



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

**The Effect of Relative Income in the Dynamics of Migration:
Evidence from the VHLSS Panel Data**

M.S Plan-B Paper

Department of Applied Economics

Advisor: Paul Glewwe

Defense Date: August 31 2012

WON FY LEE

Table of Contents

I.	Introduction	1
II.	Literature Review	1
	II.A Literature Review on the Concept of Relative Deprivation	1
	II.B Literature Review of Empirical Studies of Relative Deprivation	3
III.	The Theory of Relative Deprivation	5
	III.A Quantifying Relative Deprivation	6
	III.B The Individual-Level Yitzhaki Index	9
IV.	Data	11
	IV.A Constructing the 2004 – 2006 VHLSS Panel Data	13
	IV.B Reference Group	15
	IV.C Assumptions about the Reference Group	16
	IV.D Identification of Migrants	18
	IV.E Calculation of Household Income	23
V.	The Econometric Model	25
VI.	Results	27
	VI.A Reference Group (1) District (2) District and Income	27
	VI.A.1 Results by different age group	31
	VI.B Reference Group (3) District and Age	32
	VI.C Reference Group (4) District and Age and Education	33
	VI.D Reference Group (5) District and Age and Income	34
VII.	Conclusion	35
VIII.	References	38
	APPENDIX A: Figures	41
	APPENDIX B: Tables	48

List of Figures

Figure 1	Percent of Panel Mismatch by Province	41
Figure 2	Origin of Out-Migrants and Cities with Population Greater than 200,000	42
Figure 3	Origin of Long- and Short-Term Out-Migrants.....	43
Figure 4	Origin of Out-Migrants and Poverty Rate	44
Figure 5	Per Capita Income Comparison Between GSO Calculation and Author's Calculation	45
Figure 6	Relationship Between Income and the Frequency of Migrants.....	46
Figure 7	Income distribution of Entire Sample	46
Figure 8	Duration of Migration	47

List of Tables

Table 1	Reason for Moving	48
Table 2	Summary Statistics of Migrants and Non-Migrants	48
Table 3	Occupations of Long-term Migrants, Short-term Migrants and Non-Migrants	49
Table 4	Summary Table of Household Income in 2004	49
Table 5	Reference Group 1 and 2 Regression Results: Total Migrants	50
Table 6	Reference Group 1 and 2 Regression Results: Long-term Out-Migrants	51
Table 7	Reference Group 1 and 2 Regression Results: Short-term Out-Migrants	52
Table 8	Reference Group 1 and 2 Regression Results by age groups: Long-term Out-Migrants	53
Table 9	Reference Group 1 and 2 Regression Results by age groups: Short-term Out-Migrants	54
Table 10	Reference Group 3 Regression Results: Long-term Out-Migrants	54
Table 11	Reference Group 4 Regression Results: Long-term Out-Migrants.....	55
Table 12	Reference Group 5 Regression Results: Long-term Out-Migrants	55

I .Introduction

This paper examines, within the Vietnamese context, the role of relative deprivation, as proposed by Stark (1984) and Stark and Taylor (1989, 1991), in the dynamics of internal migration. The conventional wisdom in the economics of migration literature, such as the model of Harris and Todaro (1970), holds that as the rural-urban income gap widens, the influx of migrants from rural to urban areas increases until equilibrium is reached. The absolute income gap is of the utmost importance in the decision-making process concerning migration in that model. On the other hand, Stark (1984) hypothesizes that household members undertake migration not necessarily to increase the household's absolute income but rather to improve the household's position (in terms of relative deprivation) with respect to a specific reference group.

II. Literature

II.A. Literature Review on the Concept of Relative Deprivation

The study of relative income has long been a focus of many social scientists; hypotheses on how it functions and shapes the lives of people have been proposed by numerous prominent scholars in various fields.

Jean-Jacques Rousseau (1754), in his treatise on the origin of inequality, saw the genesis of inequality as due to the establishment of agricultural society. New aspects began to arise as people no longer roved in the woods and took on a more settled manner of life. Life in neighboring huts brought about a consciousness of differences as well as preferences and comparisons. He writes: "Each one began to consider the rest, and to wish to be considered in

turn; and thus a value came to be attached to public esteem. Whoever sang or danced best, whoever was the handsomest, the strongest, the most dexterous, or the most eloquent, came to be of most consideration; and this was the first step towards inequality.”

In Rousseau’s mind the consciousness of the gaze of others is the most important drive behind human activities in the modern world. Even though Rousseau’s perspective may put too much emphasis on the dismal side of what the agricultural society has brought us, the problem of inequality he posed and its influence on people’s lives has been given extensive attention in various academic fields. A hundred years later, Marx noted: “Our desire and pleasures spring from society: we measure them, therefore, by society and not by the objects which serve for their satisfaction. Because they are of a social nature, they are of a relative nature” (Tischler 2006, 426). Veblen (1899) was the first economist to argue that the process of individual decision-making is significantly affected by human interaction. Veblen’s concept of conspicuous consumption, which argues that consumption of luxury goods takes place so that people can show off, seems to be closely related to Rousseau’s description of the origin of inequality.

Duesenberry (1949), in his book about consumption and saving behavior, argues that the self-contained independence of individual preference that marginal utility theory postulates lacks a firm basis for its justification, and so he tries to come up with a theory that can better account for the root of consumption desire. In what he calls the “demonstration effect”, people’s desire for consumption rises with increasing frequency as they come into contact with superior goods, and that contact with superior goods increases primarily as the consumption expenditure of “others or associates” increases. The implication is

important because the consumption behavior is now not only a function of income and prices of goods and individual preference, but also a function of the consumption of others. He distinguishes his concept of demonstration effect from Veblen's concept of conspicuous consumption, in that his concept applies to general human consumption behavior.

Sen (2006), in his conceptual discussion of income, capability and poverty, also claims the importance of including relative concepts in the discussion of poverty, because the capability to do something is not solely dependent on the absolute income but, rather, determined in the social interaction where people find themselves positioned. These conceptual grounds of the relative deprivation have led many researchers of different backgrounds to an empirical examination of the extent to which hypothesized relative property of human nature its role in our society.

II.B. Literature Review of Empirical Studies of Relative Deprivation

Much attention has been given to the relative position of people and its relationship to their subjective happiness. Easterlin (1974) conducted a comprehensive empirical study that examined the relationship between income and happiness utilizing data from 19 countries. His study distinguishes within-country and cross-country relationships. For the within-country relationship, he found, using 29 separate surveys from 19 countries, an unequivocally positive relationship between income and happiness, which corresponds to the conventional economic theory that a person's utility increases as his or her level of income increases, even though happiness could be a function of many other aspects of life. For the cross-country relationship, using average personal happiness ratings for 14 countries, Easterlin found no clear relationship

between the average GDP per capita of a country and its average happiness rating, which one might expect based on the within-country relationship he found. Moreover, he found no clear relationship in US time-series data from 1946 to 1970. While there were ups and downs in the happiness rating over the periods, the proportion of people who answered that they are very happy in 1970 was about the same as in 1947, although GDP per capital almost doubled between those two years. This study is important because it clearly shows the relative nature of subjective happiness as well as the spatiotemporal nature of the frame of reference.

Ferrer-i-Carbonell (2005), using German panel data, showed that individuals are happier the larger their income is in comparison with the income of the reference group. His empirical result suggests that the income of the reference group is about as important as the one's own income for one's happiness. In a recent comparative study using the General Social Survey on the relationship between happiness and relative income in China, Japan, and Korea, Oshio et al. (2010) confirmed a positive correlation between the two even after controlling for the family income.

Relative deprivation has also been studied as an underlying mechanism of public health. Eibner and Evans's (2005) study of the effect of relative deprivation on the probability of death due to health-compromising behaviors found a positive and statistically significant influence of relative deprivation on the probability of death after controlling for reference group effects and individual income. Naoki et al.'s (2008) empirical study using Japanese data on 47,114 individuals aged between 25-64 found a similar correlation between relative deprivation and public health.

Few studies in the migration literature have also found significant effects of relative deprivation on migration behavior. Bhandari (2004) used the size of land holding as a proxy for relative deprivation in Nepal and found a significant effect of relative deprivation on migration behavior. Quinn (2006) also found that relative deprivation was a significant motivating factor for internal migration in Mexico.

The literature on the issue of relative deprivation seems to support an inseparable characteristic of human behavior: concern for one's relative position. The primary question being asked in this study is the extent to which relative deprivation, measure by proxy variable using income, affects the decision to migrate in Vietnam.

III. The Theory of Relative Deprivation

The most well-known definition of relative deprivation comes from Runciman (1966), who claims that an individual is relatively deprived if:

- (1) He does not have X, (2) he sees some other person or persons, which may include himself at some previous or expected time, as having X (whether or not this is in fact the case), (3) he wants X, and (4) he sees it as feasible that he should have X.

The economist's interpretation of the concept of relative deprivation in its current form used in this study was first introduced by Yitzhaki (1979, 1982) and Stark (1984), who made the conceptual distinction between conventional utility and a new relative deprivation concept of utility. Starting from the four characteristics of relative deprivation that Runciman, a sociologist, suggested, Stark (1984) contends that the deprivation concept and the utility concept are two sides of the same coin. The deprivation concept comes from (1) and (3). X can be represented

by a bundle of commodities that one has and $U(X)$ is the utility one obtains from having a bundle of commodities. Deprivation is simply $-U(X)$, the disutility of having no more than X . That is, as utility increases, $-U(X)$ decreases, so deprivation decreases. The fact that maximizing utility also minimizes deprivation to the same extent, convinces Stark and Taylor (1989) to see the two concepts as inversely inter-changeable. However, due to the relative characteristic of relative deprivation approach that comes from (2) and (4), differences arise between *relative* deprivation and utility.

The difference derived from the relative characteristic of relative deprivation approach is not trivial, whereas the marginal utility of income is a function of absolute income alone in the income approach, the relative deprivation approach is a function of absolute income as well as income of others. This difference necessitates the existence of a reference group in the relative deprivation approach. The feeling of deprivation that arises from not having X is an increasing function of the number of individuals in the reference group who own it.

III.A. Quantifying Relative Deprivation

Yitzhaki (1979), Stark and Yitzhaki (1988), and Stark and Taylor (1989, 1991) proposed a practical method to measure relative deprivation. They assume that x can be measured by income and start with continuous income distribution. Let $F(x)$ denote the cumulative distribution of income in a reference group, say, a village. Then $1 - F(x)$ is the relative frequency of individuals with income higher than x . By hypothesis, the feeling of deprivation is an increasing function of the percentage of individuals who have income larger than x , that is, $1 - F(x)$. Let $g[1-F(x)]$ be the deprivation from having only x , in other words, not having more than x ,

where $g(0)=0$ and $g' >0$. A household with income x is deprived of all units of income above x . Thus, we can represent the relative deprivation of household i , whose income is y_i , as

$$RD^i = \int_{y_i}^{y^h} g[1 - F(x)] dx \quad (1)$$

where y^h indicates the highest income in a village. The integration part of the equation (1) is meant to account for the sensitivity of the measure in capturing the income level variation of those who are richer than the individual i , without the integration, the measure is blind to the increase in the income of richer individuals, in other words, $F(x)$ only accounts for the ordering of the income distribution, and is insensitive to the size of the income of others.

Stark assumes that $g[1-F(X)] = 1-F(X)$ for simplicity, as will the following discussion. With algebraic manipulations, the equation (1) can be decomposed into the product of the mean excess income of individuals richer than individual i and the proportion of individuals in the village that are richer than the individual i ,

$$RD^i = [1 - F(y_i)]E(z - y_i|z > y_i) \quad (2)$$

where z is the income of individuals with income higher than individual i . Thus, $E(z - y_i|z > y_i)$ is the average gap between individual i 's income and the average income of all people with an income greater than person i . The important feature of this formulation is that, although the order of the income distribution remains the same when everyone's income whose income higher than person i increases by the same proportion, individual i 's feeling of relative deprivation increases.

Stark's proposed method of quantifying relative deprivation incorporates only income for the calculation. There could be many more factors at work in the formation of one's feeling of deprivation compared to others, such as total assets, ability to do something, a social network, having more children, and many others, but the difficulty of capturing all these factors in a survey often leads us to consider income as the only variable in the calculation.

The relative deprivation hypothesis proposed by Stark and Taylor (1991) is that, controlling for expected income gains from migration, household members would decide to migrate if $U(RD_1^i) > U(RD_0^i)$, where RD_1 is the relative deprivation associated with postmigration and RD_0 is the relative deprivation in the absence of migration. As mentioned in the previous section, the relative nature of relative deprivation theory requires a reference group, and reference group substitution can be an important issue in relative deprivation theory. This will be discussed in more detail in a later section of this paper.

In this study, two variants of the Yitzhaki Index, as a legitimate proxy for relative deprivation measurement, are used in an econometric analysis. The first Yitzhaki Index is the original index proposed by Yitzhaki (1979) based on a concept suggested by Runciman (1966). The second relative deprivation index is a modified version of the Yitzhaki Index, which uses log of income rather than absolute income for the sum of differences. This form of the Yitzhaki Index can be formalized similarly to original Yitzhaki Index and its theoretical background and its implication is discussed in more detail in the following section of the paper.

III.B. The Individual-Level Yitzhaki Index

Two versions of the Yitzhaki index will be used in this paper. The original Yitzhaki Index for person i is calculated as:

$$Yitzhaki\ Index_i = \frac{1}{N} \sum_j (y_i - y_j) \forall y_j > y_i \quad (3)$$

The modified Yitzhaki Index for person i is calculated as:

$$Modified\ Yitzhaki\ Index_i = \frac{1}{N} \sum_j (\ln(y_i) - \ln(y_j)) \forall y_j > y_i \quad (4)$$

where individual i 's income is compared to that of every other individual, j , whose income is higher than i and belongs to the same reference group as i . The measure assumes that individual i feels deprived only compared to those whose income level is higher than i . The original measure suggested by Yitzhaki used absolute income for the sum of differences, whereas in the modified Yitzhaki Index, the log of income is used, following the suggestions made by Eibner and Evans (2005), who followed the suggestions made by Podder (1996) and Hey and Lambert (1980). In rationalizing the use of log of income rather than nominal income, Podder (1996) started from the concavity of the utility function and derived the following properties:

- 1) The function RD_i must be an increasing function of y_j and a decreasing function of y_i , where person i 's magnitude of relative deprivation with respect to the person j is $U(y_j) - U(y_i)$. In other words, the first derivative of RD_i with respect to y_j is positive whereas with respect to y_i is negative.

- 2) The function RD_i must be concave with respect to y_j . In other words, the second derivative of RD with respect to y_j is negative. One implication of this property is that if a unit of income is transferred from a rich man, j_2 to a poorer man, j_1 , both having income higher than the person i , the magnitude of RD for person i would increase.

The distribution of income of people whose incomes are higher than individual i affects the magnitude of the relative deprivation (RD) as captured both in the modified and original Yitzhaki index. Yet, the original Yitzhaki index, which utilizes absolute income, is insensitive to the distribution of income among those whose income is higher than the income of individual i . Unlike the original Yitzhaki index, the modified Yitzhaki index utilizes the log of income to account for the tendency of people to be concerned more about others with the similar range of income than people with income far different.

Podder (1996) points out the important fact that the transfer of income from a rich person to a poor person may result in an increase in aggregated RD in a society, which may be inconsistent with the Pigou-Dalton condition. The Pigou-Dalton condition, one of the central characteristics of conventional inequality measures, states that an income transfer from a poorer to a richer person in a society increases inequality in that society. This inconsistency with the Pigou-Dalton condition in relative deprivation is related to the non-monotonic relationship between the relative deprivation and conventional inequality measure as mentioned in an earlier section of this paper.

Podder gives an example of this by supposing six income units in ascending order:

$$x_1 < x_2 < x_3 < x_4 < x_5 < x_6$$

If income is transferred from x_3 to x_5 , this would increase the RD for x_3 and x_4 but decrease the RD for x_1 , x_2 , and x_5 , and thus, the total amount of RD in the society depends on the net effect. The total sum of individual relative deprivation, in this case, could decline as a result of this transfer from x_3 to x_5 , which increases inequality via the Pigou-Dalton condition of conventional inequality measure. In the extreme case, the aggregate relative deprivation of a society could be at its minimum where there are only two groups of people with large income differences assuming that reference groups include only people in one's own group, where the conventional inequality measure such as Gini Index would diagnose such a society as highly unequal.

The increased feeling of relative deprivation of individual i is captured in both original and modified Yitzhaki index, however the modified index give less weight to the people whose income is in the far end of the income distribution than people of similar income in the calculation of the index. However, it should be noted that there is no theoretical reason to prefer one Yitzhaki Index over the other. Both indices satisfy the properties of equation (2), and both decline with improved relative position in the income distribution within a reference group and rise with increased level of income of those who are richer than the referenced individual. Only empirical testing will provide guidance as to which version is more consistent with actual behavior assumed by the theory of relative deprivation.

IV. Data

The 2004 and 2006 Vietnam Household Living Standards Survey (VHLSS) datasets are used in this study. The VHLSS is a biennial core and rotating module household survey designed

to look at living standards at the national level. It uses a multi-topic household questionnaire designed to study multiple aspects of household welfare and behavior. The survey contains information on demographics, employment, income, health, and education as well as community characteristics variables. The VHLSS is considered to be the most representative nationwide survey that provides extensive information on community and household level characteristics in Vietnam (Nguyen et al, 2008).

For every new survey, half of the previous households are replaced by new households, and this structure forms the panel dataset over the two-year period. Expenditure data are usually a more accurate indicator of household welfare than income data in developing countries. However, in order to calculate household-level relative deprivation more precisely one needs the sample size for each district to be as large as possible, so the income-only data, with a bigger sample size for each district, are used in the study. Using the 45,900 sample households leads to an average of 73 households per district in the income-only survey, rather than an average of 15 households per district when one uses income and expenditure survey. The comparative advantage of allowing for a more precise household-level relative deprivation index outweighs the advantage of having expenditure data.

In 2004, the nationally representative VHLSS was carried out in all 64 provinces of Vietnam, and cluster sampling was conducted in 630 of its 670 districts and 3,061 of its 10,511 communes.

IV.A. Constructing the 2004 - 2006 VHLSS Panel Data

One advantage of using panel data is that one can study the characteristics of the sending community before migration takes place. Panel datasets enable one to study the effect of sending community characteristics on migration decisions in a manner that is less vulnerable to the potential endogeneity problems that are common in cross-sectional studies (Nguyen et al, 2008). In this section, the steps taken to set up a panel data used in this study are described. The M1_2_3.dta dataset, which contains demographic information of people who participated in the 2004 survey, is comprised of 202,716 people in 45,944 households. Each individual is assigned a unique identification number. This identification number enables the matching of various datasets within the current survey year and across survey years. Even though identification numbers in the VHLSS 2006 are different from those in the VHLSS 2004, the samples that make up the panel in 2006 contain the identification numbers used in both surveys.

When the 2004 income variable was merged with the demographic variables, households with negative incomes were dropped from the dataset. In this negative income category were 352 households with 1,396 people, which is only about 0.7 % of the sample. The information on migration status can be found in the 2006 muc1b.dta dataset, which contains the 21,844 households and 96,375 people who comprise the 2004 – 2006 panel data set. Of the households that participated in the VHLSS 2004, 48% also participated in the VHLSS 2006. The final dataset used in the regression analysis was created by going through further data cleaning

based on mismatches in demographic information found in the M1_2_3.dta dataset of 2004 and muc1b.dta dataset of 2006.

There were 9,926 people in 2518 households whose personal identification numbers did not match between the two survey years. These were the people who reported in the 2006 survey that they participated in 2004 survey, yet, both their personal identification numbers and household identification numbers for that year were not found in the 2004 survey. These people, 10.3 % of the 2006 VHLSS data who should have been in the 2004 - 2006 VHLSS panel, had to be dropped from the dataset. Figure 1 shows bar graphs for each province representing the percentage of people in the 2006 survey who also should have been participants in the 2004 survey, yet who could not be matched with the identification codes in the 2004 survey. Among 9,926 people, 14.5 % of them are from a province coded as 601, for which 1,436 people fall into such a category. The few high spikes in the graph indicate that the miscoding took place in few particular provinces and from there, one can speculate that there were coding errors involved in conducting the survey in those provinces.

Gender discrepancies between the two surveys resulted in another 2,488 people, 2.9 % of 85,817 people, being excluded. There were 7,324 people whose reported year of birth is different between the two years. Due to the high frequency of such cases, only those with an absolute year difference of more than 5 years are excluded from the dataset, of which there were 2,299 cases, so an additional 2.8 %. Finally, 879 people who passed away between 2004 and 2006 survey were excluded in the final dataset.

The panel dataset consists of 17,276 households and 80,151 people. However, only people who were 13 - 55 years of age in 2004 were included in the analysis and final dataset, which consists of 50,936 people.

IV.B. Reference Group

Determining a legitimate reference group is the key problem in the study of relative deprivation, yet it is often given little attention in the economics literature. Sherif (1953) defined a reference group as: "Those groups to which the individual relates himself as a part or to which he aspires to relate himself psychologically." The concept of a defined reference group is obviously too ambiguous to define in empirical studies in the field of applied economics. Podder (1996) emphasizes that the formal definition of reference groups is nonexistent in empirical studies. Silber and Verme (2012) note that the specific proposals on how to define the reference group have been made only very recently. However, empirical studies on relative deprivation consistently regard individuals with proximate characteristics and similar opportunities in life to serve as points of comparison (Kuegler, 2009; Stark and Taylor, 1991; Ferrer-i-Carbonell, 2005; Eibner and Evans, 2005; Deaton, 2001). The geographical proximity is often assumed as a reference group (Stark and Taylor, 1991; Vickstrom, 2010; Deaton, 2001). Socio-demographic factors such as age, race, sex, marital status and educational level have been suggested and employed for empirical studies (Ferrer-i-Carbonell, 2005; Eibner and Evans, 2005;). Mangyo and Park (2011), using the nationally representative sample of 3,267 Chinese adults collected from China Inequality and Distributive Justice Survey project conducted in 2004, argues that relatives and classmates are salient reference groups for urban residents and neighbors play the same role for rural residents.

This paper uses the most popular and easily available variable, geographical proximity, as the basis for the reference group. For robustness of the analysis, five different reference groups are defined:

- (1) People in the same administrative boundary.
- (2) People of a similar income range in the same administrative boundary.
- (3) People of a similar age in the same administrative boundary.
- (4) People of a similar age with similar education in the same administrative boundary.
- (5) People of a similar age with similar income in the same administrative boundary.

The administrative boundary used in this study is the district. There are 64 provinces in Vietnam and each of the provinces have 8 – 9 districts on average. There are 670 districts in Vietnam, of which 630 are included in the 2004 VHLSS survey. The area and population of the districts are not uniform, varying between 6.9 square miles and 1419.7 square miles in terms of area, and between 11,650 people and 572,132 people in terms of population. There are 10,511 communes in Vietnam, which are the next smallest administrative areas after districts. Commune-level analysis would be the ideal scenario for the purpose of this study; however, due to the small sample size at the commune level, district-level analysis was used.

IV.C. Assumptions about the Reference Group

Reference group substitution can take place after people migrate to a new community. This scenario makes the empirical test of the hypothesis, which says $U(RD_1^i) > U(RD_0^i)$ would hold after migration, difficult to interpret. When a reference group is not stable and subject to change, there is the possibility that the utility they derive from the migration declines due to

the new and lower relative position they find in the new reference group, even though migrants enjoy increased absolute income in the receiving community. Stark and Taylor (1989), in their empirical study of Mexican internal and international migrants, find that international migration has the built-in protection against the reference group substitution due to an entirely different socio-cultural milieu in the new community, so that their home reference group remains the same, whereas internal migration is more subject to reference group substitution.

One of the key assumptions made in this study is that the reference group is stable for internal migrants in Vietnam. This assumption is has some supporting grounds. The Vietnam household registration system significantly limits and adversely impacts the internal migrants' access to social services and basic rights in big cities, which suggests that the big cities foster detachment and social distance from current big city residents for the migrants who arrive in search of better jobs. The household registration system monitors changes in people's residence in Vietnam by classifying them into different residential categories with differing rights and obligations. It is required for certain administrative procedures, such as buying land or building a house, registering a motor vehicle, borrowing money, obtaining access to medical care, water and electricity as well as government programs for poverty reduction (Marx and Fleischer, 2010). These structural barriers would impede reference group substitution among migrants to big cities in Vietnam.

Another supporting ground for the assumption could be the regular return of migrants to their home town. For Tet, or "Vietnamese Lunar New Year", most migrants return to their

home town to visit their families. This regular return to the hometown would maintain the social cohesion and bond to their original reference group.

While it is very likely that reference group substitution will take place eventually, especially for permanent migrants, because this paper focuses on the determinants of recent out-migration, the problem of reference group substitution should be minimal.

IV.D. Identification of Migrants

In this study, two types of out-migration are examined: long-term migration and short-term migration. Migrants here refer to those people who leave their hometown for some period of time for work purposes. Households that participated in both the 2004 and 2006 VHLSS surveys were asked in the 2006 VHLSS to provide reasons for the absence of household members in 2006 who were present in 2004. In other words, family members of those who were household members in the 2004 survey, yet are no longer in the household in 2006, are asked to explain why those who left are not there anymore. Between the 2004 and 2006 surveys, 5,556 people moved out of the panel households. Table 1 shows the reasons given for those individuals who moved out of their households between 2004 and 2006. Of the 5,556 people who left their households, the 1,612 people ages 13 - 55 whose reason for moving out of the household is "to work" are defined as long-term migrants.

Short-term migrants are defined as someone 13 - 55 years old who was a member of the same household in both 2004 and 2006, but was absent for at least one month during the past 12 months. There are 2,543 people in this category. Among these potential short-term migrants, 1,039 people who answered that they had not been employed for the last 12 months

prior to the survey date were dropped. Attending school was the most popular reason (over 91 %) for having no work experience in the last 12 months. The other reasons include, housework, retired, ill, disabled, and unable to find a job. There were 1,469 people defined as short-term migrants. The definition of short-term migrant used in this study may not capture what it is ideally designed to capture. A person's reason for staying out of a household for few months can vary, including family events. However, the selected samples are the best candidates for short-term migrants who moved for work purposes and are assumed here to represent the people who seasonally migrate to improve the socio-economic condition of their households. Note that Long- and short-term migrants are mutually exclusive.

Long-term migrants and short-term migrants are analyzed both together as well as separately. The summary statistics for both types of migrants, and for non-migrants, are shown in Table 2.

As shown in Table 2, there are slightly more males than females among long-term migrants. The average age of these migrants is 21.6 years old, and 86.9 % of them are single, which seems to be a sensible number considering the fact that young and unmarried people are more willing to take risks and migrate to other places for work. The average years of schooling of long-term migrants between ages 13 - 55 is 8.9, which is higher than the 7.8 years of education for non-migrants of the same age group. Similar characteristics for internal migrants in Vietnam were also found in the 2004 Vietnam Migration Survey (VMS) and the 2009 census. The VMS and the census reported that most migrants are young adults, more than 50% of them are under 25 years old, and most are single males from rural areas. However, reports based on

both of the surveys also report that female migrants are increasing as the demand for female workers in the industrial sector has risen (UNFPA, 2007; Marx and Fleischer, 2010). The 2004 VHLSS is also similar to the 2004 VMS regarding the level of education of migrants and non-migrants; contrary to the common perception that migrants are the least educated and least qualified workers, both surveys show that migrants are typically young people with higher levels of education (UNFPA, 2007). It should also be noted that the qualified migrants are not from wealthy families. They are from, on average, economically worse-off households. Almost 90% of migrants' positions in the household were children of the household head, indicating that children who are old enough to work migrate in order to improve the welfare of their families.

It is interesting to see the distinct characteristics of long-term and short-term migrants. Short-term migrants are more likely to have their own families and to be heads of households. The obligation to take care of the family at home and the higher cost of the migration of entire families may be important reasons for the inability to undertake long-term migration. Short-term migrants are poorer than long-term migrants and have slightly less education, but still more than non-migrants.

Table 3 shows the composition of occupations of long-term and short-term migrants, as well as non-migrants. There are 10 categories of occupation which break down into a total of 34 subcategories of different occupations in the employment section of the 2004 and 2006 VHLSS surveys. After combining a few of the categories that are similar, six different occupation groups were used to measure the occupational composition of migrants and non-migrants.

There is no notable difference among the three groups except that a large portion of long-term migrants had no work experience in 2004. This might be simply because long-term migrants are younger than the other two groups. Most workers in Vietnam are employed as unskilled agricultural workers. Subgroups of unskilled workers according to the type of industry are included in the table. Unskilled workers in agriculture, sylviculture and aquaculture constitute more than 65% of unskilled workers across all groups. Types of industry that employ unskilled workers do not vary across long-term and short-term migrants. However, when compared to non-migrants, migrants are slightly more likely to come from mining, construction, manufacturing and transportation and slightly less likely to come from agriculture, sylviculture and aquaculture. Potential explanations for this phenomenon are that people who work exclusively in primary industry such as agriculture have smaller networks in urban areas than people with exposure to secondary industry or that the higher cost of migration for a greater distance of is prohibitive.

Figure 2 shows the total migrants' geographical origins at the province level. The map on the left illustrates the percentage of migrants for each province. For example, in Thanh Hoa province, there are 135 migrants who left the province for some period or longer after 2004 survey, and they form approximately 7% of the whole provincial population in the sample. The migrants here are the sum of both long-term and short-term migrants. The map illustration is at the province level because the district-level geographical information that links the VHLSS district code and the district-level GIS map frame is not available.

The map on the right shows cities with populations over 200,000. Ho Chi Minh City is the most populous city in Vietnam, followed by Hanoi and Hai Phong. As one can see, most migrants seem to come from provinces that are adjacent to the populous cities, rather than provinces that are farther away. Because long-term migrants and short-term migrants may decide on the duration of their migration differently depending on the proximity of big cities, a comparison is made and shown in figure 3.

In Figure 3, the map on the left presents the origin of long-term migrants and the proportion of migrants in each province. The map on the right does the same for short-term migrants. While there are differences, it seems that both short- and long-term migrants come from provinces near the populous cities rather than from provinces that are farther away from the populous cities.

Figure 4 presents the map illustration of the number and percentage of total migrants in each province and percentages of people who live below the national poverty threshold in each province as well as in the populous cities. The poverty line is set at 1,300,000 VND per capita per year, which is equivalent to 87 USD per year. This income level was the amount earned by people in the 10th percentile of the income distribution according to the 2004 VHLSS calculation. The map on the right illustrates the, percentage of people who live under the poverty line in each province. Poverty is highly concentrated in the Northwest whereas poverty rates in provinces with populous cities are the lowest. The spatial correlation between the poverty rate and the migration rate seems to be less significantly related than neo-classical models of migration would suggest.

IV.E. Calculation of Household Income

For the calculation of household income, all sources of reported income were summed and then reported business expenses were subtracted to calculate net income. The reported income from primary or secondary jobs in agriculture, forestry, and aquacultural production activities, livestock breeding, agricultural services, and hunting, as well as other income from remittances and the value of in-kind gifts from people overseas or in Vietnam who are not household members, and all other incoming money that is considered income, is included in the calculation. However, “other” income, such as money from the one-time sale of property or borrowing, is not included because it implies the rearrangement of existing wealth. The survey also includes questions about operating expenditures of income generating activities other than salary-paying jobs. All the operating expenditures are subtracted to derive the net income of each household. The entire population who participated in the 2004 survey, 201,320 people in 45,592 households, is included in the calculation. Summary statistics for income are shown in Table 4.

Figure 5 shows the per capita income at the province level. Two maps are presented for the purpose of comparison between the Author’s calculation of the per capita income from the 2004 VHLSS and the GSO’s calculation of the per capita income, also from the 2004 VHLSS. The correlation plot shows a clear positive relationship between the two calculations. Derivation of Income is pivotal factor in this study because the per capita income variable is used to calculate the measure of relative deprivation. Per capita income was calculated by dividing total household income by the number of household members in 2004. Many scholars, such as Mincer (1978), have criticized the atomistic focus of migration analysis, arguing that the

migration decision is closely interrelated with family ties. Thus, household income rather than individual income is used, based on the assumption that Vietnam is a family-centered society and that the household as a unit plays an important role in the migration decisions of its members. This assumption is supported by the 2004 VMS, in that about two-thirds of male migrants and 80% of female migrants reported that other people were involved in their decision to migrate (UNFPA, 2007). Another reason, perhaps more importantly, for choosing household income is the fact that self-employment income is reported only at the household level rather than the individual level. This limits the use of income at the household level.

The map on the left in Figure 5 shows the VHLSS 2004 average per capita income for each province, as reported in the General Statistics Office 2004 Report (GSO 2006), whereas the map on the right represents per capita income for each province following the calculation described above. There is no notable deviation between the two calculations. The calculation used in this study for household income is thus assumed to be a good approximation of the income variable used by Vietnam's General Statistics Office.

Scanning the map, one may observe the geographical clustering of the level of income. Spatial autocorrelation was examined by using Global Moran's I test, a measure of spatial autocorrelation offered by the ArcGIS software. The Global Moran's I is an extension of the Pearson product-moment correlation coefficient and evaluates whether each of the geographical units, which are the provinces in this paper, exhibit any pattern that is clustered, dispersed or random. (ESRI, 2012) The Global Moran's I also calculates a Z score that indicates the validity of the null hypothesis that there is no spatial correlation. Given that Moran's index

ranges from -1 to 1, meaning perfect dispersion for -1 and perfect correlation for 1, the index of 0.49 with a significance level below 1% indicates non-trivial positive spatial correlation of the level of income.

V. The Econometric Model

A logistic regression model is used to estimate the long- and short-term migration decision. The dependent variable is a binary variable that equals one if a participant in the 2004 survey migrated to another place by 2006. One advantage of using the 2004 and 2006 VHLSS panel data is that socio-economic status can be measured prior to the migration decision as reported in the 2006 survey. This significantly reduces the possibility of reverse causation bias that is found in studies that use cross-sectional data.

A two-stage cluster sampling technique was employed for the VHLSS. Households that are drawn randomly from within the same cluster, such as communes or enumeration areas in the VHLSS, are more likely to share similar characteristics than randomly selected households without clustering (Deaton, 1997). Thus household and individual characteristics are not completely independent from each other when two-stage cluster sampling is used, and this correlation increases the variance of the estimates, which reduces their precision. In the VHLSS, there are two choices of primary sampling units (PSU), either communes or enumeration areas, where each commune in the sample contains three enumeration areas and each enumeration area is a clusters of 15 households (GSO, 2004). The commune was chosen as the PSU cluster for the purpose of adjusting the standard errors for clustering.

Because generalized linear models such as the logistic model assume the independence of the observations, this could lead to serious and significant underestimation of the standard errors. In order to account for this problem, the cluster-robust standard error correction offered in STATA was used for the conditional logistic regression model.

The Yitzhaki Index, which is a proxy for relative deprivation, is calculated using five different assumptions about reference groups. For the baseline case, each district is assumed to be a reference group, so that people are assumed to compare their relative position to everyone within the district in which they live. For the other four cases, the reference group is people in similar income ranges, or with the same socio-demographic characteristics, within the district in which they live. The details of these other groupings were described in section IV.B.

For the baseline results, the following logistic regression model is used:

$$Pr(Migrate_{ij} = 1 | RD_{ij}, X_{ij}) = G(\beta_0 + \beta_1 RD_{ij} + X_{ij}\beta + \gamma + \varepsilon_{ij})$$

where the dependent variable, *migrate*, equals 1 if the person migrated since the 2004 survey and 0 otherwise. Subscripts *i* and *j* indicate the individual and the reference group, respectively. RD is the measure of relative deprivation measured by the Yitzhaki index. The term X_{ij} is the vector of variables that accounts for personal and socioeconomic characteristics such as age, marital status, education level, gender, household size, insurance status and per capita household income. The health insurance status variable is included in the model as a proxy for the safety net people have in their home town. The term γ is a district fixed effect that is designed to control for person-invariant differences across people in the same districts. The regression analysis for all migrants, that is, long- and short-term migrants combined, is

presented in Table 5. Analyses for long-term and short-term migrants are also conducted separately, and the results are presented in Tables 6 and 7, respectively.

VI. Results

VI.A. Reference Groups: (1) District and (2) District and Income

The reference group in the first four columns of Table 5 is assumed to be anyone who lives in same district. Different specifications are included to examine the impact of RD on the migration decision. For specifications in columns 1 and 2, every observation is assumed to be perfectly random and the standard logistic regression is used. The sign of RD is negative and statistically significant at the 1% level when clustered standard errors are used, which is exactly the opposite of what the theory of relative deprivation predicts. When district fixed effects are added, as shown in specifications 3 and 4, the negative sign of RD persists, but it is much smaller in magnitude and no longer statistically significant.

Consistent with the summary statistics of migrants, regression results also support that those who left their hometown after the 2004 survey are more likely to be male, younger, educated, and single. Two types of Yitzhaki Indices, as described in the Quantifying Relative Deprivation section, are included in the first row of the Table 5 under the variable name “RD”. The analogous regression was run separately for the modified Yitzhaki index using the same model with the same covariates, however, the coefficients for the covariates are not presented and only the coefficient for the modified Yitzhaki index was included in the same table that contains the results for the regression with original Yitzhaki Index for the sake of comparison. The differences in the covariate coefficients between the two models are negligible.

The regression results for income-stratified reference groups in the same district are presented in the last four columns of Table 5. It is assumed that people are more likely to have social ties to people within a similar income range. This reduced the size of available observation samples. Any districts with fewer than 15 households were dropped from the index calculation in order to have a reliable index measure.

The sign of RD is positive with the exception of the people in the top 25% income range, although the effect is not statistically significant. The relative deprivation hypothesis predicts that the sign of RD is positive. The RD has the strongest influence on the migration decisions for group 3, those whose earnings are in the 50–75% range of the earnings distribution. For group 3, people whose level of RD is one standard deviation above the mean are 1.54 times more likely to migrate than people at the mean, with statistical significance at the 10% level when original Yitzhaki Index was used as a proxy for relative deprivation and 5% level when modified Yitzhaki index was used as a proxy for relative deprivation. The degree of effect is less for groups 1 and 2. However, the effect is still non-trivial. It is interesting to note that a one standard deviation increase in log of income is associated with a roughly 44% decrease in being a migrant for people in the top 25% of income distribution. It seems reasonable that there is no significant RD effect for people in upper 25 % of income distribution because there might be many other factors in the migration decision for the wealthiest people, particularly the amount of assets owned by the wealthiest in a society. The inclusion of asset variables in the regression model would allow testing the effect of assets for people of higher income distribution. Even though household-level asset information is available from the VHLSS income and expenditure data, the income-only dataset data has no information on the assets is used in this study for the

benefit of greater sample size. Income-related factors including relative deprivation may have a smaller effect for those in the top income distribution. The quadratic term for income is not used in the income stratified case as the range of income is small. Figure 6 presents the frequency of total migrants according to the level of income in the whole sample as well as the income distribution of entire sample. One should note that the concave-shaped distribution is made possible by including the quadratic term in the regression. However, the frequency distributions of migrants for income stratified groups are nearly linear, and thus regressions do not require the inclusion of quadratic term. Figure 7 shows the Income distribution of entire sample in the survey. Note that the mean and median incomes of migrants are smaller than the sample population mean and median.

Tables 6 and 7 give regression results for long- and short-term out-migration, respectively. There is no significant effect of RD for either long- and short-term migration when the district as a whole reference group is the base of the analysis. One major difference that should be noted is that gender has no bearing on long-term migration, whereas being a male doubles the chance of short-term migration. Being single has a large impact on the long-term migration decision, and more education correlates with long-term migration more than with short-term migration.

For income-stratified reference group analysis, there is a sharp difference between the RD effects on long- and short-term migration. RD proxied by the Original Yitzhaki index is strongly and positively associated with long-term out-migration for people in lower 25% and 50 – 75% of income distribution, whereas it has virtually no effect on short-term migration. One

standard deviation increase in RD doubles the likelihood of long-term migration for people in lower 25% and 50 – 75% of income distribution but has no significant effect on short-term migration. RD proxied by modified Yitzhaki index have similar effect on the people in the 50 – 75% of income distribution, yet no correlation was found for the people in lower 25% of income distribution.

The distinctive difference in migration dynamics for short- and long-term migration is quite interesting. Given that there are presumably higher economic and psychological costs of migration for long-term than for short-term, it could be assumed that the degree of relative deprivation is an important factor in the decision to migrate with higher risks.

Figure 8 shows the time periods in months for which short-term migrants were absent from their hometown during the 12 months prior to the survey date. There is no clear pattern in the distribution.

The relative position in a reference group is not an important factor that drives short-term migration. One possible explanation for this phenomenon could be that, for short-term migrants, relative deprivation is not a major concern, as they are older, less willing to take risk and psychologically settled with what they already have with their family, and thus it has no effect on their consideration to migrate temporarily to other places in search of better jobs. This possibility is somewhat supported by the characteristics of short-term migrants as shown in Table 2. Short-term migrants are relatively less educated, economically worse off than long-term migrants and mainly older males with families to look after. Another explanation along similar lines is that short-term migration does little to improve income and thus, to reduce less

relative deprivation. It would give less incentive for potential short-term migrants to undertake the difficult journey of finding a temporary job that pays just a little more than what they earn at their hometown, away from their home and family. This background of short-term migrants gives them less reason to move in order to improve their relative position in the hometown.

Another reason could be partly due to the loose definition of short-term migration. The VHLSS does not ask specific questions on either short-term migration status or reasons for the migration. It only asks the participant's duration of absence from the household between 2004 and 2006. The reason for an absence could be not only for the work purposes, but also for many other reasons such as family gatherings or etc. The short-term migrant depicted in this study is not as clearly defined as the long-term migrant, which may create noisiness in the data.

Although there is a definitional problem with the short-term migration in this study, it is still assumed that the relative position in a reference group matters more for long-term migration than for short-term migration based on the result of the study.

VI.A.1 Results for Different Age Groups

The dataset used in the previous section includes people ages 13–55 at the time of the 2004 survey. The econometric results show that migration behavior is highly correlated with age. In this section, the effect of relative deprivation for three different age groups, in their teens, 20s and 30s, are separately studied utilizing the same econometric model used in the previous section. For the income-stratified reference group, only people in their 20s and 30s are included as they are the ones who migrate the most among all age groups. It should be noted that the reference group remains the same as before; the age factor is used merely as a

disaggregating tool for the regression analysis, and is not considered in the calculation of the Yitzhaki index.

The three columns of the left-side in Tables 8 and 9 show the regression output for different age groups whose reference group is assumed to be all the people of the district in which they reside. For either long- and short-term migration, no effect of relative deprivation is observed when district reference group is grouped by different age.

The Income-stratified reference group analysis includes only people in their 20s and 30s with the expectation that one would observe a larger effect of relative deprivation on out-migration for those in their 20s and 30s than for larger age groups. Tables 8 and 6 compare the results between the 20s and 30s and the larger age range. The effect of relative deprivation is greater for people in their 20s and 30s, especially those people in lower 25 % and 50 – 75 % of income distribution. The effect is twice as large for people in lower 25 % and three times as large for people in 50 – 75% as the effect for the whole age range.

There is still no effect of relative deprivation on short-term migrants. It is assumed that the same explanation that was given in the previous section that they are older, married and have children, applies to this regression results as well. For further analysis with different reference group assumptions, only long-term migration is considered.

VI.B. Reference Group: (3) District and Age

Perhaps only people of similar ages in the same district are assumed to be an individual's reference group. To check this, the Yitzhaki Index was calculated for each age group.

The regression output in Table 10 shows that in this case, there is no meaningful effect of relative deprivation on long-term migration. Given that a significant effect of relative deprivation was found on long-term migration when a reference group was defined as people of similar income, the results may indicate that people do not choose their reference group just because they are the same age and reside in geographically close places. To further explore which people constitute a reference group, another section takes the education variable into consideration in calculation of the Yitzhaki index.

VI.C. Reference Group: (4) District and Age and Education

Next, consider results when people in their 20s and 30s with similar levels of education in the same district are assumed to be a reference group. “Primary or none” is defined as people who have never attended school or who have completed between one and five years education (the primary school cycle). Lower secondary is defined as people who have completed between six and nine years of education (the lower secondary education cycle). Upper secondary is defined as people who have completed between ten and twelve years of education. Of the 2004 VHLSS participants aged 20–39, 95% reported that their level of education falls into one of the three categories defined above. The remaining 5% are the people who have post-secondary education.

Approximately 7% of people who have upper secondary education chose to migrate, whereas only 3% and 2% in lower secondary and “primary or none” categories, respectively, chose to migrate. Defining a specific reference group considerably reduces the size of the dataset available for analysis. Yet, the sign of the relative deprivation coefficient is positive

across all categories as shown in Table 11, most notably for people with upper secondary education.

VI.D. Reference Group: (5) District and Age and Income

Finally, perhaps only those in their 20s and 30s with similar income ranges who live in the same district are considered to be the reference group. To check this, the Yitzhaki Index was calculated accordingly.

The results are shown in Table 12. The effect of relative deprivation is significant for all groups except people in top 25 % of income distribution in case with the original Yitzhaki index, as was the case with the whole income-stratified reference group in Table 8. The size of the effect is also very similar to that found in Table 8, which is understandable considering the fact that per capita income is derived from the household income (household income was normalized by the size of the household members) and all the members of the household have the same amount of income. This means that the Yitzhaki Index calculated in this section is the index for a household that contains a person aged 20–39. This index would decrease the sample size included in the regression analysis, yet the results do not vary much from the results in Table 8.

The advantage of defining the reference group used in this section is that the results are more reliable, given the assumption that people in the age range are more likely to react to the psychological deprivation one might feel in the reference group, as the Yitzhaki Index excludes the households without members in their 20s and 30s in the calculation. The effect of relative deprivation measured by the original Yitzhaki Index is significant at a 5% level for people in

lower 25 % and 25 – 50 % of income distribution and at a 1% for people in 50 – 75% of income distribution.

VII. Conclusion

This paper examines the effect of relative deprivation on the migration decisions of Vietnamese people. Many scholars in the social sciences have pointed out that the motivation for migration is multi-dimensional, going beyond the simple pursuit of higher income, and thus migration theorists must take a multi-disciplinary approach to formulate a complete model (Bodarsson and Van Den Berg 2009). The theory of relative deprivation, which is rooted in the psychology and experimental economics literatures, predicts that, holding all demographic and socioeconomic characteristics constant, a deteriorating relative position within a reference group will result in migration for economic betterment and improvement of one's relative position.

The empirical evidence from the VHLSS 2004 -2006 panel data provides some support to the hypothesis that relative deprivation is a factor that influences migration decisions in Vietnam. The definition of a reference group is the most important consideration in the empirical study of relative deprivation. The sizable VHLSS dataset allows the testing of the effect of relative deprivation using various reference groups. Among the five reference groups used in this study, the effect of relative deprivation is significant in three groups, whereas no effect were found in the other two groups. However, this significant result applies only to the long-term migration, not short-term migration. Given the importance of defining a reference group and somewhat arbitrary results of this study for different definitions of reference groups,

a careful study of what constitutes the reference group in a specific context should precede an empirical study of the theory of relative deprivation.

Two versions of the Yitzhaki Index are used as proxies for relative deprivation in the empirical testing. The results show that the original Yitzhaki index always dominates the modified Yitzhaki index in terms of statistical significance as well as the magnitude of the effect on migration decision. The theoretical justifications made as described on page 9 - 10 for the modified Yitzhaki index are meant to be captured in mathematical formulation of the modified Yitzhaki index, but it is hard to mathematize people's psychological weight in their formation of relative deprivation.

While there is a comparative advantage of using panel data, relative to cross-sectional data, in the study of migration, the endogeneity problem still remains the most troublesome drawback of this study. It is possible that the measure of relative deprivation used in this study could be a function of accumulative past migration within a family or a community. A study by McKenzie and Rapoport (2007) shows, for the Mexican case, that the relationship between wealth and migration is affected by the level of accumulated migration that had taken place in a community. The inability of this study to effectively trace down the level of accumulated past migration in each district leaves room for inaccurate interpretation of the empirical results. It would be the best approach to use multiple panel of surveys and employ household fixed effect estimation methods to assess the causal relationship between relative deprivation and migration behavior. This is not possible here because the sample size shrinks drastically when multiple surveys are used.

Since the analysis was carried out at the individual level, suitable instruments for relative deprivation that would allow one to solve the endogeneity problem cannot be employed.

Given the importance of a reference group, as well as people's relative positions within them, inequality in the sending community should be given increased attention in order to prevent abrupt influx of people to urban areas that associate unwanted social instability. Although consideration of all dimensions that influence the migration behavior of people in a particular country is an impossible task, empirical research and policy decisions that neglect the context-specific role that relative position plays in human activity, especially the economic-driven ones, might result in social dysfunctions that are caused by neglected attention to the increased discontent among relatively deprived groups. It is often the discontent arising from the interwoven nature of human relations that drives individuals' activities, either economic or political. Increased attention to the magnitude of relative deprivation of a community and further study on what factors determine the individuals' reference groups in a society would be required to adequately design policies to achieve the balanced development that Vietnam is now pursuing.

REFERENCES

- Bhandari, Prem. 2004. "Relative Deprivation and Migration in an Agricultural Setting of Nepal." *Population and Environment* 25 (5): 475-499.
- Bodvarsson, Orn Bodvar, and Van Den Berg Hendrik. 2009. "The Economics of Immigration: Theory and Policy." Springer Chapter 2: 54
- Deaton, Angus. 1997. "The Analysis of Household Surveys: a microeconomic approach to development policy." Washington, D.C: The World Bank.
- . 2001. "Relative Deprivation, Inequality, and Mortality." NBER Working Paper.
- Duesenberry, James. "Income, Saving and the Theory of Consumer Behavior." Cambridge, MA: Harvard University Press., 1949.
- Easterlin, R A. 1974. Does Economic Growth Improve the Human Lot? Some Empirical Evidence. Ed. Paul A David and Melvin W Reder. *Nations and households in economic growth*. New York and London: Academic Press.
- Eibner, Christine, and William N. Evans. 2005. "Relative Deprivation, Poor Health Habits, and Mortality." *Journal of Human Resources* 40 (3): 591-620.
- ESRI Developer Network. "Spatial Autocorrelation (Morans I)." Web. 24 June 2012
<[http://edndoc.esri.com/arcobjects/9.2/net/shared/geoprocessing/spatial_statistics_tools/spatial_auto_correlation_morans_i_spatial_statistics .htm](http://edndoc.esri.com/arcobjects/9.2/net/shared/geoprocessing/spatial_statistics_tools/spatial_auto_correlation_morans_i_spatial_statistics.htm)>.
- Ferrer-i-Carbonell, Ada. 2005. "Income and Well-Being: An Empirical Analysis of the Comparison Income Effect." *Journal of Public Economics*, 89 (5–6): 997–1019.
- General Statistics Office (GSO). "Vietnam Household Living Standards Survey (VHLSS) 2002 and 2004: Basic Information." 2006.
- Harris, John R., and Michael P. Todaro. 1970. "Migration, Unemployment, and Development: A Two-Sector Analysis." *American Economic Review* 60: 126–42.
- Hey, John D., and Peter J. Lambert. 1980. "Relative Deprivation and the Gini Coefficient: Comment." *The Quarterly Journal of Economics*, 95 (3): 567–73.
- Kuegler, Alice. 2009. "A curse of comparison? Evidence on reference groups for relative income concerns." World Bank.
- Marx, Veronique, and Katherine Fleischer. 2010. "Internal Migration: Opportunities and challenges for socio-economic development in Viet Nam." UN papers.

- Mangyo, Eiji, and Albert Park. 2011. "Relative Deprivation and Health: Which Reference Groups Matter?" *Journal of Human Resources*, 46 (3): 459–481.
- McKenzie, David, and Rapoport Hillel. 2007. "Self-selection patterns in Mexico-U.S. migration: the role of migration networks." Policy Research Working Paper Series 4118, The World Bank
- Mincer, Jacob. 1978. "Family Migration Decisions." *Journal of Political Economy* 86: 749-73
- Naoki ,Kondo., Ichiro Kawachi., S. V. Subramanian., Yasuhisa Takeda., and Zentaro Yamagata. 2008. "Do Social Comparisons Explain the Association between Income Inequality and Health? Relative Deprivation and Perceived Health among Male and Female Japanese Individuals." *Social Science & Medicine*, 67 (6): 982–87.
- Nguyen Thu Phuong, Tran Ngo Thi Minh Tam, Nguyen Thi Nguyet and Remco Oostendorp. "Determinants and Impacts of Migration in Vietnam." Working Papers01, Development and Policies Research Center (DEPOCEN), Vietnam, 2008.
- Oshio, Takashi, Kayo Nozaki, and Miki Kobayashi. "Relative Income and Happiness in Asia: Evidence from Nationwide Surveys in China, Japan, and Korea." PIE/CIS Discussion Paper 487, Institute of Economic Research, Hitotsubashi University, 2010.
- Phung Duc Tung and Nguyen Phong. 2004 "Vietnam Household Living Standards Survey (VHLSS), 2002 and 2004: Basic Information." Vietnam General Statistics Office.
- Podder, Nripesh. 1996. "Relative Deprivation, Envy and Economic Inequality." *Kyklos* 49 (3): 353–76.
- Quinn, Michael. 2006. "Relative deprivation, wage differentials and Mexican migration." *Review of Development Economics* 10: 135–53.
- Rousseau, Jean-Jacques. *A Discourse on the Origin of Inequality*. Translated by G. D. H. Cole. A Digireads.com Book, 2005.
- Runciman, Walter G. *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-Century England*. Berkeley, CA: University of California Press, 1966.
- Sen, Amartya. 2006. Conceptualizing and Measuring Poverty. *Poverty and Inequality: Studies in Social Inequality (Chapter 2, 30 - 46)*. Stanford. Stanford University Press.
- Sherif, Muzafer. *Group Relations at the Crossroads*. New York: Harper, 1953.
- Silber, Jacques and Verme, Paolo. 2012. "Relative deprivation, reference groups and the assessment of standard of living," Policy Research Working Paper Series 5930, The World Bank.
- Stark, Oded. 1984. "Rural to Urban Migration in LDCs: A Relative Deprivation Approach." *Economic Development and Cultural Change* 32: 475–86.

- Stark, Oded, and J. Edward Taylor. 1989. "Relative Deprivation and International Migration," *Demography* 26 (1): 1–14.
- . 1991. "Migration Incentives, Migration Types: The Role of Relative Deprivation." *The Economic Journal* 101: 1163–78.
- Stark, Oded, and Yitzhaki, Shlomo. 1988. "Labour Migration as a Response to Relative Deprivation," *Journal of Population Economics* 1 (1): 57–70.
- Tischler, Henry L. *Introduction to Sociology*. 9th edition. Wadsworth Publishing, 2006.
- UNFPA. Internal migration in Vietnam: The current situation. 2007.
- Veblen, Thorstein. *The Theory of the Leisure Class: An Economic Study in the Evolution of Institutions*. New York, Macmillan, 1899.
- Vickstrom, Erik. 2010. "Keeping up with the Diallos? Household Wealth, Relative Deprivation, and Migration from Senegal to Europe." Unpublished manuscript, Princeton University, Princeton, NJ.
- Yitzhaki, Shlomo. 1979. "Relative Deprivation and the Gini Coefficient." *The Quarterly Journal of Economics* 93: 321–24.
- . 1982. "Relative deprivation and economic welfare." *European Economic Review* 17 (1): 99 - 113

APPENDIX A: FIGURES

Figure 1. Percent of Panel Mismatches by Province

Percent of people with mismatched ID number at each provinces

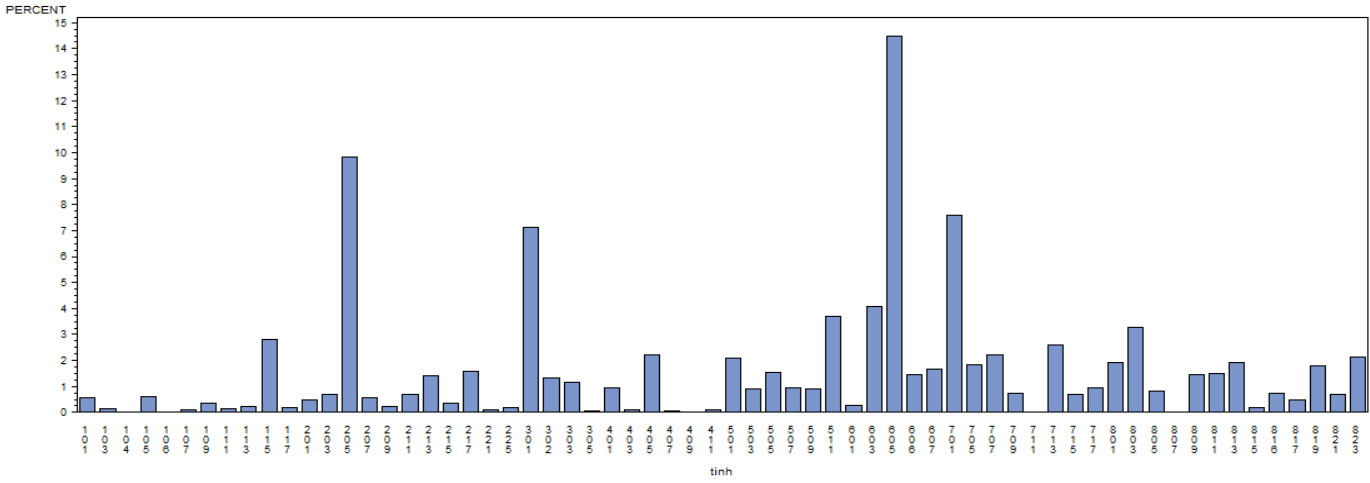


Figure 2. Origin of Out-Migrants and Cities with Population Greater than 200,000

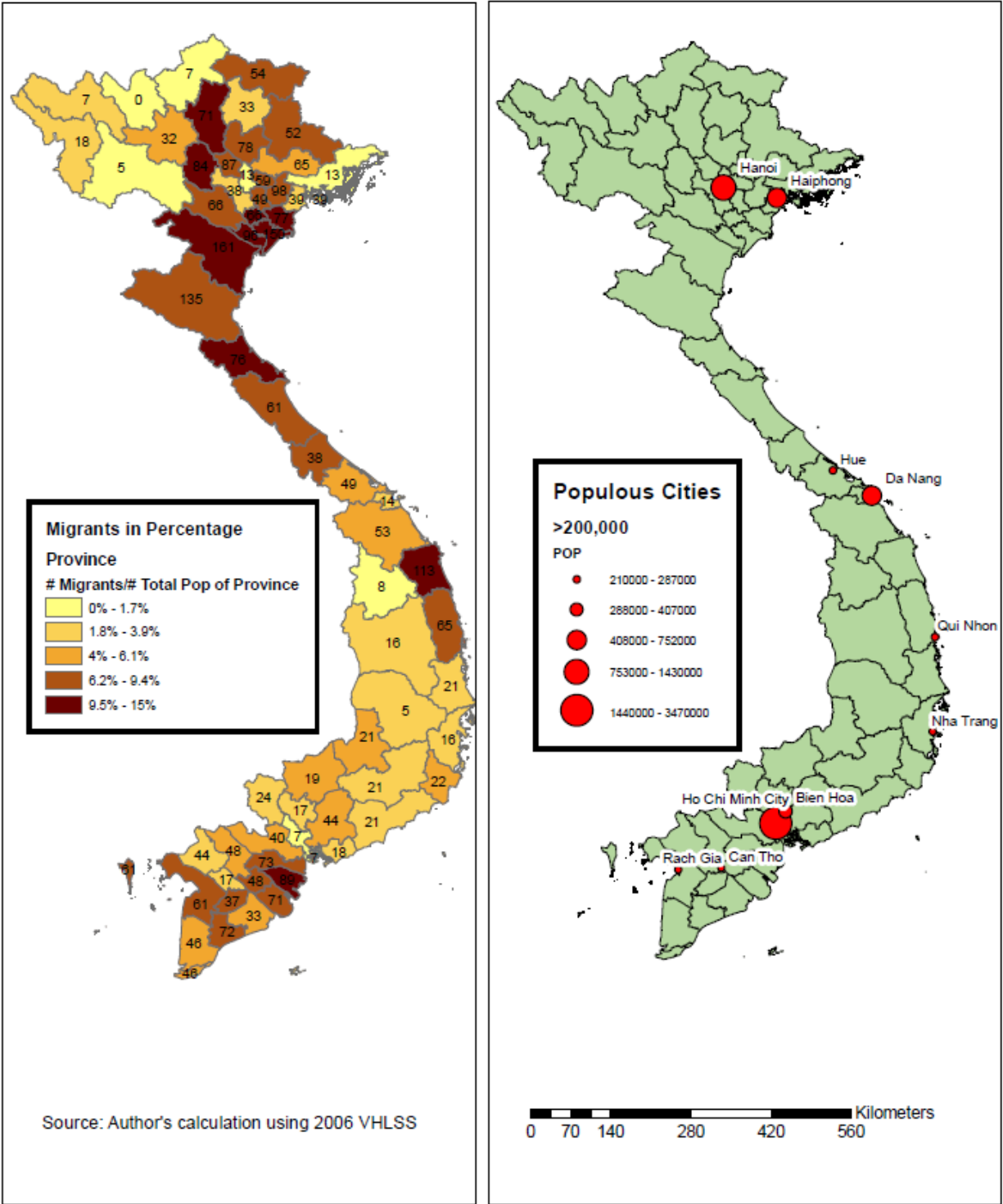


Figure 3. Origin of Long- and Short-Term Out-Migrants

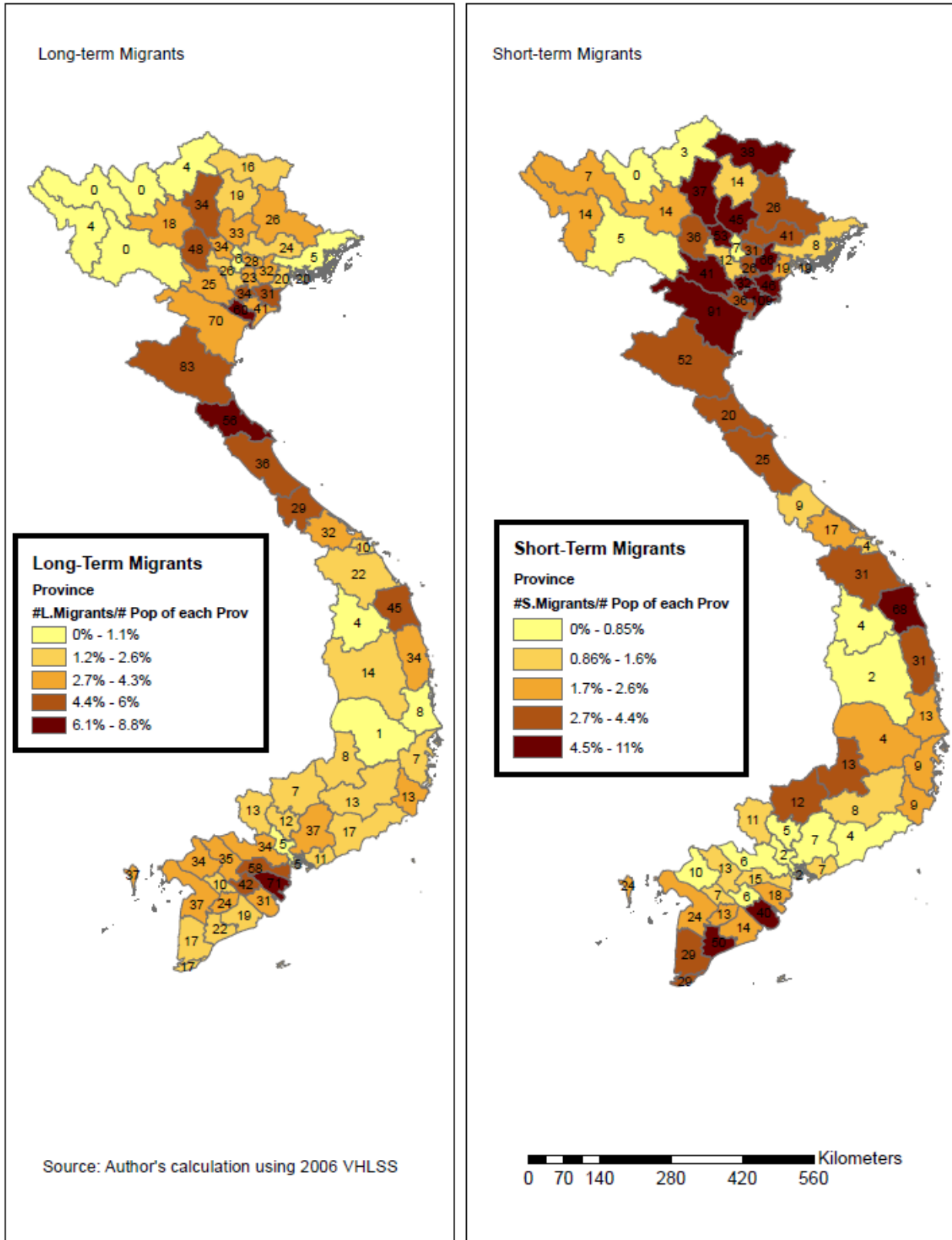


Figure 4. Origin of Out-Migrants and Poverty Rate

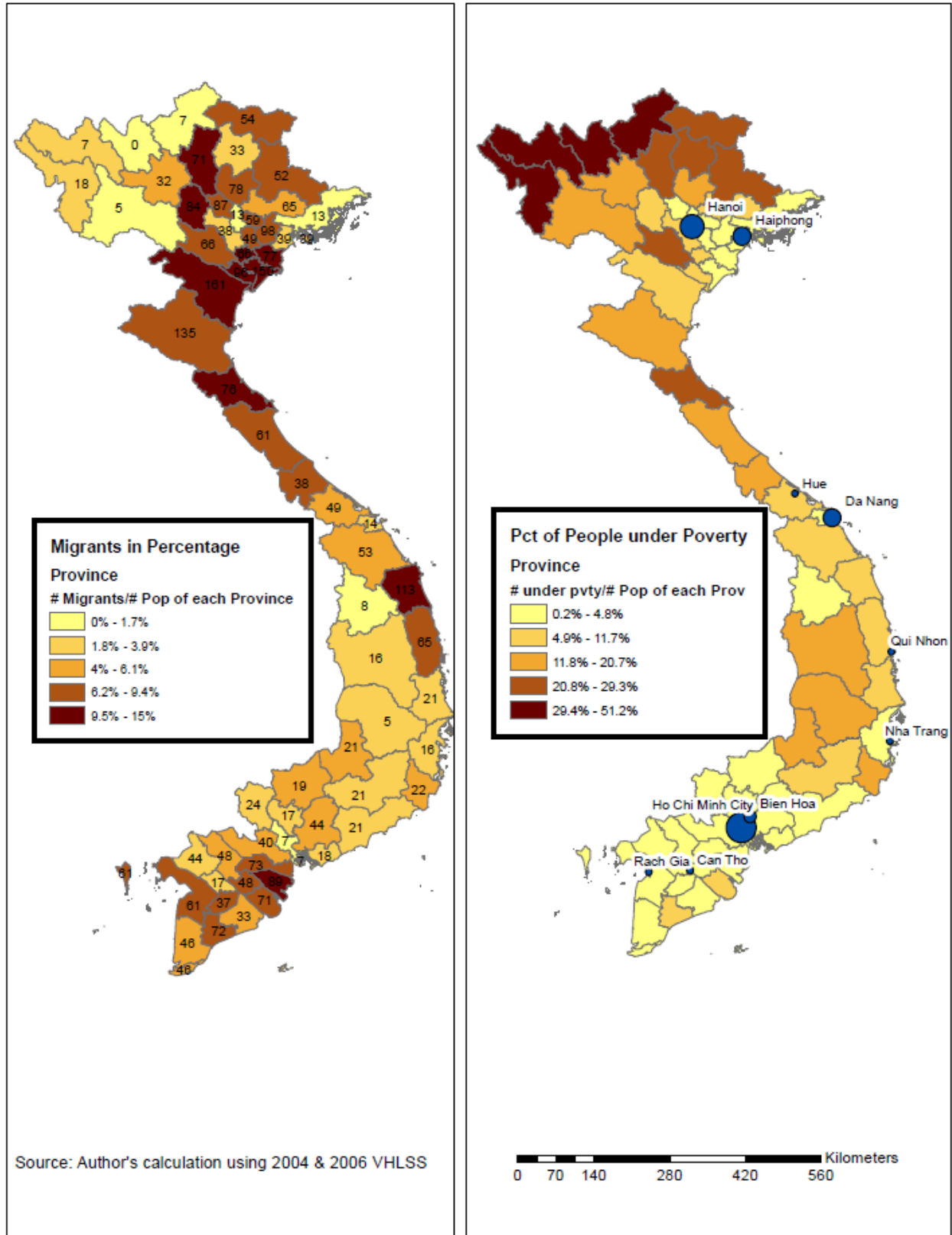
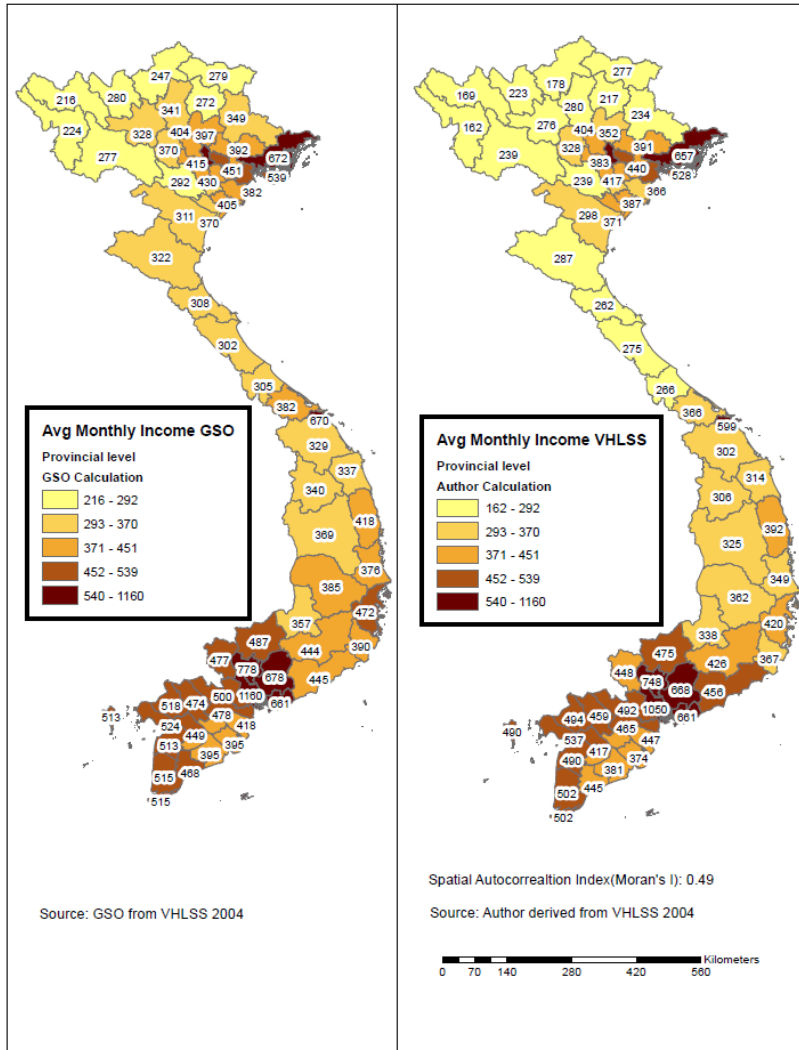


Figure 5. Per Capita Income Comparison between General Statistics Office Calculation and Author's Calculation



2004 Monthly Per Capita Income Correlation Plot

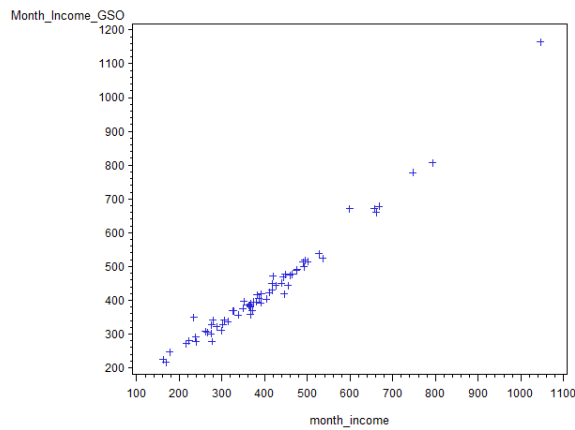


Figure 6. Relationship Between Income and the Frequency of Migrants

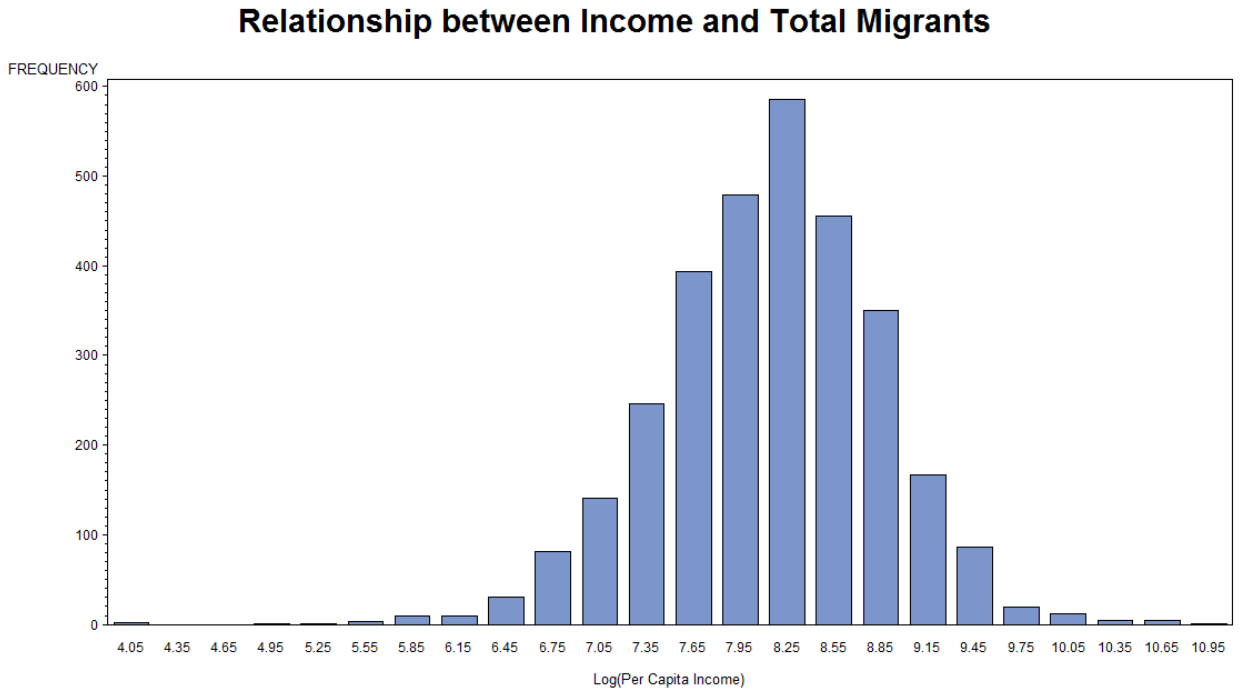


Figure 7. Income distribution of Entire Sample

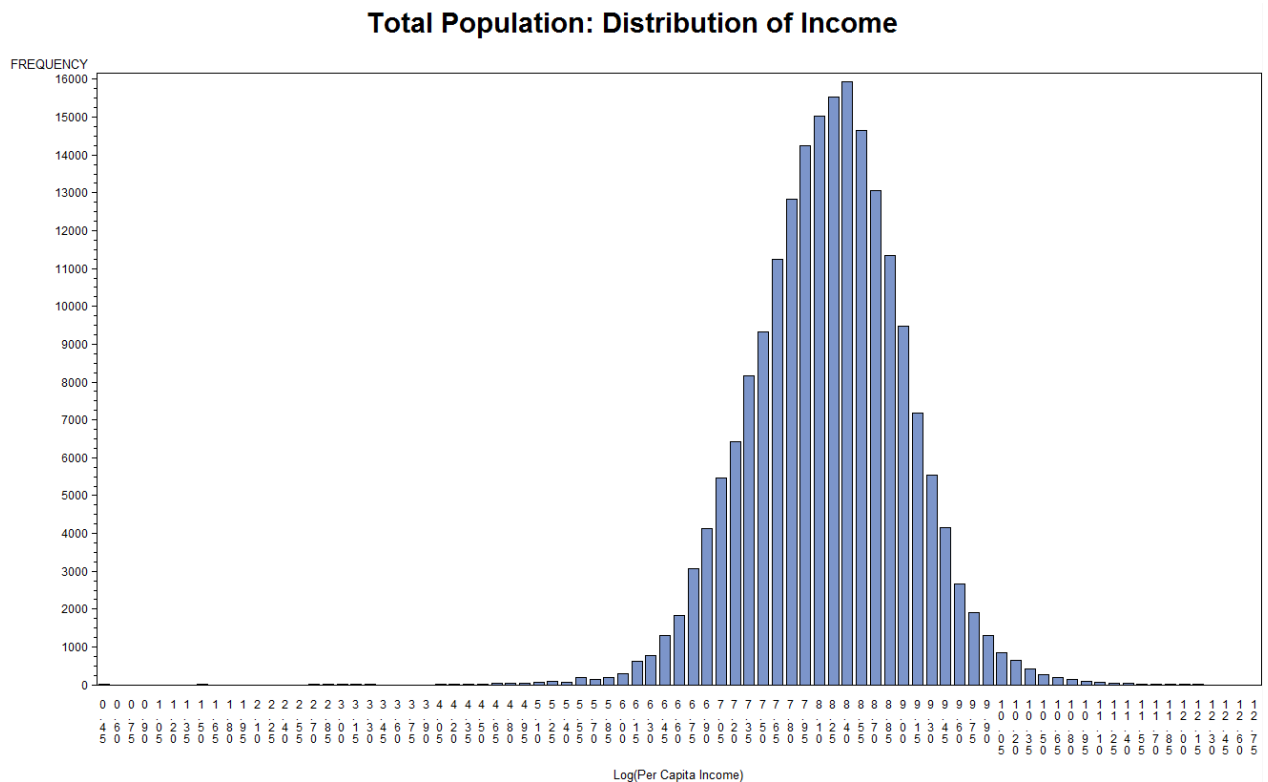
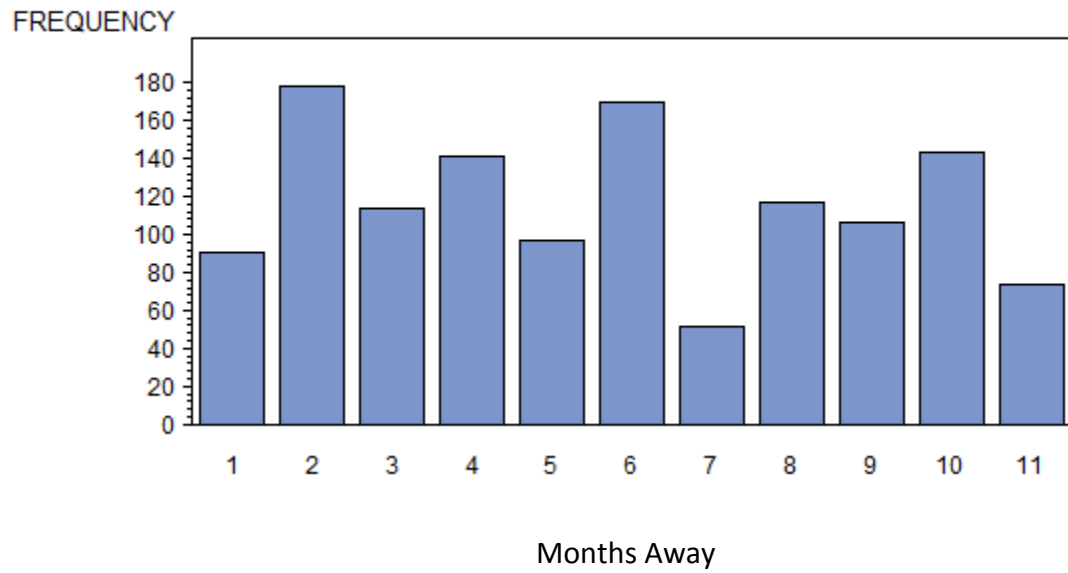


Figure 8. Duration of Migration

Frequency of Short-Term Migration by Duration in Month

APPENDIX B: TABLES

Table 1. Reason for Moving

Reason for moving	N	%
Work	1612	29.0
Married	1716	30.9
Household Split	1683	30.3
Study	188	3.4
Go with Family	129	2.3
Other	228	4.1
Total	5556	100

Table 2. Summary Statistics of Migrants and Non-Migrants

	Long-term Migrant(N=1612)		Short-term Migrant(N=1469)		Non Migrant (N=47855)	
	%	Mean	%	Mean	%	Mean
Age in 2004	-	21.6	-	26.7	-	31.4
Total Household Income/Number of Household Members	-	4521.6	-	4216.2	-	5240.3
Education(years of schooling)	-	8.9	-	8.4	-	7.8
From Rural Area	85	-	86.4	-	76.1	-
Health Insurance	31.7	-	28.6	-	33.6	-
Male	55.5	-	67.7	-	49.1	-
Relationship to Head						
Head	1.9		23.55	-	25.8	-
Wife/husband	2.2		8.9	-	24.66	-
Child	89.7		65.2	-	44.9	-
Marital Status						
Single	86.9		58.9	-	36.8	-

Table 3. Occupations of Long-term Migrants, Short-term Migrants and Non-Migrants

Description		Long-Term Migrants		Short-Term Migrants		Non-Migrants	
		N	%	N	%	N	%
High Level	Leaders, Top-Level professionals	21	1.3	28	1.9	1449	3.0
Mid-Level	Mid-Level Professionals, White-collar personnel	35	2.2	33	2.2	1910	4.0
Manufacturing	Assemblers and Machine Operators	29	1.8	36	2.4	899	1.9
Skilled Workers	Salesmen, Skilled Worker in Agriculture, Handicraftsmen	232	14.4	235	16.0	5844	12.2
Unskilled Workers	Sale and Service unskilled workers	95	12.5	100	11.4	4067	14.3
	Unskilled workers in agriculture, sylviculture, and aquaculture	512	67.7	583	66.2	20815	72.9
	Unskilled workers in mining, construction, manufacturing and transportation .	149	19.8	198	22.4	3659	12.8
	Total	756	46.9	881	60.0	28541	59.6
No Work Experience missing	Never worked before	533	33.1	256	17.4	9183	19.2
Total		1612	100	1469	100	47855	100

Table 4. Summary Table of Household Income in 2004

	Mean	Median	SD
Per Capita Income	5187	3928	5680
Log(Per Capita Income)	8.247	8.276	0.784

VND 000s. Exchange rate VND/USD in 2003: 1VND=0.000067 USD; Average per capita of 5,063,000 VND = 339 USD.

Table 5. Reference Group 1 and 2 Regression Results: Total Migrants

Variable	District as Reference Group				Income Stratified Reference Groups			
	(1)	(2)	(3)	(4)	Lower 25% (0<x≤2362)	25%<x<50% (2362<x≤3928)	50%<x<75% (3928<x≤ 6330)	Upper 25% (6330≤x)
RD Original Yitzhaki Index	-0.166*** (.026)	-0.188*** (.026)	-0.027 (.055)	-0.072 (.063)	.093 (.110)	.080 (.091)	.131* (.079)	-.020 (.020)
	.846	.827	.973	0.888§	1.297§	1.206§	1.544§	0.642§
Modified Yitzhaki Index				-.241 (.288)	.654 (.745)	2.693 (2.409)	6.210** (3.255)	-1.221 (1.237)
				0.917§	1.240§	1.239§	1.542§	0.726§
Male	.359*** (.047)	.358*** (.047)	.387*** (.050)	.387*** (.050)	.367*** (.113)	.306*** (.109)	.448*** (.113)	.582*** (.115)
	1.433	1.430	1.473	1.474	1.443	1.358	1.566	1.789
Age	.143*** (.015)	.143*** (.015)	.180*** (.015)	.180*** (.015)	.133*** (.036)	.179*** (.033)	.224*** (.040)	.232** (.050)
	1.154	1.154	1.197	1.198	1.142	1.196	1.251	1.261
Age ²	-.002*** (.000)	-.002*** (.000)	-.003*** (.000)	-.003*** (.000)	-.002*** (.000)	-.003*** (.000)	-.003*** (.000)	-.003*** (.000)
	.997	.997	.996	.996	.997	.996	.996	.996
Education	.069*** (.010)	.068*** (.010)	.038*** (.009)	.039*** (.009)	.057*** (.022)	.036* (.022)	-.019 (0.025)	.107*** (.033)
	1.072	1.071	1.039	1.040	1.059	1.037	.980	1.114
Single	1.278*** (.079)	1.282*** (.079)	1.336*** (.081)	1.335*** (.081)	1.292*** (.191)	1.422*** (.162)	1.393*** (.205)	1.019** (.285)
	3.591	3.605	3.804	3.802	3.643	4.148	4.029	2.772
Size of Household	-.023 (.016)	-.021 (.016)	.050*** (.017)	.049*** (.017)	.045 (.034)	.053 (.041)	.091 (.045)	.126** (.050)
	.976	.978	1.051	1.050	1.046	1.055	1.095	1.134
Insurance	-.426*** (.063)	-.433*** (.063)	-.248*** (.058)	-.243*** (.058)	-.178 (.133)	-.290** (.128)	-.040* (.145)	-.363** (.149)
	.653	.648	.780	.783	.836	.747	.960	.694
log(per capita income)	-.387*** (.043)	-.380*** (.042)	-.124 (.081)	.636 (.447)	.612 (.766)	1.432 (1.510)	3.797* (1.862)	-1.112* (.765)
	.678	.683	.883	1.890	1.346§	1.235§	1.693§	0.565§
log(per capita income) ²				-.051* (.029)				
				.949				
Gini Index		1.547*** (.598)						
		1.114§						
Control for district fixed Effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Account for Cluster Robust Standard Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint Significance Test(Chi-sq)	-	-	-	-	<.0001	<.0001	<.0001	<.0001
Pseudo R-Square	0.100	0.102	0.107	0.107	0.112	0.115	0.115	0.118
N	50932	50932	50932	50932	10010	10191	10272	10478
Number of Migrants	3081	3081	3081	3081	742	776	644	419

*significant at 10%, ** significant at 5%, *** significant at 1%

The figures in the second row of each variable represents odds ratios

§: One Standard Deviation Change

Table 6. Reference Group 1 and 2 Regression Results: Long-Term Out-Migrants

Variable	District as Reference Group				Income Stratified Reference Groups			
	(1)	(2)	(3)	(4)	Lower 25% (0<x≤2362)	25%<x<50% (2362<x≤3928)	50%<x<75% (3928<x≤ 6330)	Upper 25% (6330≤x)
RD Original Yitzhaki Index	-.091*** (.029)	-.117*** (.030)	.047 (.073)	-.003 (.083)	.330*** (.121)	.090 (.122)	.260** (.108)	-.014 (.029)
	.912	.889	1.048	0.995§	2.527§	1.236§	2.385§	0.656§
Modified Yitzhaki Index				.182 (.426)	.887 (.888)	2.865 (3.214)	11.424*** (4.450)	-1.568 (1.621)
				1.070§	1.339§	1.260§	2.239§	0.771§
Male	.019 (.054)	.016 (.054)	.020 (.054)	.021 (.054)	-.192 (.140)	.006 (.133)	-.016 (.125)	.309** (.156)
Age	1.019 .369*** (.040)	1.016 .367*** (.039)	1.020 .407*** (.041)	1.021 .408*** (.041)	.825 .457*** (.104)	1.358 .407*** (.097)	.983 .427*** (.091)	1.364 .523*** (.129)
Age ²	1.446 -.007*** (.000)	1.444 -.007*** (.000)	1.503 -.008*** (.000)	1.504 -.008*** (.000)	1.579 -.009*** (.002)	1.196 -.008*** (.002)	1.525 -.007*** (.001)	1.687 -.009*** (.002)
Education	.992 .066*** (.013)	.992 .066*** (.013)	.991 .045*** (.013)	.991 .046*** (.013)	.990 .063* (.035)	.996 .031 (.031)	.992 .018 (0.037)	.990 .077* (.046)
Single	1.069 1.658*** (.124)	1.068 1.663*** (.124)	1.046 1.696*** (.127)	1.047 1.694*** (.128)	1.065 1.476*** (.309)	1.037 1.551*** (.322)	1.018 1.783*** (.297)	1.081 1.238*** (.383)
Size of Household	5.248 -.016** (.021)	5.276 -.014 (.021)	5.453 .049*** (.023)	5.441 .048*** (.024)	4.377 -.001 (.063)	4.148 .041 (.059)	5.948 .170*** (.060)	3.437 .167*** (.064)
Insurance	.983 -.368*** (.073)	.985 -.375*** (.074)	1.050 -.214*** (.075)	1.049 -.208*** (.075)	.998 -.159 (.228)	1.055 -.336** (.165)	1.186 .166 (.197)	1.179 -.397** (.197)
log(per capita income)	.691 -.282*** (.055)	.686 -.275*** (.054)	.806 .007 (.118)	.811 1.182** (.612)	.852 2.188*** (.809)	.747 2.304 (1.835)	1.181 7.204*** (2.519)	.673 -1.207 (1.124)
log(per capita income) ²				-.076** (.038)				
Gini Index		1.809** (.754)		.925				
Control for district fixed Effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Account for Cluster Robust Standard Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-Square	0.178	0.146	0.178	0.178	0.211	0.201	0.196	0.210
N	50932	50932	50932	50932	10010	10191	10272	10478
Number of Migrants	1612	1612	1612	1612	321	429	356	256

*significant at 10%, ** significant at 5%, *** significant at 1%
The figures in the second row of each variable represents odds ratios
§: One Standard Deviation Change

Table 7. Reference Group 1 and 2 Regression Results: Short-Term Out-Migrants

Variable	District as Reference Group				Income Stratified Reference Groups				
	(1)	(2)	(3)	(4)	Lower 25% (0<x≤2362)	25%<x<50% (2362<X≤3928)	50%<x<75% (3928<x≤ 6330)	Upper 25% (6330≤x)	
RD	Original Yitzhaki Index	-.242***	-.258***	-.072	-.113	.093	.029	-.014	-.029
		(.038)	(.039)	(.073)	(.088)	(.110)	(.128)	(.105)	(.023)
		.784	.772	.973	0.844§	0.817§	1.073§	0.952§	0.438§
	Modified Yitzhaki Index				-.380	.532	1.539	.379	-1.221
					(.353)	(.975)	(3.213)	(4.399)	(1.237)
					0.952§	1.192§	1.131§	1.025§	0.853§
	Male	.683***	.683***	.728***	.727***	.685***	.599***	.950***	.946***
		(.061)	(.061)	(.068)	(.068)	(.157)	(.137)	(.168)	(.197)
		1.981	1.979	1.473	2.070	1.984	1.821	2.587	2.576
	Age	.058***	.058***	.094***	.095***	.086**	.149***	.102**	.101
		(.016)	(.016)	(.017)	(.017)	(.043)	(.042)	(.045)	(.062)
		1.060	1.060	1.197	1.099	1.089	1.160	1.107	1.107
	Age²	-.001***	-.001***	-.001***	-.001***	-.001**	-.002***	-.001***	-.001*
		(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
		.998	.998	.996	.998	.998	.997	.998	.998
	Education	.057***	.056***	.017***	.018	.037	.025	-.069**	.126**
		(.014)	(.014)	(.013)	(.013)	(.023)	(.027)	(.031)	(.051)
		1.058	1.057	1.039	1.018	1.037	1.025	.932	1.135
	Single	.672***	.673***	.712***	.712***	1.003***	.948***	.679***	.288
		(.107)	(.108)	(.108)	(.108)	(.251)	(.209)	(.249)	(.350)
		1.958	1.961	3.804	2.039	2.727	2.582	1.972	1.334
	Size of Household	-.044**	-.043**	.023***	.022	.060	.033	-.065	.008**
		(.021)	(.022)	(.022)	(.022)	(.043)	(.053)	(.056)	(.084)
		.956	.957	1.051	1.023	1.062	1.034	.936	1.008
	Insurance	-.432***	-.438***	-.238***	-.236***	-.177	-.194	-.292	-.231
		(.086)	(.086)	(.077)	(.077)	(.142)	(.185)	(.212)	(.199)
		.648	.645	.780	.789	.837	.823	.746	.793
	log(per capita income)	-.476***	-.471***	-.234**	.213	-.471	.026	.015	-1.004
		(.052)	(.050)	(.098)	(.490)	(.886)	(2.074)	(2.578)	(.920)
		.621	.624	.883	1.238	0.796§	1.005§	1.000§	0.671§
	log(per capita income)²				-.031*				
					(.033)				
					.968				
	Gini Index		1.042						
			(.808)						
			1.075§						
	Control for district fixed Effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
	Account for Cluster Robust Standard Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Pseudo R-Square	0.056	0.056	0.049	0.049	0.066	0.048	0.071	0.057
	N	50932	50932	50932	50932	10010	10191	10272	10478
	Number of Migrants	1469	1469	1469	1469	421	347	288	256

*significant at 10%, ** significant at 5%, *** significant at 1%
The figures in the second row of each variable represents odds ratios
§: One Standard Deviation Change

Table 8. Reference Group 1 and 2 Regression Results By Age Groups: Long-Term Out-Migrants

Variable	District as Reference Group			Income Stratified Reference Groups(AGE 20-39)				
	Teens	20s	30s	Lower 25% (0<x≤2362)	25%<x<50% (2362<x≤3928)	50%<x<75% (3928<x≤ 6330)	Upper 25% (6330≤x)	
RD	Original	-.048	.125	.039	.569***	.287*	.496***	-.020
	Yitzhaki	(.149)	(.108)	(.360)	(.213)	(.173)	(.133)	(.026)
	Index	0.924§	1.229§	1.101§	5.212§	1.885§	5.800§	0.528§
	Modified	-.195	.904	.904	.820	8.013*	21.743***	-1.168
	Yitzhaki	(.574)	(.610)	(1.509)	(1.675)	(4.700)	(5.475)	(2.039)
	Index	0.931§	1.380§	1.110§	1.336§	1.861§	5.055§	0.817§
log(per capita income)		.516	2.722**	7.759*	4.648***	5.980**	12.318***	-1.713
		(.846)	(1.066)	(4.169)	(1.477)	(3.060)	(3.209)	(1.081)
					11.573§	2.321§	6.220§	0.481§
log(per capita² inc)		-.058	-.149 **	-.463*				
		(.061)	(.063)	(.262)				
Control for other Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Control for district fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Account for Cluster Robust Standard Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Pseudo R-Square	0.055	0.084	0.201	0.272	0.296	0.236	0.161	
N	663	808	103	135	184	218	190	
Number of Migrants	12113	12208	11656	4716	4673	4785	4939	

*significant at 10%, ** significant at 5%, *** significant at 1%

The figures in the second row of each variable represents odds ratios

§: One Standard Deviation Change

Table 9. Reference Group 1 and 2 Regression Results By Age Groups: Short-Term Out-Migrants

Variable	District as Reference Group			Income Stratified Reference Groups			
	Teens	20s	30s	Lower 25% (0<x≤2362)	25%<x<50% (2362<X≤3928)	50%<x<75% (3928<x≤ 6330)	Upper 25% (6330≤x)
RD Original Yitzhaki Index	.018 (.166)	-.040 (.163)	-.244 (.170)	.040 (.151)	.006 (.205)	-.145 (.155)	-.039 (.035)
	1.032	0.943	0.654	1.122§	1.018§	0.612§	0.302§
Modified Yitzhaki Index	-.634 (.803)	-.247 (.728)	-.763 (.736)	-2.044 (1.207)	1.556 (5.136)	-5.350 (6.249)	-3.511 (2.877)
	0.783	0.923	0.732	0.481§	1.141§	0.683§	0.546§
log(per capita income)	-.654 (1.401)	-.225 (.996)	1.131 (1.208)	.372 (.954)	.662 (3.112)	-3.250 (3.929)	-1.209 (1.375)
				1.209§	1.110§	0.631§	0.595§
log(per capita ² inc)	.022 (.092)	.004 (.065)	-.088 (.079)				
Control for other Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control for district fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Account for Cluster Robust Standard Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-Square	0.037	0.057	0.054	0.117	0.057	0.083	0.043
N	488	472	271	101	94	100	68
Number of Migrants	12113	12208	11656	2199	2330	2726	2815

*significant at 10%, ** significant at 5%, *** significant at 1%

The figures in the second row of each variable represents odds ratios

§: One Standard Deviation Change

Table 10. Reference Group 3 Regression Results: Long-Term Out-Migrants

Variable	Age			
	Teens	20s	30s	20s and 30s
RD Original Yitzhaki Index	-.001 (.017)	.002 (.012)	-.032 (.041)	.001 (.011)
	0.988	1.039	0.561	0.996
Modified Yitzhaki Index	-.299 (.762)	.397 (.675)	-.695 (1.697)	.151 (.591)
	0.901	1.159	0.780	1.060
log(per capita income)	1.071 (1.312)	1.624 (1.070)	7.035 (4.957)	2.827*** (1.070)
log(per capita income) ²	-.086 (.085)	-.092 (.064)	-.452 (.314)	-.167*** (.065)
Control for other Covariates	Yes	Yes	Yes	Yes
Control for district fixed Effects	Yes	Yes	Yes	Yes
Account for Cluster Robust Standard Errors	Yes	Yes	Yes	Yes
Pseudo R-Square	0.068	0.113	0.222	0.198
N	8750	10585	10772	22939
Number of Migrants	456	615	88	847

*significant at 10%, ** significant at 5%, *** significant at 1%

The figures in the second row of each variable represents odds ratios

§: One Standard Deviation Change

Table 11. Reference Group 4 Regression Results: Long-Term Out-Migrants

		Age (20-39), Education		
		Primary or None	Lower Secondary	Upper Secondary
RD	Original Yitzhaki Index	.048 (.041)	.002 (.027)	.025* (.013)
		1.862	1.036	1.790
	Modified Yitzhaki Index	2.841* (1.659)	.566 (1.124)	1.551* (.887)
		2.793	1.242	1.749
log(per capita income)		1.528 * (2.868)	3.242 (2.107)	2.267 (1.458)
log(per capita income) ²		-.048 (.176)	-.193 (.134)	-.104 (.087)
Control for other Covariates		Yes	Yes	Yes
Control for district fixed Effects		Yes	Yes	Yes
Account for Cluster Robust Standard Errors		Yes	Yes	Yes
Pseudo R-Square		0.203	0.244	0.178
N		4971	9307	5813
Number of Migrants		116	265	386

*significant at 10%, ** significant at 5%, *** significant at 1%

The figures in the second row of each variable represents odds ratios

§: One Standard Deviation Change

Table 12. Reference Group 5 Regression Results: Long-Term Out-Migrants

		Age(20-39), income			
		Lower 25% (0<x≤2362)	25%<x<50% (2362<x≤3928)	50%<x<75% (3928<x≤ 6330)	Upper 25% (6330≤x)
RD	Original Yitzhaki Index	.597** (.242)	.423** (.210)	.302*** (.126)	-.017 (.032)
		5.512	2.776	2.881	0.581
	Modified Yitzhaki Index	1.807 (1.826)	10.672** (5.492)	13.169*** (5.267)	-1.507 (2.188)
		1.985	2.417	2.635	0.778
log(per capita income)		4.878*** (1.769)	6.532** (3.175)	7.762*** (2.904)	-1.453 (1.135)
		13.293	2.758	3.189	0.534
Control for other Covariates		Yes	Yes	Yes	Yes
Control for district fixed Effects		Yes	Yes	Yes	Yes
Account for Cluster Robust Standard Errors		Yes	Yes	Yes	Yes
Pseudo R-Square		0.269	0.298	0.205	0.134
N		4021	3848	3772	4114
Number of Migrants		119	148	165	141

*significant at 10%, ** significant at 5%, *** significant at 1%

The figures in the second row of each variable represents odds ratios

§: One Standard Deviation Change