DISCUSSION PAPER SERIES

IZA DP No. 3458

Selective Migration and Health

Timothy J. Halliday
Michael C. Kimmitt

April 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Selective Migration and Health

Timothy J. Halliday
University of Hawaii at Manoa
and IZA

Michael C. Kimmitt
University of Hawaii at Manoa

Discussion Paper No. 3458
April 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3458
April 2008

ABSTRACT

Selective Migration and Health*

Using data from the Panel Study of Income Dynamics, we investigate the impact of health on domestic migration within the United States. We find that, for men below 60 years of age, a move from the middle to the bottom of the health distribution reduces mobility by 32-40%. Non-random attrition from the panel implies that these are lower bounds. By contrast, we find evidence that, among older men, there is higher mobility at the top and bottom of the health distribution than there is in the middle. For women, we find no evidence of a relationship between their own health and mobility, although spousal health does affect the mobility of married women.

JEL Classification: J61

Keywords: migration, health, selection, attrition

Corresponding author:

Timothy J. Halliday
Department of Economics
University of Hawai'i at Manoa
2424 Maile Way
Saunders Hall 533
Honolulu, HI 96822
USA
E-mail: halliday@hawaii.edu

* This paper was originally the third chapter of my dissertation at Princeton University. I would like to thank my advisors, Chris Paxson and Bo Honoré, for comments on an earlier draft. I would also like to thank Randy Akee for useful discussions. The usual disclaimer applies.

1 Introduction

Social scientists have long recognized that health can have an impact on both the costs and the benefits of migration. As pointed out by Norman, Boyle, and Rees (2005), this observation dates back at least to Farr (1864) and Welton (1872). Numerous mechanisms are possible. For example, illness may increase the costs of executing all the activities that are necessary to migrate. In addition, poor health often necessitates an infrastructure to facilitate and maintain activity, doctors to monitor illnesses, and medications and insurance to pay for it all. On the other hand, there are plausible scenarios in which illness may affect the net benefits of migration. For example, sick people may have a greater incentive to migrate to locations where there is a stronger familial support network or better health care. Indeed, Scribner (1994, 1996) has suggested that this is one of the primary reasons Latino immigrants return to their homelands from the United States. Overall, there are good reasons to expect health to influence a person's migration propensity, although the sign and magnitude of this effect are not a priori obvious.

We use data from the Panel Study of Income Dynamics (PSID) to investigate how health affects domestic migration within the United States. One advantage of focusing on internal U.S. migration is that it enables us to observe the migrant's health prior to migration and thus provides us with a direct test of selection. We ask the following questions. What is the magnitude and direction of the effect of health on migration? Is this effect altered in any meaningful ways by the age, gender, or marital status of the migrant? Are our estimates affected by our choice of health measure? How does panel attrition affect our results?

The balance of this paper is organized as follows. In the next section, we summarize some of the existing literature on migrant selectivity. In section 3, we describe the data. In section 4, we discuss our empirical methodology. In section 5, we discuss our empirical findings. In section 6, we discuss some sources of bias in our estimates including systematic biases in self-reported health status and non-random attrition from the panel. Section 7 concludes.

2 Literature Review

There has been substantial research on migrant selection in both health and labor market outcomes in the social sciences. In the economics literature, research has focused on labor market outcomes such as wages. In the demography and public health literatures, research typically focuses on health differentials between international migrants and non-migrants. These studies often cite selection as one possible reason for migrant/non-migrant health differentials, but usually do not test for it. There are also studies on domestic migration which argue that health-based selection can affect the relationship between economic deprivation and health over time. However, by and large, neither the domestic nor the international migration studies directly test for health-based selection using regression analysis. Our study helps to fill this void.

Studies in economics have focused on whether international migrants are positively or negatively selected from the wage distributions of the sending region. In one of the earliest studies on the topic, Chiswick (1978) observed that in the period immediately after their arrival in the United States, immigrants tended to earn less than natives; however, after spending ten to fifteen years in the United States, migrant earnings overtook native earnings. The explanation that was given for this was that migrants are positively selected from their home countries. Subsequently, Chiswick's findings were contested by Borjas (1987), who pointed out that Chiswick's estimates may have been biased by omitted cohort effects. Borjas contended that these omitted cohort effects may matter if there has been a progressive deterioration in migrant quality with successive cohorts, and he then went on to present evidence of these cohort effects. However, it is important to point out that Borjas's findings have also been questioned by Jasso and Rosenzweig (1990) in a comment on his work.

There has been further economic research on migrant selectivity since the original work of Chiswick and Borjas. For example, in a study that is methodologically similar to ours, Gabriel and Schmitz (1994) tested for positive selection in internal U.S. migration. They showed that, prior to moving, migrants earned higher wages than demographically comparable non-migrants. More recently, Chiquiar and Hanson (2005) corrected many of the shortcomings of the previous work and used Mexican and U.S. Census data to calculate counterfactual wage densities. They showed that Mexican immigrants in the United States would have occupied the middle to upper parts of the Mexican wage distribution had they remained in Mexico. Overall, our reading of the economics literature is that both international and intranational migrants are either positively selected or at least not adversely selected on labor market outcomes.

There is also a substantial and related literature in demography and epidemiology that investigates the relationship between health and migration. By and large, this literature focuses on international migration and documents a correlation in which migrants tend to be in better health than non-migrants. Our study differs from this literature in focusing on domestic migration within the United States and explicitly tests for selection on health status.

Interestingly, the literature on health differentials between migrants and non-migrants has posited several underlying reasons for the correlation other than positively selective in-migration. First, some have speculated that there is adversely selective out-migration. For example, Pablos-Mendéz (1994) and Scribner (1994 and 1996) speculated that the observed lower mortality rates among Latinos in the United States, or the "Latino Paradox," is a consequence of Latino migrants returning to their home countries to die once they fall ill–the so-called "Salmon Bias." Such return migration would mean that many Latino deaths would never get recorded in the U.S. death records, thereby rendering many immigrants "statistically immortal" (Pablos-Mendéz 1994). However, the evidence on the validity of this hypothesis is mixed. Palloni and Arias (2004) provide evidence that the Salmon Bias is important in explaining the Latino Paradox; however, Abriado-Lanza et al. (1999) provide evidence that mortality differentials exist for Puerto Ricans, whose deaths do get recorded in the National Death Index, and Cuban migrants, who for political reasons do not return home. A second explanation, suggested by Marmot and Syme (1976) and Markides and Coreil (1986), is that more favorable cultural and behavioral factors are responsible for better health outcomes among migrants.

A third explanation posits that, just as migrants are positively selected on their ability to perform in the labor market, they are also positively selected on their health status. There is some evidence of this in the demography and epidemiology literatures. For example, Jasso, Massey, Rosenzweig, and Smith (2004), using the National Health Interview Survey, show that international migrants in the United States have a lower prevalence of many chronic conditions upon arrival. In another study, Marmot, Adelstein, and Bulusu (1984) compared mortality rates of migrants in the United Kingdom to the corresponding mortality rates of non-migrants in the sending countries; they found a strong tendency for mortality rates to be substantially lower among migrants than among non-migrants. Swallen (2002) has found similar findings in the United States. However, Newbold (2005) used Canadian data and did not find any differences in self-rated health between migrants and non-migrants upon arrival. Generally speaking, while there is some suggestive evidence for positive health selectivity in international migration, there is a dearth of rigorous tests of health-based selection in international migration. This probably stems from data limitations.

Better data have made it possible for researchers to investigate health-based selection in domestic migration. Many of these studies investigate how selective migration alters the relationship between indices of a geographic location's economic and health deprivation over time. For example, Brimblecome, Dorling, and Shaw (1999), Boyle, Norman, and Rees (2002), Norman, Boyle, and Rees (2005), and Cox, Boyle, Davey, and Morris (2007) show that selective migration can intensify the relationship between material and health deprivation over time. One study, by Findley (1988) directly investigates health-based selectivity. She finds that "there is little temporal connection between migration and health status changes" for people under 45, but she does find evidence for negative selection at older ages.

Our study makes the following contributions to the literature on migrant selectivity. First, we provide additional evidence that just as working-age male migrants tend to be positively selected on labor market outcomes, they are also positively selected on health outcomes. Second, we provide one of the first investigations into migrant selectivity on health status that makes extensive use of regression analysis. Third, we investigate how this selection varies across important demographic groups. Fourth, in contrast to Findley, who uses cross-sectional data, we use ten years of panel data, which provide us with good measures of migration and the evolution of health status over time and, more importantly, enough observations to provide us with adequate power. Fifth, we highlight some important issues that arise when testing for selection with panel data.

3 The Data

We use data from the PSID spanning the years 1984 to 1993 on geographic mobility, health status, and other control variables, including age, income, gender, education, race, and marital status. The PSID only collects data on health status for heads of household and their spouses. Consequently, in this analysis we restrict our attention to these people. Our migration measure, Moved, is an indicator of whether an individual moved from one state to another during the survey years; this indicator is turned on if the individual lived in a different state in the previous time period. This definition of migration is common in the literature on internal migration within the United States (see, e.g., Borjas, Bronnars, and Trejo 1992; Gabriel and Schmitz 1994). We use Self-Reported Health Status (SRHS) as our primary measure of health. SRHS is a categorical variable that takes on integer values between 1 and 5, with 1 being the most healthy category and 5 the least healthy. While these measures are subjective, there is an extensive literature that has shown a strong link between SRHS and health outcomes such as mortality and the prevalence of disease (Mossey and Shapiro 1982; Kaplan and Camacho 1983; Idler and Kasl 1995; Smith 2003). We break the SRHS variable into two binary indicator variables: Healthy, which is turned on when SRHS is either 1 or 2, and Unhealthy, which is turned on when SRHS is either 4 or 5. In addition, we employ the variable Disability, which is turned on if a respondent claimed to have a "physical or nervous condition that limits the type of work or the amount of work" that he or she could do. The survey defined work as including housework, so this variable should be pertinent to women and retired people. All other variable definitions and descriptive statistics for this sample can be found in Table 1. Finally, the PSID contains an over-sample of economically disadvantaged people called the Survey of Economic Opportunities (SEO). In this paper we include the SEO. We do so because dropping it would have substantially reduced the number of moves in our data and potentially compromised the power of our study.

4 Empirical Methods

Our primary estimation equation is

$$m_{i,t} = 1(g_{i,t-1}\gamma + b_{i,t-1}\beta + \mathbf{X}'_{i,t}\boldsymbol{\theta} + \varepsilon_{i,t} \geq 0) \tag{1}$$

where $m_{i,t}$ is Moved, $g_{i,t}$ is Healthy, $b_{i,t}$ is Unhealthy, and $\mathbf{X}_{i,t}$ is a column vector of individual characteristics. Note that in the above specification, the middle SRHS category (i.e., SRHS equal to 3) is omitted. We omitted the middle category because this makes it easier to see what types of selection are present. For example, if there is positive selection then $\beta < 0$ and $\gamma > 0$. On the other hand, if there is negative selection then $\beta > 0$ and $\gamma < 0$. The control variables that we include in $\mathbf{X}_{i,t}$ are age, education, race, state, and marital status dummies, as well as a quadratic function of (lagged) income. All models are estimated separately by gender using Probit regressions, which assume normal residuals. Given this, we can write the migration propensity as

$$P\left(m_{i,t} = 1|\mathbf{Z}_{i,t}\right) = \Phi\left(\mathbf{Z}'_{i,t}\boldsymbol{\lambda}\right) \tag{2}$$

where $\mathbf{Z}_{i,t} \equiv \left(g_{i,t-1}, b_{i,t-1}, \mathbf{X}'_{i,t}\right)'$, $\boldsymbol{\lambda} \equiv (\gamma, \beta, \boldsymbol{\theta}')'$ and $\Phi(.)$ is the CDF of a standard normal random variable. All tables report marginal effects, which were calculated either by differentiating equation (2) with respect to a continuous variable and evaluating the derivative at the mean of the right-hand side variables or, in the case of binary variables, differencing the probabilities evaluated at each value of the binary variables while evaluating the remaining variables at their means. All standard errors are adjusted for correlations within individuals.

A crucial part of our identification strategy is to include lags of the health status variables. We do this because a proper test of health selectivity involves a comparison of the health of movers and stayers prior to the occurrence of any migration, since there are numerous reasons to expect migration to have an impact on a person's health. For example, if a sick person moves to be closer to a family member who is able to take care of him or her, then we might expect the move to improve the person's health. On the other hand, Kasl and Berkman (1983) and Findley (1988) have speculated that the stress of moving may induce a deterioration in health.

5 Empirical Results

5.1 Results Using SRHS

Table 2 displays our results for men and women younger than age 60 estimated separately by gender. The first three columns show the results for men and the last three for women. In all six columns we see substantial evidence that migrants are positively selected on health. The F-tests show that the health variables are always jointly significant. For men, being unhealthy decreases the probability of migration by between 1.2 and 1.5 percentage points. For women, being unhealthy decreases the probability of migration by between 0.4 and 0.6 percentage points. Since the average probabilities of migration for men and women under 60 are 3.7% and 3.3%, respectively, these marginal effects constitute rather large effects in percentage terms. Indeed, moving from good health, which is the omitted SRHS category in these regressions, to worse health lowers the probability of migration by 32-40% for men and 12-18% for women. In columns three and six, we add a comprehensive set of controls, including a complete set of state dummies. While the coefficients on the health variables are attenuated somewhat, they still remain jointly significant at levels higher than 95%. The coefficients on the additional control variables all have the expected signs.

In Table 3 we estimate the models using a sub-sample of people older than 60. We consider the same specifications that we considered in Table 2. The effects of health on migration are substantially different for older people than they are for younger people. For men, we see that being healthy and being unhealthy both have positive effects on migration probabilities. Thus health appears to have a non-monotonic effect on mobility. By contrast, we do not see any evidence that health affects the mobility of women over 60. The positive effect of being healthy on migration may indicate that good health reduces migration costs. The positive effect of being unhealthy on migration suggests that illness increases the benefits to migration. For example, older people who are exceptionally ill may migrate to be closer to family members who can care for them. We might expect this effect to be especially large among a population of older people, many of whom are widowed and may lack a household member to care for them in old age.

5.2 Results Using Disability Status

In Table 4 we employ information on disability to estimate the impact of health on migration. We employ these disability measures instead of SRHS because they are ostensibly a more direct measure of functional impairment and thus may be more closely associated with migration propensities than SRHS. First, to verify that the disability measures in the PSID are correlated with health status, we estimate Probit models in which we predict disability status using functions of SRHS and other control variables. These results are reported in the first panel of the table. We see that there is a very strong association even after we partial out the effects of confounding factors such as age, education, income, race, state, and marital status. Next, to estimate the impact of disability on migration, we run the same specification as columns 3 and 6 of Table 2 except with the disability indicator in lieu of functions of SRHS. The results are reported in the bottom panel of the table. Interestingly, we do not see any association between disability and migration despite the fact that it is strongly associated with SRHS and that 19% of our sample reports some work-limiting disability. One possible reason for this is that many ill people may not consider themselves disabled, but still may not be able to move because of the costs of changing insurance and finding new physicians. Moreover, many people may erroneously claim to be disabled to justify being out of the labor force. Overall, we believe that there is substantial measurement error in the disability statistics.

5.3 Results by Marital Status

We now explore the effects of an individual's own health and spousal health on migration for a sub-sample of married couples. In Table 5 we report the results for people under age 60. The results for married men, which we report in the first two columns, show that his own health matters but the health of his spouse does not. The results for women, which are reported in the last two columns, show the opposite. In Table 6 we report the results for people age 60 or older, and we see the same pattern, namely, that married men's own health matters and that the spouse's health matters for married women. One explanation for this intriguing finding is that, in married couples, illness in a husband has a larger impact on moving costs than illness in a wife. This would be true, for example, if many of the physical burdens of migration, such as lifting, are borne by men. Another explanation (which is related) is that the set of conditions associated with low SRHS for men is more debilitating than the conditions associated with low SRHS for women.

5.4 Additional Results

In the first two columns of Table 7 we investigate the impact of having a family support network on the migration of older people. We do this by investigating whether being married attenuates the impact of poor health on migration. The rationale behind this question is that married individuals who are in poor health can be taken care of by their spouses and would thus have fewer incentives to relocate to a state where they can be cared for by another family member. To test this "family support network" hypothesis, we estimate modifications of the regressions from columns 3 and 6 of Table 3 where we interact the unhealthy indicator with the marriage indicator. If this hypothesis is true, then we would expect the interaction between bad health and the marriage indicator to be negative. We see that the interaction term is negative and significant for men, but is negative and insignificant for women. The evidence for this hypothesis is mixed.

The next two columns of Table 7 look for additional interactions between age and health beyond the break at age 60 that we have already considered. We do this by interacting the variables Healthy and Unhealthy, with dummies for being between ages 40 and 49 and between 50 and 59. Only one of the interactions is significant. The only thing of note in the table is that the coefficients on the interactions between Unhealthy and the dummy for being between age 50 and 59 is positive. This is indicative of the transition from the negative effects of poor health on mobility that we see at younger ages to the positive effects that we see at older ages. We conclude that the relationship between health and migration is fairly stable roughly until age 60, when retirement becomes more common and health starts to decline more rapidly.

The last two columns of Table 7 compare the relationship between health and migration for married and single women. Interestingly, we see a significant relationship for married women, but not for single women. Because Tables 5 and 6 show that the effects for married women operate entirely through their husbands, there does not appear to be any evidence that a woman's own health affects her migration. Note, however, that Halliday (2007) presented results from a specification similar to that in the last column of Table 7 but used contemporaneous health in lieu of lagged health and found significant effects. We conclude that, while significant effects of health on mobility have been found for single women in the PSID, they are sensitive to specification.

6 Potential Sources of Bias

In this section, we explore some factors that may bias our estimates. The next sub-section investigates the role of reporting bias in SRHS. After that, we look into the role of non-random attrition from the panel.

6.1 Systematic Reporting Bias in SRHS

While it true that SRHS has proven to be a reliable measure of health, it is still subject to biases and errors, which in some contexts have been shown to be systematically correlated with socioeconomic variables. If similar errors exist in our data that are systematically correlated with migrant status, then our results will be biased. For example, if movers are systematically more optimistic about their health than stayers, then this would result in estimates that look as if illness raises the costs of migration when, in fact, there may be no actual relationship between health (in the objective sense) and migration.

To investigate this issue, we employ the PSID's mortality file and estimate the relationship between SRHS and mortality using Cox-Proportional Hazard Models while adjusting for age. This model can be written as

$$\lambda_i(t) = \lambda(t) \exp(w_i' \delta)$$

where t denotes years survived and w_i includes age and functions of SRHS. We employ a partial likelihood procedure that imposes no assumptions $\lambda(t)$. For the technical details of this procedure, we refer the reader to pp. 449-50 of Amemiya (1985).

To see if there were any systematic biases across movers and stayers, we split the sample by migration status. One sample contained people who never moved while they were in our sample, and the other contained people who moved at least once. We then estimated the hazard models on both samples by gender and age using the 1984 wave of the PSID. Each cell reports the hazard ratio for the relevant variable and its 95% confidence interval. If the hazard ratio is above (below) unity, then that variable has a positive (negative) effect on mortality. Results are reported in Table 8.

We take three points away from these results. First, we observe a statistically significant relationship between SRHS and mortality for both movers and stayers in all specifications except for women under 60 years old. However, it is important to mention that we do see a statistically significant relationship for the sample of all women. Second, the samples of movers can be quite small, and thus the 95% confidence intervals for the hazard ratios in these samples are wide. Accordingly, the confidence intervals on the SRHS hazard ratios always overlap for movers and stayers. Third, while the hazard ratios for movers and stayers do differ somewhat, the difference is not systematic. In other words, there is no evidence of a systematically weaker (or stronger) relationship between SRHS and mortality among movers than among stayers. Overall, we take this as evidence that there are probably not any systematic biases in our SRHS measures that affected our findings.

6.2 Non-Random Attrition

A potentially more important source of bias is non-random attrition, since two of the most common reasons for attrition are migration and death. Indeed, Fitzgerald, Gottschalk, and Moffitt (1998) examined the reasons for attrition from the PSID and showed, for the 40% of people for whom the cause of attrition was known, approximately two-thirds was due to mortality and one-third to a move that could not be followed. Unfortunately, because non-random attrition is one of the least understood areas in econometrics, there are few solutions proposed in the literature to address it. Some of the solutions that have been proposed, such as Inverse Probability Weighting (discussed in Moffitt, Fitzgerald, and Gottschalk 1999), are only valid under restrictive conditions. Our approach, therefore, is to provide a discussion of the bias that will result from attrition and argue that many of our estimates constitute a lower bound of the true effect of health on migration.

We consider a linear version of equation (1):

$$m_{i,t} = g_{i,t-1}\gamma + b_{i,t-1}\beta + \mathbf{X}'_{i,t}\boldsymbol{\theta} + \varepsilon_{i,t}. \tag{3}$$

Working with this linear model greatly facilitates the exposition. Owing to non-random attrition, our data do not constitute an *i.i.d.* sample from this population. Instead, we only have observations for individuals who "survive" in the sample across survey years. We let $s_{i,t}$ denote the survival indicator. If the individual has survived from time *0* to time t, then the indicator equals unity; otherwise, it is zero. For the sake of simplicity, we assume that attrition is an absorbing state. As a consequence of panel attrition, the econometrician does not observe the vector $(m_{i,t}, g_{i,t-1}, b_{i,t-1}, \mathbf{X}'_{i,t})$. Instead, she observes $(m^*_{i,t}, g^*_{i,t-1}, b^*_{i,t-1}, \mathbf{X}^{*\prime}_{i,t})$, where we have adopted the notation that $z^*_{i,t} = s_{i,t}z_{i,t}$, and she attempts to consistently estimate the parameters in

$$m^*_{i,t} = g^*_{i,t-1}\gamma + b^*_{i,t-1}\beta + \mathbf{X}^{*\prime}_{i,t}\boldsymbol{\theta} + \varepsilon^*_{i,t}. \tag{4}$$

However, OLS will result in inconsistent parameter estimates if the attrition is systematic since this implies that the residual in equation (4) will be correlated with the right-hand side regressors, particularly the health variables. This, in turn, implies that the orthogonality conditions that are required for identification will be violated and that OLS will not recover γ or β.

The direction of the bias of the OLS estimates of these parameters will depend on the signs of two expectations: $E[s_{i,t}g_{i,t-1}\varepsilon_{i,t}]$ and $E[s_{i,t}b_{i,t-1}\varepsilon_{i,t}]$. We argue that the former is negative and the latter is positive. The reason is that survival in the panel is positively (negatively) correlated with being healthy (unhealthy) and negatively correlated with migration. Intuitively, what this means is that as people exit the sample it will change in systematic way with each successive survey year consisting of people who are less likely to move and less likely to die (and thus also healthier). Accordingly, we expect that the OLS estimate of γ to be biased downwards and the estimate of β to be biased upwards.

In Table 9 we give the reader a sense of how attrition rates in our PSID sample vary by health status. Each cell of the table reports the rate of attrition on a PSID wave across survey years. We break the calculations down by gender, health status, and age. Two points should be taken away from the table. First, attrition rates are substantially higher in the bottom two SRHS categories than they are when calculated for the entire health distribution. Second, attrition rates are substantially higher among people who are 60 years of age or older.

These points have several implications for our findings. First, they suggest that the estimates of the effects of health on migration from Table 2 are conservative. In other words, the true impact of health on migration is probably greater than what we have estimated. Second, they suggest that if the attrition bias is great enough, the OLS estimate of β may actually be positive even if the true parameter is negative. Consequently, non-random attrition may also be responsible for the observed non-monotonicity for older men from Table 3, especially since we would expect the biases from attrition to be higher among older people for whom mortality-induced attrition is higher.

7 Conclusion

In this paper, we used the PSID to investigate how health influences domestic migration. Our results indicate that among men younger than age 60, poor health is associated with less geographic mobility. Among men older than age 60, the results suggest that there is greater mobility at both the top and bottom of the health distribution. For women, there is no evidence that their own health affects mobility, but there is evidence that spousal health affects women's mobility. Owing to the bias induced by non-random attrition, the estimated effects of ill health on migration constitute a lower bound.

One important avenue for future research is to better understand the relationship between positive migrant selection on both health and labor market outcomes. One would expect these two types of selection to be intimately related, as there is a large literature showing that poor health has large causal effects on the labor supply (Smith 1999; Rust and Phelan 1997) and on educational attainment (Miguel and Kremer 2004; Bobonis, Miguel, and Puri-Sharma 2006; Case, Fertig, and Paxson 2005). Accordingly, an interesting (and ambitious) topic for future research would be to investigate how much of the observed positive selection on labor market outcomes is the result of the impact of health on labor supply and educational attainment.

References

[1] Abraído-Lanza, A. F., B. P. Dohenwend, D. S. Ng-Mak, and J. B. Turner. 1999. "The Latino mortality Paradox: a test of the 'Salmon Bias' and healthy migrant hypotheses." American Journal of Public Health 89: 1543-1548.

[2] Amemiya, T. 1985. Advanced Econometrics. Cambridge, MA: Harvard University Press.

[3] Bobonis, G., Miguel, E., and Puri-Charma, C. 2006, "Anemia and school participation," Journal of Human Resources 41: 692-721.

[4] Borjas, G. J. 1987. "Self-selection and the earnings of immigrants," American Economic Review 77: 531-553.

[5] Borjas, G. J., S. G. Bronnars, and S. J. Trejo. 1992. "Assimilation and the earnings of young internal migrants," Review of Economics and Statistics 74: 170-175.

[6] Boyle, P. J., P. Norman, and P. H. Ress. 2002. "Does migration exaggerate the relationship between deprivation and limiting long-term illness?" Social Science and Medicine 55: 21-31.

[7] Brimblecombe, N., D. Dorling, and M. Shaw. 1999. "Mortality and migration in Britain, first results from the British Household Survey," Social Science and Medicine 49: 981-988.

[8] Case, A., A. Fertig, and C. Paxson. 2005. "The lasting impact of childhood health and circumstance," Journal of Health Economics 24: 365-389.

[9] Chiquiar, D., and G. H. Hanson. 2005. "International migration, self-selection, and the distribution of wages: evidence from Mexico and the United States," Journal of Political Economy 113: 239-281.

[10] Chiswick, B. R. 1978. "The effects of Americanization on the earnings of foreign-born men," Journal of Political Economy 86: 897-921.

[11] Cox, M., P. J. Boyle, P. Davey, and A. Morris. 2007. "Does health-selective migration following diagnosis strengthen the relationship between Type 2 Diabetes and deprivation?" Social Science and Medicine 65: 32-42.

[12] Farr, W. 1864. Supplement to the 25th Annual Report of the Registrar General. London: HMSO.

[13] Findley, S. E. 1988. "The directionality and age selectivity of the health-migration relation: evidence from sequences of disability and mobility in the United States," International Migration Review 22: 4-29.

[14] Fitzgerald, J., P. Gottschalk, and R. Moffitt, R. 1998. "An analysis of sample attrition in panel data: the Michigan Panel Study of Income Dynamics," Journal of Human Resources 33: 251-299.

[15] Gabriel, P. E., and S. Schmitz. 1995. "Favorable self-selection and the internal migration of young white males in the United States," Journal of Human Resources 30: 460-471.

[16] Halliday, T. 2007. "Business cycles, migration and health," Social Science and Medicine 64: 1420-1424.

[17] Idler, E. L., and S. V. Kasl. 1995. "Self-ratings of health: do they also predict change in functional ability?" Journal of Gerontology 50, S344-S353.

[18] Jasso, G., and M. R. Rosenzweig. 1990. "Self-selection and the earnings of immigrants: comment," American Economic Review 80, 298-304.

[19] Jasso, G., D. S. Massey, M. R. Rosenzweig, and J. P. Smith. 2004. "Immigrant health – selectivity and acculturation," unpublished manuscript, Institute for Fiscal Studies.

[20] Kaplan, G.A., and Camacho, T. 1983. "Perceived health and mortality: A nine-year follow-up of the Human Population Laboratory cohort," American Journal of Epidemiology 177: 292-304.

[21] Kasl, S. V., and L. Berkman. 1983. "Health consequences of the experience of migration," Annual Review of Public Health 1983(4): 69-90.

[22] Markides, K. S., and J. Coreil. 1986. "The health of Hispanics in the southwestern United States: an epidemiological paradox," Public Health Reports 101: 253-265.

[23] Marmot, M. G., A. G. Adelstein, and L. Bulusu. 1984. "Lessons from the study of immigrant mortality," Lancet, June 30: 1455-1457.

[24] Marmot, M. G., and L. Syme. 1976. "Acculturation and coronary heart disease in Japanese-Americans," American Journal of Epidemiology, 104: 225-247.

[25] Miguel, E., and M. Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," Econometrica 72: 159-218.

[26] Moffitt, R., J. Fitzgerald, and P. Gottschalk. 1999. "Sample attrition in panel data: the role of selection on observables," Annale d'Economie et de Statistique 55-56: 129-152.

[27] Mossey, J. M., and E. Shapiro. 1982. "Self-rated health: a predictor of mortality among the elderly," American Journal of Public Health 72: 800-808.

[28] Newbold, K. B. 2005. "Self-rated health within the Canadian immigrant population: risk and the healthy immigrant effect," Social Science and Medicine 60: 1359-1370.

[29] Norman, P., P. Boyle, and P. Rees. 2005. "Selective migration, health and deprivation: a longitudinal analysis," Social Science and Medicine 60: 2755-2771.

[30] Pablos-Mendéz, A. 1994. "Mortality among Hispanics," Journal of the American Medical Association 271: 1237.

[31] Palloni, A., and E. Arias. 2004. "Paradox lost: explaining the Hispanic adult mortality advantage," Demography 41: 385-415.

[32] Scribner, R. 1994. "Mortality among Hispanics," Journal of the American Medical Association 271: 1238-1239.

[33] ——. 1996. "Paradox as paradigm – the health outcomes of Mexican Americans" (editorial), Journal of the American Medical Association 86: 303-305.

[34] Rust, J., and C. Phelan. 1997. "How Social Security and Medicare affect retirement behavior in a world of incomplete markets," Econometrica 65, 781-831.

[35] Smith, J. 1999. "Healthy bodies and thick wallets: the dual relation between health and economic status," Journal of Economic Perspectives 13, 145-166.

[36] ——. 2003. "SES and health over the life-course," unpublished manuscript, RAND.

[37] Swallen, K. C. 2002. "Mortality in the U.S.: comparing race/ethnicity and nativity," unpublished manuscript, University of Wisconsin.

[38] Welton, T. A. 1872. "On the Effect of migrations in distributing local rates of mortality as exemplified in the statistics of London and the surrounding country for the years 1851-1860," Journal of the Institute of Actuaries 16: 153.

Table 1: Variable Definitions and Statistics

Variable	Definition	Mean (Standard Deviation)
Moved	Indicator of whether or not individual has moved across two survey years	0.03 (0.17)
Self Reported Health Status (SRHS)	1=excellent; 2 = very good; 3 = good 2=fair; 1 = poor	2.49 (1.15)
Healthy	SRHS = 1 or 2	0.53 (0.50)
Unhealthy	SRHS = 4 or 5	0.19 (0.39)
Age	Individual's Age	42.49 (15.91)
Labor Income	Individual's Labor Income in 1982 dollars	11928.21 (15843.05)
Sex	=1 if female	0.54 (0.50)
No College Experience	= 1 if the individual never attended college	0.65 (0.48)
College Degree	= 1 if the individual has a college degree	0.19 (0.39)
White	= 1 if the individual is white	0.66 (0.48)
Black	= 1 if the individual is black	0.29 (0.46)
Married	= 1 if the individual is married	0.71 (0.45)
Disability	=1 if the individual is disabled	0.19 (0.39)

Table 2: Lagged Period Health - Under Age 60

	(1)	(2)	(3)	(4)	(5)	(6)
	Men			Women		
Unhealthy at t-1[1]	-0.014**	-0.015**	-0.012**	-0.006*	-0.006*	-0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Healthy at t-1[2]	0.008**	0.008**	0.004	0.007**	0.008**	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Income at t-1	-	-0.001	0.001	-	0.003*	0.004**
		(0.002)	(0.001)		(0.001)	(0.001)
Income at t-1 ˆ2	-	0.00008	-0.0002	-	-0.0003*	-0.0005**
		(0.0002)	(0.0001)		(0.0001)	(0.0001)
No College Experience	-	-	-0.012**	-	-	-0.014**
			(0.003)			(0.003)
College Degree	-	-	0.016**	-	-	0.007*
			(0.004)			(0.003)
White	-	-	0.009	-	-	0.004
			(0.005)			(0.004)
Black	-	-	0.002	-	-	-0.005
			(0.006)			(0.004)
Married	-	-	-0.019**	-	-	-0.007**
			(0.003)			(0.002)
State Dummies	No	No	Yes	No	No	Yes
F-test[3]	40.69	41.59	23.18	25.51	27.95	6.39
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.041]
R ˆ 2	0.0361	0.0362	0.0863	0.0352	0.0358	0.0731
NT[4]	39679	39677	39363	45291	45289	44906

[+]This table reports marginal effects of Probit models where the dependent variable is moved. Standard errors of the marginal effect are in parentheses. Standard errors allow for clustering within individuals. All regressions include a complete set of age dummies.

[*] Denotes 95% significance.

[**]Denotes 99% significance.

[1]Refers to SRHS equal to 4 or 5.

[2]Refers to SRHS equals to 1 or 2.

[3]$F-$ test of the null that the health variables are zero. p-values are in brackets.

[4] NT refers to individual/time observations.

Table 3: Lagged Period Health - Age 60 and Over

	(1)	(2)	(3)	(4)	(5)	(6)
	Men			Women		
Unhealthy at t-1[1]	0.014**	0.015**	0.016**	-0.002	-0.002	-0.000
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.002)
Healthy at t-1[2]	0.024**	0.022**	0.020**	0.001	0.000	-0.001
	(0.007)	(0.007)	(0.006)	(0.003)	(0.003)	(0.002)
Income at t-1	-	-0.0008	-0.0009	-	-0.0003	-0.0014
		(0.002)	(0.001)		(0.002)	(0.002)
Income at t-1 ˆ2	-	0.0001	0.0001	-	-0.00004	0.0002
		(0.0002)	(0.0002)		(0.0003)	(0.0002)
No College Experience	-	-	-0.003	-	-	0.002
			(0.006)			(0.003)
College Degree	-	-	0.005	-	-	0.010
			(0.006)			(0.007)
White	-	-	0.004	-	-	0.001
			(0.005)			(0.005)
Black	-	-	-0.004	-	-	-0.006
			(0.005)			(0.004)
Married	-	-	-0.005	-	-	-0.007**
			(0.004)			(0.002)
State Dummies	No	No	Yes	No	No	Yes
F-test[3]	20.37	19.61	24.19	0.59	0.55	0.23
	[0.000]	[0.000]	[0.000]	[0.7436]	[0.7593]	[0.8928]
R ˆ 2	0.0400	0.0408	0.1064	0.0255	0.0255	0.1069
NT[4]	7591	7591	6299	10895	10895	10455

[+]This table reports marginal effects of Probit models where the dependent variable is moved. Standard errors of the marginal effect are in parentheses. Standard errors allow for clustering within individuals. All regressions include a complete set of age dummies.

[*] Denotes 95% significance.

[**]Denotes 99% significance.

[1]Refers to SRHS equal to 4 or 5.

[2]Refers to SRHS equals to 1 or 2.

[3]$F-$ test of the null that the health variables are zero. p-values are in brackets.

[4] NT refers to individual/time observations.

Table 4: Disability Results

	(1)	(2)	(3)	(4)
	Age < 60		Age >= 60	
	Men	Women	Men	Women
	Dependent Var = Disability			
Unhealthy at t[1]	0.223**	0.270**	0.412**	0.416**
	(0.011)	(0.010)	(0.017)	(0.015)
Healthy at t[2]	−0.096**	−0.100**	−0.228**	−0.230**
	(0.005)	(0.005)	(0.019)	(0.015)
$R\hat{}2$	0.2767	0.2497	0.2726	0.2633
NT[3]	39292	44802	8125	11046
	Dependent Var = Moved			
Disability at $t-1$	−0.003	0.003	0.003	0.002
	(0.003)	(0.003)	(0.003)	(0.002)
$R\hat{}2$	0.0841	0.0729	0.0832	0.1076
NT[3]	39345	44803	6296	10446

[+]This table reports marginal effects of Probit models. Standard errors of the marginal effects are in parentheses. Standard errors allow for clustering within individuals. All regressions include the same set of controls as columns 3 and 6 of Table 2.

**Denotes 99% Significance.

[1]Refers to SRHS equal to 4 or 5.

[2]Refers to SRHS equal to 1 or 2.

[3]NT refers to individual/time observations.

Table 5: Own and Spousal Health - Married Couples - Under Age 60

	(1)	(2)	(3)	(4)
	Men		Women	
Unhealthy at $t-1$ - Own[1]	-0.012^{**} (0.002)	-0.010^{**} (0.003)	-0.003 (0.003)	-0.001 (0.003)
Healthy at $t-1$ - Own[2]	-	0.003 (0.002)	-	0.001 (0.002)
Unhealthy at $t-1$ - Spousal[1]	-0.001 0.003	-0.000 (0.002)	-0.012^{**} (0.002)	-0.009^{**} (0.003)
Healthy at $t-1$ - Spousal[2]	-	-0.000 (0.002)	-	0.006^{**} (0.002)
F-Test on Own Health[3]	-	14.93 [0.001]	-	0.75 [0.686]
F-Test on Spousal Health[3]	-	0.03 [0.985]	-	22.60 [0.000]
$R\hat{}2$	0.0897	0.0900	0.0747	0.0764
NT[4]	30173	30173	30238	30238

[+]This table reports marginal effects of Probit models where the dependent variable is moved. Standard errors of the marginal effect are in parentheses. Standard errors allow for clustering within individuals. All regressions include the same set of controls as columns 3 and 6 of Table 2.

[**]Denotes 99% significance.

[1]Refers to SRHS equal to 4 or 5.

[2]Refers to SRHS equal to 1 or 2.

[3]$F-$ test of the null that the health variables are zero. p-values are in brackets.

[4]NT refers to individual/time observations.

Table 6: Own and Spousal Health - Married Couples - Age 60 and Over

	(1)	(2)	(3)	(4)
	Men		Women	
Unhealthy at $t-1$ - Own[1]	0.002 (0.003)	0.013** (0.005)	-0.003 (0.003)	-0.004 (0.003)
Healthy at $t-1$ - Own[2]	-	0.020** (0.006)	-	-0.003 (0.003)
Unhealthy at $t-1$ - Spousal[1]	-0.003 (0.003)	-0.002 (0.003)	-0.001 (-0.38)	0.005 (0.004)
Healthy at $t-1$ - Spousal[2]	-	0.001 (0.003)	-	0.015** (0.006)
F-Test on Own Health[3]	-	18.84 [0.000]	-	1.92 [0.384]
F-Test on Spousal Health[3]	-	0.94 [0.625]	-	11.15 [0.004]
Rˆ2	0.1089	0.1360	0.1320	0.1555
NT[4]	4446	4446	3090	3090

+This table reports marginal effects of Probit models where the dependent variable is moved. Standard errors of the marginal effect are in parentheses. Standard errors allow for clustering within individuals. All regressions include the same set of controls as columns 3 and 6 of Table 2.

**Denotes 99% significance.

[1]Refers to SRHS equal to 4 or 5.

[2]Refers to SRHS equal to 1 or 2.

[3]$F-$ test of the null that the health variables are zero. p-values are in brackets.

[4]NT refers to individual/time observations.

Table 7: Additional Results

	(1)	(2)	(3)	(4)	(5)	(6)
	Men	Women	Men	Women	Married Women	Single Women
	Age >= 60			Age < 60		
Unhealthy at $t-1$[1]	0.040*	0.002	-0.011**	-0.003	-0.004	-0.003
	(0.015)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Healthy at $t-1$[2]	0.019*	-0.001	0.004	0.002	0.004**	-0.001
	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Unhealthy at $t-1$ * Indicator for Age >= 40 & < 50	-	-	-0.011	-0.004	-	
			(0.007)	(0.006)		
Healthy at $t-1$ * Indicator for Age >= 40 & < 50	-	-	-0.001	-0.002	-	
			(0.005)	(0.004)		
Unhealthy at $t-1$ * Indicator for Age >= 50 & < 60			0.016	0.013	-	
			(0.015)	(0.011)		
Healthy at $t-1$ * Indicator for Age >= 50 & < 60			0.007	0.027**	-	
			(0.010)	(0.013)		
Unhealthy at $t-1$*Married	-0.012*	-0.005	-	-	-	
	(0.005)	(0.003)				
Rˆ2	0.1137	0.1086	0.0867	0.0740	0.0740	0.0987
NT[3]	6299	10455	39363	44906	31133	13715

[+]This table reports marginal effects of Probit models where the dependent variable is moved. Standard errors of the marginal effect are in parentheses. Standard errors allow for clustering within individuals. All regressions include the same set of controls as columns 3 and 6 of Table 2.

* Denotes 95% significance.

**Denotes 99% significance.

[1]Refers to SRHS equal to 4 or 5.

[2]Refers to SRHS equal to 1 or 2.

[3]NT refers to individual/time observations.

Table 8: SRHS and Mortality by Migration Status

< 60 Years

	Men		Women	
	Stayers	Movers	Stayers	Movers
Age	1.079**	1.103**	1.087**	1.102**
	[1.067, 1.092]	[1.071, 1.136]	[1.073, 1.100]	[1.062, 1.143]
Unhealthy[1]	2.067**	1.314	1.755**	2.307
	[1.542, 2.772]	[0.498, 3.470]	[1.306, 2.359]	[0.761, 6.993]
Healthy[2]	0.632**	0.361**	0.706**	0.709
	[0.474, 0.843]	[0.180, 0.722]	[0.514, 0.972]	[0.232, 2.169]
N	3329	601	4010	649

All Ages

	Men		Women	
	Stayers	Movers	Stayers	Movers
Age	1.084**	1.092**	1.090**	1.100**
	[1.076, 1.092]	[1.073, 1.111]	[1.084, 1.095]	[1.083, 1.118]
Unhealthy[1]	1.398**	1.260	1.685**	0.924
	[1.159, 1.687]	[0.687, 2.310]	[1.419, 2.002]	[0.517, 1.652]
Healthy[2]	0.611**	0.411**	0.839	0.427**
	[0.507, 0.735]	[0.232, 0.728]	[0.691, 1.020]	[0.234, 0.777]
N	3952	647	4904	731

+This table reports estimates of Cox proportional hazard models. The dependent variable is mortality. Each cell reports the hazard ratio and its 95% confidence interval. All estimates adjust for clustering within individuals. If the hazard ratio is greater (less) than unity then the variable has a positive (negative) effect on mortality.

**Denotes 99% significance.

[1]Refers to SRHS equal to 4 or 5.

[2]Refers to SRHS equal to 1 or 2.

Table 9: Panel Attrition Rates by Age, Gender, and Health Status

< 60 Years

Men		Women	
All	SRHS = 4 or 5	All	SRHS = 4 or 5
4.32%	5.87%	4.19%	6.03%

>= 60 Years

Men		Women	
All	SRHS = 4 or 5	All	SRHS = 4 or 5
5.17%	7.35%	5.10%	7.68%