

# **Test of the ‘Healthy Migrant Hypothesis’: A Longitudinal Analysis of Health Selectivity of Internal Migration in Indonesia**

Yao Lu

(Final version published in *Social Science and Medicine*)

Lu, Yao. 2008. “Test of the ‘Healthy Migrant Hypothesis’: A Longitudinal Analysis of Health Selectivity of Internal Migration in Indonesia.” *Social Science and Medicine* 67(8):1331-1339.

## Contact information

Yao Lu

Department of Sociology, Columbia University

501 Knox Hall

New York, NY 10027, USA

Phone: 212-854-5442

Email: [yao.lu@columbia.edu](mailto:yao.lu@columbia.edu)

## **Acknowledgements**

The author would like to thank Donald J. Treiman, William M. Mason, Elizabeth Frankenberg, and Naihua Duan for their contributions at earlier stages of this work.

## **Abstract**

Previous studies show migrants are generally healthier than the populations in receiving societies, a result generally attributed to the positive selection of migrants on health. This hypothesis, however, has not been adequately evaluated due to lack of adequate data. In this article, using high-quality longitudinal data from Indonesia, the health selectivity hypothesis, also referred to as the healthy migrant hypothesis, is examined with respect to internal migration. Specifically, this study explores whether pre-migration health status affects the likelihood of migration by comparing those from the sending population who do and do not move. Results show that migrants in Indonesia tend to be selected with respect to health and that this selection is robust to household unobserved heterogeneity. However, the strength and direction of the health-migration association vary by types of migration and dimensions of health.

## **Introduction**

Migration and health have each received a significant amount of attention in social sciences. However, only recently have researchers examined the link between these two population processes. Evidence indicates that migration and health are intertwined in complex ways. Health itself can impact the decision to move and migration may affect the health of those who move, those who stay, and perhaps even those who host migrants (Hull, 1979). Studying health in the context of migration thus offers a better understanding of the complexity and diversity of the migration process, which is critical, as migration has become a widespread and persistent phenomenon that is changing the structure of family units, communities and societies in our modern world.

Previous work on migration and health largely compares the health of immigrants to that of the native population at destination. Such studies usually find that immigrants are generally healthier than the native-born populations as indicated by mortality rates, chronic conditions, mental health, etc., though the advantage enjoyed by immigrants tends to deteriorate over time (Anson, 2004; Feranil, 2005; Marmot, Adelstein, & Bulusu, 1984a, 1984b; Palloni & Morenoff, 2001). This is often referred to as the “epidemiological paradox”, as immigrants usually face disadvantages in many aspects that have negative implications for health.

One of the most frequently offered explanations is the “healthy migrant hypothesis”. It states that migrants represent a selectively healthy group that are not representative of all potential migrants from origin societies (Palloni & Morenoff, 2001). As a result, their health advantage stands out when they are compared with the general population at destination.

However, this theory has been purely speculative and inadequately tested. This is in clear contrast to the abundant research on the selection of migration with respect to demographic and socioeconomic characteristics. In fact, in almost none is there an attempt to compare the health of migrants with that of populations in sending communities other than the U.S. This is largely due to the lack of adequate data, which require collecting information on the population from the home region prior to migration.

In addition, although similar theories pertaining to internal migration have been advanced, there is surprisingly limited evidence on how the health status differs between migrants and non-migrants, and almost none on how health considerations may come into play in the decision-making for internal migration. This is unfortunate since the vast majority of migration streams involve movements within national boundaries. Take Indonesia, the study setting, for example, over 23 million Indonesians migrate each year, but only 10% of these migrants move internationally. This situation calls for the need for more attention on internal migration.

This research represents one of the first attempts to examine the potential influence of health on the likelihood of and reason for migration. An explicit test of the “healthy migrant hypothesis” is crucial for a sound understanding of health disparities between migrants and native population, because it enables us to disentangle the impact of migration on health from the health selectivity effects. By contrast, most earlier work establishing an “epidemiological paradox” inevitably confounds these two aspects due to their exclusive focus on health status after migration.

I first propose an integrated theoretical framework for understanding the implications of

health for internal as well as international migration. I then explicitly assess the health selectivity of migration by comparing the pre-migration health status between those from the sending population who do and do not move. I focus on the important but inadequately researched phenomena, internal migration, for which data on comparable non-migrant population are more readily available. Data used are from the Indonesia Family Life Survey (IFLS), a national representative longitudinal sample survey conducted in 1993, 1997, and 2000. The longitudinal structure and the detailed migration histories facilitate the examination of my research questions. I further distinguish different types of migration to take account of substantial heterogeneity among migrants, and various dimensions of health given that they may have different implications for migratory decisions. Adjustment for household unobserved heterogeneity, which may contaminate the relationship, is built into the analysis, via household fixed-effects.

Indonesia, the fourth most populous nation in the world, is drawn on as a case study. The country has enjoyed rapid economic growth over the past three decades, along with concomitant improvements in access to health care and common measures of health status such as life expectancy and infant mortality rate (Frankenberg & Thomas 2000; Muhidin, 2002). With respect to migration, Indonesia is recognized as one of the world's major sources of unskilled migrant workers to Southeast Asian countries (Hugo, 2005; Sukamdi & Brownlee, 1998). As the industrialization process has intensified, geographical mobility within the country also has increased in recent years (Hugo, 2005; Muhidin, 2002).<sup>1</sup> In the most recent census, one in ten Indonesians was classified as a migrant. This stream is largely characterized by rural to urban and economically motivated migration, with the largest cities

as the main destinations. A few studies have examined the characteristics of internal migrants in Indonesia (Muhidin, 2002; Speare & Harris, 1986), which suggest that migrant workers are drawn disproportionately from young adult males who are better educated and from relatively poor households.

### **Analytic Framework**

Since migration often involves disruption of individual's life and adaptation to a new environment, migrants tend to be selected for personal characteristics that foster their ability to handle change and adapt to new environments. This has been demonstrated in both the internal and international migration literature (De Jong et al., 1983; Massey, 1988). Aside from a variety of demographic and socioeconomic characteristics, physical health is likely to be one of the selection factors (Evans, 1987; Findley, 1988; Freedman, 1947; Hull, 1979; Jasso, Massey, Rosenzweig, & Smith, 2004; McKinlay, 1975). People with poor health may be less likely to move because they are less capable of moving or managing the difficulties and stress associated with migration. This is especially true for long-distance and work-related migration, which requires considerable level of physical stamina to endure the demanding journey and achieve economic gains.

In contrast, people in good health may be more likely to move since health does not place an impediment to migration, and good health may even foster their ability to make migration a gainful experience. Previous studies suggest that migrants appear to be healthy because they are favorably selected on socioeconomic and demographic characteristics that foster good health (McKinlay, 1975). However, I argue that even when adjusting for demographic and

socioeconomic factors, migrants are selective of those with superior health for reasons detailed above.

Although the limited empirical studies have focused on health selection of international migration, most theoretical work makes little distinction between internal and international migration flows but often suggests that the basic underlying mechanisms apply equally to both groups (Evans, 1987; Findley, 1988; Freedman, 1947; Hull, 1979; McKinlay, 1975). This is because both migration flows are in response to similar forces and tend to exhibit similar patterns: migration primarily involves labor movement from less developed to economically advantaged regions. Migration is often challenging, and most of the challenges that confront migrants are the same (i.e., disruption of family life, adaptation of new life, hardship associated with labor market conditions), regardless of whether the move is across or within national borders. As a result, migration usually attracts the most ambitious and competent people, intellectually as well as physically. There have been some efforts toward developing a unified framework to understand international and internal migration (Pryor, 1981), under the notion that both streams involve a similar decision-making process. A case study in Philippines, a setting comparable to Indonesia, lends support to this view (De Jong et al., 1983).

Given the above discussion, I expect a similar association between health and migration for both migration streams, and exploit the better-developed theories on international migration to examine the “healthy migrant hypothesis” regarding internal migration. This, however, does not imply that no difference exists between internal and international migration. One main difference is that international movements are often associated with a

longer distance and involve crossing national borders. This indicates greater difficulties and therefore a stronger selection effect. Some studies suggest that the degree of immigration selection may be inversely associated with distance from the destination (Jasso et al., 2004). In addition, one aspect of international migration that may explain the healthy migrant effect, namely immigrant health screening, usually does not apply to internal migration. However, health screening is unlikely to completely account for the favorable selection of immigrants with respect to health, because the selection is usually observed for a host of health outcomes, whereas the screening almost always takes place for a few infectious diseases, mainly tuberculosis. Also, such screening is usually required before adjustment of immigration status (i.e. permanent residency) rather than upon entry into the country (Centers for Disease Control and Prevention, 1998), which means not all cross-nation migrants are screened for health conditions (i.e., undocumented, short-term, or non-permanent immigrant labor are often not screened). All of the above is to illustrate that health screening is not the major explanation for the “healthy migrant effect”; rather, health selection for internal and international movements tends to operate in broadly similar ways through the channels discussed earlier in this section, though the strength of the selection may be greater for cross-national migration.

Beyond the general patterns, the complexity of the health-migration association is stressed in the present study. First, the direction and strength of the association may depend on individual traits such as age, given that the migration behavior of older people tends to differ substantively from that of younger adults (Findley, 1988; Freedman, 1947). Older people are relatively less mobile and less likely to migrate for labor market considerations;



thus, their decisions tend to result primarily from life events in which changing the place of residence is necessary. Health-related considerations, such as seeking healthier environments, better health care, or the aid of family members to cope with health and functional deficits, are likely the common motivations. Therefore, while younger adult migrants may be healthier in comparison to non-migrants, the elderly migrants tend to be less healthy. In addition, for the elderly health may have a causal impact on migration decisions, especially when they experience poor health or expect future deterioration of their health (Patrick, 1980). For the young, in contrast, health might not be directly related to migration decisions, but rather increases or reduces the likelihood of moving because of the difficulties associated with the migration process.

Second, the degree of the selection may vary by different groups of migrants, particularly with respect to the reasons for moving, which reflects different levels of uncertainty and hardship (Evans, 1987; Findley, 1988; McKinlay, 1975). Family-related migration and work-related migration are the most common type in Indonesia, especially for young people. I expect that those moving for work-related reasons are particularly favorably selected with respect to health, because labor migration represents the most demanding form of movement—it often requires certain capabilities, in particular physical ability, to perform physically demanding jobs and grasp other opportunities at the destination. Family-related migration, by contrast, largely reflects a continuation of family life, with much less uncertainty and hardship. So family migrants do not necessarily need superior health to settle in at the destination.

Furthermore, given that the concept of health is multidimensional, I expect migration

behavior to be differentially responsive to various aspects of health. More chronic and severe conditions may have a greater influence on migration decisions because they lead to physical weakness that is highly salient for the victims. Such health deficits usually create a strong impediment to physical activities, which tend to make migration and subsequent adaptation difficult and even impossible. In contrast, acute and mild health conditions may not play a crucial role because they are likely to heal in a short time interval and the impact may not be strong enough to impede normal life and may not even be acutely felt. This is especially true for illnesses that have no noticeable symptoms and develop slowly and silently. The asymptomatic nature of these conditions implies that they require a professional diagnosis to be recognized. This, however, is less likely to occur in Indonesia due to the limited use of health care services, especially when it comes to preventive health care such as regular medical examination (Chernichovsky & Meesook, 1986).

The hypotheses can be summarized as follows, which should be understood as net effects, after taking count of correlated covariates:

- H1. For the young, migrants are positively selected with respect to health.
- H2. For the elderly, migrants are negatively selected with respect to health.
- H3. Work-related migrants are particularly favorably selected with respect to health.
- H4. Selectivity is especially salient with respect to chronic and severe conditions.

## **Data and Methods**

### ***Data***

Data used are from the 1997 and 2000 waves of the Indonesia Family Life Survey (IFLS),

a high-quality panel survey of individuals, households and communities. To maximize representation of the population, the IFLS was conducted in 13 out of 27 provinces in Indonesia, representing 83% of the population. The first round (IFLS1) was collected in 1993 using multi-stage probability sampling and it included interviews with 7224 households and with 22,347 individuals (Frankenberg & Karoly, 1995). In each household representative members (typically household heads) provided detailed household demographic and economic information. In addition, several household members were randomly selected and interviewed on a broad range of topics including migration experience, socioeconomic conditions, and health status.

In 1997, a resurvey (IFLS2) was conducted of the IFLS1 individuals, households and communities. It attempted to reinterview all IFLS1 households and respondents (and also to interview all these not interviewed in 1993) (Frankenberg & Thomas, 2000). The IFLS has very low sample attrition, and represents one of the first efforts in social surveys in developing settings to track respondents who had moved. Excluding the households in which everyone had died, IFLS2 succeeded in interviewing 94% of the IFLS1 households and 91% of target individuals, including about 1500 respondents who had moved out from the original household and were successfully tracked in a new household (Frankenberg & Thomas, 2000). Following the practice of IFLS2, IFLS3, which was conducted in 2000, attempted to reinterview all households and all members in previous rounds. Again, over 94% of the households in IFLS1, and over 90% of the households in both IFLS1 and IFLS2 were reinterviewed (Strauss et al., 2004). Given the high follow-up rate, sample attrition bias is not likely to be a major concern using IFLS.

The IFLS is a multi-purpose survey. It collected a broad array of demographic, socioeconomic and health information on individuals, households, and communities. Much of these information were repeated across waves of the survey. Importantly, the IFLS contains a detailed migration history and a wide range of health indicators through self-reports as well as physical assessments. In the migration history module, information on birth place, place of residence at age 12, and each trip longer than six months since age 12 and a number of characteristics associated with each trip (e.g., date, purpose, and destination) was gathered. In all three waves, self-reported health and measures of height and weight were available. IFLS2 and IFLS3 included additional physical assessments such as hemoglobin level and blood pressure.

### ***Measurement of variables***

Health status is difficult to measure and no agreement exists as to what measures are good indicators of health. This is partly because that the concept of health is complex and is, in fact, multi-dimensional. Strauss and Thomas (1998) have argued that health is composed of distinct components that must be measured and interpreted separately. For this reason, various indicators of health are examined individually in association with migration. The analysis includes several physical assessments as well as self-report of physical functioning. The self-report of general health is not used, as it has been shown in Indonesia and other settings to suffer from systematic biases (Sadana, Mathers, Lopez, Murray, & Iburg, 2000; Thomas & Frankenberg, 2002a).

Activities of daily living (ADL) is used as an indicator of physical functioning. The ADL scale generally comprises of items measuring functionality, such as the ability to kneel, to

carry heavy load, and to dress and stand up without help. ADL measures long-term health, in particular social limitations resulting from severe chronic disease and disability (Johnson & Wolinsky, 1993). Since the data did not collect information on chronic diseases, it is the single measure available that reflect long-term health conditions. Although ADL measures were collected through self-reports, they are less subject to respondent bias than self-reported general health because the questions are well-defined and specific to individual's ability to function at particular tasks (Bound, 1991). Sensitivity analysis using IFLS shows that ADL correlates well with other health measures, but self-reported general health does not. Many of the standard ADL items are of the greatest salience for the elderly. In IFLS, a few items relevant for prime-age adults were included, which measure the respondent's capacity to perform physically strenuous activities, such as carrying a heavy load and walking long distance. These ADLs have been utilized effectively in investigating the health of prime-age adults in previous work (Frankenberg & Jones, 2004). I constructed a dichotomous variable, coded 1 if the respondent reported having difficulties with any of the nine tasks included in IFLS.

Similarly, I created a dichotomous measure indicating self-reported acute disease conditions to study short-term health: whether the respondent had any acute morbidity symptoms in the last four weeks, such as headache, cough, stomach ache, etc. Although self-report measures are sometimes contaminated by recall bias, it is less of a concern using IFLS. This is because the ADL measures ask about severe physical impediments and the acute morbidity questions refer to recent events, both of which are not difficult to remember.

Body Mass Index (BMI), which reflects adult nutritional status and is thought to be correlated with physical capacity, is used frequently as an indicator of adult health. The BMI mainly captures short-term dimensions of nutrition and is subject to changes over a relatively short period of time. I focus on low levels of BMI (underweight), which is more prevalent than overweight in a developing setting like Indonesia. I used a dichotomous measure of whether an individual has a low BMI based on the WHO cut-off 18.5 (NHLBI, 1998).

High blood pressure, or hypertension, is a major health concern in adult population with multiple contributing risk factors. It has no noticeable symptoms, and can be heightened in a relatively short period or develop over many years. I created a dichotomous indicator of whether the respondent has hypertension using cut-offs defined by the American Medical Association—a systolic blood pressure of at least 140 or a diastolic blood pressure of at least 90.

Anemia is a common health problem in developing countries, not only among children but also among adults (WHO, 2002). It is measured by hemoglobin level, and is connected to greater susceptibility to diseases, reduced cognitive development and lower levels of productivity (Thomas & Frankenberg, 2002b). Anemia mainly captures nutritional status for children; for adults, it is also a disease marker that captures blood loss due to chronic conditions such as ulcers or cancer. Similar to hypertension, it has few noticeable symptoms and is usually unknown to people until formal diagnosis. I constructed a dichotomous measure indicating whether the respondent is anemic, based on the WHO cut-off of 13g /dl for women and 13g/ dl for men.

### ***Statistical methods***

To test the healthy migrant hypotheses identified above, I estimated a set of logistic regressions predicting whether an individual moved between 1997 and 2000 from each health indicator along with other control variables, all measured in 1997. All analysis was performed using Stata statistical software (StataCorp, 2005). I used the 1997 and 2000 IFLS, for which longitudinal health data were available. The outcome variables, migration status, were created using migration history information. I first created a binary measure indicating whether a person moved between 1997 (after the completion of the IFLS2 interviews) and 2000. I further differentiated three types of migration by purpose: migration for work, for family-related reasons (including moving for marriage, moving with a family member or to live with a family member), and for all other purposes.

Other control variables include age (an indicator variable defined by 10-year age groups ranging from 18 to 75, except for the youngest group), gender, education (years of completed schooling), log per capita annual household income, place of residence (rural vs. urban residence), marital status (never married, married and living with spouse, married not living with spouse, and all other status such as divorced, separated, and widowed), and previous migration experience (a dichotomous variable indicating whether an individual ever moved before the 1997 interview). All covariates were measured in 1997, prior to the migration measure.

These covariates were incorporated throughout the analyses because they tend to be important factors affecting both migration decisions and health. The control of previous migration experience captures some unmeasured aspects of individual's propensity to move (such as previous life exposure and personality), which works similarly as a lagged dependent

variable. The key test in all models is the statistical significance of the health indicator. In the analysis, respondents older than 75 were excluded to reduce bias due to selective mortality, similarly for those younger than 18, for whom migration is likely to be motivated by decisions of adult family members.

Using binary migration status as the outcome, I estimated separate equations for the prime-age adults (age 18-45 in 1997) and older respondents (age 46-75 in 1997), given the distinctive hypotheses regarding each group. The cut-off (45 years old) is based on the most productive age range (prime-age adults) considered in Indonesia as well as many other developing countries (Frankenberg & Thomas, 2001; Oey, 1982). I conducted sensitivity analysis by using different cut-off points, up to 55 years old, which has almost no impact on the results. The models can be formulated as:

$$\log\left(\frac{p_{ii}}{1 - p_{ii}}\right) = \alpha + \beta hlt_{(t-1)i} + \gamma X_{(t-1)i} \quad (1)$$

where  $p_{ii}$  is the probability that  $i$ th individual moved between 1997 and 2000;  $hlt_{(t-1)i}$  is the dichotomous health status in 1997; and  $X_{(t-1)i}$  is a column vector of individual and household characteristics in 1997 as described above.

I next studied migration for different purposes focusing on the younger group, which give adequate sample sizes in each disaggregated migration category. I used multinomial logistic regressions to model the probability of being in each of the three migration groups relative to the non-migrant group. The models can be formulated as:

$$\log\left(\frac{p_{ij}}{p_{i1}}\right) = \alpha + \beta hlt_{(t-1)i} + \gamma X_{(t-1)i} \quad (2)$$

where  $\frac{p_{ij}}{p_{i1}}$  is the relative probability for  $i$ th individual to choose alternative  $j$  rather



than alternative 1 (stay); and the left-hand side variables are similar to those in equation (1).

Given the nature of the IFLS, the sample included individuals from the same household. Therefore, I used the Huber-White robust estimator in all models to adjust standard errors for clustering of individuals within households to take account of homogeneity within the same family (White, 1980).

In addition, I took into account potential confounding stemming from economic shocks in all models. Any socioeconomic shocks external or internal to the household, such as economic crisis, crop failure, layoffs and natural disaster, may be an impetus for migration and may also cause health problems. This can lead to a spurious association between health and migration. I first included a measure indicating whether the household experienced any economic shocks in the past five years, which is directly available in the 1997 IFLS.

Socioeconomic shocks external to the household are also pertinent because Indonesia entered a period of financial crisis in 1997-1998. The impact of the crisis has been shown to vary by region and household economic conditions (Frankenberg, Thomas, & Beegle, 1999). Therefore, I controlled for household per capita income as an indicator of economic conditions, and province of residence as an indicator of regional heterogeneity to provide more conservative estimates.

Moreover, I addressed potential biases due to unobserved heterogeneity that occurs when unmeasured factors are associated with both health status and the propensity of migration (examples include previous life exposure, genetics, etc.). In this situation, the observed relationship is not necessarily evidence of an association between health and migration, but rather indications of highly selective characters of migrants that affect both the decision to

move and the health status. I used a household fixed-effect (FE) model to absorb the unobserved heterogeneity constant at the household-level. Here I assumed that the unobserved heterogeneity, such as genetic disposition and family background, is stable across adults within the household, which is a relatively reasonable assumption. FE models were utilized only in binary logistic regressions, because multinomial FE logistic regressions require estimation algorithms not widely available. I thus used results from the binary models to assess the possibility that unobserved heterogeneity affects the multinomial estimates.

The FE approach contrasts eligible household members who moved with those who stayed with respect to their differences in health status and other factors. This approach provides further insight into the health selectivity of migration because it conceptualizes migration as a household decision-making process, which is often documented in developing countries (Bhattacharyya, 1985; Mincer, 1978). It examines not only the question of whether migrants tend to be healthier on average than non-migrants, but also whether within a household, healthier members are more likely to move than less healthy members. To estimate a household FE model, equation (1) can be reformulated as:

$$\log\left(\frac{P_{tj}}{1 - P_{tj}}\right) = \alpha + \beta hlt_{(t-1)ij} + \gamma X_{(t-1)ij} + \varepsilon_j \quad (3)$$

where  $\varepsilon_j$  represents all unobserved differences between households (stable within each household  $j$ ). It is regarded as a set of fixed parameters, one per household. The basic idea of fixed-effect logit models is to reformulate the likelihood function using conditional maximum likelihood so that the household-specific parameter  $\varepsilon_j$  is purged out of the equation. This is in contrast to multi-level random-effect models, where  $\varepsilon_j$  is assumed to have a random distribution and is explicitly estimated. A more detailed description can be

found in Wooldridge (2002).

## **Results**

### *Descriptive statistics*

Table 1 presents descriptive statistics for the analytic sample with complete information on all variables, separately for two age groups. There are substantial differences by age. The overall rate of migration is significantly higher for the younger group (11%) than for the older group (3%). Note that the number of migrants in the older group is small, which may limit the statistical power of this analysis.

There are also noticeable differences by purpose of migration, with younger people more likely to move for work-related reasons and older respondents for other and family-related reasons. After further disaggregating the response categories, I find that among the older group, a large fraction reported health-related reasons, in particular sickness or moving to live with family members. This provides direct evidence for the speculation that older people tend to migrate to seek better care. Additionally, Indonesia is characterized by high rates of lifetime migration, as more than 50% of the respondents ever moved before 1997.

As for other attributes, the younger group received three years more schooling on average than the older group. They also tend to enjoy better economic conditions. The per capita income averaged 100,000 Rupiah higher per year in households largely comprising young adults (in 1997, 1 US dollar  $\approx$  2,500 Indonesian Rupiah). The pattern of marital status mainly reflects the aging of the population: the majority of adults above age 45 were once married but they were also more likely to experience marriage dissolution of various

kinds.

All health measures exhibit sufficient variability, even for younger respondents. But the disease rates for younger people average half of those for the older group in most cases. The extremely high rate of short-term morbidity symptoms is unsurprising, because they measure mild and acute conditions that are commonly experienced in everyday life.

The attrition rate for the analytic sample is 8.7%, substantially lower than the majority of longitudinal studies. The concern here is whether panel respondents and people lost to follow-up are different with respect to health. Additional analysis suggests that this is not the case, as only anemia reveals a significant difference. Ideally, I would like to distinguish those who dropped out for various reasons (death, migration, refusal, etc.). However, the data do not provide sufficient information for further distinction.

### ***Health selectivity of migration***

Tables 2 and 3 present logistic regressions predicting whether an individual moved between 1997 and 2000. I estimated separate models for each health indicator, and for the young and the old respectively. Because information on ADLs and morbidities was gathered via self-reports, they have fewer missing data in comparison to physical assessments. To preserve information, the analyses were based on two slightly different samples.

We see from Table 2 that age is negatively associated with the propensity to migrate for both age groups. Gender does not significantly affect the likelihood of moving. Among younger respondents education increases the likelihood of migration whereas household income diminishes the probability. Among the elderly, neither education nor income is related to migration. Previous migration experience is a strong predictor of future migration

behavior for both groups. Importantly, health impairment as measured by ADLs is negatively associated with the likelihood of migration for the younger group, but positively for the older group. The results are consistent with Hypotheses 1 and 2.

Results for other health indicators are reported in Table 3. Coefficients of other control variables reveal similar patterns to those in Table 2, and are thus not shown for the rest of the tables. We see that the remaining health measures do not seem to matter for either age group, though most of them are in the expected direction. These results lend support to Hypothesis 4 that health deficits with a short course or minimal symptoms are unlikely to enter the migration decision-making process. Hypertension, the “silent killer,” shows no symptoms until there is a catastrophic vascular event. Anemia must be quite severe before its symptoms become salient. Low BMI may be regarded as an individual trait rather than a symptom of poor health. And morbidity in the month before the survey presumably mainly refers to acute conditions, from which individuals recover. By contrast, ADL deficits are clear and unambiguous signals that something is amiss. People who experience shortness of breath or have difficulty in walking a distance or carrying a load know that they are not strong enough to risk migration. Thus, it is unsurprising that ADL deficits are associated with a reduced probability of migrating, but that the remaining health indicators show no relationship.<sup>2</sup> This conclusion holds even when all health measures are entered simultaneously into the model.

Table 4 disaggregates migration by purpose, and repeats the analysis of Tables 2 and 3. I restricted the analysis to the younger group because they have sufficient cases. Here we see that ADL deficits affect only migration for work but not that for other reasons. This is consistent with Hypothesis 3 that health is mainly implicated in labor migration decisions.

Health deficits as measured by three biometric indicators also seem to be negatively related to labor migration, though they lack sufficient statistical significance. The patterns for other types of migration are mixed, suggesting a different decision-making process. As for other covariates not shown, most of them have similar impacts on labor migration as in Table 2, with the exception of gender and place of residence. Males and rural residents clearly have a much higher propensity of moving for work-related reasons, a pattern that is made unambiguous when I subdivide by types of migration.

Table 5, in which models are corrected for unobserved heterogeneity, gives consistent results. ADL deficits but not other health conditions reduce the probability of moving. This suggests that the association between migration and health is relatively robust to unobserved confounding factors at the household-level. This result can also be interpreted as indicating that even within a household, healthier members are more likely to move than are others.

## **Discussion**

This paper examines whether health conditions are associated with migration decisions in the context of a growing migration population in Indonesia. It represents one of the first attempts to formally test the “healthy migrant hypothesis” by incorporating pre-migration health information on people who move and who stay.

I find support for the conjecture that, in addition to demographic and socioeconomic factors, health considerations also play a role in migration decision-making. But the relationship between health and migration is complex because migrants are a heterogeneous group: younger migrants are positively selected with respect to health whereas older migrants

are negatively selected. This is not surprising, especially since it turns out that the positive selection of younger migrants is mainly restricted to labor migrants. Among older people health problems appear to be a major reason for migrating, as they often move to seek improved medical care, and the support of relatives who can provide care. Younger migrants, and labor migrants in particular, tend to be negatively selected for chronic health conditions and disabilities, reflected in the inability to perform one or more “Activities of Daily Living”. No other health indicators seem to matter, presumably because they either show little or no symptoms that may affect physical activities (e.g., anemia and hypertension) or because they are regarded as temporary (e.g., acute morbidities). Moreover, the health selectivity of migration is robust to household unobserved heterogeneity. I find not only that migrants are healthier on average than are non-migrants, but within households, migrants are healthier than non-migrant household members, net of individual factors included in the models.

A few limitations should be noticed. The survey only captures relatively long-term internal migration (moves that last more than six months), thereby limiting the generalizability of the results to short-term and international migrants. Also, information on the purpose of migration is not very finely-tuned, and was gathered after the moves. Therefore, a fair amount of error may be introduced in the distinction.

Despite these limitations, the data quite convincingly provide support to the long-standing presumption—the “healthy migrant effect”—and help explain the “epidemiological paradox” that associates migrants with lower mortality and morbidities. While earlier work disproportionately focused on international migration, the present study examines internal migration, an equally important yet understudied phenomenon. Results

clearly demonstrate that the general theoretical framework usually proposed for international migration operates in similar ways for internal migration. As speculated, health is also implicated in internal migration choices. These findings contribute to development of a unified framework in understanding factors that generate both internal and international migrant flows. They also extend the literature on population health, suggesting the need for conceptualizing health not only as an outcome resulting from various socioeconomic processes, but as a potential determining factor of such processes.

However, it is not to suggest that there is no difference in health selection effects between two migration streams. Immigrants may be more rigorously selected on health because international moves are especially physically demanding: it implies crossing national boundaries and overcoming various cultural and legal barriers; and it often involves health screening for infectious diseases. The extent to which this conjecture is true needs further investigation. Some other future directions include studying additional aspects of health, and temporal change of the selectivity. They will yield valuable information, as migrants may be selected on dimensions of health not studied here, and over time they may be increasingly drawn from a less selective pool as more people enter the flow and migration gains momentum.

Importantly, findings regarding the health selectivity of migration can be used to better understand the implications of migration for health. Early work has illustrated either a protective or a disruptive effect of the migration experience on health over the course of migrants' stay. However, most of them fail to address potential health selection of migration. As the next stage of this project, I will explore this aspect of the migration-health association



by incorporating health conditions prior to and after migration to correct for the health selection effects.

### **Endnotes**

1. The analysis excludes involuntary migration associated with government programmes for population redistribution (the “Transmigration Program”) and refugee migration, which form only a small proportion of total migration.

2. I also examined two additional health measures, self-reported general health and high levels of BMI (overweight). Neither is significantly associated with migration decisions, which is not surprising for reasons stated above.

## References

- Anson, J. (2004). The migrant mortality advantage: A 10 month follow-up of the Brussels population. *European Journal of Population*, 20, 191-218.
- Bhattacharyya, B. (1985). The role of family decision in internal migration: The case of India. *Journal of Development Economics*, 18(1), 51-66.
- Bound, J. (1991). Self-reported versus objective measures of health in retirement models. *Journal of Human Resources*, 26(1), 106-38.
- Centers for Disease Control and Prevention. (1998). Recommendations for prevention and control of tuberculosis among foreign-born persons. *Morbidity Mortality Weekly Report*, 47, 1-26.
- Chernichovsky, D., & Meesook, O.A. (1986). Utilization of health services in Indonesia. *Social Science and Medicine*, 23(6), 611-20.
- De Jong, G.F., Abad, R.G., Arnold, F., Carino, B.V., Fawcett, J.T., & Gardner, R.W. (1983). International and internal migration decision making: A value-expectancy based analytical framework of intentions to move from a rural Philippine province. *International Migration Review*, 17(3), 470-84.
- Evans, J. (1987). Introduction: migration and health. *International Migration Review*, 21(3), v-xiv.
- Feranil, A. B. (2005). Anaemia among migrant and non-migrant mothers in disadvantaged areas in the Visayas, the Philippines. In S. Jatana, M. Toyota, & B. Yeoh (Ed.), *Migration and health in Asia* (pp. 100-15). London: Routledge.
- Findley, S. E. (1988). The directionality and age selectivity of the health-migration relation: Evidence from sequence of disability and mobility in the United States. *International Migration Review*, 22, 4-29.
- Frankenberg, E., & Jones, N. R. (2004). Self-rated health and mortality: Does the relationship extend to a low income setting? *Journal of Health and Social Behavior*, 45(4), 441-52.
- Frankenberg, E., & Karoly, L. (1995). The 1993 Indonesian Family Life Survey: Overview and field report. Publication No. DRU-1195/1-NICHD/AID. Santa Monica, CA: Rand Corporation.
- Frankenberg, E., & Thomas, D. (2000). The Indonesia Family Life Survey (IFLS): Study design and results from waves 1 and 2. Publication No. DRU-2238/Volume1/

- NIA/NICHD. Santa Monica, CA: Rand Corporation.
- Frankenberg, Elizabeth, and Duncan Thomas. 2001. "Women's health and pregnancy outcomes: Do services make a difference?" *Demography* 38(2):253-65.
- Frankenberg, E., Thomas, D., & Beegle, K. (1999). The real costs of Indonesia's economic crisis: Findings from the Indonesia Family Life Surveys. Publication No. DRU-2064-NIA/NICHD. Santa Monica, CA: Rand Corporation.
- Freedman, R. (1947). Health differentials for rural-urban migration. *American Sociological Review*, 12(5), 536-41.
- Hugo, G. (2005). Population movement in Indonesia: Implications for the spread of HIV/AIDS. In S. Jatrana, M. Toyota, & B. Yeoh (Ed.), *Migration and health in Asia* (pp. 17-40). London: Routledge.
- Hull, D. (1979). Migration, adaptation, and illness: A review. *Social Science and Medicine*, 13A, 25-36.
- Jasso, G., Massey, D.S., Rosenzweig, M.R., & Smith, J.P. 2004. Immigrant health: Selectivity and acculturation. In N.B. Anderson, R.A. Bulatao, & B. Cohen (Ed.), *Critical Perspectives on Racial and Ethnic Differences in Health in Late Life* (pp. 227-66). Washington, DC: National Academies Press.
- Johnson, R. J., & Wolinsky, F. D. (1993). The structure of health status among older adults: Disease, disability, functional limitation, and perceived health. *Journal of Health and Social Behavior*, 34(2), 105-21.
- Marmot, M. G., Adelstein, M. A., & Bulusu, L. (1984a). Immigrant mortality in England and Wales 1970-78. OPCS Studies of Medical and Population Subjects No. 47. London: HMSO.
- Marmot, M. G., Adelstein, M. A., & Bulusu, L. (1984b). Lessons from the study of immigrant mortality. *Lancet*, 323, 1455-7.
- Massey, D.S. (1988). Economic development and international migration in comparative perspective. *Population and Development review*, 14, 383-14.
- McKinlay, J. B. (1975). Some issues associated with migration, health status and the use of health services. *Journal of Chronic Disease*, 28, 579-92.
- Mincer, J. (1978). Family migration decisions. *Journal of Political Economy*, 86, 749-773.
- Muhidin, S. S. (2002). The population of Indonesia: Regional demographic scenarios using a

- multiregional method and multiple data sources. Ph.D. thesis at the University of Groningen. Amsterdam: Rozenberg Publishers.
- National Heart, Lung, and Blood Institute (NHIBI). (1998). Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults. Maryland: NHIBI.
- Oey, M. (1982). The transmigration program in Indonesia. In G.W. Jones, & H.V. Richter (Ed.), *Population Resettlement Programs in Southeast Asia* (pp.27-51). Australian National University
- Patrick, C. H. (1980). Health and migration of the elderly. *Research on Aging*, 6(2), 233-41.
- Palloni, A., & Morenoff, J. D. (2001). Interpreting the paradoxical in the Hispanic paradox: Demographic and epidemiologic approaches. *Annals of the New York Academy of Sciences*, 954(1), 140-74.
- Pryor, R.J. (1981). Integrating international and internal migration theories. In M.M. Kritz, C.B. Keely, & S.M. Tomasi (Ed.), *Global Trends in Migration: Theory and Research on International Population Movements* (pp. 110-129). Center for Migration Studies, Staten Island, New York.
- Sadana, R., Mathers, C. D., Lopez, A. D., Murray, C. J. L., & Iburg, K. (2000). Comparative analysis of more than 50 household surveys on health status. GPE Discussion Paper No 15. Geneva: World Health Organization.
- Speare, A. Jr., & Harries, J. (1986). Education, earnings and migration in Indonesia. *Economic Development and Cultural Change*, 34(2), 223-44.
- StataCorp. 2005. Stata Statistical Software: Release 9.0. College Station, TX: StataCorp.
- Strauss, J., & Thomas, D. (1998). Health, nutrition, and economic development. *Journal of Economic Literature*, 36, 766-817.
- Strauss, J., Beegle, K., Sikoki, B., Dwiyanto, A., Herawati, Y., & Witoelar, F. (2004). The third wave of the Indonesia Family Life Survey (IFLS3): Overview and field report. Publication No. WR-144/1-NIA/NICHD. Santa Monica, CA: Rand Corporation.
- Sukamdi, A. H., & Brownlee, P. (1998). Labour migration in Indonesia: Policies and practices. Population Studies Center Working Paper, Gadjah Mada University. <http://www.unesco.org/most/apmrlabo.htm>, accessed August, 2007.
- Thomas, D., & Frankenberg, E. (2002a). The measurement and interpretation of health in social surveys. In C. Murray, J. Salomon, C. Mathers, & A. Lopez (Ed.), *Measurement of*

- the Global Burden of Disease* (pp. 387-420). Geneva: World Health Organization.
- Thomas, D., & Frankenberg, E. (2002b). Health, nutrition, and economic prosperity: A microeconomic perspective. *Bulletin of the World Health Organization*, 80, 106-13.
- White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test of heteroskedasticity. *Econometrica*, 48, 817-38.
- Wooldridge, J. (2002). *Econometric analysis of cross section and panel data*. Cambridge, Mass.: MIT Press.
- World Health Organization (WHO). (2002). *The world health report 2002: Reducing risks, promoting healthy life*. Geneva: World Health Organization.

## Tables

Table 1.

Means and percentages of variables used in the analysis by age group, Indonesian adults age 18-75 (standard deviations in parentheses).

Variables <sup>a</sup>	Age 18-45	Age 46-75
<i>Migration status</i>		
Migration between waves* (%)	11.1	3.1
Migration by purpose (for migrants only) *	(N=1074)	(N=137)
Labor migrant (%)	21.5	16.7
Family migrant (%)	42.4	34.3
Other migrants (%)	36.1	49.0
Ever moved before* (%)	49.4	55.3
<i>Health measures</i>		
Problems with ADL* (%)	19.2	44.4
Morbidity in last month* (%)	78.1	83.1
Low BMI* (%)	13.2	22.8
Hypertension* (%)	18.7%	46.7%
Anemia* (%)	28.5%	38.0%
<i>Control variables</i>		
Age*	31.3 (8.0)	57.2 (7.7)
Male* (%)	42.2	47.0
Years of completed schooling*	7.3 (4.4)	4.1 (4.2)
Per capita annual income (in thousands of 1997 Rupiah)*	693.3 (939.0)	579.3 (906.6)
Marital status*		
Never married (%)	22.5	0.7
Married, living with spouse (%)	70.9	77.2
Married not living with spouse (%)	3.0	1.2
Other (%)	3.7	21.0
Rural residence* (%)	51.9	57.1
Economic shock in past five years (%)	41.4	41.3
N	9666	4455

<sup>a</sup> All variables were measured in 1997, with the exception of migration between waves and by purpose, which were gathered in 2000.

\*  $p$ -value < 0.05. P-values were obtained from chi-square tests or t-test, depending on whether the variable is continuous or categorical.

Table 2.

Logistic regression of migration between 1997 and 2000 on problems with ADL and other predictors measured in 1997, IFLS 1997-2000 (robust standard errors in parentheses).<sup>a</sup>

Logits	Age 18-45	Age 46-75
Problems with ADL	-0.243** (0.093)	0.373* (0.187)
Age	(ref. 18-25)	(ref. 46-55)
Second decile (26-35 or 56-65)	-0.601*** (0.089)	-0.175 (0.193)
Third decile (36-45 or 66-75)	-1.318*** (0.121)	-0.514† (0.284)
Male	-0.001 (0.060)	0.344* (0.164)
Years of schooling	0.071*** (0.010)	0.021 (0.023)
Log per capita annual income	-0.016† (0.008)	-0.011 (0.018)
Marital status (ref. never married)		
Married with spouse	-0.717*** (0.106)	-1.153† (0.659)
Married not with spouse	-0.075 (0.187)	-0.431 (0.822)
Other	-0.126 (0.201)	-0.815 (0.682)
Ever moved before 1997	0.741*** (0.072)	1.105*** (0.204)
Rural residence	-0.206* (0.087)	-0.555** (0.211)
Economic shocks	-0.095 (0.080)	0.357† (0.196)
Constant	-1.507*** (0.220)	-3.195*** (0.752)
Log-likelihood	-3371.5	-653.7
<i>N</i>	10520	4787

<sup>a</sup> Estimates for province of residence are not shown.

\*\*\*  $p$  value < 0.001; \*\*  $p$  value < 0.01; \*  $p$  value < 0.05; †  $p$  value < 0.1

Table 3.

Logistic regression of migration between 1997 and 2000 on health status and other predictors measured in 1997, IFLS 1997-2000 (robust standard errors in parentheses).<sup>a</sup>

Logits	Morbidity last month	Low BMI	Hypertension	Anemia
<i>Age 18-45</i>				
Health measures	-0.014 (0.082)	-0.081 (0.099)	0.059 (0.093)	-0.134 (0.082)
Log-likelihood	-3375.1	-3001.9	-3002.1	-3000.9
<i>N</i>	10520	9672	9672	9672
<i>Age 46-75</i>				
Health measures	0.130 (0.228)	0.021 (0.232)	-0.025 (0.184)	0.066 (0.185)
Log-likelihood	-655.6	-578.2	-578.2	-578.1
<i>N</i>	4787	4459	4459	4459

<sup>a</sup> Estimates for other predictors (same as in Table 2) are not shown.

\*\*\*  $p$  value < 0.001; \*\*  $p$  value < 0.01; \*  $p$  value < 0.05; †  $p$  value < 0.1



Table 4.

Multinomial logistic regression of migration by purpose between 1997 and 2000 on health status and other predictors in 1997, adults age 18-45, IFLS 1997-2000 (robust standard errors in parentheses).<sup>a</sup>

Logits	Labor migrants	Family migrants	Other migrants	Log-likelihood	<i>N</i>
ADL	-0.847*** (0.238)	-0.145 (0.132)	-0.120 (0.138)	-4552.1	10520
Morbidity last month	0.003 (0.029)	0.028 (0.021)	-0.006 (0.023)	-4558.9	10520
Low BMI	-0.164 (0.189)	-0.035 (0.139)	-0.063 (0.158)	-4039.2	9672
Hypertension	-0.109 (0.192)	0.029 (0.140)	0.168 (0.141)	-4038.8	9672
Anemia	-0.262 (0.178)	-0.184 (0.123)	-0.017 (0.121)	-4037.5	9672

<sup>a</sup> The base category is non-migrants. Estimates for other predictors (same as in Table 2) are not shown.

\*\*\*  $p$  value < 0.001; \*\*  $p$  value < 0.01; \*  $p$  value < 0.05; †  $p$  value < 0.1

Table 5.

Household fixed-effect logit models of migration between 1997 and 2000 on health status and other predictors in 1997, adults age 18-45, IFLS 1997-2000 (standard errors in parentheses).<sup>a</sup>

Logits	ADL	Morbidity last month	Low BMI	Hypertension	Anemia
Health measures	-0.508* (0.236)	-0.184 (0.185)	0.339 (0.256)	-0.198 (0.228)	-0.022 (0.202)
Log-likelihood	-325.5	-327.4	-274.3	-274.9	-275.3
<i>N</i>	1256	1256	1083	1083	1083

<sup>a</sup> Estimates for other predictors (same as in Table 2) are not shown.

\*\*\* *p* value < 0.001; \*\* *p* value < 0.01; \* *p* value < 0.05; † *p* value < 0.1