Malaysian Journal of Economic Studies 53(2): 195–209, 2016 ISSN 1511-4554

Internal Migration in the Context of Trade Liberalisation in Vietnam

Binh Bui Quang[a]
Thi Thu Ha Nguyen[b]
Chuong Ong Nguyen[c]
The University of Danang, University of Economics

Abstract: This study aims to examine the determinants of internal migration in the context of trade liberalisation in Vietnam. The methods of instrumental variables (IV) and two stage least squares (2SLS) estimations were employed on panel data on gross province-to-province migration flows from 2002-2005 to solve the problem of endogeneity of explanatory variables. The empirical analyses showed that although overall internal migration was medium in Vietnam in this period, it did depend on distance, density of population, income (GDP per capita), employment (working time rate), quality of labour (rate of unskilled workers), living conditions (hospital beds per capita) and degree of openness.

Keywords: Gravity model, internal migration, inter-provincial migration, trade liberalisation, two stage least squares
JEL classification: O15; O24.

1. Introduction

Vietnam has experienced great changes and has also experienced many socio-economic achievements, particularly sustained high levels of economic growth and remarkable poverty reduction rates. Some of these changes have been associated with the shifts in population distribution and labour mobility.

Migration across provinces in Vietnam, like other countries, is a socio-economic phenomenon which shows distinct patterns in the course of economic development. During the country's process of transformation into a market economy, migration has been taking place at a higher pace due to unequal economic growth between provinces. Although Vietnam has implemented a household registration system with the explicit aim of directly managing population distribution, rural-urban migration, largely temporary or unregistered, is currently the main factor contributing to urbanisation and industrialisation.

Since the economic renovation known as *Doi moi*, the spontaneous migration flows have notably increased in both rural-urban and rural-rural directions. In the period 1991-1999, there were more than 1.2 million people and 550,000 labourers moving from the North to the Mekong River Delta region, the South-east and the Central Highlands. The South-east and Central Highlands regions experienced the highest rates of urban migration from other regions (more than 8%).

[a] Faculty of Economics, The University of Danang, University of Economics, 71 Ngu Hanh Son District, Danang, Vietnam. Email: *binhbq@due.edu.vn*

[b] Faculty of Economics, The University of Danang, University of Economics, 71 Ngu Hanh Son District, Danang, Vietnam. Email: *ha.ntt@due.edu.vn, nguyenthithuhaktdn@gmail.com* (Corresponding author)

[c] Faculty of Economics, The University of Danang, University of Economics, 71 Ngu Hanh Son District, Danang, Vietnam. Email: *chuongon@due.edu.vn*

Table 1. Migration by regions in the period 2002-2005

Region	In-migration (Person)	In-migration rate (%)	Out-migration (Person)	Out-migration rate (%)	Difference (Person)	Net migration rate (%)
Red River Delta	416,292	23.65	433,330	24.62	-17,038	-0.97
North East	118,108	12.89	158,538	17.30	-40,430	-4.41
North West	24,417	10.01	26,338	10.80	-1,921	-0.79
North Central Coast	123,958	11.95	208,172	20.07	-84,214	-8.12
South Central Coast	84,572	12.30	148,247	21.56	-63,675	-9.26
Central Highlands	101,055	22.12	114,930	25.16	-13,875	-3.04
South East	636,550	49.20	284,581	22.00	351,969	27.20
Mekong River Delta	212,265	12.60	343,081	20.36	-130,816	-7.76
Total	1,717,217	21.25	1,717,217	21.25	0	0.00

Source: Authors' calculation using General Statistical Office (GSO) data of Vietnam

In recent years, industrialisation and open door policies which have had an impact on income, living standards, employment opportunities and ability to access socio-economic services, induced great internal migration flows. During the 2000-2005 period, the migration flows across provinces took place at a higher pace and took on a different pattern.

As indicated in Table 1, in the 2002-2005 period, the migration flows remained large. In terms of in-migration, the South East had the highest rates of in-migration, followed by Red River Delta and Central Highlands, with Ho Chi Minh City, Ha Noi, Binh Duong and Dak Lak being the best destinations for migrants (accounting for 46.57% total in-migration). Of the 8 regions, only the South East had a net inflow from other regions (351,969 persons equivalent to 27.2%), with Ho Chi Minh City and Binh Duong being the two highest net migration provinces (both in terms of scale and rate - 319,209 persons and 58,800 persons corresponding to 259.1% and 44.89% respectively). This can be attributed to the high economic growth rate, high pace of urbanisation and the development of industrial zones; in fact the South East region requires more labour and is currently the most attractive destination for the migrants due to its dynamic development. In contrast, the Central Coasts experienced the highest net out-migration rate (the figures for North Central and the South Central Coasts were 8.12% and 9.26% respectively), with Thanh Hoa and Quang Nam being the two highest net out-migration provinces (43.38% and 19.43% respectively). Income differences provide a strong incentive for people to move out of their homelands.

Tables 2 shows migration trends by region in terms of gross migration in 2002-2005. The out and in-migration flows of the Red River Delta, South East and Mekong River Delta are the largest in terms of scale and rate. This demonstrates that the people living in these regions have a higher tendency to move; hence, migration is highly correlated to the levels of economic development of regions.

Migration drives up the reallocation of population and labour, hence it contributes to economic development in regions; however, the huge flows of rural-urban migration can get out of control resulting in increasing pressure on infrastructure and services in the cities. Therefore, migration has both benefits and costs that have to be considered and linked to the wider context of national and regional development strategies.

Table 2. Gross migration rates from region to region in the period 2002-2005 (%)

	Red River Delta	North East	North West	North Central Coast	South Central Coast	Central Highlands	South East	Mekong River Delta
Red River Delta	37.78	11.36	7.31	8.68	1.64	7.89	13.49	0.68
North East	21.82	12.28	3.26	2.39	1.14	6.56	4.66	0.15
North West	16.82	3.90	16.84	2.54	0.55	3.28	1.84	0.14
North Central Coast	8.03	1.15	1.02	20.63	5.67	15.74	21.57	0.57
South Central Coast	2.28	0.83	0.33	8.56	22.00	18.46	28.57	0.98
Central Highlands	12.46	5.39	2.25	26.85	20.87	19.82	19.25	1.06
South East	13.35	2.40	0.79	23.07	20.25	12.05	45.69	22.79
Mekong River Delta	0.81	0.09	0.07	0.73	0.84	0.80	27.36	23.02

Source: Authors' calculation using GSO data of Vietnam

This paper is organised as follows. In the next section, a selected review of the literature on migration and the determinants of migration in countries are given. This is followed by the formulation of a statistical model that is to be estimated. Theoretical underpinnings for the inclusion of explanatory variables are included in this section. Statistical results are reported in the subsequent section. A final section offers concluding remarks as well as policy implications.

2. Literature Review

The earliest migration models are rooted in the theory of Ravenstein (1885; 1889). It is concerned with the relation between distance and the propensity to move. He proposed 11 laws of migration based on the observation of migration patterns in Great Britain and, later, the United States. Ravenstein's Laws of Migration explain that although most migrants travel short distances, longer-distance migrants prefer to go to centres of commerce or industry; each stream of migration produces a counter-stream; large towns owe more of their growth to migration than to natural increase; the volume of migration increases with the development of industry and commerce and as transportation improves; most migration is from agricultural areas to centers of commerce and industry; and the main causes of migration are economic.

Migration theories were initially based on the dual economy theory introduced by Lewis (1954). In the Lewis's model, subsistence areas refer to rural - the agricultural sector where the labour force suffers from unemployment and underemployment, and modernised areas refer to the industrial sector where many employment opportunities are generated and suffer from a labour shortage. In the subsistence sector, the marginal productivity of labour is very low and employees are paid subsistence wages, thus wages in this sector are usually lower than the marginal products they create; while wage rates in the industrial sector exceed these rates. The difference in wage stimulates migration from the subsistence to the modern urban sector. However, wage-driven migration models only explain a small part of changes in migration flows and do not work in the context of high proportions of urban unemployment. Kirkpatrick and Barrientos (2004) point out that

Lewis's open economy model the capitalists can avoid the turning point by encouraging immigration or by exporting their capital to countries where there is still abundant labour at a subsistence wage. In addition, in Lewis's model, institutional factors play an important role in explaining increased wages in urban areas in some developing countries.

One important strand of the theory of migration was developed by Sjaastad (1962), and then later by Borjas (1990) and Borjas, Freeman and Katz (1991). It is considered as the standard neoclassical model of migration decision-making. Sjaastad (1962) explained that under perfect information, a migration decision is made based on how an individual perceives the expected value of migration. Within the cost-benefit model, the individual compares the current value of the expected present and future costs and returns of migration. The returns of migration include monetary returns which refer to changes in labour incomes, and non-monetary returns which relates to facilities, for example, a good climate and a favourable environment. The costs of migration comprise (i) monetary cost, for example, travel expenditure and costs for seeking a new job, (ii) non-monetary costs, for example, foregone incomes (opportunity costs) and (iii) psychological costs that stem from an unfamiliar environment and a different culture. Migration would be of value if the expected returns exceed the costs of migration. However, Faist (2000) denotes that even if wage or income levels between origin and destination constitute the most important factor, the reasoning is faulty if the case of North African/Turkish-West European migration in the 1960s is considered.

In addition, there have been several contributions to the field of migration since the Harris-Todaro model (1970). This original model focuses on the situation of a developing country in which the migration flow is from rural to urban areas, but where both inter-regional and international migration exist. The neo-classical model of Harris and Todaro also explains that geographic differences in the supply and demand for labour cause (international) migration. The differences in wage and employment are the key factors driving up the migration between regions. Regions with a limited supply of labour in relation to capital will usually have a higher wage that will attract a large inflow of labour from low wage regions. According to Harris and Todaro (1970), the migration decision incorporates awareness of the potential migrant of a potentially higher urban wage which would give them a better earning, and thus a better life. Consequently, the migrants are willing to be unemployed or underpaid to wait for a better job opportunity that gives them good compensation in the future. It means that there usually exists a situation where there is a high flow of migrants from rural to urban areas, but with migrants ending up unemployed. However, a major limitation of Harris-Todaro's model is its assumption that potential migrants are homogenous in respect to skills and attitudes and have sufficient information to work out the probability of looking for a job in an urban area; furthermore, the assumption on expected incomes is also unrealistic in that the migrants are perceived as having enough information to project and make a decision.

Other authors, such as Mincer (1978) and Stark (1992) considered migration as a household decision rather than an individual decision and developed the theory of new economics of migration. This theory focuses on the household and views migration as a collective strategy to ensure the capability of the domestic individual through the allocation of their resources. Besides geographic and economic factors, social factors such as family characteristics and family ties have to be taken into account in migration decisions. Mincer (1987) states the family characteristics relevant to migration decisions affect the

probability of migration. In addition, Stark (1992) explains that the migration decision is a portfolio strategy of potential migrants' household that attempts to diversify the sources of income. Nevertheless, Faist (2000) argues that offering rational choice as the solution to extend rationality beyond individual actors to collective units is not sufficient. The main problem with social choice approaches to migration has been that they are only concerned with transactions between the collective units and not with what goes on within them; it does not consider the social attributiveness of potential migrants' behaviour.

Given that the Harris-Todaro model accounts for a macro-economic model, Borjas (1990) and Borjas et al.(1991) introduced a micro-economic model to explain migration. In this model, they described that migrants compare the cost of moving to alternative locations while maximising the expected discounted net return over a certain time horizon. He figured out that results in the macro-economic model are slightly different from the micro-economic model. More specific, in the macro model, the differences in wages and unemployment rate determine the movement of labour among regions, while the expected net return of migration is primarily the key factors related to a household's decision to move to a new region. Another difference is that in the micro model, high-skilled workers are expected to move more than general workers. However, Huddle (1992) criticised that the conceptual approach in Borjas' model looks simple and straightforward and the econometric methodology itself is somewhat weak. He pointed out serious flaws in the analyses of Borjas' typification of the econometric evidence on job displacement and wage depression. In particular, the average immigrant, especially the illegal, most probably constitutes a drain and not a gain for the unskilled native worker in the economy.

A modification of the Harris-Todaro model has been suggested by Pissarides and McMaster (1990) to explain net migration defined as total inflows in a region less total outflows from a region. Pissarides and McMaster (1990) state that the gain from migration depends on relative wages and regional unemployment rates. They also explained that migration is also affected by regional unemployment in two ways. First, because unemployed workers have less to lose than the employed workers, they are more likely to move out than the employed ones. Regional unemployment differentials approximate the differences in employment probabilities in regions. The Harris-Todaro's framework indicates that expected income depends on the possibility of employment, which is based on the assumption that the individuals are risk neutral and not quantity constrained and both relative wages and unemployment differentials can be converted to a single 'expected income' variable. Second, aggregate unemployment may impact on the gain from migration: if the unemployment rate is high, employed workers may find security in their jobs and if they have achieved seniority rights in their current jobs, their gains from moving might be reduced; net migration is influenced by unemployment duration; unemployed workers may be discouraged from moving if they face liquidity constraints, therefore the gains from moving might be reduced.

Massey et al. (1993) extended Borjas's micro-economic model by focusing on the household as the unit of decision-making. The households' risk of a decrease in income can be diminished by diversifying the allocation of resources such as family labour. Massey et al. (1993) established that if the regional differences in wages or unemployment rate no longer exist, migration still takes place. Furthermore, as mentioned in the network theory, social factors are related to the migration decision. They explained that migration networks include sets of interpersonal relations owing to kinship, friendship or shared

communities of origin, which help to increase the migration of labour between regions as the networks decrease the asymmetry in information among employees living in the destination and other regions; these networks diminish the costs and risks of migration.

Other migration theories have been considered in macro approaches. The political, economic and cultural structures denote an array of factors in the migration processes. Faist (2000) reviews these factors and states that these have a considerable degree of variation in affecting migration decisions such as international power, administrative capacity, efficiency and political stability. The fundamental factor of migration is the differences in economic condition between origin and destination areas, such as wages, living standards, jobs, working conditions, and unemployment rates.

Tinbergen (1962) introduced the gravity model to economics by using the equation to explain international trade flows. Thereafter, the equation has been employed to a wide range of social interactions, for example, migration, tourism and foreign direct investment (FDI).

Lee (1966) assumed that both the origin and the destination have characteristics that attract or resist migrants and the perceptions of these characteristics differ among migrants. Poor economic conditions at the origin are the main push factors in emigration, whereas the possibility of improved economic conditions at the destination is a critical pull factor in immigration.

Later, migration models took on greater complexity when studies were carried out to deal with questions related to migration flows. Lowry (1966) also established the migration model which explains that migration from *i* to *j* is directly related to high wages at *j*, low relative unemployment at *j* and a large civilian labour force at either origin and/or destination. It is inversely related to high wages at *i*, low unemployment at *i* and increasing distance between *i* and *j*. His model is instituted by the equation:

$$M_{ij} = k \left[\frac{U_i}{U_j} \times \frac{W_i}{W_j} \times \frac{L_i L_j}{D_{ij}} \right] e_{ij} \tag{1}$$

where M is number of migrants, L is the number of person in the labour force, U is the unemployment rate, W is the hourly wage in manufacturing, D is the distance between provinces, k is the gravitational constant, and e is the error term.

The study shows that a place experiencing economic stagnation is likely to have a high rate of out-migration, with the highest proportion being young adults entering the labour force, while a place experiencing high economic growth is likely to attract in-migrants from the whole country (Lowry, 1966).

Most studies concerning the determinants of migration usually use the linear regression model or the log-linear regression model, in which all variables of a linear model are transformed into the logarithmic form. There are many alternative specifications of the gravity model.

In its standard form, the gravity model of migration is

$$M_{ij} = \ln \beta_0 D_{ij}^{\beta_1} X_i^{\beta_1} X_j^{\beta_1} u_{ij} \tag{2}$$

where, M_{ij} is the gross migration flow from province *i* to province *j*; D_{ij} is the distance between provinces; X_i and X_j are characteristics of province *i* and province *j,* and u_{ij} is the error term.

In the literature, there is a long-standing tradition of log-linearising equation (2) and estimating the parameters by the Ordinary Least Squares method (OLS), using the equation

$$\ln M_{ij} = \ln \beta_0 + \beta_1 \ln D_{ij} + \beta_i \ln X_i + \beta_j \ln X_j + \ln u_{ij} \quad (3)$$

Thus, equation (3) is a constant-elasticity model. Researchers often prefer the log-linear model to the linear model because the relation between variables is not linear. For instance, Greenwood (1969), and Greenwood and Sweetland (1972) employed this method to investigate the determinants of inter-state labour migration in the United States in the period 1955 to 1960. However, this method requires some restrictive assumptions such as the elasticity of migration is constant with any change of an independent variable, of the error terms u_{ij}, and of constant variance (homokedasticity). Therefore, $\ln u_{ij}$, is statistically independent of the regressors. Their findings show that people tend to migrate from low income to high income areas and from areas having low public expenditures to areas providing more public benefits.

Recently, Karemera, Oguledo and Davis (2000) used the gravity model to investigate the impacts of demographic, economic and political factors on the size and composition of migration flows to North America. A modified gravity model was used and was adjusted to include the regulations and characteristics of immigration specific to the origin and destination countries. They found the population of origin countries and the income of destination countries to be the two main determinants of migration to North America; in addition, domestic restrictions on political and civil freedom in origin countries had significant effects on migration to North America.

The study of Hatton and Williamson (2003) provides a quantitative measurement of the demographic and economic fundamentals that have driven international migration, across different historical eras and in the world. They exerted the pooled OLS regression method to investigate how the standard theories of migration perform; how inequality and poverty influence world migration; and the distinction between migration pressure and migration ex-post. They concluded that demographic and economic variables strongly affect migration in the world, but these facts do not reduce the importance of migration policy.

Andrienko and Guriev (2004) also focused on the gravity model to examine the determinants of internal migration in Russia, including economic, political and social factors. Their modified gravity models were specified in a logarithmic form using the fixed-effects and between-effects estimations. They found a significant and positive relationship between the differences in income and unemployment rate of the origin/ destination and the gross migration flows. It meant that people move from poorer and job scarce regions to regions which are richer and better both in terms of employment prospects and public services. In addition, the presence of high government subsidies in social and economic infrastructure in lower economic development areas could decrease incentives for the population to move.

Mayda (2005) empirically investigated economic and non-economic determinants of bilateral immigration flows across the origin and destination countries (fourteen OECD countries) using panel data approach. She focused on both supply and demand determinants of migration patterns by looking at the probability that an individual chosen randomly from the population of country *i* will migrate to country *j*. The probability is proxied by the emigration rate which is established as the function of economic and non-

economic factors; the empirical analysis is based on econometric models using aggregate data. This study found statistical significant and positive impacts on immigrant inflows of improvements in the mean income opportunities in the host country; in addition, there are the various impacts on emigration rates of geographical, cultural, and demographic factors as well as the role played by changes in destination countries' migration policies.

While Peeters (2008) adapted three gravity-model specifications to investigate the characteristics of the inter-state migration in Mexico from 1995 to 2000: log-linear model, basic Poisson model and flexible Poisson model. According to him, the gravity model in levels (Poisson model) provides a more realistic description of the migration process than a model in logs. He also developed a more flexible model by relaxing the standard assumption of spatial invariance (constancy) of the distance-decay parameter, allowing to account for the impacts of unique ties between origin and destination countries. The empirical findings show that inter-state migration in Mexico is inelastic in relation to geographical distance measure after correcting for the impacts of some structural elements of migration, for example, migrant stocks.

There have been some studies that provide the features and determinants of migration in Vietnam. Anh, Goldstein and McNally (1997) employed the Vietnam Population Census data in 1989 to assess how population movements respond to market opportunities as a baseline estimate. This study shows that national reunification resulted in population redistribution and a rural resettlement policy while market reforms and macro structural changes affected labour markets and contributed indirectly to more spontaneous and voluntary spatial mobility. In addition, their findings from the multivariate analysis reveal that more developed provinces attracted more in-migrants; however, the government policy did not stimulate out-migration from sending provinces or cities. Later, Anh et al. (2003) basically focused on the statistical analysis to assess the determinants of internal migration by a making distinction between macro-socio-economic factors and the specific mechanisms (wage and unemployment differentials, etc) through which structural factors operate. They exploited the secondary data sources from the Vietnam Population Census in 1989 and the Population and Housing Census in 1999 covering all 61 provinces in Vietnam (GSO, 2001). They argued that economic factors, for instance, incomes and employment opportunities, have a larger influence than non-economic factors in determining the current migration pattern in Vietnam; in particular, the migration decision reflects the income differences rather than demographic variables (for example, age, gender, marital status).

However, there have been a few econometric studies of inter-regional migration in Vietnam. Sang (2004) used statistical analysis and adapted both log-linear and Box-Cox models in the empirical analysis of urban-ward migration from 1984 to 1989 and from 1994 to 1999. The results from the regressions indicate that, in particular, geographic distance is one of the strongest barriers to inter-provincial urban migration; the proportion of urban population in a province positively affects the urban in-migration rate while land pressures significantly affects out-migration over the two periods. He argued that, in the pre-reform era, migration was regulated by the government; thus the push and the pull factors across provinces had a insignificant impact on the migration flows. However, in the 1990s, the productivity of non-state industries emerged as the major force attracting migrants from rural to urban areas, reflecting the pull factor in urban areas.

Based on several studies in Vietnam, it is noted that the most important factor contributing towards migration in recent times is the economic factor. People move to other

regions because they want to have better job opportunities and a higher income. Places with good living conditions will attract more people than less developed places. In contrast, the costs of migration (such as distance, travel expenses) are migration-limiting factors. Furthermore, macro changes such as industrialisation and urbanisation as well as residential relocation have created gaps in terms of income, employment opportunities and living conditions between provinces or regions in Vietnam. This also contributes to migration within Vietnam.

The ASEAN countries are in the process of establishing a Common Economic Area (CEA). This will create positive conditions for workers moving between the ASEAN countries. There are existing differences in the economic development between countries in the ASEAN. This will generate migration flows from countries with a lower level of economic development such as Vietnam, Laos, Cambodia, Myanmar to countries with the higher level of economic development such as Thailand, Singapore, Malaysia and Indonesia. This will represent major issues for further studies in the ASEAN countries in particular and Asia in general.

3. Empirical Models for the Study

Motivation behind migration flows from origin to destination comes from the attractiveness of the receiving counterpart such as higher income, higher standards of living, more job opportunities, more land, and more enterprises. In Vietnam, a high growth rate alongside increasingly deep integration into the world economy has been followed by resource allocation and movement between the regions in the country.

In this study, the econometric models are built based on the Gravity model. Characteristics of provinces will be analysed in addition to traditional factors affecting migration flows in the literature. Equation (2) is used to formulate the relations between numbers of migrants M_{ij} from province i (leaving province) to province j (receiving province) with characteristics of provinces as follows:

$$M_{ij} = \beta_0 D_{ij}^{\beta_1} \left(\frac{DEN_j}{DEN_i}\right)^{\beta_2} \left(\frac{X_j}{X_i}\right)^{\beta_3} e^{\beta_4 (X_j - X_i)} u_{ij} \tag{4}$$

where, M_{ij} is the gross migration from province i to province j; D_{ij} is the distance between provinces; DEN_i and DEN_j are population density of province i and province j respectively; X_i and X_j are characteristics of province i and province j respectively, and u_{ij} is the error term. Taking logarithm of Equation (4), we have

$$\ln M_{ij} = \ln \beta_0 + \beta_1 \ln D_{ij} + \beta_2 (\ln DEN_j - \ln DEN_i) + \beta_3 (\ln X_j - \ln X_i) + \beta_4 (X_j - X_i) + \ln u_{ij} \tag{5}$$

The differentials in economic factors between leaving and receiving provinces would have significant effects on migration flows. However, in terms of degree of openness, international trade and foreign investment could affect inter-provincial migration through other factors, for instance higher growth, or reallocation of resources, industries and thus workers. Trade could contribute to creating more jobs and widening the gaps of living standards among provinces, hence causing an increase in migration from poorer to richer provinces. In any case, the impact of trade liberalisation on migration would be more indirect. Therefore, taking into consideration endogeneity of the employment factor, the method of instrumental variables (IV) is used to solve the problem:

$$(RWT_j - RWT_i) = \eta_0 + \eta_1 (InLandpcj - InLandpc_i) + \eta_2 (SLabour_j - SLabour_i) \eta_3 (InFDIpc_j - InFDIpc_i) + \eta_4 (InTradepc_j - InTradepc_i) + \nu_{ij} \quad (6a)$$

$$In\ Mij = In\beta_0 + \beta_1 InD_{ij} + \beta_2 (InDEN_j - InDEN_i) + \beta_3 (\ln GDPpc_j - \ln GDPpc_i) + \beta_4 (\ln Hosbed_j - \ln Hosbed_i) + \beta_5 (RWT_j - RWT_i) + \ln u_{ij} \quad (6b)$$

In equation (6a), RWT_i is an instrumental variable that represents the working time rate of province *i* (the proportion of the working time of rural workers which is calculated by the total time rural workers are employed divided by the average working time per year or 2200 hours)*;* ln $Landpc_i$ is the per capita area of agricultural land of province *i* in log; $SLabour_i$ is the proportion of skilled workers of province *i;* ln $FDIpc_i$ is the per capita FDI of province *i* in log; ln $Tradepc_i$ is the per capita trade of province *i* in log; and v_{ij} is the error term. In equation (6b), ln M_{ij} is the gross migration from province *i* to province *j* in log; lnD_{ij} is the distance between province *i* to province *j* in log; ln DEN_i is the population density of province *i* in log; ln $GDPpc_i$ is the per capita GDP of province *i* in log; ln $Hosbedpc_i$ is the number of hospital beds per 1000 person of province *i* in log; WTR_i is the fitted values of the working time rates of province *i* and province *j*; and ln u_{ij} is the error term.

This study uses Vietnam's provincial panel data in the period 2002-2005 to investigate the effects of geographic, demographic and economic factors on internal migration. The provincial level data are extracted mainly from the yearly Population Survey in the period 2002-2005 published by General Statistical Office (GSO) and combined with the Socio-Economic Statistical Data of 64 provinces and cities. This dataset covers only a short time period, because provincial data is lacking. It should be noted that from 1997 to 2003, Vietnam had only 61 provinces but from 2004 it had 64 provinces because some provinces were divided into two (for convenience, all will be referred to as provinces hereafter).

For each pair of provinces, the number of people migrating from one province to another within a year aggregates to a migrant flow which is taken to account for the explained variable. Such data are officially collected and published by GSO. This is the only available dataset on province-to province migration flows in Vietnam and it is found that there are many zero migration observations. In the empirical analysis, given the concern about the large number of missing values, the zero-flows observations are kept in the data set. To acquire a balanced panel data set of 64 provinces, several approaches to data processing have been adopted. First, the values of all observations of Dien Bien, Dak Nong and Hau Giang in 2002 and 2003 are duplicated by those in 2004. Second, migration flows data set include only formal migrants who are officially registered or surveyed; consequently, many observations correspond to truly zero flows, while others are likely to correspond to very small flows that are informal or not surveyed; therefore, a number of observations were randomly generated ranging from 1 to 11 to replace the missing values; subsequently all observations which are 1 will be assigned to 2 in order to change to its logarithm form. Third, zero observations on FDI and trade are also processed in this manner.

4. Empirical Results

Because the analysis data captures the period 2002-2005, pooled data OLS regression could be applied. It is realistic and significant to analyse and to quantify the factors influencing migration flows between the provinces using Gravity model. Several specifications are exploited and combined with sensitivity analysis.

Table 3. First stage regression

Instrumental variables	Instrumented variable: difference of working time rate (%)
Difference of per capita area of agricultural land in log	-0.117***
Difference of rate of unskilled workers (%)	-0.025***
Difference of per capita FDI in log	-0.016*
Difference of per capita trade in log	0.245***
Constant	-0.005
Number of observations	15,384
Adjusted R^2	0.066

Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Taking into account endogeneity problems of employment, first stage regression is employed to examine the factors (instrumental variables) affecting the working time rates between provinces. The regression results in Table 3 shows that the coefficients of the instrument variables are statistically significant.

In addition, the Hausman test shows that the variable of the difference in working time rates between provinces is endogenous (Table 4).

Subsequently, the second stage regression for the structure equation is estimated by replacing the fitted values of difference in working time rate for the actual values; the time dummy variables are incorporated at this stage (Table 5).

From Table 5, several interpretations of estimated results using the method of instrumental variables (IV) and two-stage least squares (2SLS) are made. These interpretations are clarified based on partial effect of the determinants (*ceteris paribus* assumption).

First, the elasticity of migration with respect to distance between provinces is negative and significant, with its absolute value being approximately equal to 0.69. This figure demonstrates that if the distance between provinces is longer than 1 %, it will reduce 0.69% of migration flow. This suggests that, to a certain extent, the longer the distance from province *i* to province *j*, the lower the mobility of migrants in Vietnam.

Second, the variable difference of the density of population in log has a significant negative estimated coefficient (nearly -0.1) to migration from leaving province *i* to receiving province *j* (in log). This means that if the difference of the density of population between receiving and leaving provinces changes 1%, it will diminish 0.1% of migration. This finding shows that the higher the difference in density of population between provinces, the fewer the movement of migrants. This inverse relationship between migration and the density of population indicates that the characteristics of migration between provinces of Vietnam in the period 2002-2005 are different compared to the previous period. In the previous period, migration flows were from 'crowded populated areas' to 'sparsely populated areas'; however, in recent times, the development of many large-scale industrial zones in the South East region (Ho Chi Minh, Binh Duong and Ba Ria-Vung Tau city), Red River Delta (Hanoi), which are the provinces with a high population density, has attracted migrants from other provinces that also have a high population density such as Thanh Hoa, Nam Dinh, Ha Tay, Hai Duong, Thai Binh, Quang Nam. Therefore, the coefficient of

Table 4. Hausman test for endogeneity

Explanatory variables	Dependent variable: Migration from leaving province *i* to receiving province *j* (in logarithms)
Distance between province *i* and province *j* in log	-0.692***
Difference in density of population in log	-0.034
Difference in per capita area of agricultural land in log	0.069
Difference in per capita GDP in log	0.190***
Difference in working time rate (%)	0.123***
Difference in number of hospital beds per 1000 persons in log	0.115***
v_hat	-0.107***
Constant	7.111***
Number of observations	15,384
Adjusted R^2	0.186
H0: vij is uncorelated with uij (v_hat = 0)	
F (1, 15377) = 17.25	
Prob > F = 0.0000 < 0.05 (Reject H0)	

Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

'the population density' variable cannot explain the gravity effect of 'big city' in the case of Vietnam in this period.

These findings are identical to the conclusions from a literature review of the gravity model.

Third, in terms of income, the elasticity of migration regarding the difference of GDP per capita between receiving and leaving provinces is positive and significant (roughly 0.21). This figure points out that if the difference in GDP per capita between provinces increases 1 %, it will lead to 0.21% of migration. This explains the greater need for migrating to the provinces with a higher income.

Fourth, in terms of employment, the elasticity of migration with regard to the variable of the difference in working time rates between provinces as the endogenous variable explained by other factors-instrument variables (such as difference in per capita area of agricultural land, difference in rate of unskilled workers (%), difference in per capita FDI, difference in per capita trade) is positive and significant (0.11). Table 3 shows that in terms of level of openness, international trade positively affects the difference in working time rates (with elasticity being approximately 0.26) and also affects positively inter-provincial migration; however, differences in per capita area of agricultural land and rate of unskilled workers negatively affect the difference in working time rates and then negatively inter-provincial migration. The results explain that trade liberalisation contributes to creating more jobs, therefore causing an increase in migration while lower quality of labour constrains migration flows. These findings show that the higher the chance of having a job at the receiving province, the more attractive the province is to migrants.

Fifth, in terms of living conditions with the proxy being the number of hospital beds per 1000 persons, the elasticity of migration with respect to difference in number of

Table 5. Second stage regression

Explanatory variables	Dependent variable: Migration from leaving province *i* to receiving province *j* (in logarithms)		
	All	Central Highlands	South East
Distance between province *i* and province *j* in log	-0.691***	-0.614***	-0.657***
Difference in density of population in log	-0.103***	-0.241***	-0.138***
Difference in per capita GDP in log	0.205***	0.897***	0.370***
Fitted values of difference in working time rate (%)	0.110***	-0.428***	0.354***
Difference in number of hospital beds per 1000 persons in log	0.084***	-0.697***	0.900***
Year 2003	-0.303***	-0.248	-0.245*
Year 2004	-0.364***	-0.401***	-0.430***
Year 2005	-0.238***	-0.301**	-0.303**
Constant	7.337***	6.827***	7.769***
Number of observations	15,384	1,110	1,968
Adjusted R^2	0.191	0.101	0.195

Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

hospital beds per 1000 persons between receiving and leaving provinces is positive and significant but the value is moderate (0.08). This figure reveals that if the difference in number of hospital beds per 1000 persons between provinces improves by 1 %, it will contribute towards increasing migration flow by 0.08%. This indicates that migrants are likely to move to provinces with better living conditions.

In addition, all estimated coefficients of time dummy variables are negative and significant with the elasticity of migration being the highest in 2002. The values of these coefficients vary and diminish in the period 2002-2005.

Finally, Table 5 specifies that the determinants affect differently migration flows among regions. For instance, for the two highest in-migration regions such as the Central Highlands and the South East, there is a strong significant association between the elasticity of migration with regard to the difference in working time rates for migration flows to the South East while there is a negative significant association for migration flows to the Central Highlands (0.354 and -0.428 respectively). This clarifies the positive effects of trade liberalisation on creating more jobs hence causing an increase in migration flows into the South East. It also explains that the South East region requires more labour due to its dynamic development and therefore is currently the most attractive for migrant arrivals

5. Conclusion and Policy Implications

This study attempts an empirical analysis of the determinants of internal migration in Vietnam by employing the panel dataset of gross migration flows and factors between provinces. The methodologies of instrumental variables (IV) and two-stage least squares

(2SLS) are used to distinguish the effects of geographical, demographical and economical factors. The empirical analyses show that although overall internal migration is medium in Vietnam in the period 2002-2005, it does depend on distance, density of population, income (GDP per capita), employment (working time rate), quality of labour (rate of unskilled workers), living conditions (hospital beds per capita) and level of openness. Some policy implications can be drawn from these conclusions.

Firstly, empirical findings suggest that there is a positive relationship between migration and income. Provinces with higher income attract more migrants. Therefore, increasing the income in 'leaving' provinces will reduce the population and infrastructure pressure on 'receiving' provinces. Difference in income is the key factor explaining an increase in the migration flow between provinces; however, the largest cities in Vietnam such as Ho Chi Minh and Hanoi are facing the pressures of an infrastructure overload and unemployment due to excessive migration from other provinces. Consequently, these large cities have used administrative measures to restrict excessive immigration. Meanwhile, the labour shortage in the rural areas is increasing as a result of massive migration to the big cities.

Second, promoting exports and imports and foreign investment via trade liberalisation and open door policies will create more jobs and higher income thus narrowing the income gap among provinces. This policy implies that in order to lessen the issue of massive migration among provinces, promoting trade liberalisation is an appropriate way of significantly driving inter-province migration flows in Vietnam. In other words, with respect to 'receiving' province, the more jobs that are created, the more the in-migration will be; with regard to 'leaving' province, in order to reduce out-migration flows, the province needs to generate more employment.

Third, besides the need to reduce the pressures of excessive migration, it is necessary to have policies aimed at an efficient allocation of labour with a competitive rate among the provinces. More specifically, with regard to provinces that are facing a high rate of out-migration resulting in a shortage of labour, investing in infrastructure and improving the living conditions in the 'leaving' provinces could help in the reallocation of labour among provinces. Moreover, in order to attract highly qualified migrants, rural areas should develop programs aimed at directly facilitating migrants such as granting land and credit to enterprises which invest in building houses for their migrant workers or in providing information on jobs and housing to migrant workers.

References

Andrienko, Y. & Guriev, S. (2004). Determinants of interregional mobility in Russia: Evidence from panel data. *Economics of Transition, 12*(1), 23–27.

Anh, D.N, Goldstein, S., & McNally, J. (1997). Internal migration and development in Vietnam. *International Migration Review, 31*(2), 26-29.

Borjas, G.J. (1990). *Friends or strangers: The impact of immigrants on the US economy*. New York: Basic Books.

Borjas, G.J, Freeman, R.B., & Katz, L. F. (1991). *On the labour market effects of immigration and trade*. National Bureau of Economic Research Working Paper No. 3761, p.33.

Faist, T. (2000). *A review of dominant theories of international migration. The volume and dynamics of international migration.* Retrieved from http*://fds.oup.com/www.oup.co.uk/pdf/0-19-829391-7.pdf*

Greenwood, M.J. (1969). An analysis of the determinants of geographic labour mobility in the United States. *Review of Economics and Statistics*, *51,* 189-194.

Greenwood, M.J., & Sweetland, D. (1972). The determinants of migration between standard metropolitan areas. *Demography*, *9,* 665-681.

General Statistical Office (GSO) (2000). *Vietnam living standard survey 1997-1998.* Hanoi: Statistical Publishing House.

General Statistical Office (GSO) (2001). *Population and housing census in Vietnam in 1999.* Hanoi: Statistical Publishing House.

General Statistical Office (GSO) (2005). Statistical handbook in 2002, 2003, 2004, 2005 (Summary) Hanoi: Statistical Publishing House.

Harris, J.R., & Todaro, M.P. (1970). Migration, unemployment and development: A two sector analysis. *American Economic Review*, *60,* 12-18.

Hatton, T.J., & Williamson, J.G. (2003). *What fundamentals drive world migration?* WIDER Discussion Paper Discussion Paper No.2003/23, UNU/WIDER 2003. Retrieved from: http*://www.wider.unu.edu/publications/working-papers/discussion papers/2003/en_GB/dp2003-23/.*

Huddle, L. D. (1992). *Migration, jobs and wages: the misuses of econometrics*. Retrieved from *http://dieoff.org/page53.htm.*

Karemera, D., Oguledo, G., & Davis, B. (2000). A Gravity model analysis of international migration to North America. *Applied Economics*, *32*(13), 45-48.

Kirkpatrick, C., & Barrientos, A. (2004). T*he Lewis Model after fifty years.* Development Economics and Public Policy, Working Paper Series 9.

Lee, E. S. (1966). A theory of migration. *Demography 3,* 47–52.

Lewis, A.W. (1954). Economic development with unlimited supplies of labour. *The Manchester School, 22*(2), 139-191.

Lowry, I.S. (1966). *Migration and metropolitan growth: Two analytical models*. San Francisco: Chandler Publishing.

Massey, D.S. *et al*. (1993). Theories of international migration: A review. *Population and Development Review*, *19,* 431-466.

Mayda, A.M. (2005). *International migration: A panel data analysis of economic and non-economic determinants.* Institute for the Study of Labour (IZA), Discussion Paper No. 1590.

Mincer, J. (1978). Family migration decisions. *Journal of Political Economy*, *86,* 749-773

Peeters, L. (2008). *Inter-state migration in Mexico, 1995-2000: A flexible Poisson Gravity-modeling approach*. Retrieved from: *h*ttp*://uhdspace.uhasselt.be/dspace/bitstream/1942/7973/1/Internal%20migration%20in%20Mexico%2015-01-2008.pdf.*

Pissarides, C. A. & McMaster, I. (1990). Regional migration, wages and unemployment: empirical evidence and implications for policy. *Oxford Economic Papers*, *42,* 812-831.

Ravenstein, E.G. (1885). The laws of migration. *Journal of the Royal Statistical Society*, *48,* 36-45.

Ravenstein, E.G. (1889). The laws of migration. *Journal of the Royal Statistical Society*, *52,* 72-78.

Sang, L.T (2004). *Urban migration in pre- and post-reform Vietnam: Macro patterns and determinants of urbanward migration, the 1984-1989 and 1994-1999 periods.* Retrieved from http*://apmrn.usp.ac.fj/conferences/8thAPMRNconference/15a.Urbanward%20Migration%20in%20Vietnam.pdf.*

Sjaastad, l. (1962). The costs and returns of human migration. *Journal of Political Economy*, *70* (Supplement): 80-93.

Stark, O. (1992). *The migration of labour*. Cambridge, Mass: Blackwell Publishers, pp.82-86.

Tinbergen, I. (1962). *Shaping the world economy: suggestion for an international economics policy*. New York: Twentieth Century Fund.