Inequality in mortality in Vietnam during a period of rapid transition*

D Granlund[1], NT Chuc[2] , HD Phuc[3] and L Lindholm[4]

[1]Department of Economics, Umeå University, Sweden, and The Swedish Retail Institute (HUI), Stockholm, Sweden; [2]Faculty of Public Health, Hanoi Medical University, Vietnam; [3]Institute of Mathematics, Hanoi, Vietnam; [4]Department of Public Health and Clinical Medicine, Umeå University, Sweden

Post-refereeing-version The final version is published in: Social Science & Medicine 70, 232-239, 2010.

*The final publication is available at http://dx.doi.org/*10.1016/j.socscimed.2009.10.019.

Abstract

The associations between socioeconomic variables and mortality for 41,000 adults Vietnamese followed from January 1999 to March 2008 are estimated using Cox's proportionally hazard models. Also, we use decomposition techniques to investigate the relative importance of socioeconomic factors for explaining inequality in age-standardized mortality risk. The results confirm previously found negative association between mortality and income and education, for both men and women. We also found that marital status, at least for men, explain a large and growing part of the inequality. Finally, estimation results for relative education variables suggest that there exist positive spillover of education, meaning that that higher education of one's neighbors or spouse might reduce ones mortality risk.

Keywords: Health inequality; Socioeconomic status; Mortality risk; Decomposition; Vietnam

JEL classification: D30; D63; I10; I30

---

*We are grateful for comments and suggestions by Kurt Brännäs, Carl Lönnbark, two anonymous reviewers and the editor Stephen Birch. Corresponding author: (D Granlund) david.granlund@econ.umu.se.

1 Introduction

The literature about socioeconomic inequality in health is large and growing (Wildman, 2003; Wagstaff & van Doorslaer, 2004; Charasse-Pouélé & Fournier, 2006; Balia & Jones, 2008). The problem has got even more attention during the last year since the Commission on Social Determinants of Health launched their final report in 2008 (Marmot, 2008). The commission suggested three principles: improve the conditions of daily life, tackle the inequitable distribution of power, money and resources in general, and finally, measure the problem and evaluate actions (Marmot, 2008). This particular concern for health has been motivated by, among others, Amartya Sen: "In any discussion of social equity and justice, illness and health must figure as a major concern" (Sen, 2002, p. 659). There are many aspects on equity in health and health care, but we believe that Wagstaff and van Doorslaer were right when they argued that the most fundamental concern must be about the ultimate upstream variable – health, in this case measured as mortality (Wagstaff & van Doorslaer, 2000).

The majority of the literature about socioeconomic inequality in health is based on European and US data (Macintyre & Hunt, 1997). There is, however, also a growing literature of health inequality in Vietnam, including studies of the association between socioeconomic status and mortality (Huong, Minh, Janlert, Van, & Byass, 2006); mortality from cardiovascular diseases (Minh, Byass, & Wall, 2003; Minh, Huong, Wall, Chuc, & Byass, 2006); prolonged cough (Khe, Phorson, Hoa, Diwan, & Eriksson, 2004) and height-for-age (Wagstaff, van Doorslaer, & Watanabe, 2003). The mentioned studies used one or a few of the following measures of socioeconomic status: education, occupation and household income or expenditure.

Using Vietnamese data, Khe, Eriksson, Phuong, Höjer, and Diwan (2003) showed that different socioeconomic variables are correlated, but only weakly so. This tells us that different socioeconomic variables describe different aspect of socioeconomic status. That the socioeconomic variables are correlated also tells us that it is preferable to, at once, access the relationship between them and health, rather than separately analyzing the association between each aspect of socioeconomic status and health. Hence, the primary purpose of this paper is to study the associations between mortality and several socioeconomic measures, including education, occupation, income and also marital status. Marital status is neglected in the previous Vietnamese studies, but found to be related to mortality elsewhere (see e.g. Gardner & Oswald, 2004, and references therein).

Mortality is chosen as dependent variable since it is an objective measure. Objectivity is important since several studies have found evidence of state-dependent bias in subjective self-reported health measures, meaning that people in disadvantage groups generally over-estimate their health relatively to others (Charasse-Pouélé & Fournier, 2006, and references therein discuss this).

Following Balia and Jones (2008) we calculate Gini coefficients as measures of inequality in mortality risks, and decompose them in order to show how large parts of the inequality that are explained by the different determinants of mortality. Similarly to Wildman (2003) we also study how the determinants' contribution to inequality has changed over time. We do this under the standard assumption that the slope coefficients remain constant, since we based on a Wald test can not reject that this is true. The main effect of assuming that the slope coefficients remain constant is that it increases the precision of the estimates (Greene, 2003, Chapter 4, discusses effects of including irrelevant variables, which this relates to). Our study is based on a sample of 41,000 adult men and women from the rural Bavi district in northern Vietnam, which are followed from January 1999 to March 2008. Following Wildman, we concentrate on the results for the first and last year under study, since the time between these years are likely to show the largest changes.

As a result of a transition from a planned agriculture based economy to a more market oriented one, Vietnam has during the study-period experienced a rapid economic growth, making it especially interesting to study the development over time. The transition stared in 1986 when the party congress approved broad economic reforms ("Doi Moi" or renovation), including a greater role for markets and foreign investments (U.S. State Department, 2009). Vietnam became one of the Asian "Tiger economies", and GDP grow annually with above 7% during the study-period. Agricultural production nearly doubled between 1990 and 2007, transforming Vietnam to the world's second largest rice exporter. This shift from a centrally planned economy to a more market oriented economy improved the living conditions for many Vietnamese. Per capita income rose from \$ 220 in 1994 to \$ 1024 in 2008.

Like most papers in this field, the purpose of this paper is a descriptive one; to describe the association between socioeconomic status and health, not to access the causal effect of one of them on the other. The motivations of this are twofold. First, this approach gives a clearer picture of the double burden of having both low socioeconomic status and poor health. Second, several econometric studies have demonstrated the limitations of the alternative approach, instrumental-variable regression. For example, Heckman, Urzua, and Vytlacil (2006) demonstrated the limitations when the independent variables might have heterogeneous effects on the dependent variable and Stock and Yogo (2005) showed that the instruments must indeed be quite strong in order to obtain good estimates using instrumental-variable regression.

A secondary purpose is to assess the importance of relative socioeconomic variables. We therefore estimate one specification including variables describing households' income in relation to the income in its cluster (village) and also corresponding measures for individuals' education. We also make an attempt to shed some light of the importance of gender-roles in explaining the well-known difference between men's and women's mortality. To this end, we in another specification include variables describing the individuals' age and education relative to their partner's dito. Our prior expectations were that these variables could serve as proxies for gender equality and those affect the mortality risk in different directions for men and women. This approach was inspired by Månsdotter, Lindholm, Lundberg, Winkvist, and Öhman (2006) who studied the association between gender equality and mortality and morbidity in Sweden.

It should be made clear that the approaches regarding our secondary purpose suffer from some weaknesses. The clusters (villages) do, of course, not exactly match with the groups that individuals in this area of Vietnam in reality compare themselves to, which reduces the possibilities of finding significant effects. Compared to using larger areas, there is, however, likely an advantage to use the clusters when creating these relative variables, since individuals have more interactions with those living close by and therefore probably are more affected by their status relative those, than their status relative others that they seldom or never interact with. The proposed proxies of gender equality are not perfect proxies since they might capture also other effects than those of gender equality itself. Still, we believe it is important to also report these results, since little is written about this, and so that our findings can serve a starting point for further research. Our results also demonstrate how the inclusion of these relative measures affects the estimates for the absolute variables.

2 Methods

In this section we first present the model used to estimate functions for mortality risk. We then discuss how the results from these estimations can be used to calculate and decompose Gini-coefficients in order to reveal the importance of different socioeconomic variables for explaining inequality in individuals' predicted mortality risk.

2.1 Estimating mortality-risk

Since mortality is censored by migration out of the Bavi district and by the length of the follow-up period, we use Cox's (1972) proportional hazard model to estimate the effect of the independent variables on mortality risk. The hazard rate at time t for individual i is written

$$h_i(t) = h_0(t)e^{\boldsymbol{X}_{it}\boldsymbol{\beta}}, \tag{1}$$

where $\boldsymbol{X}_{it}$ is a vector of covariates and $\boldsymbol{\beta}$ a vector of parameters. The first part of the function, $h_0(t)$, is called the baseline hazard and only depends on time in days from that the individual came under observation. The second part is a function of explanatory variables, which all are assumed to have the same proportional influence of the hazard irrespective of study time.

2.2 Measuring and decomposing inequality

We analyze inequality in predicted mortality risk, which gives sufficient degree of individual-level variation to use Gini coefficients as measures of socioeconomic inequality across individuals. In contrast to for example a concentration index (Wagstaff, van Doorslaer, & Paci, 1989; van Doorslaer, Wagstaff, Bleichrodt, Calonge, Gerdtham, Gerfin, et al., 1997) it reflects also inequalities not associated with income. That Gini coefficients depend less on income than concentration indexes is an advantage in this study, since there is a risk that the incomes measures we use suffers from measurement errors. Still, the decomposing techniques allow us to relate the inequalities to income and also to other relevant socioeconomic factors.

For any measure of health a Lorenz curve plots the cumulative proportion of the population - ranked by increasing health – against the cumulative proportion of health. The Gini coefficient is defined as the ration of the area between the

Lorenz curve and the uniform distribution line (the 45° line), to the area below the uniform distribution line (e.g. Le Grand, 1989; Wagstaff, Paci, & van Doorslaer, 1991). Thus, the Gini coefficient can take a value between 0 and 1, where 0 corresponds to perfect equality and 1 corresponds to perfect inequality. Lerman and Yitzhaki (1989) showed that the Gini coefficient can be expressed as

$$G = \frac{2}{\bar{y}} cov[y, F(y)], \tag{2}$$

where $\bar{y}$ is the mean of health and $F(y)$ is the cumulative distribution of health.

Van Doorslaer and Jones (2003) showed that for a linear regression model

$$y_i = \alpha + \sum_k \beta_k x_{ki} + \varepsilon_i$$

the Gini coefficient for y can be written as

$$G = \sum_k \left(\frac{\beta_k \overline{x}_k}{\bar{y}} \right) C_k + \frac{C_\varepsilon}{\bar{y}} = \sum_k \eta_k C_k + \frac{C_\varepsilon}{\bar{y}}. \tag{3}$$

C_k is regressor k's concentration index, which shows its relationship with the cumulative distribution of health and is calculated analogous with the Gini coefficient (2).

The first part of (3) is the explained part of inequality, consisting of a weighted sum of the concentration index of the regressors, where the weights are η_k - the elasticities of the regressors with respect to y_i. This part can be decomposed to show each determinants contribution to the inequality: the larger η_k and C_k are, the larger part of the inequality is explained by the determinant x_k. The second component is a residual or unexplained part, consisting of a generalized concentration index for the error term divided by the mean of y. When, as here, the Gini coefficient is calculated for a predicted variable, the second component vanishes.

To be able to decomposed inequality in predicted mortality, we follow van Doorslaer and Gerdtham's (2003) and linearized (1) as

$$\ln h_i(t) = \ln h_0(t) + \boldsymbol{X}_{it}\boldsymbol{\beta}. \tag{4}$$

Consequently, we decompose the inequality in the log of 1-day hazard rates, instead of the inequality in survival probability.

The decomposition is informative since it gives an easily understandable description of variables importance for explaining inequality. It should, however, be made clear that the approach suffers from two shortcomings. Firstly, like previous calculations of measures of inequality in predicted mortality risk (or some other predicted health status) reported in the literature (e.g. van Doorslaer & Gerdtham, 2003; Balia & Jones, 2008), the Gini coefficients reported in this paper will be a function not only of the "true" inequality, but also of the specific function chosen to estimate the mortality risk. Therefore, it is not correct to state that, for example, the socioeconomic inequality in health is larger among women, only because the Gini coefficients are larger for them than for men, since this also could be due to that the specification chosen is better on explaining mortality for women. Secondly, decompositions based on a linearization of (or a linear prediction from) a non-linear model, are not unique. For example, the approach used here and by van Doorslaer and Gerdtham (2003) provides an exact decomposition only for an individual with a mortality risk equal to the average in the sample, and are thus only an approximation for the general population.

3 Data and variable definitions

This study is conducted using household survey data from the Bavi district, Ha Tay province in northern Vietnam, 60 kilometer west of Hanoi. Bavi is a rural district in which a field laboratory, FilaBavi, was established in January 1999 (Chuc & Diwan, 2003). The baseline survey in the beginning of 1999 was followed by quarterly surveillance of vital events and complete re-surveys every two years.

At start, the FilaBavi sample consisted of 11,089 households with 49,893 individuals belonging to 69 clusters, but it has changed over time due to migration, births and deaths. Administrative, the Bavi district is divided into communes, which in turn are divided into villages. The clusters mainly consist of one village, except some cases of small villages merged into one cluster or some big village divided into 2 or 3 clusters.

After excluding 2% of the observation that lacked data on income or number of household members, data of 18,776 men and 22,085 women above 20 years of age, belonging to 14,422 households, remains in the study. Of these 1,137 men and 953 women died during the follow-up. Some descriptive statistics are presented in Table 1, separately for men and women and for the first and last survey, while definitions of the variables are provided in the next two subsections. Means are presented for the continues variables and for the indicator variables the percentage of observations in each category are presented. The descriptive statistics reveal the rapid development that this part of Vietnam has experienced between 1999 and 2007, with for example a considerable increase in income and a decrease in share of farmers. The table also shows that the crude mortality rates have decreased in the sample during the study-period.

3.1 Baseline specification

The dependent variable is censored *survival time* and the *survival status* at the latest quarterly surveillance. For 73% of those still living in the Bavi district the last quarterly surveillance was conducted between January 1 and March 19, 2008, when the data for this study was collected, whereas it was conducted in the 4th quarter of 2007 for the remaining 27%. During the follow-up, 7,826 individuals in the study population migrated out of the district, without migrating back.

Mortality is estimated separately for women and men and as a function of individuals' education, occupation, marital status, age, and the households' income and size at the latest of the five surveys. The estimations are performed without weights using Cox's (1972) proportional hazard model as implemented in Stata 10.0, and the error terms are allowed to be correlated within households. The choice of parameterization and grouping of the variables is based on the Akaike information criterion (Greene, 2003, Chapter 8). In order to select the same specification for the sexes and hence facilitate comparisons, this was done when the samples of men and women were pooled, but the coefficients and baseline hazard functions were allowed to differ between the sexes.

Table 1 about here.

The individual's educational level is classified as either *no schooling*, *primary school*, *secondary school* (omitted), or *high school/college*. People are categorized according their main occupation as either *farmers* (omitted), *business/ students*, *workers*, *government staff*, *retired*, *homemaker/jobless* or *others* and according to their marital status as either *married* (omitted), *separated*, *widowed* or *single*. The occupational groups are obtained by aggregating together 20 original groups based on similarities between the jobs or in the groups' mortality risk and the choice of number of groups was guided by the Akaike information criterion.

The income for many of the households in this sample, for example, for the farmers, can be unstable and come irregularly. This increases the risk of measurement errors of income. To reduce this possible problem, two measures of the households' income are included in the model. The first is a set of four dummy-variables - *poor*, *average* (omitted), *rich* and *not classified*. According to standards set by the Ministry of Labor, Invalids and Society, local leaders have classified most of the households into five categories based on the income per person expressed in kilogram of rice or Vietnamese dong (Minh, Byass, & Wall, 2003). Due to few observations in the lowest and highest categories, the two lowest and the two highest categories are aggregated together to *poor* and *rich*, respectively. The second measure of income is $ln(income/hhmemb)$ - the logarithm of the household's total yearly income per household members. The two income measures will likely complement each other in describing the association between income and mortality, since there might be different measurement errors in the two.

The number of household members is controlled for using the squared polynomial $hhmemb$ and $hhmemb^2$. These variables are included partly to allow for direct effects, but also to equivalise household income. Age is controlled for using the quartic polynomial age, age^2,age^3 and age^4. Cluster specific fixed effects are included to control for time-invariant heterogeneity among the clusters. Finally, fixed effects for each re-surveys are included to control for changes in mortality over time, which the baseline hazard not fully can control for since not all individuals were followed during the same time. Since there is no variation in the re-survey fixed effects on a given year, these variables will not affect the Gini coefficients. In the chosen specification we do not control for whether the household was situated at riverside and islands, highlands and mountains, since this was not found to affect the mortality rate significantly and since it worsened the Akaike information criterion.

3.2 Extended specifications

In specification 2 we also include the variables *relative income*, *low relative income*, *high relative income*, *low relative education*, *high relative education* which all describe the households' and individuals' situation in relationship to the cluster to which it belongs. *Relative income* is defined as *income/hhmemb* divided by the average income per household member in the cluster to which the household belongs. The next four variables take the value one if the household's economic classification or the individual's educational level is below the 25th, or above the 75th, percentile in the cluster, respectively. These variables are included in order to study if social position affects mortality risk, but might also capture other effects. The three measures describing the households' relative income, might capture direct effect of neighbors income, for example that you might be more likely to have access to drinking water of high quality if your neighbors afford to drill a well. Similarly, the estimates for low and high relative education might indicate how individuals are affected by the knowledge of their neighbors. The five original economic classifications were used when creating the variables *low relative income* and *high relative income*. Likewise, high school and college were treated as separate educational levels when creating the relative educational variables discussed here, and the two discussed below.

Specification 3 differ from the baseline specification by including the variables *spousé's relative age*, *spouse lower education*, *spouse higher education*, and *no spouse identified*. *Spousé's relative age* is defined as the partners' age subtracted with the individuals' own age, and the next two take the value one if the partners' education is lower or higher than the individuals' own education. These three variables are included as proxy-variables for equality between the sexes in a household. As can be seen in Table 1, men in this area of Vietnam are on average older and more educated than their wives. Our prior expectations were that the gender inequality would be lower in couples that deviate from this traditional pattern and that this would result in lower mortality rates for men and higher mortality rates for women belonging to these couples, given that men's higher mortality is at least partly explained by gender roles. The estimates for these variables might, however, also capture for example that those who marries a younger partner differ systematically in health from others of the same age, and that one can benefit from one's spouse's education.

No spouse identified is included in specification 3 to control for systematic differences between individuals who have an identified spouse and those that have not. This variable take the value one for individuals who are not married or lack information of spouse, and *spousés relative age*, *spouse lower education* and *spouse higher education* are assigned the value zero for these observations.

4 Results

Our choice of estimating the mortality functions separately for men and women are supported by a Wald test showing that the coefficient estimates for men and women are jointly significantly different and by that the log likelihood value is increase when men and women are allowed to have different baseline hazard functions. The estimated hazard ratios with standard errors corrected for heteroskedasticity and correlation within households are reported in Table 2 for men and Table 3 for women.

Tables 2 and 3 about here.

The results show that the education variables have the expected gradient for both sexes, with highest hazards for those with lowest education. The hazard ratios for *high school/college* are significantly lower for women than for men, except in specification 2. For men, all occupational groups except *business/students* have significantly higher hazards than the omitted category *farmers*. A similar pattern is found for women, but for them only the hazard ratios for *retired* and *others* are significantly different from unity. For both sexes, *singles* and *separated* have higher hazards than married, whereas no significant association is found for *widowed*.

The income-class variables have the expected gradient for both sexes, as do *ln*(*income/hhmemb*) for men, but the hazard ratios for *rich* are not significant for men in any of the three specifications. The hazard for women is found to increase with the number of household members up to seven members (10% live in households with 8 or more members), but no significant association is found for men.

The additional variables included in specification 2 - describing the households' income and the individuals' education relative to the cluster (village) they belong to - are not significantly different from unity, with the exception of *low relative education* for men. The negative association between this variable and mortality can be interpreted as that men with low education benefit from neighbors' education. The results also indicate that the inclusion of the relative variables in specification 2 increases the importance of the absolute educational variables for men, whereas it reduces the importance of the absolute income classes: for women the income-classes even lose their statistical significances.

The lack of significance for most of the relative variables in specification 2 cannot be seen as a proof of that relative income and education is unimportant for peoples mortality. The clusters do of course not exactly match with the groups that individuals in this area of Vietnam in reality compare themselves with. In other words, the relative variables in this specification are measures with errors of the relative variables that people do care about, which leads to bias towards unity. Similar results are obtained when the relative variables are based on commune level, instead of cluster level.

The additional variables included in specification 3 - describing individuals' age and education relative to their spouses - are not significantly different from unity at the 10% level, with two exceptions. *Spousés relative age* is significantly above unity for men, meaning that having an older wife is associated with higher mortality for men. The hazard ratio for *spouse lower education* is significantly above unity for women (and nearly so for men, p-value =0.12), indicating that having a spouse with lower education increase one's mortality risk.

Table 4 shows that the variables in each variable group in specification 1 are jointly significant for women and for men and women together, but that the year-effects and the number of household member variables are not jointly significant for men. The additional variables included in specification 2 are not jointly significant for men or for women, and the additional variables included in specification 3 are only jointly significant for men and for men and women together. Table 4 also shows that the variable groups: occupation, number of household members, age, and the cluster specific effects, are significantly different between the sexes.

Table 4 about here.

We have also tested for presence of trends in the slope coefficients by including interaction-variables between the variables in the baseline specification and the year of the re-census. A Wald test showed, however, that these interaction-variables are not jointly significant. Since the estimates from this specification also suffer from high multicollinearity, we do not report them.

4.1 Decomposition

Table 5 reports age-standardized Gini coefficients and variables' and variable groups' percentage contribution to predicted age-standardized inequality in mortality for men in 1999 and in 2007 for all three specification, while the corresponding figures for women are reported in Table 6. An indirect age-standardization is employed, which is recommended for individual data by O'Donnell, van Doorslaer, Wagstaff, and Lindelow (2007, Chapter 5).

Tables 5 and 6 about here.

We have chosen to report age-standardized values of two principal reasons. Firstly, we view aging (measured in time units) as fair, simply because the law of aging is the same for people in all socioeconomic groups: for each year that passes you get one year older. Secondly, since age is the major determinant of mortality risk it would overshadow all other results. Age explain over 60% of non age-standardized Gini coefficients for women and over 70% for men, resulting in that other variables' contribution is largely influenced by their correlation to age. For example, being a single was found to have a negative influence on non age-standardized inequality, simply because its negative correlation with age.

As discussed in the method-section, the decomposition is only a linear approximation, and should thus be interpreted with some caution. Still, the decomposition results indicate that, besides the cluster effects, occupation is the variable group that contribute most to inequality in mortality risk, except for men in 2007. This is mainly due to the considerable negative association between being retired and mortality, conditioned on age and the other variables. For both sexes the importance of this variable group have decreased between 1999 and 2007, mainly due to a decrease of nearly five percentage points of the share of retired.

For men, *single* makes marital status the most important variable group in 2007, while it was only the third most important variable group in 1999. *Singles* contribution have increased over time as the share of men that are single have increased with four percentage points. For women, the third most important variable group in all years is education. The decomposition also show that the income variables together explained up to 14% of the inequality among men, but less than 4% among women, and that the additional variables included in specification 2 and 3 have limited contributions to predicted inequality in mortality. Finally, the Gini coefficients are found to be lower for men than for women, and to have decreased marginally between 1999 and 2007.

5 Discussion

Most of the studied variables indicate improved welfare in Bavi during the studied period. Incomes have significantly increased, the proportion of poor has decreased, and the numbers of well educated have increased. Crude mortality rates are substantially lower in the end of the period while inequality in mortality seems to be stable. Thus the transition from a planned agriculture based economy to a more market oriented economy with growing industrial production has brought improved welfare and population health to Bavi.

Compared to neighboring countries, Vietnam has achieved an outstanding level of health considering economic resources. Life expectancy at birth in 2005 was in the same magnitude (around 70) as in Thailand and China, despite Thailand has three times and China two times larger GDP per capita. Cambodia and Lao is approximately on the same GDP level as Vietnam but achieve 10 years less in life expectancy (Son, 2009). Even a comparison of mortality patterns indicated the rapid transition in Vietnam, where cardiovascular diseases account for 24% of deaths, while corresponding figures for Cambodia is 11%, Lao 13% and Thailand 13% (WHO, 2006).

The estimated hazard ratios confirm the results reported by Huong et al. (2006): that the mortality risk in Vietnam decrease with income and education - a pattern observed also in many other countries. However, the decomposition indicates that the income and education variables together explain only 15-30% of the predicted age-standardized inequality for each sex, suggesting that it is important to also consider other dimension of socioeconomic status, for example related to occupation and marital status.

The association between occupation and different aspect of health have been studied previously based on Vietnamese data (e.g. Minh, Byass, & Wall, 2003) and are also commonly included in analyses of socioeconomic inequality in other countries. The decomposition results reported in this study suggest that this is indeed important, since this variable group explain 25-35% of the inequality. Marital status has, however, been ignored in previous studies of inequality in health in Vietnam, but our results indicate that marital status capture an important aspect of socioeconomic inequality. The hazards for singles are more than two and a half times as high than those for married, both for men and women, and the decomposition analyses indicate that this variable alone explain 20-30% of the inequality for men and approximately 10% of that for women. Similar results have been reported from Europe. A Swedish register study showed a 3,7 times increased risk for total mortality for men and 2,5 times for women, compared to cohabiting with children (Ringbäck Weitoft, 2003). Possible mechanism behind the high hazard rates for singles could be associated with them potentially being more socially isolated, having lower hierarchical status or more unhealthy lifestyle than married people. Another pathway could be selection: that people with poor health are less fortune at the marriage market (Gardner & Oswald 2004). Based on the data used for this study, we can not discriminate between these possible mechanisms, but this is something than warrants further research.

To live alone seems to be an emerging risk factor for poor health, and a possible policy response is greater efforts to promote a health life style, e.g. interventions towards excessive alcohol consumption and smoking. The extended, traditional family are likely exposed to growing strains in this economic, social and cultural transition and polices aimed at family support could perhaps even further reduce inequalities in the population health as well as reduce gender inequalities.

Among the results for the relative variables included in the second and third specification, the most interesting are those regarding education. Conditioned on their own education and other variables, men are found to benefit from having low education relative to the village they belong to, and the estimates for both sexes indicate that having a spouse with lower education than you increases your mortality risk. In other words, the hazards for men decreases if the education level in their village increases, and the point estimates also indicate that both men's and women's hazard would decrease if their spouse had as high education as them, instead of a lower education. Together these results suggest that positive spillover effects of education exists, that is, you benefit not only from your own education but also from that of those around you. This warrants further research, however, since only two of these hazard ratios are significantly different from unity at conventional levels.

References

Balia, S., & Jones, A.M. (2008) Mortality, lifestyle and social-economic status, *Journal of Health Economics*, 27, 1-26.

Charasse-Pouélé, C., & Fournier, M. (2006) Health disparities between racial groups in South Africa: A decomposition analysis. *Social Science & Medicine,* 62, 2897-2914.

Chuc, N.T.K., & Diwan, V.K. (2003) FilaBavi, a demographic surveillance site, an epidemiological field laboratory in Vietnam. *Scandinavian Journal of Public Health*, 62, 3-7.

Cox, D. (1972) Regression models and life tables. *Journal of the Royal Statistical Society Series B*, 34, 187-220.

Gardner, J., & Oswald, A. (2004) How is mortality affected by money, marriage and stress? *Journal of Health Economics*, 23, 1181-1207.

Greene, W.H. (2003). *Econometric analysis*, fifth edition. Prentice Hall: New York.

Heckman, J.J., Urzua, S., & Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics*, 88, 389-432.

Huong, D.L., Minh, H.V., Janlert, U., Van, D.D., & Byass, P. (2006) Socioeconomic status inequality and major causes of death in adults: A 5-year follow-up study in rural Vietnam, *Public Health*, 120, 497-504.

Khe, N.D., Eriksson, B., Phuong, D.N., Höjer, B., & Diwan, V.K. (2003) Faces of poverty: sensitivity and specificity of economic classifications in rural Vietnam, *Scandinavian Journal of Public Health*, 31, 70-75.

Khe, N.D., Phorson, A.E.K., Hoa, N.P., Diwan, V.K., & Eriksson, B. (2004) An economic analysis of persons with prolonged cough in rural northern Vietnam, *International Journal of Tuberculosis and Lung Disease*, 8, 424-427.

Le Grand, J. (1989) An international comparison of distributions of ages-at-death. In Fox, J. (Ed.) Health inequalities in European countries. Brookfield: Gower, Aldershot.

Lerman, R.I., & Yitzhaki, S. (1989) Improving the accuracy estimates of Gini coefficients, *Journal of Econometrics*, 42, 43-47.

Macintyre, S., & Hunt, K. (1997) Socio-economic position, gender and health: How do they interact, *Journal of Health Psychology*, 2, 315-334.

Marmot, M., Friel, S., Bell, R., Houweling, T.A.J, & Taylor, S. (2008) Closing the gap in a generation:health equity through action on the social determinants of health. *Lancet*, 372, 1661-1669.

Minh, H.V., Byass, P., & Wall, S. (2003) Mortality from cardiovascular diseases in Bavi district Vietnam, *Scandinavian Journal of Public Health*, 31, 26-31.

Minh, H.V., Huong, D.L., Wall, S., Chuc, N.T.K., & Byass, P. (2006) Cardiovascular disease mortality and its association with socio-economic status: findings from a population-based cohort study in rural Vietnam, 1999-2003, *Preventing Chronic Disease*, 3, 1-11.

Månsdotter, A., Lindholm, L., Lundberg, M., Winkvist, A., & Öhman, A. (2006) Parental share in public and domestic spheres – a population study on gender equality, death and sickness, *Journal of Epidemiology and Community Health*, 60, 616-620.

O'Donnell, O., van Doorslaer, E., Wagstaff, A., & Lindelow, M. (2007) Analyzing health equity using household survey data: A guide to techniques and their implementation, World Bank Publications.

Ringbäck Weitoft, G. (2003). Lone parenting, socioeconomic conditions and severe ill-health: Longitudinal register-based studies. PhD-thesis, Umeå university, Sweden, Print and Media.

Sen, A. (2002) Why health equity? *Health Economics*, 11, 659-666.

Son, H.H. (2009) A cross-country analysis of achievements and inequalities in economic growth and standard of living. ADB Economic Working Papers Series, 159, Asian Development Bank.

Stock, J.H., & Yogo, M (2005) Testing for weak instruments in linear IV regression. In: Andrews, D.W.K. and Stock, J.H. (eds.) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg.* Cambridge University Press: Cambridge.

U.S. State Department (2009). www.state.gov/r/pa/ei/bgn/4130.htm, 2009-06-22.

Van Doorslaer, E., & Gerdtham, U-G. (2003) Does inequality in self-assessed health predict inequality in survival by income? Evidence from Swedish data, *Social Science & Medicine,* 57, 1621-1629.

Van Doorslaer, E., & Jones, A.M. (2003) Inequalities in self-reported health: validation of a new approach to measurement, *Journal of Health Economics*, 22, 61-87.

Van Doorslaer, E., Wagstaff, A., Bleichrodt, H., Calonge, S., Gerdtham, U-G., Gerfin, M., Geurts, J., Gross, L., Hakkinen, O., Leu, R., O'Donnell, O., Propper, C., Puffer, F., Rodriuez, R., Sundberg, G., & Winkelhake, O. (1997) Income-related inequalities in health some international comparisons, *Journal of Health Economics*, 16, 93–112.

Wagstaff, A., Paci, P., & van Doorslaer, E. (1991) On the measurement of inequalities in health, *Social Science & Medicine,* 33, 545-557.

Wagstaff, A., & van Doorslaer, E. (2000) Equity in health care finance and delivery. In; Handbook in health economics. Eds: Culyer, A.J., & Newhouse, J.P. North-Holland.

Wagstaff, A., & van Doorslaer, E. (2004) Overall vs. socioeconomic health inequality: A measurement framework and two empirical illustrations, *Health Economics*, 13, 297-301.

Wagstaff, A., van Doorslaer, E., & Paci, P. (1989) Equity in the finance and delivery of health care: some tentative cross-country comparisons, *Oxford Review of Economic Policy*, 5, 89-112.

Wagstaff, A., van Doorslaer, E., & Watanabe, N. (2003) On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam, *Journal of Health Economics*, 112, 207-223.

Wildman, J. (2003) Income related inequalities in mental health in Great Britain: analysing the causes of health inequality over time. *Journal of Health Economics*, 22, 295-312.

WHO (2006) Mortality Country Fact Sheet. World Health Statistics.

Table 1. Descriptive statistics

Variable	Men		Women	
Dependent variable				
Survival time (in years)	6.87		6.78	
Survival status (=1 if dead)	6.09		4.37	
Independent variables	1999	2007	1999	2007
No schooling	1.93	1.25	9.30	5.47
Primary school	17.56	13.00	25.18	21.56
Secondary school (omitted)	51.01	51.45	48.79	50.16
High school/college	29.50	34.31	16.74	22.80
Farmers (omitted)	60.39	38.72	68.83	57.19
Business/students	3.90	6.94	4.11	9.53
Workers	3.20	31.44	1.50	8.65
Government staff	4.64	5.11	4.53	4.82
Retired	13.36	8.93	15.03	10.50
Homemaker/jobless	1.62	2.49	2.80	6.57
Others	12.89	6.36	3.21	2.73
Married (omitted)	84.85	81.80	72.33	72.14
Separated	0.92	0.72	2.83	2.21
Widowed	2.68	2.15	17.07	15.80
Single	11.55	15.33	7.77	9.85
Poor	17.36	10.68	19.24	13.67
Average (omitted)	52.69	64.73	51.95	62.96
Rich	24.18	22.67	22.85	21.27
Not classified	5.77	1.92	5.97	2.09
*Ln(income/hhmemb)**	0.80	1.46	0.76	1.39
Hhmemb	4.66	5.64	4.54	5.51
Age	40.22	42.11	43.49	44.99
*Relative income**	1.03	1.03	0.97	0.98
Low relative income	9.26	7.52	10.68	10.27
High relative income	14.75	10.48	14.43	10.08
Low relative education	5.63	11.06	15.58	22.49
High relative education	21.14	19.05	10.76	11.88
Spouse's relative age	-2.36	-2.41	1.82	1.99
Spouse lower education	23.39	20.08	10.27	11.39
Spouse higher education	12.94	13.60	19.06	17.12
No spouse identified	17.78	20.17	34.03	32.90
Crude mortality	8.77	6.25	6.89	4.17

Note: *Income is measured in 1000 Vietnamese dong and expressed in 2007 years prices. For the independent variables the total number of observations are 71,967 for men and 83,425 for women, meaning that the individuals on average were observed at 3,8 surveys. Fewest observations (19.44%) originate from the 1999 survey and most (20.73%) from the 2005 survey.

Table 2. Hazard ratios (H.R.) and robust standard errors (S.E.) for men

	1		2		3	
Variable	H.R.	S.E.	H.R.	S.E.	H.R.	S.E.
No schooling	1.35**	0.20	1.73***	0.33	1.39**	0.22
Primary school	1.17*	0.10	1.31***	0.13	1.20**	0.11
High school/college	0.87	0.08	0.76	0.13	0.85	0.08
Business/students	0.95	0.26	0.95	0.26	0.97	0.26
Workers	1.35**	0.19	1.33**	0.19	1.36**	0.20
Government staff	1.61**	0.38	1.59**	0.37	1.63**	0.38
Retired	2.70***	0.31	2.68***	0.31	2.66***	0.30
Homemaker/jobless	2.00***	0.35	1.99***	0.35	1.99***	0.35
Others	1.84***	0.24	1.84***	0.24	1.83***	0.24
Separated	1.81**	0.49	1.78**	0.48	2.05**	0.61
Widowed	1.09	0.10	1.09	0.10	1.18	0.19
Single	2.57***	0.57	2.53***	0.57	2.90***	0.75
Poor	1.60***	0.15	1.49***	0.21	1.61***	0.15
Rich	0.91	0.07	0.96	0.09	0.91	0.07
Not classified	0.87	0.13	1.02	0.20	0.87	0.13
Ln(income/hhmemb)	0.88***	0.04	0.88**	0.05	0.88**	0.04
Hhmemb	1.03	0.06	1.03	0.06	1.04	0.06
$Hhmemb^2$	1.00	0.00	1.00	0.00	1.00	0.00
Relative income			1.01	0.02		
Low relative income			1.14	0.19		
High relative income			0.85	0.11		
Low relative education			0.80**	0.09		
High relative education			1.18	0.21		
Spouse's relative age					1.02**	0.01
Spouse lower education					1.14	0.10
Spouse higher education					1.08	0.14
No spouse identified					0.87	0.14
Log likelihood	-9,592		-9,589		-9,586	
AIC (Akaike)	19,372		19,376		19,368	

Notes: The asterisks ***, ** and * denote significance at the 1%, 5% and 10% levels. Estimation results for age-variables and year- and cluster-specific fixed effects are suppressed but available from the authors upon request. In all three specifications the numbers of observations are 71,967, individuals 18,776 and deaths 1,137.

Table 3. Hazard ratios (H.R.) and robust standard errors (S.E.) for women

	1		2		3	
Variable	H.R.	S.E.	H.R.	S.E.	H.R.	S.E.
No schooling	1.27	0.23	1.19	0.27	1.33	0.25
Primary school	0.96	0.16	0.94	0.17	1.00	0.17
High school/college	0.50**	0.15	0.34**	0.18	0.46***	0.14
Business/students	0.75	0.27	0.76	0.27	0.77	0.27
Workers	1.09	0.47	1.11	0.48	1.11	0.48
Government staff	2.12	1.09	2.05	1.05	2.14	1.09
Retired	2.85***	0.43	2.86***	0.43	2.87***	0.44
Homemaker/jobless	1.32	0.25	1.32	0.25	1.34	0.26
Others	3.82***	0.78	3.81***	0.78	3.84***	0.79
Separated	1.59*	0.43	1.57*	0.42	1.99**	0.62
Widowed	0.89	0.08	0.89	0.08	1.09	0.20
Single	2.60***	0.50	2.58***	0.49	3.31***	0.85
Poor	1.31***	0.13	1.13	0.18	1.31***	0.13
Rich	0.84**	0.08	0.89	0.09	0.84*	0.08
Not classified	1.14	0.17	1.33	0.29	1.14	0.17
Ln(income/hhmemb)	1.04	0.06	1.04	0.06	1.04	0.06
Hhmemb	1.26***	0.07	1.27***	0.07	1.26***	0.07
$Hhmemb^2$	0.98***	0.00	0.98***	0.00	0.98***	0.00
Relative income			1.00	0.03		
Low relative income			1.23	0.22		
High relative income			0.86	0.13		
Low relative education			1.07	0.14		
High relative education			1.71	0.95		
Spousés relative age					1.00	0.01
Spouse lower education					1.52*	0.37
Spouse higher education					0.98	0.14
No spouse identified					0.79	0.16
Log likelihood	-7,731		-7,729		-7,729	
AIC (Akaike)	15,650		15,657		15,653	

Notes: The asterisks ***, ** and * denote significance at the 1%, 5% and 10% levels. Estimation results for age-variables and year- and cluster-specific fixed effects are suppressed but available from the authors upon request. In all three specifications the numbers of observations are 83,423, individuals 22,085, and deaths 953.

Table 4. Wald test of joint significance: P-values

	Men			Women		
Variable group	1	2	3	1	2	3
Education	0.0204	0.0073	0.0224	0.0006	0.0295	0.0003
Occupation	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Marital status	0.0001	0.0001	0.0001	0.0000	0.0000	0.0000
Income class	0.0000	0.0419	0.0000	0.0037	0.1043	0.0041
Hhmemb	0.7706	0.7303	0.6845	0.0000	0.0000	0.0000
Age	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Year	0.2329	0.2184	0.2287	0.0104	0.0124	0.0103
Clusters	0.0271	0.0313	0.0242	0.0090	0.0116	0.0090
Add. variables		0.2262	0.0398		0.5856	0.2904

	Men and Women			Sex-differences		
Variable group	1	2	3	1	2	3
Education	0.0001	0.0017	0.0001	0.2460	0.2487	0.1838
Occupation	0.0000	0.0000	0.0000	0.0102	0.0113	0.0097
Marital status	0.0000	0.0000	0.0000	0.4781	0.4553	0.9149
Income status	0.0000	0.0249	0.0000	0.1850	0.3039	0.1883
Hhmemb	0.0000	0.0000	0.0000	0.0121	0.0113	0.0206
Age	0.0000	0.0000	0.0000	0.0103	0.0134	0.0196
Year	0.0170	0.0185	0.0167	0.4512	0.4319	0.4584
Clusters	0.0002	0.0004	0.0002	0.0063	0.0060	0.0066
Add. variables		0.3773	0.0623		0.5862	0.4721

The figures reported in the lower half of the table is obtained when pooling the samples for men and women allowing the coefficients and baseline hazard functions to differ between the sexes. Under the heading "Men and Women" the joint significance of the coefficients for both sexes are reported, while tests for jointly significant differences between the sexes are reported under "Sex-differences".

Table 5. Decomposition results for men: % contribution to inequality

	1		2		3	
Variable	1999	2007	1999	2007	1999	2007
No schooling	1.13	0.65	2.01	1.23	1.24	0.72
Primary school	2.93	2.04	5.34	2.95	3.40	2.46
High school/college	1.33	0.26	2.60	0.72	1.37	0.23
Sum :	5.38	2.95	9.95	4.90	6.01	3.41
Business/students	0.04	0.22	0.04	0.24	0.03	0.15
Workers	-0.21	4.58	-0.24	4.17	-0.21	4.54
Government staff	-0.85	-0.30	-0.85	-0.22	-0.84	-0.33
Retired	24.88	15.34	24.71	15.27	23.42	14.47
Homemaker/jobless	2.00	2.26	1.95	2.17	1.91	2.08
Others	8.00	5.01	7.82	4.98	7.53	4.72
Sum :	33.87	27.11	33.43	26.60	31.83	25.63
Separated	0.91	0.75	0.85	0.72	1.09	0.91
Widowed	0.42	0.30	0.44	0.31	0.76	0.55
Single	19.60	27.89	18.86	27.63	21.69	30.79
Sum :	20.93	28.95	20.15	28.66	23.54	32.25
Poor	9.47	8.35	7.68	7.02	9.45	8.33
Rich	0.77	0.94	0.29	0.36	0.84	1.01
Not classified	0.39	0.01	-0.06	0.00	0.38	0.01
Sum :	10.63	9.30	7.91	7.37	10.68	9.35
ln(income/hhmemb)	3.18	3.44	3.05	3.43	3.01	3.19
Hhmemb	0.66	0.74	0.63	0.75	1.03	0.92
$Hhmemb^2$	-0.40	-0.57	-0.36	-0.55	-0.72	-0.81
Sum :	0.26	0.17	0.26	0.20	0.32	0.11
Relativeincome			-0.24	-0.30		
Low relative income			1.68	1.52		
High relative income			1.33	1.27		
Low relative education			-1.38	-1.99		
High relative education			-0.58	1.21		
Sum :			0.81	1.71		
Spousés relative age					3.43	3.78
Spouse lower education					-0.37	-0.51
Spouse higher education					-0.03	-0.07
No spouce identified					-3.62	-4.49
Sum :					-0.59	-1.29
Sum : Clusters	25.75	28.09	24.43	27.12	25.19	27.35
Gini coefficient	0.3479	0.3415	0.3519	0.3399	0.3515	0.3454

Table 6. Decomposition results for women: % contribution to inequality

	1		2		3	
Variable	1999	2007	1999	2007	1999	2007
No schooling	3.17	2.02	2.20	1.37	3.61	2.32
Primary school	-0.34	-0.39	-0.54	-0.73	-0.01	-0.01
High school/college	14.10	15.04	21.06	24.14	15.05	16.64
Sum :	16.93	16.68	22.72	24.78	18.65	18.95
Business/students	1.74	3.06	1.59	2.66	1.58	2.82
Workers	-0.13	-0.03	-0.16	-0.08	-0.15	-0.06
Government staff	-1.65	-0.45	-1.43	-0.70	-1.79	-0.62
Retired	25.92	18.00	25.17	16.80	25.28	17.64
Homemaker/jobless	0.13	0.37	0.10	0.44	0.10	0.34
Others	6.31	7.76	5.63	6.85	6.07	7.53
Sum :	32.32	28.71	30.90	25.97	31.09	27.65
Separated	0.82	0.91	0.74	0.81	1.12	1.28
Widowed	-1.42	-1.16	-1.37	-1.16	0.98	0.79
Single	10.59	11.27	9.98	9.92	12.55	12.58
Sum :	9.99	11.02	9.35	9.56	14.66	14.65
Poor	2.84	3.13	1.12	1.36	2.86	3.16
Rich	0.82	1.15	0.48	0.71	0.89	1.22
Not classified	0.01	0.13	-0.01	0.23	0.03	0.13
Sum :	3.67	4.41	1.59	2.30	3.78	4.51
ln(income/hhmemb)	-0.65	-0.55	-0.78	-0.67	-0.73	-0.59
Hhmemb	11.81	1.91	11.88	1.99	11.60	1.36
$Hhmemb^2$	-7.78	-0.29	-7.84	-0.60	-7.53	0.22
Sum :	4.02	1.63	4.04	1.39	4.07	1.58
Relative income			-0.02	-0.01		
Low relative income			1.76	1.75		
High relative income			0.73	0.52		
Low relative education			1.02	1.24		
High relative education			-4.13	-1.87		
Sum :			-0.64	1.62		
Spouse's relative age					-0.04	-0.04
Spouse lower education					-0.36	-0.46
Spouse higher education					0.10	0.09
No spouce identified					-3.99	-3.80
Sum :					-4.29	-4.21
Sum : Clusters	33.72	38.11	32.82	35.04	32.79	37.47
Gini coefficient	0.3889	0.3821	0.3942	0.3928	0.3920	0.3849