Journal of Development Economics 91 (2010) 100–112

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Inter-provincial migration and inequality during Vietnam's transition[☆]

Diep Phan [a], Ian Coxhead [b,*]

[a] Department of Economics, Beloit College, 700 College Street, Beloit, WI 53511, United States
[b] University of Wisconsin, 413 Taylor Hall, 427 Lorch St., Madison, Wisconsin 53706, United States

ARTICLE INFO

Article history:
Received 9 August 2007
Received in revised form 11 November 2008
Accepted 22 June 2009

JEL classification:
J61
O15
R23

Keywords:
Labor
Migration
Inequality
Vietnam

ABSTRACT

Vietnam's economic boom during the transition to a market economy has centered on very rapid growth in some sectors and provinces, yet poverty has diminished across the entire country. With capital investments highly concentrated by province and sector, geographic labor mobility may be critical in spreading the gains from growth. Conversely, rising income inequality may be attributable in part to impediments to migration. We first use census data to investigate migration patterns and determinants. We then examine the role of migration as an influence on income ratios between pairs of provinces. The former analysis robustly confirms economic motives for migration but also suggests the existence of poverty-related labor immobility at the provincial level. Examination of income ratios between pairs of provinces reveals that the impact of migration on inequality can be either negative or positive. A robust inequality-reducing impact of migration is found for migration flows into provinces where most of Vietnam's trade-oriented industrial investments are located.

1. Introduction

Economic growth in diversified economies is inherently unbalanced, because sectors have different factor intensities and factor endowments grow at differing rates. In the developing world, where capital endowments are typically highly concentrated by location, any new investments, price or productivity shocks, or policy interventions that alter the value marginal products of factor inputs at a sectoral level thus induce a spatial reallocation of the relatively mobile factor, labor. In this way, economic growth and internal migration are complements: growth stimulates migration, and migration facilitates growth.

The growth–migration relationship is a source of many empirical questions with normative and policy implications. In this paper we address two such questions, using data on interprovincial migration from a rapidly-growing low-income economy, Vietnam. First, we inquire into the determinants of inter-provincial migration during an era of rapid growth, testing the extent to which labor flows between provinces can be explained by distance, income of sending and receiving provinces, and past migration. These variables have been found to be significant in other studies, but equally important, an investigation of this type can also yield information about impediments to migration. This is important to our second inquiry, into the links between migration and inter-provincial income inequality.

Since the 1986 adoption of the economic reform package known as doi moi, Vietnam's economy has experienced rapid growth, averaging 7% per year, accompanied by impressive poverty reduction. But growth and poverty reduction have been unevenly distributed, with regions such as the Northern Mountains, North Central Coast, and Central Highlands falling behind.[1] A large share of industrial capital in Vietnam is concentrated in a very small number of urban centers, mainly Ho Chi Minh City (HCMC) and surrounding provinces (Ba Ria-Vung Tau, Dong Nai, and Binh Duong) in the South, and Hanoi, Hai Duong, Hai Phong and Quang Ninh in the North.[2] As a result, labor market adjustments and migration have become vital in spreading the benefits of growth from urban centers to hinterland areas.

Despite the apparent importance of the topic, the literature on labor mobility and migration in Vietnam only started to receive attention very recently. Earlier studies (Guest, 1998; Dang, 2003; Nguyen, 2003) are mostly descriptive and are useful only in giving migrants' characteristics and correlates of migration. More recent studies, including Hoang et al. (forthcoming), deBrauw and Harigaya (2007), ADB (2007), and Phan

☆ The authors thank seminar participants at the University of Wisconsin and the University of Minnesota, and two anonymous JDE reviewers, for helpful comments on previous drafts. Remaining errors are ours alone. This research was supported by a grant from the Vietnam and Thailand Office of the Ford Foundation.
* Corresponding author. Tel.: +1 608 262 6390; fax: +1 608 262 4376.
E-mail addresses: phand@beloit.edu (D. Phan), coxhead@wisc.edu (I. Coxhead).

[1] Please see Fig. 1 for a map of Vietnam.
[2] In 2002, the four provinces in the HCMC cluster made up 50% of the country's industrial output (HCMC alone accounted for 22%), while the northern cluster made up 14%. According to Mekong Economics (2002), the key areas in the South attract as much as 60% of all licensed foreign direct investment projects, and 53% of total registered capital.

0304-3878/$ – see front matter
doi:10.1016/j.jdeveco.2009.06.008

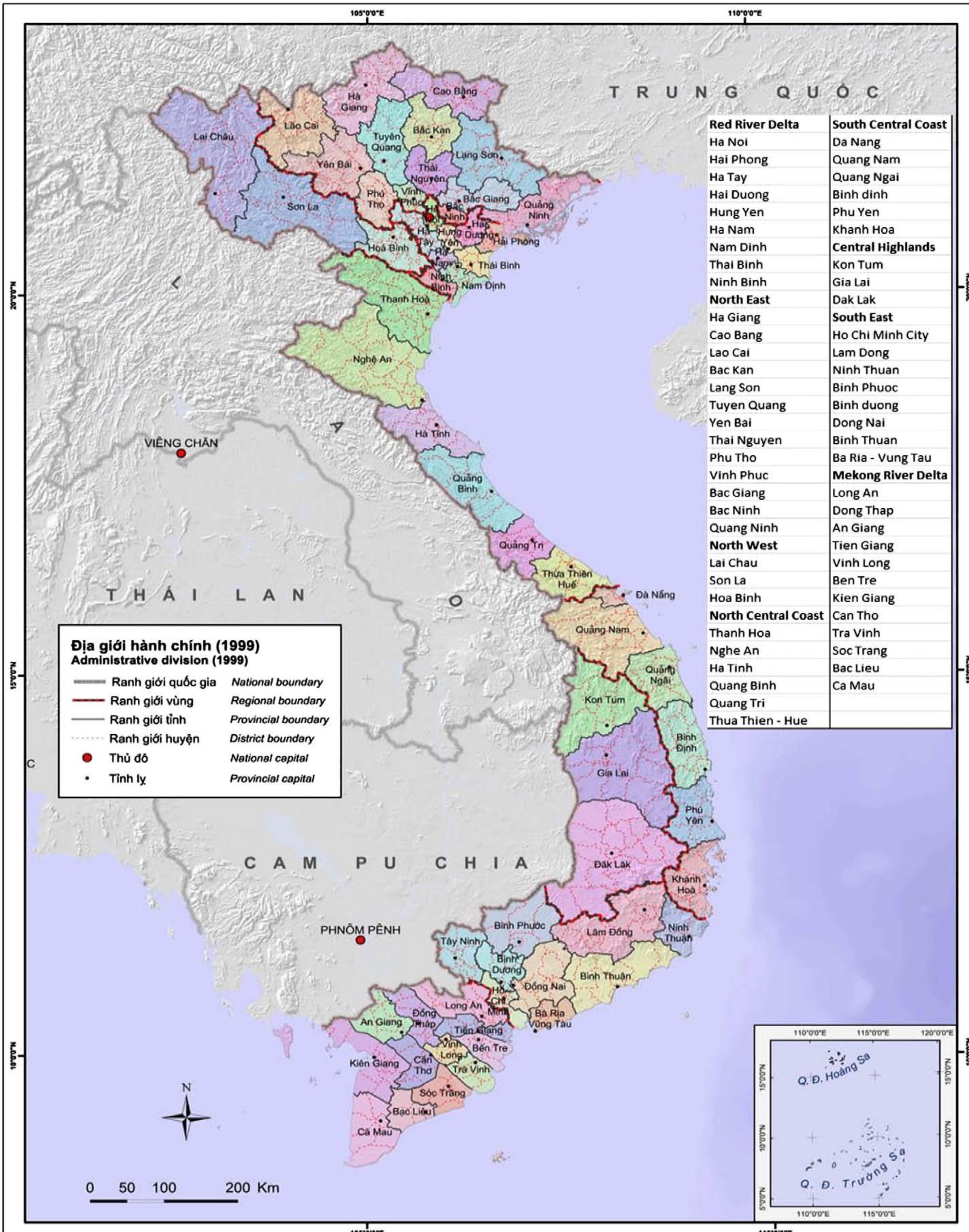

Fig. 1. Map of Vietnam's regions and provinces. Source: http://www.agro.gov.vn/map/mapitem_E.asp?ChangeMap=1&Map=1.02_admin.jpg.

(2008) study the impact of migration on the migrant's family and the source community, and begin to unravel the link between migration and poverty and inequality. These studies generally indicate that internal migration benefits not only migrants but also their families and source communities, and that internal migration tends to be pro-poor although it might increase inequality within the origin.

Our paper's contribution to the literature is twofold. First, while we employ the popular empirical gravity model, which hypothesizes that the flow of migrants between locations is a function of population, distance, wage/income differentials, differences in unemployment rate, and other variables (Greenwood, 1997; Fan, 2005; Dhar, 1984; Andrienko and Guriev, 2004; Mueser, 1989), we construct a theoretical framework that gives the model a solid structural interpretation. Furthermore, the framework incorporates a subsistence constraint and thus yields hypotheses regarding the impact of a sending province's income on migration. This allows us to examine the implications of liquidity constraints on poverty-related labor immobility.

The second contribution is to unravel the link between internal migration and regional income inequality. The empirical literature typically addresses this question with a standard convergence/growth regression, but this approach has also been subjected to methodological and empirical critiques (see Section 4). We examine the migration–inequality relationship using a simple yet novel approach: we relate the impact of the out-migration rate from province i to province j in one period to the change in income differential between the two provinces in the next period.

This second contribution also relates to the existing literature on Vietnam's migration and spatial inequality. Increasing regional inequality is becoming a concern in Vietnam, as evidenced by a steady increase in the standard deviation of per capita GDP across provinces (see Fig. 2). To study the impact of migration on regional or provincial inequality, micro data sets with national coverage are needed. Yet there is a serious lack of such data sets, making it difficult, if not impossible, to relate macroeconomic changes at the regional and sectoral levels with economic decision-making and welfare changes at the micro level. For an empirical study, then, there is a trade-off between national-level coverage at the provincial level and much more limited coverage at the level of the household. Most of the Vietnam migration studies cited above forgo the former, and hence cannot draw economy-wide inferences on migration issues. Our study uses aggregate data on interprovincial migration. The limitations of this data set for the purpose of obtaining insights into household-level migration decisions and outcomes are obvious. Unlike household level data, however, it does yield economy-wide inferences on migration. In this way our approach is complementary to the more usual household level analysis.

Our econometric analysis robustly confirms economic motives for migration, but also suggests the existence of poverty-related labor immobility at the provincial level. This in turn may imply persistence of poverty in certain regions if labor mobility is indeed a major channel through which the benefits of growth are distributed. We find that the impact of migration on income ratios between pairs of provinces can be positive, negative or zero, depending on the destination. A particularly interesting result is the robust and negative impact on income differentials of migration flows going to HCMC and surrounding provinces, where most of Vietnam's labor-intensive manufacturing growth and investments are concentrated. This lends support to the importance of the labor market and migration in distributing the benefits of trade-driven manufacturing growth.

The rest of the paper proceeds as follows. Section 2 provides a theoretical framework for the migration decision of a representative household, and links migration with inter-provincial income inequality. Section 3 investigates determinants of inter-provincial migration flows, and tests the hypothesis that poorer people/provinces have low migration propensity because of their inability to finance migration costs. Section 4 examines the relationship between migration and inter-provincial income differentials; and Section 5 offers concluding thoughts and directions for future research.

2. Theoretical framework

Lucas (1997) provides an excellent review of the literature on inter-provincial migration in developing countries. Empirical work in this literature is most often based on a gravity equation:

$$m_{ij} = m\left(y_i, y_j, d_{ij}, A_i, A_j, e\right)$$

where m_{ij} is the migration rate from i to j; y_i, y_j are mean incomes (or wages) in i and j respectively, and d_{ij} is the distance between the two. A_i and A_j are vectors of characteristics other than income or wages. This equation assumes an underlying theory of individual utility or income maximization in which migration is driven by differences in

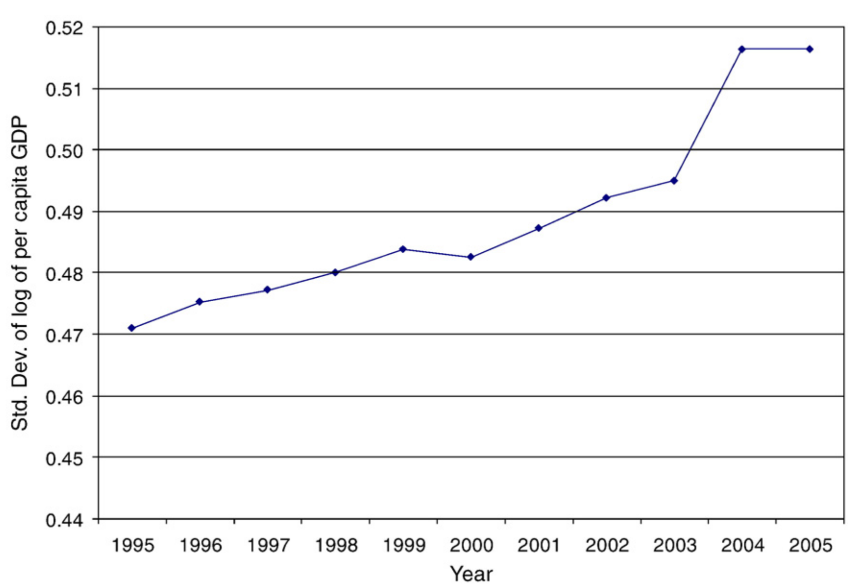

Fig. 2. Trend in provincial income disparity in Vietnam, 1995–2005.

non-market clearing regional wages or incomes; estimates using this framework typically include measures of differences in income, wages, unemployment, urbanization, etc.[3] Some authors have also used data on past migration flows, as proxies for migration networks. Of the large number of such migration equations estimated for both developed and developing countries, most find a positive effect for destination income or wage, a negative effect for origin income or wage, and a negative effect for distance (Lucas, 1997).

In our empirical analysis later in this paper, we also employ the gravity equation. In this section we present a simple theoretical model describing the inter-provincial migration decision of a representative household. The model extends that of Yang (2004) to include the possibility of liquidity constraints as barriers to migration. With an eye to the data constraint, our model assumes that characteristics affecting migration decisions are identically and independently distributed across households and individuals within each observational unit (province).

2.1. Motivations for migration

Consider a pair of provinces, *s* (sending) *and r* (receiving). Assume that in the initial period $t=0$, per capita income in *r* is higher than in *s*, or $Y_{r,0}>Y_{s,0}$. At the end of $t=0$, a representative household in province s must decide whether or not to allocate some labor to migration in order to maximize its expected income in $t=1$:

$$Y_{s,1} = \mathrm{MAX}_{m_{sr,1}}\left\{\left(1-m_{sr,1}\right)Y_{s,0}+m_{sr,1}Y_{r,0}-C\left(m_{sr,1},m_{sr,0},D_{sr}\right)\right\}$$

where $C(m_{sr,1}, m_{sr,0}, D_{sr})=\overline{C}^{*}m_{sr,1}^{\sigma}m_{sr,0}^{\theta}D_{sr}^{\gamma}$ is the assumed structure of the migration cost function; $m_{sr,0}$ is the past migration rate, and D is the distance between provinces. Total labor supply of a representative household is normalized to one, so $m_{sr,1}$ is also the out-migration rate from *s* to *r*.

We assume $\sigma>1$, $\theta<0$, and $\gamma>0$. $\sigma>1$ implies that migration cost is convex in migrant labor supply; that is, migration costs rise at an increasing rate with the number of migrants. This assumption is widely used in the regional economics literature (Yang, 2004). Convexity can arise from sources such as increasing urban rents or the reluctance of remaining family members to migrate. In the Vietnamese case, the motive of retaining control over farmland is another strong reason to assume convex costs. As we will see shortly, the assumption on σ is important, because it may determine the relationship between migration rate and sending province's per capita income.

The literature recognizes both past migrations and distance as important determinants of migration costs. In particular, distance serves as a proxy for the costs of relocating, job search and information acquisition, as well as for psychological costs. Greater distance should increase these costs and hence deter migration, i.e., $\gamma>0$. The past migration rate, on the other hand, measures the stock of existing migrants and proxies for a migration network, which can significantly reduce these costs and thus encourage migration, i.e., $\theta<0$. The impact of past migration, or migration network, is especially important when the destination is a major city with high housing and living costs, such as HCMC.

[3] This theoretical perspective, in which migration is viewed as an equilibrating force is called the disequilibrium approach. Another theoretical perspective, the equilibrium approach, also assumes that migration is motivated by spatial variations, but this approach holds that the existence of a wage/income differential does not reflect market disequilibrium, but simply reflects in regional amenities. Both migration and wages thus act as equilibrating forces in response to exogenous changes in amenity demand. This framework gives the rationale for including a wide variety of regional amenities in the gravity equation, such as average humidity, number of hot days, public services, health facilities, etc.

The first-order condition of the migration function yields the optimal migration rate:[4]

$$m_{sr,1}^{*} = \left(\frac{Y_{r,0}-Y_{s,0}}{\sigma\overline{C}m_{sr,0}^{\theta}D_{sr}^{\gamma}}\right)^{\frac{1}{\sigma-1}} \tag{1}$$

Taking logs on both sides, we get:

$$\ln\left(m_{sr,1}^{*}\right) = \frac{-\ln\left(\sigma\overline{C}\right)}{\sigma-1}+\frac{1}{\sigma-1}\ln\left(Y_{r,0}-Y_{s,0}\right)-\frac{\theta}{\sigma-1}\ln\left(m_{sr,0}\right)-\frac{\gamma}{\sigma-1}\ln(D_{sr}) \tag{2a}$$

or

$$\ln\left(m_{sr,1}^{*}\right) = \beta_0\ln(\sigma C)+\beta_y\ln\left(Y_{r,0}-Y_{s,0}\right)+\beta_m\ln\left(m_{sr,0}\right)+\beta_d\ln(D_{sr}) \tag{2b}$$

where

$$\beta_y = \frac{1}{\sigma-1},\quad \beta_d = \frac{-\gamma}{\sigma-1},\quad \beta_m = \frac{-\theta}{\sigma-1}$$

are the elasticities of migration rate with respect to income differential, distance, and past migration rate, respectively. Eq. (2b) has the form of a typical modified gravity migration equation, in which the migration rate is a function of the income differential, distance, and past migration (Greenwood, 1997; Fan, 2005; Dhar, 1984; Andrienko and Guriev, 2004; Mueser, 1989).

2.2. Liquidity constraint and non-linearity in migration decisions

As noted earlier, most empirical work finds a negative impact of origin income on migration. However, a few studies find a positive effect in some income ranges, suggesting a non-linear relationship between origin income and migration propensity (Connell et al., 1976; Barnerjee and Kanbur, 1981; Andrienko and Guriev, 2004). In reconciling this difference, Banerjee and Kanbur (1981) were among the first to incorporate migration costs and to investigate the role of origin income in the migration process. In their model, when migration is costly and the capital market operates imperfectly, higher income raises the capacity to finance migration. We use the same idea but model the migration cost function explicitly, positing a non-linear relationship between migration and origin income.

Suppose the household faces a subsistence constraint *S*, such that any savings from period 0 income at the start of period 1 must satisfy the constraint $Y_{s,0}-C\geq S$. By assumption, there is no borrowing due to capital market imperfections. For simplicity, and without changing key results, we drop past migration rate and distance from the migration cost function (several subscripts are also dropped for notational ease). The household's optimization problem becomes:

$$\begin{aligned}&\mathrm{MAX}_m\left\{(1-m)Y_s+mY_r-\overline{C}m^{\sigma}\right\}\\&\text{s.t.}\quad Y_s-\overline{C}\,m^{\sigma}\geq S\end{aligned} \tag{3}$$

The first-order condition yields:

$$m^{*} = \left(\frac{Y_r-Y_s}{\sigma\overline{C}(\lambda+1)}\right)^{\frac{1}{\sigma-1}} \tag{4}$$

[4] Note that when $\sigma>1$, the objective function is globally concave, and the solution is as presented in the paper. When $\sigma=1$, we clearly have a corner solution: the household compares the expected income from not migrating at all to that from sending all labor away, and choose the option that yields higher income. When $\sigma>1$, the objective function is globally convex, so no global minimum exists. However, since the choice variable m is constrained to be in the interval [0,1], we again have a corner solution as in the case of $\sigma=1$.

where λ is the Lagrange multiplier. If the subsistence constraint binds (λ>0), then the optimal migration rate, using Eq. (3), is $m^* = \left(\frac{Y_s - S}{\bar{C}}\right)^{\frac{1}{\sigma}}$. If not (λ=0), then the optimal migration rate, using Eq. (4), is $m^* = \left(\frac{Y_r - Y_s}{\sigma\bar{C}}\right)^{\frac{1}{\sigma-1}}$. Let $\overline{Y}_s$ be the solution to $\left(\frac{Y_s - S}{\bar{C}}\right)^{\frac{1}{\sigma}} = \left(\frac{Y_r - Y_s}{\sigma\bar{C}}\right)^{\frac{1}{\sigma-1}}$. This is the level of income above which the household is no longer bound by the subsistence constraint.[5] So we have:

$$m^* = \begin{cases} \left(\dfrac{Y_s - S}{\bar{C}}\right)^{\frac{1}{\sigma}} & \text{if } Y_s \le \overline{Y}_s \\ \left(\dfrac{Y_r - Y_s}{\sigma\bar{C}}\right)^{\frac{1}{\sigma-1}} & \text{if } Y_s > \overline{Y}_s \end{cases} \tag{5}$$

and

$$\frac{\partial m^*}{\partial Y_s} = \begin{cases} \dfrac{1}{\sigma}(Y_s - S)^{\frac{1-\sigma}{\sigma}}\,\bar{C}^{\frac{-1}{\sigma}} & >0 \quad \text{if} \quad Y_s \le \overline{Y}_s \\ \dfrac{-1}{\sigma-1}(Y_r - Y_s)^{\frac{2-\sigma}{\sigma-1}}\left(\sigma\bar{C}\right)^{\frac{-1}{\sigma-1}} & <0 \quad \text{if} \quad Y_s > \overline{Y}_s \end{cases} \tag{6}$$

Thus, the marginal impact of the sending province's per capita income on the out-migration rate is always positive at low income levels (when $Y_s \le \overline{Y}_s$) — as suggested by the liquidity constraint hypothesis, which maintains that the poorer the province, the lower the capacity to finance migration cost, and hence the lower the out-migration rate. But when $Y_s \le \overline{Y}_s$, the marginal impact of the sending province's per capita income on migration is negative (provided $\sigma>1$, as assumed here), as suggested by the push effect hypothesis, which maintains that poorer people want to migrate more. That the out-migration function operates in two distinct regimes suggests that a "labor mobility trap" is possible as long as migration costs are sufficiently high and the capital market is imperfect.

2.3. Migration and income inequality

Let us go back to the notation in Section 2.1. Following Yang (2004), for the whole household the realized net income gain from migration in period 1 is:

$$\begin{aligned} G &= \left(1 - m^*_{sr,1}\right)Y_{s,1} + m^*_{sr,1}Y_{r,1} - \bar{C}\left(m_{sr,1}\right)^{\sigma} m^{\theta}_{sr,0}D^{\gamma}_{sr} - Y_{s,1} \\ &= \left(Y_{r,1} - Y_{s,1}\right)m^*_{sr,1} - \bar{C}\left(m^*_{sr,1}\right)^{\sigma} m^{\theta}_{sr,0}D^{\gamma}_{sr} \end{aligned} \tag{7}$$

Substituting from Eq. (1) into Eq. (7) yields:

$$G = m^*_{sr,1}\left[\left(Y_{r,1} - Y_{s,1}\right) - \frac{1}{\sigma}\left(Y_{r,0} - Y_{s,0}\right)\right]. \tag{8}$$

Define $I_{sr,0} = Y_{r,0}/Y_{s,0}$ as relative income inequality between s and r in period $t=0$. For simplicity, assume that $Y_{r,0}$ and $Y_{s,0}$ are exogenously given and that they are the same as per capita GDP, i.e., there is no remittance flow between the two provinces in the initial period. Let $g_r(m_{sr}, Z_r)$ and $g_s(m_{sr}, Z_s)$ be the gross provincial growth rates[6] of per capita GDP from $t=0$ to $t=1$. Z_r and Z_s are exogenous province-specific characteristics that determine their long-run per capita GDP growth rates. Let τ be the measure of the propensity to remit. Relative income inequality between s and r in period $t=1$ is then:[7]

$$\begin{aligned} I_{sr,1} &= \frac{Y_{r,1}}{Y_{s,1} + \tau * m_{sr,1} * \left[\left(Y_{r,1} - Y_{s,1}\right) - \frac{1}{\sigma}\left(Y_{r,0} - Y_{s,0}\right)\right]} \\ &= \frac{g_r\left(m_{sr,1}, Z_r\right) * Y_{r,0}}{g_s\left(m_{sr,1}, Z_s\right) * Y_{s,0} + \tau * m_{sr,1} * \left[Y_{r,0} * \left(g_r\left(m_{sr,1}, Z_r\right) - \frac{1}{\sigma}\right) + Y_{s,0} * \left(g_s\left(m_{sr,1}, Z_s\right) - \frac{1}{\sigma}\right)\right]} \end{aligned} \tag{9}$$

The current-period income differential is a function of initial incomes, migration, and variables that determine provincial per capita GDP growth rates.

3. Determinants of migration flows in Vietnam, 1984–89 and 1994–99

3.1. Empirical strategy

From Eq. (2b), the system of structural equations to be estimated is:

$$\begin{aligned} \ln\left(m^{84-89}_{sr}\right) &= \beta^{84}_0 C + \beta^{84}_s \ln\left(y^{84}_s\right) + \beta^{84}_r \ln\left(y^{84}_r\right) \\ &\quad + \beta^{84}_d \ln(d_{sr}) + \eta^{84}X + \varepsilon^{84}_{sr} \end{aligned} \tag{10a}$$

$$\begin{aligned} \ln\left(m^{94-99}_{sr}\right) &= \beta^{94}_0 C + \beta^{94}_s \ln\left(y^{94}_s\right) + \beta^{94}_r \ln\left(y^{94}_r\right) \\ &\quad + \beta^{94}_d \ln(d_{sr}) + \beta_m \ln\left(m^{84-89}_{sr}\right) + \eta^{94}X + \varepsilon^{94}_{sr} \end{aligned} \tag{10b}$$

where d_{sr} is the bus distance (in kilometers) between the capital cities of the two provinces, and X is a vector of other control variables such as regional dummies for sending and receiving provinces, or province fixed effects.[8] As written, these equations also relax the implicit restriction in Eq. (2b) that the impacts on migration of sending and receiving provinces' incomes are equal but of opposite signs.

Eq. (10a) includes only pre-determined variables, so ordinary least square (OLS) provides a consistent estimator. Eq. (10b) includes pre-determined variables plus the past migration rate, which is endogenous in Eq. (10a). If we impose the assumption that ε^{84}_{sr} and ε^{94}_{sr} are uncorrelated, the system becomes fully recursive; then OLS applied to each structural equation will yield unbiased and consistent parameter estimates of the direct effects of the right-hand-side variables on the left-hand-side variables. The indirect effect of a change in a time-invariant predetermined variable, for example distance, on the left-hand-side variable through the endogenous right-hand-side variable (past migration) can be computed from these estimates by $\beta^{84}_d * \beta_m$. The total effects, or the sum of direct and indirect effects, can be computed as follows:

$$\beta_d = \beta_m\beta^{84}_d + \beta^{94}_d \quad \beta_0 = \beta_m\beta^{84}_0 + \beta^{94}_0 \quad \eta = \beta_m\eta^{84} + \eta^{94} \quad \varepsilon_{sr} = \beta_m\varepsilon^{84}_{sr} + \varepsilon^{94}_{sr}.$$

However, there may be unobserved factors that affect migration flows in both periods, causing ε^{84}_{sr} and ε^{94}_{sr} to be correlated. In that case OLS yields inefficient estimates of Eq. (10a) and inconsistent estimates of Eq. (10b). Consistent and efficient estimates require a systems estimator such as three-stage least squares (3SLS). We implement both OLS and 3SLS, but as will be seen, the two methods give similar results.

[5] Note that $\left((Y_s - S)/\bar{C}\right)^{\frac{1}{\sigma}}$ is monotonically increasing in Y_s, while $\left((Y_r - Y_s)/\sigma\bar{C}\right)^{\frac{1}{\sigma-1}}$ is monotonically decreasing in Y_s, so a unique solution $\overline{Y}_s$ exists.

[6] Defined as $g = 1 + r$, where r is the net growth rate.

[7] It is implicitly assumed that migrants are accounted as one separate group, so the defined relative income inequality only measures the income differential between natives of province r and stayers of provinces.

[8] Please refer to our working paper version for a discussion of inclusion of other provincial characteristics such as ethnicity, education, land, etc. (Phan and Coxhead, 2008).

Throughout the following analysis we will obtain estimates both for the full sample and for a sub-sample including as receiving provinces only HCMC and the three neighboring provinces. As noted earlier, this is a special urban cluster where most manufacturing capital and foreign direct investment are located, and where most recent growth, especially of export-oriented manufacturing, is concentrated. It is therefore of particular interest to examine the determinants and consequences of migration flows into these provinces.

3.2. Testing the liquidity constraint hypothesis

The theoretical model in Section 2.2 suggests that at low levels of income, the liquidity constraint effect should dominate the push effect, such that β_s has a positive sign in Eqs. (10a) and (10b). But at high levels of income, the push effect should dominate, such that β_s has a negative sign. To test this hypothesis, we need to first compute, for each inter-provincial flow from s to r, the threshold income level $\overline{Y}_s$, which is a function of many other variables, such as Y_r, $\overline{C}$ etc. Then, we could run two separate OLS regressions, one for observations in which the sending province's income is below its threshold, and another for observations in which it is above. Comparing coefficient estimates would provide a test of the liquidity constraint hypothesis. Unfortunately, such a direct test is not possible due to lack of data needed to compute threshold income levels.

An indirect test may be constructed by supposing (as seems reasonable) that the liquidity constraint effect varies with the level of migration cost, becoming material only when migration cost is material. Thus one intuitive way to test for a liquidity constraint is to break the sample into percentiles of distance and past migration rate (which affect migration costs), then examine how β_s varies across sub-samples. Another way, since we are concerned about non-linearity in the income–migration relationship, is to test for higher-order impacts by including the square and/or cube of the log of sending province income. Alternatively, we can estimate the following semi-parametric model:

$$\ln\left(m_{sr}^{94-99}\right) = f\left(\ln\left(y_s^{94}\right)\right) + \beta Z + \varepsilon \tag{11}$$

where the vector Z includes all explanatory variables other than sending province's income. Eq. (11) falls into the class of partially linear models, which consist of a linear part, βZ, and a non-parametric part, $f(.)$. No parametric assumption is imposed upon $f(.)$ except that it is a smooth function, while for the rest of the variables the usual parametric assumptions apply (for details see Hardle et al., 2004, Ch. 7). This model allows the data to freely determine the shape of the influence of origin income on migration.

3.3. Data

Table 1 gives summary statistics of variables used in the regressions. Most data on province characteristics, such as per capita income, population, etc., are from different versions of the Statistical Year Books published by the General Statistics Office (GSO) of Vietnam. Migration data are from the 1989 and 1999 Population and Housing Censuses (GSO, 1991, 2001). The censuses ask questions on place of birth, duration of residence, place of last residence, and place of residence at a fixed prior date. A respondent is identified as a migrant if he/she was at least five years of age at the time of census, and changed place of residence within the past five years. This allows estimation of inter- and intra-provincial migration flows during the prior five-year period. As noted, a major drawback of the approach is that it excludes temporary/seasonal and return migrants, as well as those who were born during the five-year interval. The exact timing of any reported move is also unknown. Thus the census data underestimate actual migration, and are more likely to reflect permanent than temporary moves. Lastly, since we do not have data on provincial per capita incomes in 1984, we use per capita industrial output in 1986 instead.

From 1984 to 1989, inter-provincial migration in Vietnam closely followed the resettlement program discussed earlier. From 1994 to 1999, rural–rural migration flows continued to be correlated with those in the 1980s: they were large and involved long-distance moves from the Red River Delta and North Central Coast regions to the Central Highlands (see Table 2). But unlike the 1980s, there also emerged new short-distance rural–urban moves from the Red River Delta into Hanoi and from the Southeast and Mekong River Delta into HCMC.

Such migration patterns accord with theoretical predictions that people tend to move from low income to high income areas, and also from land-scarce to land-abundant regions. The former is borne out by Fig. 3, which shows a strong and statistically significant relationship between per capita income and net in-migration, confirming that internal migration in Vietnam is motivated to a large extent by income differences. However a particular case against this trend is the low level of labor mobility either into or out of the North West region — the poorest and also most remote region of Vietnam. The persistence of poverty in this region might be attributable to a combination of high migration costs and household-level liquidity constraints (i.e., inability to finance migration cost due to low income). These in turn are correlated with ethnicity; the more remote regions are also those with the highest proportion of ethnic minorities, whose language and cultural barriers drive migration costs especially high.

Table 1
Summary statistics.

Variables	Unit	Mean	Std. dev.	Min	Max
94–99 migration flow	People	577	1770	0	33,590
84–89 migration flow	People	395	1091	0	21,353
Distance	km	1018	719	20	2597
1986 pc industrial output	Thousand VND, 1982 price	1.50	1.60	0.15	9.77
1994 monthly per capita income	Thousand VND, 1994 price	161.9	48.7	99.2	413.3
1999 monthly per capita income	Thousand VND, 1999 price	283.3	103.5	160.7	828.2
2002 monthly per capita income	Thousand VND, 2002 price	318.3	117.4	173.1	904.1

3.4. Results

Tables 3a and 3b show OLS and 3SLS estimates of Eqs. (10a) and (10b)) with regional fixed effects. For the 1980s data, the OLS and 3SLS estimates are almost the same, which is expected because there is no endogeneity in Eq. (10a). OLS and 3SLS estimates for the 1990s regression are not as similar, but the differences are small, suggesting that any endogeneity due to correlation of error terms does not have serious effects on the estimates. This may be because regional dummies already capture much of the unobserved influence on migration flows in both periods. To save space we will discuss only the OLS estimates.

Table 3a shows that all explanatory variables have coefficients of expected signs that are statistically significant at either 1% or 5% level. Provinces that are further apart send fewer migrants to each other, as the coefficient of distance variable is negative. Provinces with high per capita income attract more migrants. In the 1990s, a 1% increase in the per capita income of receiving province leads to a 1.56% increase in the migration rate, implying a rather high response of migration to income. Taking the absolute value of the ratio of the estimated distance elasticity to the estimated elasticity on destination income, we obtain the income–distance trade-off, a rough indicator of the cost of moving a given distance further (Greenwood, 1997). For the 1990s that value is 0.73 ($=|-.1.1/1.5|$), meaning that a .73% increase in destination income is needed to offset a 1% increase in distance. Computed at the sample averages, this means that a move 7 km further away can be offset by an increase in monthly income of 11,800 VND (approximately $US1 at 1999 prices). Although comparisons

Table 2
Gross inter-provincial migration flows by regions and major cities.

Region/city of residence in 1994	Region/city of residence in 1999										
	Red River Delta	North East	North West	North Central Coast	South Central Coast	Central Highland	South East	Mekong River Delta	Hanoi	HCMC	Other Urban
Red River Delta	47,553	59,284	16,305	8051	9261	76,398	53,543	11,191	108,717	47,605	23,209
North East	39,536	51,261	6507	6196	2524	52,494	18,832	1005	35,631	7631	6852
North West	10,691	4093	2043	1782	407	4772	888	287	6359	685	310
North Central Coast	12,521	6466	25,147	30,301	14,109	88,410	61,520	7756	26,995	48,199	21,621
South Central Coast	1836	745	123	5191	15,308	50,431	15,409	2347	1478	50,615	24,591
Central Highland	6458	1903	375	8457	10,150	17,016	9552	1283	1453	16,247	4250
South East	4246	1300	178	5679	5023	26,832	45,792	10,178	849	73,972	10,721
Mekong River Delta	2440	920	78	2266	2036	9370	49,466	167,417	965	153,292	8334
Hanoi	20,277	8548	1423	6977	1267	1661	2184	1241	–	8709	3039
HCMC	4820	1119	103	5480	6990	6535	30,930	30,588	3393	–	7441
Other Urban	6758	6166	282	5420	7451	8988	12,633	3516	10,793	26,707	2015

across studies are hard to make, we note that estimates for Canada were found in the range 0.146 to 0.439, with variation by age, education and time period (Courchene, 1970). For China, the figure is 0.35 (Fan, 2005). Thus the income–distance trade-off of 0.73% found in this study seems high by comparison with other studies. This suggests that in Vietnam the part of migration cost that is correlated with distance is greater than in other countries.

In the 1980s, the regional dummies show that the Central Highlands and the South East region received more migrants than would be expected based on their income levels and their distance to other regions. Per capita income in the Central Highlands was among the lowest in the country during that time. These regional dummies should be interpreted as capturing migration policy effects, as the Vietnamese government actively encouraged and organized migration into these regions in the 1980s. The dummy for urban areas is not statistically significant in the 1980s, due to policy bias against rural–urban migration at that time. In the 1990s the Central Highlands and South East regions continued to be major receiving areas; however, their dummies can no longer be interpreted as measuring policy effects. Rather, they now capture the impact of a structural shock, Vietnam's coffee export boom (nearly all Vietnam's coffee is grown in these regions). Also, in the 1990s the urban area dummy is positive and significant. This is presumably due to diminished uncertainty about future urban job prospects, as well as other factors associated with urban life but not captured in the current income data. In general, the signs of regional dummies are in accord with the theoretical prediction that the mobile factor (labor) flows to regions with high concentrations of immobile factors (land and capital).

The estimated coefficient of per capita income in the sending province is positive. It might be that the liquidity constraint effect is stronger than the push effect, causing migration flows to be larger when the sending province is richer. Intuitively, we would expect the liquidity constraint effect to matter more when migration costs are higher, i.e., when distance is longer or past migration rate is lower. We thus test this hypothesis as follows. First, we break the data into five subsamples by quintiles of distance and estimate the 1990s regression separately for each. The results are in Table 4a. A Chow test ($F=54$) rejects the null hypothesis of equal coefficients across the five piecewise regressions. A pattern seems to emerge: the impact of sending province's per capita income at first increases in both magnitude and statistical significance as distance becomes greater, then falls at the highest distance quintile. This unexpected pattern at

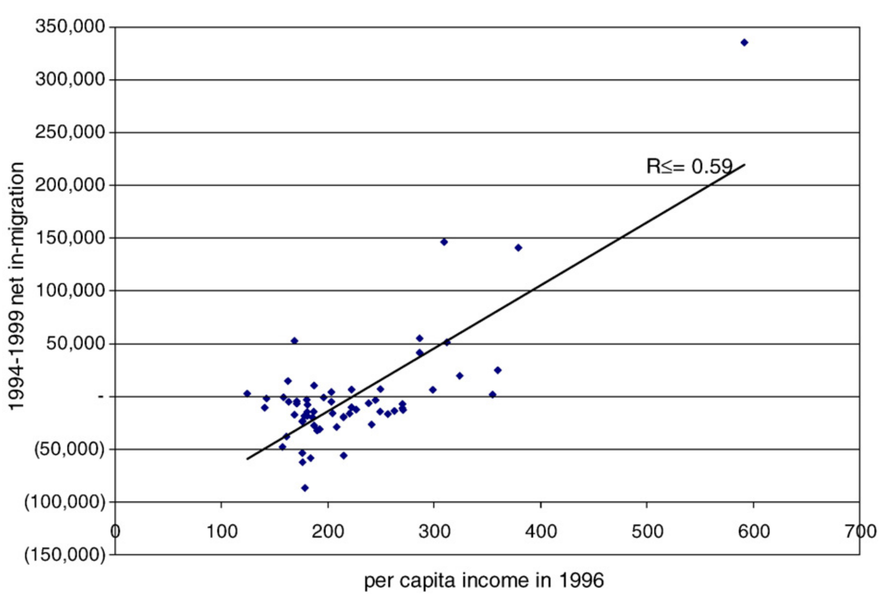

Fig. 3. Per capita income vs. net in-migration.

Table 3a
OLS regression of migration flows with regional fixed effects.

	1980s		1990s			
	Coeff.	*t*-stat	Coeff. (direct effect)	*t*-stat	Total effects	Indirect effects
ln(1984–1989 migration rate from *s* to *r*)	–	–	0.672	27.2	–	–
ln(bus distance)	− 1.131	− 48.01	− 0.437	− 14.2	− 1.198	− 0.760
ln(receiving province's per capita income, 1994)	–	–	1.560	13.2	–	–
ln(sending province's per capita income, 1994)	–	–	0.858	8.160	–	–
ln(receiving province's pc industrial output, 1986)	0.798	14.0	–	–	–	–
ln(sending province's pc industrial output, 1986)	0.093	1.84	–	–	–	–
Sending region = North East	− 0.212	− 2.6	− 0.687	− 11.6	− 0.829	− 0.142
Sending region = North West	− 0.438	− 3.5	− 0.590	− 5.3	− 0.884	− 0.294
Sending region = North Central Coast	− 0.172	− 1.9	− 0.361	− 5.5	− 0.476	− 0.115
Sending region = South Central Coast	− 1.092	− 11.1	− 0.704	− 7.7	− 1.438	− 0.734
Sending region = Central Highland	− 0.101	− 0.9	− 0.471	− 4.9	− 0.539	− 0.068
Sending region = South East	− 0.858	− 9.2	− 0.401	− 5.5	− 0.978	− 0.577
Sending region = Mekong River Delta	− 1.899	− 20.2	− 0.876	− 10.7	− 2.152	− 1.276
Sending region = urban areas	− 0.480	− 4.2	− 0.618	− 7.8	− 0.940	− 0.322
Receiving region = North East	− 1.283	− 15.0	− 0.160	− 2.1	− 1.022	− 0.863
Receiving region = North West	− 0.968	− 5.8	− 0.340	− 2.2	− 0.991	− 0.651
Receiving region = North Central Coast	0.660	7.5	− 0.303	− 4.0	0.141	0.444
Receiving region = South Central Coast	− 0.339	− 4.2	0.162	2.2	− 0.066	− 0.228
Receiving region = Central Highland	0.991	7.6	0.931	11.0	1.597	0.666
Receiving region = South East	0.813	9.2	0.266	3.3	0.812	0.546
Receiving region = Mekong River Delta	− 0.648	− 8.5	− 0.463	− 6.2	− 0.898	− 0.436
Receiving region = urban areas	− 0.022	− 0.2	0.212	2.5	0.197	− 0.015
Constant	− 1.223	− 6.9	− 12.066	− 15.2	− 12.888	− 0.822
Adj. R^2	0.60		0.73			
N	3613		3613			

Notes: a) Dependent variable = ln(m_{sr}) = log of gross migration rate from *s* to *r* for the relevant period.
b) Direct effects come from structural equation estimation.
c) *t*-statistics are for structural equation estimates, and based on White heteroskedasticity consistent covariance matrix.
d) The omitted sending and receiving region is Red River Delta.

the highest distance quintile might be explained by the fact that there is high mobility among richest urban centers, offsetting the trend that at the highest distance quintile, impact of sending province's income matters more. In brief, these results provide partial evidence that at relatively long distances (but not the longest distances), out-migration is more likely to occur from higher-income provinces.

We next break the sample into five sub-samples by quintiles of past migration rate, and again run piecewise regressions (see Table 4b). A Chow test ($F = 50.6$) again rejects the null hypothesis of constant coefficients. At lower quintiles of past migration (i.e., when migration cost is higher), the coefficient on per capita income of the sending province is positive, high, and more statistically significant. At higher

Table 3b
3SLS regression of migration flows with regional fixed effects.

	1980s		1990s	
	Coeff.	*t*-stat	Coeff. (direct effect)	*t*-stat
ln(1984–1989 migration rate from *s* to *r*)	–	–	0.599	3.6
ln(bus distance)	− 1.131	− 48.07	− 0.517	− 2.7
ln(receiving province's per capita income, 1994)	–	–	2.345	4.9
ln(sending province's per capita income, 1994)	–	–	0.981	7.110
ln(receiving province's pc industrial output, 1986)	0.798	14.9	–	–
ln(sending province's pc industrial output, 1986)	0.093	1.800	–	–
Sending region = North East	− 0.212	− 2.6	− 0.700	− 8.1
Sending region = North West	− 0.438	− 3.4	− 0.599	− 4.3
Sending region = North Central Coast	− 0.172	− 1.8	−0.368	− 3.7
Sending region = South Central Coast	− 1.092	− 10.3	− 0.785	− 3.9
Sending region = Central Highland	− 0.101	− 0.9	− 0.496	− 4.5
Sending region = South East	− 0.858	− 8.9	− 0.502	− 2.7
Sending region = Mekong River Delta	− 1.899	− 22.6	− 1.044	− 3.1
Sending region = urban areas	− 0.480	− 3.6	− 0.703	− 4.6
Receiving region = North East	− 1.283	− 15.5	− 0.235	− 1.1
Receiving region = North West	− 0.968	− 7.6	− 0.277	− 1.4
Receiving region = North Central Coast	0.660	6.7	− 0.222	− 1.5
Receiving region = South Central Coast	− 0.339	− 3.2	0.144	1.4
Receiving region = Central Highland	0.991	8.7	0.914	4.5
Receiving region = South East	0.813	8.4	0.094	0.9
Receiving region = Mekong River Delta	− 0.648	− 7.7	− 0.693	− 3.4
Receiving region = urban areas	− 0.022	− 0.2	− 0.090	− 0.7
Constant	− 1.223	− 7.3	− 16.665	− 5.6
Adj. R^2	0.60		0.72	
N	3613		3613	

Notes: a) dependent variable = ln(m_{sr}) = log of gross migration rate from *s* to *r* for the relevant period.
b) The omitted sending and receiving region is Red River Delta.

Table 4a
OLS regressions of migration flows by quintiles of distance.

	1990s				
	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile
ln(1984–1999 migration rate from *s* to *r*)	0.694	0.740	0.743	0.508	0.309
	(20.22)	(24.52)	(15.82)	(8.18)	(6.55)
ln(bus distance)	−0.572	−0.843	0.189	0.207	−9.854
	(−9.99)	(−5.76)	(1.07)	(0.38)	(−7.87)
ln(receiving province's per capita income, 1994)	0.657	1.031	1.847	2.29	51.995
	(3.49)	(6.22)	(7.26)	(6.31)	(4.67)
ln(sending province's per capita income, 1994)	0.152	0.608	1.214	1.715	0.806
	(0.85)	(3.62)	(5.6)	(5.4)	(1.89)
Adj. R^2	0.72	0.71	0.73	0.71	0.62
N	688	730	732	732	731

Notes: Regional dummies and constant terms were included but not reported.
1st quintile = shortest distance (meaning lower migration cost).
t-statistics in parentheses.

Table 4b
OLS regressions of migration flows by quintiles of past migration rate.

	1990s				
	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile
ln(1984–1999 migration rate from *s* to *r*)	0.186	0.259	0.165	0.249	0.507
	(5.00)	(6.09)	(4.46)	(4.46)	(6.24)
ln(bus distance)	−0.612	−0.396	−0.568	−0.242	−0.298
	(−2.37)	(−5.02)	(−9.56)	(−6.35)	(−7.66)
ln(receiving province's per capita income, 1994)	1.898	1.962	1.674	1.0011	.370
	(4.59)	(7.71)	(7.31)	(4.92)	(9.01)
ln(sending province's per capita income, 1994)	1.967	1.036	0.489	0.274	−0.081
	(6.11)	(3.89)	(2.15)	(1.45)	(−0.62)
Adj. R^2	0.37	0.41	0.39	0.38	0.68
N	723	723	722	723	722

Notes: Regional dummies and constant terms were included but not reported.
1st quintile = lowest past migration rate (meaning higher migration cost).
t-statistics in parentheses.

quintiles (i.e., lower migration cost) this coefficient becomes less statistically significant and also smaller. This suggests that the liquidity constraint effect is indeed stronger when migration costs are higher.

Finally, Fig. 4a and b present semi-parametric estimates for the impact of income in the sending province on migration, as specified in Eq. (11), using both the full sample and the sub-sample consisting of just HCMC and surrounding provinces as receiving areas.[9] For the full sample (Fig. 4a), there seems to be a non-monotonic impact of origin's income on the out-migration rate, but this impact is insignificant (the 95% confidence limits cover the zero axis; also, for the test of the null hypothesis that this impact is zero, $p=0.16$). However, in Fig. 4b, in which the only receiving provinces included are HCMC and environs, the null hypothesis of zero impact of origin income can be rejected (p-value $=0.012$). Furthermore, there is a clearer inverse U-shape relationship, as predicted by the liquidity constraint hypothesis: the impact of origin income on out-migration rate at first rises then diminishes with income, with a turning point of about 137,000 VND per month (mean monthly per capita income in the sample is 162,000 VND). It might be that the cost of migration into HCMC is higher than that to other regions, not because of distance and associated transport costs, but because of high housing and living costs that a new migrant might incur while searching for a job. This could explain why there are liquidity constraints on migration to the HCMC cluster but not to other regions.

In short, in the parametric estimation results, the positive sign of sending province income suggests not only that there exists a liquidity constraint effect, but also that this effect dominates the push, or income differential effect. In the full sample of receiving provinces, the semi-parametric results neither support nor reject the liquidity constraint effect. But when we restrict the sample to HCMC and surrounding provinces only, a clear inverse U-shape emerges, in accordance with the liquidity constraint hypothesis. Together, the results point to the existence of poverty-related labor immobility at provincial level, especially for migration flows into the HCMC cluster. This may in turn imply persistence of poverty in certain regions if labor mobility is indeed a major channel through which the benefits of growth are distributed. Our next task is to examine this assertion.

4. Migration and income differentials

Most studies of migration and regional inequality examine the impact of migration on the speed of convergence in per capita GDP growth rates, using methods such as in Barro and Sala-i-Martin (2004). This approach has two limitations. First, it suffers from the misspecification problem faced by any cross-section growth equation. Second, it is not just per capita GDP that interests us, but per capita income, inclusive of remittances. A sufficiently large remittance flow can offset the negative impacts of out-migration on the GDP of the origin economy. Recognizing this distinction between inequality in per capita GDP and income, in this section we relate the impact of out-migration rate from province i to province j to the change in income ratio between the two provinces.

Consider Eq. (9) in Section 2.3, which relates the current-period income ratio to migration, provincial growth rate determinants, and previous period incomes. If we had specific functional forms for the growth rates $g_r(m_{sr}, Z_r)$ and $g_s(m_{sr}, Z_s)$, then we could estimate Eq. (9) using a non-linear least squares technique. But since we neither have all the data nor the underlying determinants of provincial growth rates, we instead employ the following approach. We look at factors contributing to changes in the ratio of per capita incomes between pairs of provinces (note that these incomes include both GDP and remittances/transfers). This ratio is one measure of income inequality between any two provinces. Any change in this ratio must come from one of two sources:[10] differences in per capita GDP growth rates, or differences between the two provinces in remittances or transfer receipts. We have province-level data on incomes, but not on remittances or transfers, so we proxy these by migration rates. This leads to the following equation:

$$\Delta I_{sr} \equiv \frac{Y_{02}^{r}}{Y_{02}^{s}} - \frac{Y_{99}^{r}}{Y_{99}^{s}} = \rho_0 + \rho_1\left(\frac{g_{99-02}^{r}}{g_{99-02}^{s}}\right) + \rho_2 mig_{sr}^{94-99} + \varepsilon \qquad (12)$$

There is a concern that the per capita GDP growth rate ratio might be endogenous, because there is likely contemporaneous correlation between this variable and unobserved factors affecting the income ratio change on the left hand side. To address this, we instrument the 1999–2002 per capita GDP growth rate ratio using its lag, the ratio of per capita GDP growth rates for the period 1994–1999.[11] Since migration presumably responds to recent (rather than current) growth in sending and receiving provinces, it is unlikely that the growth rate instrument is correlated with the 1994–1999 migration rate variable.

Table 5 reports both the OLS (uninstrumented) estimates (model I), and the instrumented IV-GMM estimates (model II). The Wu–Hausman exogeneity test rejects the null hypothesis of exogeneity, indicating the need to for an instrumental variables approach. In the first stage regression (model III in Table 5), the F-statistic for the excluded instrument's relevance is 128, suggesting that there is a strong correlation between the endogenous variable and the

[9] Estimates for the parametric part are qualitatively similar to those in parametric regressions and are available from the authors upon request.

[10] Of course, there is a third source leading to the change: measurement error. Interested readers can refer to Deaton (2006) for a discussion of discrepancy between household survey data and national accounts data.

[11] We are grateful to an anonymous reviewer for comments and suggestions leading to this specification.

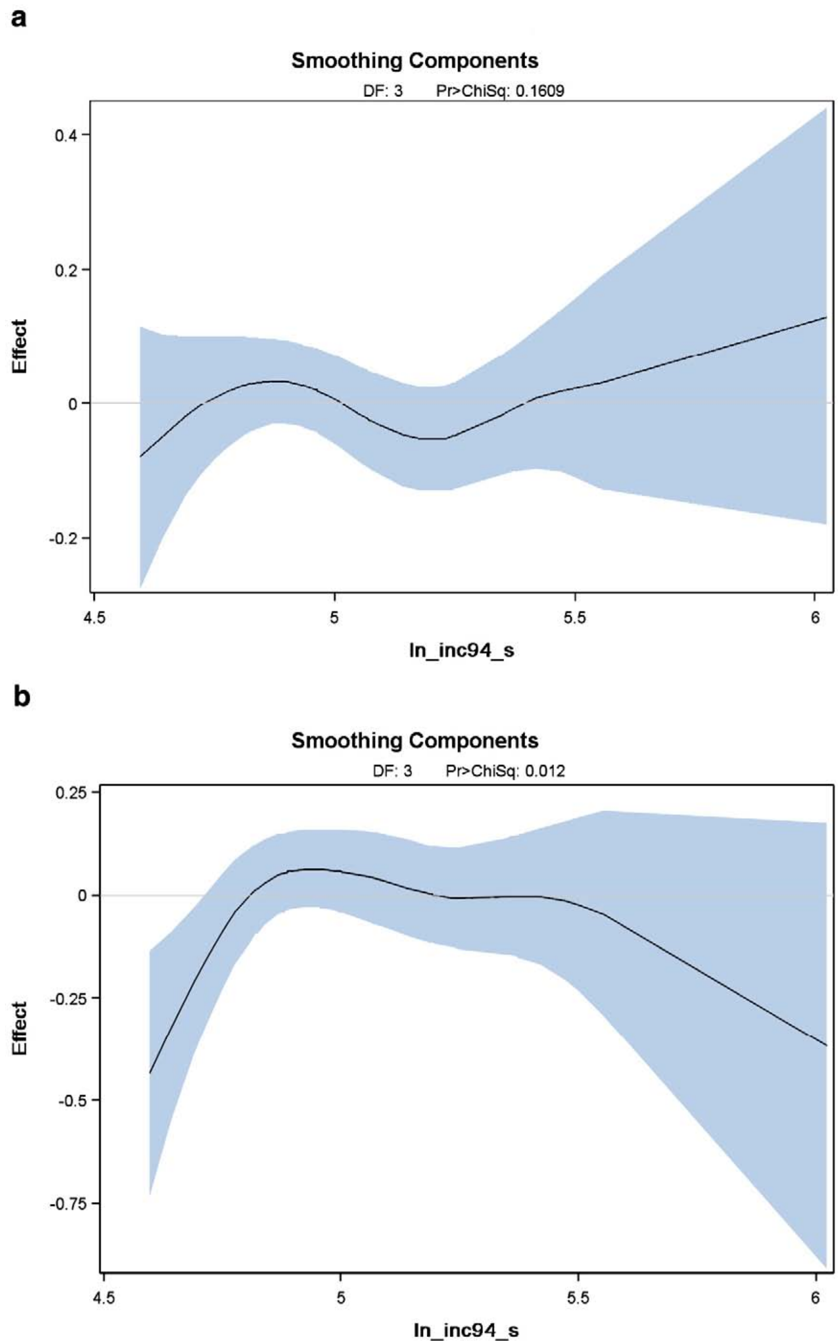

Fig. 4. a: Semi-parametric estimate (full sample result). b: Semi-parametric estimate (sub-sample includes HCMC, Ba Ria-Vung Tau, Dong Nai, and Binh Duong receiving provinces).

instrument. Furthermore, the null hypotheses of under-identification and weak identification are both rejected, again indicating that the instrument is strong enough to identify the equation.

The estimates of both regressions confirm, as expected, that a higher per capita GDP growth rate ratio leads to an increase in the income ratio from 1999 to 2002. In addition, migration is associated with a decrease in the income ratio. This result is also consistent with expectations, since migration here serves as a proxy for remittance inflows. According to these estimates, the income gap between (poorer) sending provinces and (richer) receiving provinces would diminish either if per capita GDP growth rates were relatively higher in sending provinces, or if they could send more migrants (and by implication, receive more remittance income).

In elasticity terms, the impact of migration seems much smaller compared to the impact of the per capita GDP growth rate ratio. In model I, a 1% increase in the migration rate is associated only with a 0.27% decrease in the income ratio, while a 1% increase in the per capita GDP growth rate ratio is associated with an increase of 1.35% in the income ratio (elasticities are computed at the means of the data). The IV-GMM estimates (model II) differ only in that the coefficient on the instrumented per capita GDP growth rate ratio is much larger. In model II, the percentage increase in income ratio due to a 1% increase in per capita GDP growth rate ratio is 6.7%, much higher than that in model I, and definitely much higher than the impact of migration, which stays the same for both models.

These regressions give average impacts of migration on changes in income differential when all provinces are pooled. But the analysis in Section 3 suggested that not all migration flows are similar; the liquidity constraint may be more binding for some flows than for others. In addition, we know that migrants to different regions/

Table 5
Migration's impact on income ratio.

	(I) OLS		(II) IV-GMM 2nd stage		(III) IV-GMM 1st stage	
	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
1999–2002 per capita GDP growth rate ratio	0.040	4.71	0.197	6.83	–	–
1994–2002 per capita GDP growth rate ratio	–	–	–	–	0.452	11.31
1994–1999 migration rate	−12.289	−3.25	−13.361	−2.98	13.537	0.85
Sending region = North East	0.199	12.38	0.233	12.76	−0.197	−4.32
Sending region = North West	0.261	13.28	0.238	10.91	0.109	1.85
Sending region = North Central Coast	0.182	11.28	0.188	10.73	−0.015	−0.3
Sending region = South Central Coast	−0.005	−0.27	−0.004	−0.21	−0.006	−0.11
Sending region = Central Highland	0.309	8.78	0.313	8.1	0.088	0.66
Sending region = South East	0.060	2.59	0.052	1.41	0.038	0.25
Sending region = Mekong River Delta	0.101	5.81	0.098	4.66	−0.053	−0.75
Sending region = urban areas	−0.033	−0.9	0.002	0.05	−0.169	−1.38
Receiving region = North East	−0.103	−6.94	−0.125	−8.55	0.019	0.56
Receiving region = North West	−0.219	−7.08	−0.186	−6.08	−0.201	−5.96
Receiving region = North Central Coast	−0.114	−6.38	−0.106	−6.25	−0.063	−1.97
Receiving region = South Central Coast	−0.034	−2.03	−0.026	−1.54	−0.018	−0.49
Receiving region = Central Highland	−0.586	−19.56	−0.583	−17.61	−0.154	−2.04
Receiving region = South East	−0.020	−1.16	−0.029	−1.57	−0.084	−1.52
Receiving region = Mekong River Delta	−0.112	−7.92	−0.140	−9.04	0.261	5.49
Receiving region = urban areas	0.092	3.68	0.070	2.61	0.031	0.48
Constant	−0.114	−6.26	−0.287	−8.29	0.613	11.2
N	1830		1830		1830	
Adjusted/centered R^2	0.45		0.30		0.10	
F-statistics for relevance of excluded instrument					128	
F-statistics for Wu–Hausman exogeneity test			40			
Chi-squared statistics for under-identification test			119			
F-statistics for Stock–Yogo weak identification test			127			

Note: Critical value for weak identification test is 16.38 for 10% maximal IV size.
Dependent variable in (I) & (II) is change in income ratio from 1999 to 2002.
Dependent variable in (III) is 1999–2002 per capita GDP growth rate ratio.
Instrument in model (II) is 1994–1999 per capita GDP growth rate ratio.

provinces go into different sectors, which in a developing economy may imply entirely different returns and costs. For example, migrants to the HCMC cluster tend to be temporary, young, single, and relatively well educated (high school education), and they mostly get manufacturing jobs. They are most likely from the middle part of the income distribution in the sending province. Migrants to Hanoi and the other northern cities also tend to be young and single, but they do not enter the manufacturing sector in anything like the same proportions. Hanoi is one of Vietnam's fastest growing cities, and has attracted a lot of migrants in the doi moi era, but as this is the country's political, cultural and educational center, with relatively little manufacturing sector growth by comparison with HCMC and surrounds, its growth is mostly in services, including financial sectors and government. Migrants to Hanoi are more likely to be educated professionals, and must also a large number of college students who stay after graduation in high-skill jobs. These migrants are most likely from the upper part of the income distribution in sending provinces, so the brain drain effect might be more severe for migration flows into Hanoi. Migrants to the Central Highlands, a rural region that has attracted a great deal of migration in the past three decades, tend to involve the whole family, to be long-term or permanent, and to enter the agricultural sector, especially coffee and other cash crops. They are more likely to come from the lower end of the income distribution in the origin. In short, there are various reasons to expect the impact of migration on income differential to vary with the receiving provinces.

With this motivation we re-estimate the above regressions for different groups of receiving provinces (Table 6). The results reveal strong contrasts. Migration flows to the HCMC cluster continue to have a negative impact on income differential, but surprisingly, migration to major cities in the North has a positive impact. Finally, migration flows to the Central Highland have a negative, but statistically insignificant impact on income differentials.

Summing up, we have found evidence of a negative impact of migration on provincial income inequality in the pooled data. This impact varies widely in subsets of the data, depending on the receiving provinces, and appears to be related to industrial structure in those provinces. Most importantly, for the HCMC cluster, which is by far the largest and most rapidly-growing labor market in Vietnam, we find a negative impact of migration on the income ratio between pairs of provinces.

It is worth re-emphasizing that with the data available to us we have not attempted to study inequality at the household level, nor inequality within each sending or receiving province, nor inequality for Vietnam as a whole. Given the available data and the aggregate level of analysis, we simply examine how “changes” in the income ratio of pairs of provinces vary with migration flows between them. This sheds light on inequality at the provincial level, and is meant to be complementary to household-level analysis.[12]

5. Conclusions and directions for further research

In this paper we have examined two questions: that of the determinants of inter-provincial migration flows in Vietnam, and that of the impact of these flows on inter-provincial inequality. Regarding the first question, we find that migration flows follow patterns predicted by theory: people move from low-income to high-income provinces. There is also evidence of a liquidity constraint effect which leads to poverty-related labor immobility at the provincial level.

Regarding the second question, on average migration is found to reduce inequality. Larger migration flows are associated with greater narrowing of the income ratio between pairs of provinces. But this

[12] A recent addition to this literature is Hoang et al. (forthcoming), who find that internal migration is pro-poor but also increases inequality within the sending region.

Table 6
Migration's impact on income ratio, selected receiving provinces.

	(I) Receiving provinces = HCMC, Dong Nai, Binh Duong		(II) Receiving provinces = Hanoi, Hai Phong, Quang Ninh		(III) Receiving region = Central Highland	
	Coeff.	*t*-stat	Coeff.	*t*-stat	Coeff.	*t*-stat
1999–2002 per capita GDP growth rate ratio	0.289	3.34	0.204	2.44	0.161	0.88
1994–1999 migration rate	−16.18	−1.89	48.88	5.68	−41.21	−1.57
Sending region = North East	0.329	3.96	0.409	5.84	0.183	1.76
Sending region = North West	0.381	4.11	0.489	6.37	0.164	0.98
Sending region = North Central Coast	0.314	4	0.394	6.25	0.192	1.83
Sending region = South Central Coast	0.034	0.36	0.197	2.89	0.030	0.26
Sending region = Central Highland	0.635	2.55	0.667	3.92	0.235	1.16
Sending region = South East	0.193	1.69	0.245	2.06	0.050	0.43
Sending region = Mekong River Delta	0.268	3.63	0.330	5.65	0.069	0.69
Sending region = urban areas	0.102	0.95	0.132	1.51	0.049	0.45
Constant	−0.431	−4.38	−0.420	−3.27	−0.760	−4.1
N	172		156		140	
Centered R^2	0.00		0.11		−0.10	
F-statistics for Wu–Hausman exogeneity test	11.4		6.2		1.2	
Chi-squared statistics for under-identification test	14.0		15.1		4.2	
F-statistics for Stock–Yogo weak identification test	14.5		18.0		4.0	

Note: Critical values for weak identification test is 16.38 and 8.96 for 10% and 15% maximal IV size.
Dependent variable in all regressions is change in income ratio from 1999 to 2002.
All regressions were run using 2-step IV-GMM, and the instrument is 1994–1999 per capita GDP growth rate ratio (1st stage results available upon request).

impact varies widely, depending on the receiving province. For migration flows into the manufacturing center in the South, where due to globalization-driven changes in investment and production structure, the labor market has been growing fastest, the inequality-reducing impact of migration is confirmed. But migration flows into urban centers in the North have positive effects. These results suggest that the impact of migration on income inequality between pairs of provinces is contingent on the impetus for the flow. Migration has a significant association with reduced inequality when the destination provinces are those hosting the greatest concentration of labor-intensive, export-oriented manufacturing industries. Other forms of growth, such as have occurred in the other net in-migration regions of Vietnam, do not. A deeper investigation of the reasons for this difference requires microeconomic data on migrants, jobs, and household incomes. This could permit testing of the relative magnitudes of 'brain drain' effects versus the gains from remittance income.

Everything considered, the evidence confirms that economic growth and internal migration are complements. On the one hand, migrants respond to increases in the marginal productivity of labor in sectors where Vietnam is exploiting its comparative advantage and acquiring new investments; that is, migration fuels growth. On the other hand, regression analysis indicates a robust and negative impact of migration on income inequality for those migration flows going to the largest manufacturing centers. This implies that migration helps offset some of the increase in spatial inequality caused by location- and sector-specific growth. At the same time, however, impediments to migration, such as liquidity constraints due to low incomes and imperfect capital markets, may imply persistence of poverty for populations in disadvantaged locations. Such evidence suggests that policies facilitating internal migration will be good both for economic growth and inequality reduction. It is important to note, however, that such policies must ensure broad-based access to migration across households and regions, so as to avoid increasing inequality within the sending areas.

The most serious constraint to the type of analysis employed in this paper is the restriction posed by aggregate data. Such data fail to account for different types of move, and specifically, do not do a good job of capturing return and circular migration. With these data, moreover, we observe only average characteristics and trends at provincial level; differences across individuals or among subgroups of the population are not taken into account in evaluating the decision to migrate. Finally, the gross migration rate m_{sr} for aggregates is an unbiased estimator for the underlying individual probability of migrating from s to r only if the characteristics affecting migration decisions are identically and independently distributed across households/individuals, such that behavior of a representative agent mimics that of the aggregate. This assumption is likely to be violated; conclusions on individual migration behavior using aggregate data should thus be drawn with care. The advantage of using these data, for all their limitations, is that they alone are nationally representative at a provincial scale. By gathering migration, employment and remittance data for nationally representative samples of households and individuals, future rounds of the national living standards survey could open the door to considerably richer modes of analysis.

References

Andrienko, Yuri, Guriev, Sergei, 2004. Determinants of inter-regional mobility in Russia. Economics of Transition 12 (1), 1–27.

Asian Development Bank, 2007. Rural labor markets and migration. Making Markets Work Better for the Poor Discussion Paper #18. Asian Development Bank, Hanoi.

Barro, Robert J., Sala-i-Martin, Xavier, 2004. Economic Growth. The MIT Press, Cambridge Massachusetts.

Banerjee, Biswajit, Kanbur, S.M. Ravi, 1981. On the specification and estimation of macro rural–urban migration functions: with an application to Indian data. Oxford Bulletin of Economics and Statistics 43, 7–29.

deBrauw, Alan, Harigaya, Tomoko, 2007. Seasonal migration and improving living standards in Vietnam. American Journal of Agricultural Economics 89 (2), 430–447.

Connell, J.B., Dasgupta, B., Laishley, R., Lipton, M., 1976. Migration from rural areas: the evidence from village studies. Oxford University Press, Delhi.

Courchene, Thomas, 1970. Interprovincial migration and economic adjustment. The Canadian Journal of Economics 3 (4), 550–576.

Dang, Nguyen Anh, 2003. Migration in Vietnam: review of information on current trends and patterns, and their policy implications. Paper presented at the Regional Conference on Migration, Development, and Pro-Poor Policy Choices in Asia, Dahka, Bangladesh, 22–24 June 2003.

Deaton, Angus, 2006. Measuring poverty in a growing world (or measuring growth in a poor world). The Review of Economics and Statistics 87 (1), 1–19.

Dhar, Sanjay, 1984. Inter-state and within-state migration in India. In: Binswanger, H.P., Rosenzwe ig, M.R. (Eds.), Contractual Arrangements, Employment and Wages in Rural Labor Markets in Asia. In: Economic Growth Center. Yale University, pp. 281–304.

Fan, C. Cindy, 2005. Modeling interprovincial migration in China 1985–2000. Eurasian Geography and Economics 46 (3), 165–184.

Greenwood, Michael J., 1997. Internal migration in developed countries. In: Rosenzweig, M.R., Stark, O. (Eds.), Handbook of Population and Family Economics. Elsevier Science B.V.

General Statistics Office, 1991. Population and Housing Census Vietnam 1989: Completed Census Results, General Statistics Office. Statistical Publishing House, Hanoi.

General Statistics Office, 2001. Population and Housing Census Vietnam 1999: Completed Census Results, General Statistics Office. Statistical Publishing House, Hanoi.
Guest, Philip, 1998. The dynamics of internal migration in Vietnam. UNDP Discussion Paper 1. Vietnam, Hanoi.
Hardle, W., Muller, M., Sperlich, S., Werwatz, A., 2004. Nonparametric and Semiparametric Models, Springer Series in Statistics. Springer-Verlag, Berlin.
Hoang, Thanh Ha, Nguyet, Nguyen Thi, Phuong, Nguyen Thu, Oostendorp, Remco, Tam, Tran Ngo Thi Minh "Determinants and Impacts of Migration in Vietnam," forthcoming, Market, Policy and Poverty Reduction in Vietnam, National Political Publisher (Henrik Hansen and Nguyen Thang, eds.), Hanoi (Vietnam).
Lucas, Robert E.B., 1997. Internal migration in developing countries. Handbook of Population and Family Economics, pp. 721–798. Chapter 13.
Mekong Economics, 2002. A study of trade, foreign direct investment, and labour in Vietnam. An input to the DFID, ESCOR funded projected on Globalisation, Production, and Poverty: Macro, Meso, and Micro Level Studies, Hanoi.
Mueser, Peter, 1989. The spatial structure of migration: an analysis of flows between states in the USA over three decades. Regional Studies 23.3, 185–200.
Nguyen, Duc Vinh (2003). "Urban savings and remittance in Vietnam" Migration in Vietnam: Theoretical Approaches and Evidence from a Survey, ed. Dang, Nguyen Anh.
Phan, Diep, 2008. Migration and credit constraints: evidence from four provinces in Vietnam. Manuscript. Department of Agricultural and Applied Economics, University of Wisconsin-Madison.
Phan, Diep, Coxhead, Ian, 2008. Inter-provincial migration and inequality during Vietnam's transition. Staff Working Papers No. 507. Department of Agricultural and Applied Economics, University of Wisconsin-Madison.
Yang, Liu (2004). "Unequal Provinces but Equal Families? An Analysis of Migration and Inequality in Thailand" Essays on the Determinants and Consequences of Internal Migration, PhD dissertation, University of Chicago 2004.