	Matched	0.083	0.085	-0.7	95.0	-0.1	0.916
pchild02	Unmatched	0.287	0.313	-12.0		-2.3	0.025
	Matched	0.287	0.282	2.4	79.6	0.3	0.738
rtechnical02	Unmatched	0.048	0.024	19.9		4.7	0.000
	Matched	0.048	0.038	8.0	60.0	0.9	0.356
rposecond02	Unmatched	0.031	0.011	22.0		5.5	0.000
	Matched	0.031	0.024	7.5	65.7	0.9	0.362
age02	Unmatched	47.15	46.96	1.3		0.3	0.803
	Matched	47.15	48.13	-7.0	-422.9	-1.0	0.345
headedu1	Unmatched	0.316	0.349	-6.9		-1.3	0.199
	Matched	0.316	0.332	-3.4	50.4	-0.5	0.648
headedu2	Unmatched	0.265	0.258	1.8		0.3	0.739
	Matched	0.265	0.255	2.4	-38.9	0.3	0.745
headedu3	Unmatched	0.204	0.262	-13.9		-2.5	0.012
	Matched	0.204	0.223	-4.4	68.0	-0.6	0.542
headedu4	Unmatched	0.102	0.070	11.2		2.3	0.021
	Matched	0.102	0.118	-5.7	49.0	-0.7	0.494
headedu5	Unmatched	0.078	0.044	14.1		3.1	0.002
	Matched	0.078	0.056	9.0	36.5	1.1	0.254
headedu6	Unmatched	0.035	0.017	11.4		2.6	0.010
	Matched	0.035	0.016	11.8	-4.2	1.6	0.113
anualand02	Unmatched	0.212	0.386	-27.3		-5.1	0.000
	Matched	0.212	0.301	-13.9	49.1	-2.0	0.047
pereland02	Unmatched	0.119	0.143	-3.5		-0.9	0.366
	Matched	0.119	0.111	1.2	64.3	0.2	0.860
aqualand02	Unmatched	0.002	0.027	-16.1		-2.2	0.028
	Matched	0.002	0.002	0.1	99.5	0.1	0.905
region1	Unmatched	0.214	0.211	0.9		0.2	0.871
	Matched	0.214	0.198	3.9	-356.8	0.5	0.597
region2	Unmatched	0.080	0.155	-23.4		-3.9	0.000
	Matched	0.080	0.086	-1.7	92.8	-0.3	0.796
region3	Unmatched	0.011	0.051	-23.6		-3.5	0.000
	Matched	0.011	0.013	-1.6	93.4	-0.3	0.744
region4	Unmatched	0.080	0.128	-15.5		-2.7	0.007
	Matched	0.080	0.059	7.0	54.7	1.1	0.262
region5	Unmatched	0.123	0.095	9.0		1.8	0.074
	Matched	0.123	0.118	1.7	80.8	0.2	0.827
region6	Unmatched	0.024	0.069	-21.4		-3.4	0.001
	Matched	0.024	0.029	-2.6	88.1	-0.4	0.659
region7	Unmatched	0.166	0.109	16.6		3.4	0.001
	Matched	0.166	0.164	0.8	95.3	0.1	0.923
region8	Unmatched	0.300	0.181	28.2		5.8	0.000
	Matched	0.300	0.332	-7.6	73.0	-0.9	0.357
urban02	Unmatched	0.469	0.110	86.1		20.9	0.000
	Matched	0.469	0.469	0.0	100.0	0.0	1.000

Source: Estimation from the 2002 VHLSS