Migration Enclaves, Schooling Choices and Social Mobility

Mario Piacentini

University of Geneva, Department of Economics

1. March 2008

Migration Enclaves, Schooling Choices and Social Mobility

Mario Piacentini*(University of Geneva)

April 22, 2008

Abstract

PRELIMINARY DRAFT

This paper investigates the presence of a network externality which might explain the persistence of low schooling achievements among internal migrants. A simple analytical framework is presented to show how an initial human capital disparity between migrants and non migrants can translate into persistent skill inequality if origin shapes the composition of social networks. We test empirically whether young migrants' schooling decisions are affected by the presence of covillagers at destination, using data on life-time histories of migration and education choices from a rural region of Thailand. Different modelling approaches are used to account for the self-selection of young migrants, for potential endogeneity of the network size, and for unobserved heterogeneity in individual preferences. The size of the migrant network is found to negatively affect the propensity of young migrants to pursue schooling while in the city. This finding suggests that policies seeking to minimising stratification in enclaves might have a socially multiplied impact on schooling participation, and, ultimately, affect the socio-economic mobility of the rural born.

Keywords: education, networks, migration

JEL No: I21,L14,O15

*I am grateful to Jaime de Melo for valuable advice. Helpful comments from participants at the graduate seminar of the department of economics at the University of Geneva, from participants at the Young Swiss Economist Meeting at the University of Bern, from Richard Upward and participants at the 7th GEP conference at the University of Nottingham are appreciated as well. I am also thankful to Eik Leong Swee for sharing his rain data and programs. All errors are mine. *Corresponding address*: Department of Economics, Uni Mail, Bd du Pont d'Arve 40, 1211 Genève 4, Switzerland. *E-mail:* mario.piacentini@ecopo.unige.ch

1 Introduction

"Human capital accumulation is a social activity, involving groups of people in a way that has no counterpart in the accumulation of physical capital" Robert E. Lucas (1989)

"*People in the city and people in the village aren't the same. City people, Bangkok people, you can't trust them, they only think of themselves. In the city people don't know each other. I've lived in this room for many months now and I still don't know the neighbors. In the village I know everyone. We grow up together, we're all relatives and friends together. I know where they come from, their background. I can trust them.*" Daeng, a 20-year-old textile operative quoted in Mills (1997)

Disparities in growth between urban and rural areas, as well as reduction in migration costs, act as powerful pull factors augmenting the demographic pressures on overcrowded cities of developing countries. It is tempting to assert that the access to more remunerative employment opportunities in urban agglomerates increases the expected returns on education and therefore acts as an incentive to invest in human capital for those planning to migrate. However, young migrants take the decision to pursue schooling beyond a literacy level jointly with their family and influenced by the broader social network they belong to. It is increasingly recognized that differences in the composition and exclusion mechanisms of these networks affect opportunity costs of higher education.

Economic research on network effects and welfare of the migrant population has produced contrasting results. One strand of the literature emphasizes that the reliance of migrants on origin-specific social capital can be associated with a lower rate of assimilation of destination-specific skills. Among others, Borjas (1995) has shown that ethnic neighborhoods have detrimental effects on the educational attainment of migrants in the US. A competing hypothesis is that ethnic or origin-based concentration of migrants is a source of opportunities for gainful interactions in the labor market, for example by disseminating information on job opportunities. Banerjee (1983)'s research on rural-urban migration in India documents extensively the importance of networks of covillagers for explaining migration flows, success of initial job search, and duration of urban employment. With evidence in support of both positive and negative effects, it is unclear how segmentation along enclaves affects economic performance of migrants.

This paper contributes to this literature by testing the hypothesis that migrant networks might act as important externalities in the education process of young migrants, representing a potential determinant of the observed low educational attainments. A simple analytical framework, derived from recent contributions on group inequality (Bowles & Sethi (2006), Bowles

et al. (2008)), shows how a family living in a disadvantaged community can optimally decide to invest relatively less in schooling when network externalities are in place. In this framework, belonging to the group with the lower average human capital is associated with an higher cost of investment in education; inequality in educational attainment of individuals from different groups is shown to persist across generations if segregation in enclaves is sufficiently high. In our application of this framework to the issue of migrant assimilation, an important innovation is the identification of two mechanisms through which segregation can endogenously emerge. The first draws on Benabou (1996)'s research on stratification in the urban space, leading migrants to concentrate in neighborhoods less favorable to human capital investments. The second argues that migrant families care about being accepted among members of their own origin community, and that they might be willing to reduce their investments in human capital to avoid the costs of exclusion.

We test the relevance of this network externality using unique data from Thailand, the Nang Rong Project database. This dataset provides direct individual and family information on both migrants and stayers, through longitudinal surveys in the villages and migrant follow-ups in the main urban destinations. It also includes rich retrospective information on migration and education decisions. We can thus work with a panel data set of individual location decisions and schooling outcomes, from multiple communities and to multiple destinations, over a long period of time. Differently from the majority of contributions in the migration literature, we can take advantage of this balanced representation of migrants and stayers to control for the fact that those choosing to migrate and to join the migrant network at destination are not a random sample of the rural young population.

For the empirical analysis, we follow an approach close to Munshi (2003)'s, who provides a solid empirical illustration of the positive effects networks can play for labor market outcomes of mexican migrants. His instrumental variable analysis succeeds in controlling for the effects of correlated shocks among migrants from the same origins, using regional random rainfall variation from the area of the migrants to identify the network effects. Differently from Munshi (2003), we focus on the schooling decision of young internal migrants from the age of 13, as higher education is believed to be an important proxy for social mobility of migrants within the urban labor market[1]. The young migrant's network is measured by the number of sampled individuals in his village who are located at the destination at each point in time. Identification of network effects is based on the fact that each village has a different history of migration, so that the rural young coming from different villages rely on differently developed networks of contacts.

[1] An increasing number of studies shows the importance of education for migrant assimilation and long-term earning prospects. Yamauchi (2003), examining wage dynamics of migrants in Bangkok, shows evidence of a complementarity between upon-arrival human capital and labor market experience at destination. His results imply that more-educated migrants have higher learning efficiency and can perform tasks of greater complexity, ultimately yielding higher wage growth in the destination market. In the context of international migration, some important contributions have shown that the source of human capital matters, the one acquired at destination being more rentable on the labor market (see Friedberg (2000) and Eckstein & Weiss (1998))

The aggregation of networks at the tiny village level is another departure from Munshi[2]: data evidence validates this measure as a good approximation of the true composition of the migrant's reference group[3]. We instrument network size by the exogenous source of variation associated with random, ballot-based, assignment of young to the military service. The forced movement to the cities of the rural young balloted to serve in the military tends in fact to be accompanied by increased voluntary migration of these same young and others villagers in later years. Additional instruments for network size are the exposure of the village to rain shocks, and a lagged measure of the incidence of return migration to the village.

The estimates show that a larger network of co-villagers at destination lowers the probability that a young migrant is enrolled in higher schooling. We interpret this result as an indication that social interactions with the "origin" network matter in determining the migrant "long term" profile, in terms of willingness to acquire destination-specific human capital and, possibly, in terms of the speed at which he is able to converge to natives' performances on the labor market. The robustness of the results is confirmed when we shift to an alternative multinomial mixed logit approach for the joint migration and education decision, which accounts explicitly for individual heterogeneity.

These results suggest that network effects can act as a driving mechanism of low-mobility traps among migrants, at least in the context of this study. As a result, the important role networks play in facilitating migration and easing life at destination might come at a cost, slowing down the convergence of people of rural origin to the average level of skills of the urban born. Considering the strength of the links migrants keep with their origin villages (by remittances, return episodes and as role models), it cannot be excluded that this network effect on human capital might reinforce persistence in inequality between backward rural areas and dynamic urban poles of developing countries.

The remainder of the paper is organized as follows. In Section 2 we present a theoretical framework giving important insights on the potential effects of social networks on educational investments of migrant minorities. Section 3 describes the data and the network measure used. Section 4 illustrates the econometric analysis and summarizes the results. Section 5 concludes.

2 Modelling insights from the theory of segregation

To study the implications of segregation for differential human capital accumulation and persistent inequality consider the following setting, based on the theoretical studies of Bowles and

[2]We consider this as a significant data advantage, making us more confident that we are actually capturing social relations between migrants. On average, each village in our data has a population of less than 600 individuals, much lower than that of a community in the mexican data used by Munshi (2003). Our network measure is also specific to each urban destination, that have a geographical extension much lower than the one of US States in Munshi (2003).

[3]Thai villages were historically fairly tightly bounded social communities and traditional village life was relatively isolated, generating a common community culture and encouraging behavioral conformity (Godley (2001)). Of course, this static picture is evolving now with the increasing migration flows to the cities.

Sethi (2006), and Bowles, Loury and Sethi (2008). Individuals belong to two groups, Natives (N) and Migrants (M), and have a preference for association with individuals of their group, which is defined as segregation. There is an imperfect mixing of migrants and natives. This diversity in social network composition translates in different opportunity costs of human capital accumulation, as an externality is in place linking each training achievement of a young with those of the other agents to which he is tied. Inequality in the allocation of skills between migrants and non migrants is shown to widen and persist in time if segregation is high enough. This result is crucially linked to the assumption of a positive, exogenously given level of segregation. Drawing on the work of Bénabou (1996) on the role of skill differences in determining residential choices in the urban space, and on the work of Austen-Smith & Fryer (2005), in which agents care about social acceptance in their own group when taking decision on human capital accumulation, we "endogenize" the networking process, giving us even stronger foundations to the negative link between migrant networks and schooling enrollments observed in the data. This provides the foundation for the empirics, which emphasizes the role of elective participation in social networks, and tries to account for the endogenous development of these networks.

2.1 Basic setting

We here follow closely the development and notation in Bowles, Loury and Sethi (2008). Consider a scenario with overlapping generations and families composed by one parent and one child, living in a city for two periods. In the first period of life parents can decide to invest in the training of the child, while no human capital accumulation occurs in the second period of life when everyone works for a wage. Time is indexed by $t, t = 0, 1...T$. Agents belong by birth to one of two groups $V = N, M$, the natives of the city N and those of rural origin (migrants and descendents of migrants) M. The proportion of N and M in the total urban population is exogenously given respectively by n and $1 - n$. Agents are characterised by their skill (human capital) level s, which can take two values, being equal to l if the individual stays unskilled and equal to h if the individual acquires skills when young.

There is a complementarity between high and low skill labor in the process of production. The total output of the city economy in period t is given by the production function $q(h_t, l_t)$, where h_t and l_t are the proportions of workers employed in high-skill and low-skill jobs, respectively.Let $\overline{s_t}$ denote the proportion of skilled individuals at each period in the overall economy: this is simply a weighted average of s_t^N and s_t^M, the average education levels in the native and in the migrant group, respectively:

$$\overline{s}_t = n s_t^N + (1 - n) s_t^M \tag{1}$$

The production function q satisfies constant return to scale, diminishing marginal returns to each factor and the conditions $\lim_{\overline{s}\to 0} q_1 = \lim_{\overline{s}\to 1} q_2 = \infty$.

Competitive firms assign two levels of wages according to the skill of the worker ($w_h > w_l$): the wage differential $\delta(\overline{s}_t)$ is positive and decreasing in $\overline{s}$.

Assume that in the first period there is a higher proportion of skilled among adult natives than among adult migrants ($s_0^M < s_0^N$) : this assumption is not strong if one allows that skills acquired at the rural village are not fully transferable at destination.

Each individual's social network is composed by a mix of individuals from one's origin group and from the other group. Suppose, as in Bowles and Sethi (2006), that a proportion η of each agent's social affiliates is drawn from his own group of origin, while the remaining $1-\eta$ is randomly drawn from the overall city population. Denote by L^V the mean level of human capital in the social network of an individual belonging to either the native or migrant group. This index of the human capital quality of one's social network depends on the levels of human capital in each one of the two groups and on the extent of segregation η, as follows:

$$L_t^V = \eta s_t^V + (1-\eta)\overline{s}_t \tag{2}$$

It should be clear that in the absence of perfect integration ($\eta > 0$), L_t^M will be lower than L_t^N as long as s_t^M is lower than s_t^N.

Parents fully internalize the preferences of their children with no discounting for the future, so that training is undertaken if the benefit for the child outweighs the cost for the parent. The cost of training a young in any generation is given by the function $c(a_{t+1}, L_t)$, strictly decreasing in both the arguments a, child ability, and L, social network's quality. The distribution of ability is given by the function $G(a)$, with support $[0, \infty)$, assumed to be the same across groups. The benefit of training is simply the wage differential $\delta(s_{t+1})$, which is the same across group, i.e. no explicit labor market discrimination is in place.

2.2 Training decisions and persistence in inequality of skills

A threshold level of ability $\widetilde{a}$ can be defined for each of two groups in $V = N, M$, so that only the young of each group who have ability above this threshold are trained. This threshold level, $\widetilde{a}_{t+1}$, is implicitly defined by the value of ability satisfying

$$c(\widetilde{a}_{t+1}, L_t^V) = \delta(s_{t+1}) \tag{3}$$

The following dynamics of human capital accumulation apply to each group:

$$s_{t+1}^V = 1 - G(\widetilde{a}_{t+1}(\delta(s_{t+1}), L_t^V)) \tag{4}$$

simply meaning that only those young having ability greater than $\widetilde{a}_{t+1}(\delta(s_{t+1}), L_t^V)$ become skilled workers in period $t+1$.

It is clear that increased segregation, by affecting the quality of the social network L_t^V, raises the costs of human capital accumulation for the disadvantaged group, the internal migrants, while lowering these costs for the natives.We now aim to show that, given a positive level of segregation and the spillovers in human capital accumulation described above, optimal training decisions of the native and the migrant group can be persistently different over time.

Define a competitive equilibrium as a sequence of skill allocations across groups $\{(s_t^N, s_t^M)\}_{t=1}^{\infty}$ satisfying equations (1-4). The main result is summarized in the following proposition:

Proposition 1 *(Bowles et al.(2008)). Given any initial allocation* $(s_0^N, s_0^M) \in [0,1]^2$, *a unique equilibrium path* $\{(s_t^N, s_t^M)\}_{t=1}^{\infty}$ *exists. Along this equilibrium path, if* $s_0^M \leq s_0^N$, *then* $s_t^M \leq s_t^N$ *for all t.*

The intuition is that an initial situation of skill disadvantage of one group can not be reversed along the equilibrium path, which is unique given an initial state (s_0^N, s_0^M). Bowles, Loury and Sethi (2008)'s formal proof of the uniqueness of the equilibrium path and of persistence of inequality is reproduced in the appendix. The appendix also provides an extensive treatment of steady state behaviour of the model, a steady state being defined as an equilibrium in which $s_t^V = s_0^V$, $\forall\ t \geq 1$. The most interesting result concerns the instability of the unique *symmetric* steady state - symmetric meaning that the additional condition $s_t^N = s_t^M$ is satisfied - if the segregation level η and the size of the interpersonal spillover are sufficiently high. This indicates that convergence in skill levels among groups is not achievable even in the very long run, group inequality persisting asymptotically.

2.3 Extensions accounting for determinants of segregation and network externalities

The model identifies conditions under which the combined effect of network spillovers in the cost of acquiring human capital and an origin bias in the composition of social networks can be sufficient for internal migrants to be locked in a condition of low education achievements. In fact, persistent group inequality can arise even in the absence of discrimination in the labor market, and with equality of opportunity, that would hold if liquidity constraint for education are not binding. This conclusion rules out one-shot redistributive policies as effective instruments for equalization. However, the scenario for policy becomes less somber if one is willing to get rid of the assumption of a level of segregation η exogenously given and fixed. Below, we provide intuitions on two possible extensions of the model that endogenize η, thereby increasing the scope for policy interventions.

2.3.1 Social segregation and residential enclaving

One potential determinant of origin-biased social networking is the residential segmentation of groups in city space. Think of the utility maximization problem faced by the adult as involving also a location decision among two different neighborhoods, each parent trading off the benefit of a better environment for his child's learning with the cost of higher rents. Equilibrium in the city space is an allocation of households across neighbourhoods and rents such that no household prefers a neighborhood different from its own. Bénabou (1996) proves that this equilibrium will result in location stratification (those from one specific community preferring to locate in one specific neighborhood), if the rich in human capital are able to bid more than the poor

for living in the community with a higher endowment of human capital. Formally, Bénabou's sorting condition says that the marginal rate of substitution between the network quality and the rent price increases with the parent level of education. When this sorting condition holds, any divergence from the symmetric initial allocation of skill, $s_0^N = s_0^M$, sets in motion a cumulative process with skilled (higher income) people outbidding unskilled (lower income) people for the privilege of locating close to other skilled (higher income) people. This process ends when at least one neighborhood is completely homogeneous. At least three distinct determinants of segregation can be identified by extending the model so as to account for spatial segmentation à la Bénabou: a) complementarities between family human capital and community quality; b) imperfections in capital markets resulting in the poor having a relatively higher opportunity cost of borrowing; c) differences in family life time resources[4]. Segregation in the city space is a variable that can be affected by policy: differentiated taxes and subsidies can be effective in reallocating families across communities (Bénabou, 1996), and urban regeneration programs can raise the attractiveness of neighborhoods where there is an historically higher concentration of migrants[5].

2.3.2 A "social signalling" effect?

As spatial and social proximity are likely to be correlated, the allocation of individuals in stratified neighborhoods can alone generate segmentation. Other mechanisms can explain why a migrant can rationally *prefer* to maintain segregated social relations. We have in mind a scenario in which the young derive utility from being part of his origin group and admission to the group is selective. It seems reasonable to assume that migrants can diverge in the value they assign to group participation, and that the migrant community might preserve its cohesion by excluding families signalling little interest in participation by "deviant" behaviours. One of these behaviours can be an investment in the schooling progression of the young which is believed excessive by the group[6].

Austen-Smith and Fryers (2005) develop a two-audience signalling model in order to explain

[4] Complementarity here simply means that families with higher human capital are more sensitive to neighborhood quality than those with lower levels of human wealth.We refer to Bénabou (1996) for a detailed explanation of these determinants of the segregation process. The original model of Bénabou allows for another determinant of stratification, the level of decentralised expenditures in education financed through taxes on the local population. Basically, familiar and public inputs determine together the neighborhood-specific, per-student budget. Obvious complementarities between parental human capital and local public expenditures on education supply are enough to make the integrated equilibrium unstable (Bénabou, 1996).

[5] This issue is acknowledged by policy makers. The housing projects of the National Housing Authority (NHA) and the Bangkok Metropolitan Administration (BMA) have engaged in ensuring some mix of housing at all price range in each area. In addition, the Bangkok Plan addresses existing spatial disparities in the location of jobs and housing, and encourages balanced jobs and housing growth in each of the city planning units (see Tapananont (2004)).

[6] Various motives can explain why the migrant community can interpret higher schooling as a signal of low social attachment. One reason is that the school enrollment of young already able for work can make a migrant family temporarily unable to meet some social obligations, like sending to the village a minimal amount of remittances.

the phenomenon of "acting white", according to which young blacks in the US make a low effort in school as they are exposed to peer pressure condamning behaviors perceived to be characteristic of whites. As in their model, a network effect on training attainments can emerge in our framework if the market and the social network value differently an educational investment of a young migrant. Add then the assumption that each individual migrant's type is defined by a pair, $\tau = (\alpha, \gamma)$, where α is intrinsic ability as above and γ is the individual's social type (his innate compatibility with his own community) which takes two values, high or low, and is only privately observed[7]. Individuals have an interest in being accepted members of their community: assume, for simplicity, that utility of time not spent acquiring education is augmented by a factor $\lambda(\gamma)$ if one is accepted by his origin community. The migrant network, being imperfectly informed about the individual's attachment to his origin community (if the individual social type is high or low), can interpret an investment in destination specific human capital as a signal of divergence and sanction it with exclusion from community life[8]. In the framework above, no change occurs for the native group. However, for the migrant group, the rate of preference for origin-biased networking η is now different from zero only for an endogenously defined subset of individuals for whom the utility value of group participation outweighs the costs in terms of wage income foregone. The threshold ability level $\widetilde{a}$, needed by these individuals to be trained, increases with respect to the scenario with no need for social signalling. When the social type is private knowledge then, conflicting incentives to invest in destination specific human capital and to signal loyalty to the group can be a channel through which historical group differences in human capital levels spill over into next generation's investment behaviour.

Anything altering the trade-off between income gains and value of group participation will affect the incidence of the network externality. For example, a productivity shift raising the wages perceived by the migrants on the market would raise the opportunity costs of community participation and lower the proportion of those willing to sacrifice education for group acceptance. In the other direction, access to a larger community at destination can be expected to increase the value of group participation, and thus the relevance of the externality on individual schooling decisions.

[7] In another class of models, incomplete information plays a role in education decisions via information costs. In this case it can be rational to use the action of others as a signal instead of acquiring own information (see Streufert (2000)).

[8] The group is assumed to have interest only on individuals whose social type is high, having no direct concerns with such types' educational levels. In Austen Smith and Fryers (2005) model, this is formalized by assuming a fixed, non negative, payoff for the group from rejecting any individual. This anthropomorphization of the community group is consistent with other studies on social networking and under-development traps. Hoff & Sen (2005) show that control mechanisms at disposal of the peer network can generate a poverty trap when economic opportunities outside the origin community widen. An interesting study by Munshi & Rosenzweig (2003), with survey data from Bombay, shows tha networks of lower caste male channel boys into local language schools leading to traditional occupations, despite the substantial rise of returns to nontraditional occupations.

3 Nang Rong data and network measures

The model above predicts a direct relation between the weight of people from one's origin group in the social network and individual human capital achievements. A higher network size increases the likelihood of interactions with people from one's origin community and it is thus expected to be positively correlated with the segregation index η. If a "social signalling" effect exists, it is likely to reinforce the externality in the cost of education, as social contacts will matter more as determinants of similar behaviours within a group. We test the empirical relevance of this network effect for education decisions of young migrants using data on rural-urban migration from Thailand. The Thai economy experienced a rapid development process in the period under study, with an average growth rate of real GNP per capita of 5.7% per year between 1976 and 1996. However, there are concerns over the growing rural-urban divide in the same period, as the income Gini coefficient passed from 0.436 in 1976 to 0.515 in 1996, a level even higher than the average one for Latin America and the Caribbean in the same year (Jeong (n.d.)).

The data come from a collection of research surveys of social, economic and environmental change in the district of Nang Rong, historically one of the least-developed parts of Thailand. The Nang Rong project dataset consists of three waves of data collection - for the years 1984, 1994, 2000. A migrant follow-up survey was added to track a sample of migrants who had gone to one of the four following urban destinations: (1) metropolitan Bangkok, (2) the Eastern Seaboard, a highly dynamic area comprising the two urban centers of Rayong and Chonbury, (3) Korat, an important regional pole, and (4) Buriram, the provincial capital. The 2000 round builds on the previous data collection efforts incorporating a geo-spatial component in addition to the community, household and migrant follow-up surveys. Moreover, the 1994 and 2000 surveys undertook the innovative task to identify both social and kinship networks among the residents, households and villages of Nang Rong. These ties are measured directly through kinship, labor exchanges, and agricultural equipment exchanges.

In the estimations, we mainly exploit the life history sections of the 2000 data, collected on the sample of individuals who are resident in Nang Rong at the time of the survey, and on the sample of those residing in one of the major urban destination at that time. Life history, or retrospective, data provide long-term information on migration and schooling choices, yielding an unbalanced panel with a minimum of one observation for those aged 13-year-old in 2000, to a maximum of 13 observations for those aged 25-year-old or older in 2000. The analysis thus extends to multiple cohorts of individuals, and cover a period of more than 30 years. As mentioned in the introduction, the most salient aspect of these data comes from the merging of information on migrants and non migrants, which provides a picture of the rural community not bounded by the rural district and thus closer to its true, geographically mobile, configuration. Previous empirical research on migration has often suffered from a deep selection problem, limiting the observation to those who are migrants at a given point in time. The fact that these migrants are not a random sample of the origin population bias inference, selectivity occurring

along many and mostly unobserved characteristics[9]. Other research on migration and source communities has asked origin families to provide information on migrated members: there are limits to the level of detail and to the quality of this indirect information. Of course, quality is an issue also for retrospective data. There are reasons to suspect recall bias especially concerning individuals who are older at the time of the 2000 survey, and thus have to reconstruct how they behave in a distant past. However, Nang Rong surveys provide an opportunity to examine data quality because they contain repeated retrospective histories undertaken in two points in time, in 1994 and in 2000; a simple check of the matching of information provided by a subset of individuals present in both survey confirm that quality is quite high[10]. Moreover, there is evidence that events that are highly salient to the respondent, like schooling and migration decisions, are better recalled (Beckett et al. (1999))[11].

Network size (Net in the equations which follow) is defined by counting, for each young individual and each year, the number of co-villagers present at destination the year before[12]. Excluded from the count are migration episodes due to serving in the military service or as a monk. The network measure is thus time varying, and specific to each couple of village-destination. As already said, we assume that the village is the most relevant agora shaping social interactions outside the family, ruling out the progressive integration among villages through trade, sharing of equipment, and, most relevantly, seasonal mobility of laborers. We also exclude the migrated members of the family from the count of the network size assigned to each individual-year, as we are mainly interested in testing the effect of acquiantances and transitory social relations.

INSERT TABLE 2 "Descriptive Statistics" HERE

In table 2, descriptive statistics are provided on adults residing in Nang Rong (column 1), all the migrants (column 2), and disaggregated by destination. A quick overview suggests some interesting patterns. There is no big difference in internal migration behaviour of women and men: the only destination where male migrants are predominant is Khorat. The average level of education of parents is slightly higher for migrants than for non migrants, as well as the migrants' mean educational attainment. The average size of the migrant stock tends to be considerably higher, as expected, in traditional destinations, such as Bangkok and the provincial capital

[9] Moreover, identifying the migrant population as the people absent from the village at the time of the survey is problematic, as it puts no weight on the important issue of duration of migration, treating seasonal workers as long-term movers. Using long retrospective data, we are able to approximate better the true investment in migration of families and villages, reconstructing the full history of temporary and lasting movements to the city and of returns.

[10] Comparing life histories data on the whole sample from the 1994 and 2000 survey, we obtain a positive match of answers on migration histories (whether one spent most of the time outside Nang Rong in a given year) around 86% of the times. In our final sample the percentage of positive matches is higher given that the average age is lower (we keep information only on individuals for whom we have data on the previous generation).

[11] The authors assess the presence of recall bias in the Malaysian Family Life Survey (MFLS), finding that the quality of the long retrospective histories in the MFLS is quite high, across a range of topics.

[12] We chose to use a count with a one year lag for two reasons: the decision of whether to migrate is probably taken with some months in advance the actual departure; moreover, with a lagged value of the network size, we should partially reduce the simultaneity problem, according to which unobserved shock affect both the network and the individual choices to migrate or to study.

of Buriram, with respect to new emerging poles, such as cities in the Eastern Seaboard. At their first migration, movers rely on an extensive support from their communities, ranging from hospitality (three out of four go and live with others they know), to help for finding a job or for setting an enterpreneurial activity (61% received this kind of help). Migrants are also highly successful in finding a job at destination within one month from their arrival, and they rather move with friends or family. The last variables in the table give an interesting snapshot of the extent of origin bias in social interactions. The whole pool of migrants reveal that, at the timing of their first move, around 62% of their neighbours - people living between 100 meters where they lived - came from the region of Isan, where Nang Rong is located. This percentage lowers only to 60% when migrants are asked about their location at the year of the survey, 2000. Considering that between the year of the first move and the year 2000 there is an average interval of 11 years, this seems to indicate that residential mixing does not occur in our sample. Statistics on the origin of friends at destination, at the bottom of the table, give similar indications.

A legitimate concern is that our specific measure of network *size* at the village level might not capture well effective social network *use* by the migrants. In table 3, we undertake a simple experiment to validate our network measure. Using the 2000 sample of migrants, we inspect whether a positive correlation exists between network size and the preference for members of one's own origin network as sources of help. Migrants are asked to identify the person outside their family to which they would ask in the hypothetical situation of money problems. We run a regression linking the probability that this person is someone from the origin village of the migrant to the village-destination specific network size in 1999 and a set of controls. The highly statistically significant coefficient associated with network size gives further indications that the village of origin matters for migrants'social relations. Moreover, variations in size of origin based networks seem to be associated with variations in the importance of this informal institution for internal migrants.

INSERT TABLE 3 "Validating the network measure" HERE

Table 4 diplays a transition matrix including data from the sample used in all the estimations, including all the young individuals whose mobility and study choices are observed when they had an age between 13 and 25. It shows the number of cases - a case corrisponding to the individual-year pair- in which we observe a transition along one of the four following states: stay in Nang Rong and work, stay in Nang Rong and study, having migrated to an urban destination and work, having migrated and study. It can be seen that the sample has a fairly balanced representation of migrants and non migrants. Simply spotting the number of unchanged states on the diagonal of the matrix, one can see that there is a relatively higher number of cases of individuals studying as residents at Nang Rong than as migrants, suggesting the possibility of a schooling dispersion associated to migration.

INSERT TABLE 4 "Transition matrix for the joint schooling mobility choice" HERE

4 Econometric models and results

The econometric identification of causal effects of networks on education outcomes poses complex problems. As emphasized by Manski (1993), the fundamental problem with the research on social interactions is the necessity to control for correlated unobserved effects within the community. The empirical literature has followed different avenues for tackling confounding factors associated with omitted common variables, simultaneity or reverse causality and selection or sorting.

Here we present three modelling approaches aimed to tackle the challenges of sample selection into migration, endogeneity of changes in network characteristics, and omitted variable bias due to unobserved individual preferences for migration and schooling. The first method corrects selection bias by modelling, as a first stage, the endogenous choice of moving to the city. The second method tackles the potential endogeneity of network size in the young school enrollment by instrumenting networks through the exogenous variations in the number of twenty years old males balloted to serve in the military service and two additional instruments. As will become clear, these two approaches, based on standard non-linear binary choice models, are complementary. The third approach proposes a multinomial formulation of the problem, which explicitly controls for selection in migration status and unobserved heterogeneity. This third approach yields results broadly consistent with those obtained using the binary choice formulation.

4.1 Estimation of network effects corrected for migrant selection

We are interested in the education decision of the young once he has reached an urban destination[13]. The critical assumption for identification of migrant network effects on this decision is that the flow of young moving from rural to urban areas is sorted into groups according to their village of origin. This sorting is however far from being a random assignment from the whole pool of rural young. The availability of data on migrants and non migrants allows to model explicitly the choice of migration as a pre-requisite for participation in group interactions. By controlling for the non randomness of the migration choice, we correct sample selection and we also aim to reduce the bias associated with non-random sorting into migrant networks, in essence leading to a simultaneous selection of migration destination and associated networks. Because of the dichotomy of the enrollment variable, treating the data as a pooled cross section, I follow here Van de Ven & Van Praag (1981), whose probit sample selection model is an extension of the Heckman (1979)'s selection model for dichotomous outcome variables.

The binary enrollment equation for a young migrant i can be expressed in latent terms as follows:

$$E_i^* = \alpha' X_i + \beta' Net_i + \varepsilon_{i1} \tag{5}$$

[13] Note that the migrating and schooling decisions need not to be sequential, with the education decision being taken after the residential one, as we can think of a young that chooses to migrate in order to pursue his higher studies in city schools.

E_i^* can be interpreted as a latent variable, expressing the expected utility gain from enrollment in school. In the notation above, β' is the main coefficient of interest, expressing the impact on schooling enrollment of variation in Net_{ij} the size of the network of co-villagers, not members of the origin family, present at destination the year before the education decision is taken (NETWORK SIZE); α is a $K \times 1$ vector of parameters to be estimated and X_i is a $K \times 1$ vector of time varying and constant exogenous variables, at the individual, family and village level: the schooling enrollment model controls for the gender of the young (MALE), his age in 2000 (AGE IN 2000), the number of younger siblings (SIBLING), the average education of the parents (PARENTAL EDUCATION), origin family wealth (ASSET SCORE), the availability of an high school at the origin village (SCHOOL), whether more than a language is spoken at home (LANGUAGE), age dummies, origin village and city destination fixed effects, time controls and a constant term. The error term is assumed to be normally distributed: $\varepsilon_i^1 \sim N(0, \sigma^2)$.

Assume a reservation destination exists for each individual. This latent variable, expressing the expected utility gain from migrating, can be expressed as:

$$M_i^* = \gamma' R_i + \varepsilon_{i2} \tag{6}$$

the vector R_i includes the control variables in X_i, the total number of migrants from a village with one year lag (TOTAL NETWORK), a constant term and the following variables selected as exclusion restrictions: the distance of the origin village from the main road to Nang Rong (ROAD LENGHT), a measure of family migration experience (FAMILY MIGRATION), and a family level measure of exposure to rain shocks at the origin (RAIN). Table 1 defines how these variables have been constructed: however, a brief explanation can be useful here to clarify the identification strategy. The distance of the village from the main road is expected to be negatively correlated with habitants' capacity to leave. Information and moving cost of migration are expected to decrease with family previous migration experience. As shown in other studies of migration from rural areas in developing countries[14], scarce rains represent a serious shock to agricultural production - the culture of rice in this context - which can induce further migration. Given that there is only one meteo station collecting rain data for the whole Nang Rong region, the exposure index has been built by interacting the average yearly volume of rain in the district with the time investment of family members in agricultural activities[15]. The error term in this selection equation is assumed to follow a standard normal distribution, $\varepsilon_{i2t} \sim N(0, 1)$

The population regression function in (5), for the subsample of migrants, can be stated as:

$$E(E_i^* \mid X_i) = \alpha' X_i + E(\varepsilon_{i1} \mid X_i, Net_i, M_i^* \geq 0). \tag{7}$$

[14] The possibility of instrumenting migrant network size through rain shock has been first noted by Munshi (2003). The instrument used here has been inspired by the one applied to Nang Rong data by Swee (2007), who introduces village-level variation by interacting rain volume with an estimated proportion of rice net producers.

[15] Rain data come from the Thailand Meteorological Office, and cover all years starting 1970. The time investment in agriculture is computed by counting, for all adult family members, the number of years they spent working as peasants, starting from the age of 13, and dividing this number by their age minus 13.

the parametric assumption of the model is the bivariate standard normal distribution of ε_{i1} and ε_{i2} , with correlation coefficient ρ. With this assumption, the error term in the outcome equation has the following conditional distribution:

$$E(\varepsilon_i^1 \mid X_i, , Net_i, M_i^* \geq 0) = \rho\lambda_i$$

where $\lambda_{it} = \frac{\phi(-R_{it})}{\Phi(R_i t)}$, and ϕ, Φ are respectively the standard normal density and cumulative distribution.

As shown by Heckman (1979), the inverse Mill's ratio term λ is a control function that we need to add to the outcome equation (5), in order to estimate consistenty the parameters α :

$$E_i^* = \alpha' X_i + \beta' Net_i + \rho\lambda_i + \varepsilon_{i1} \tag{8}$$

The likelihood function on the subsample of migrants, after accounting for selection, can be written as:

$$L = \prod_{i=1}^{n^1} \Phi^2(\alpha' X_i, \beta' Net_i, \gamma' R_i, \rho) \cdot \prod_{i=n^1+1}^{n^2} \Phi^2(-\alpha' X_i, \beta' Net_i, \gamma' R_i, \rho) \cdot \prod_{i=n^2+1}^{N} \Phi(-\gamma' R_i) \tag{9}$$

α and ρ can be consistently estimated using the Heckman two step procedure, or by a full maximum likelihood approach.[16]

Given we are using pooled data with t observations for each individual, we need to correct standard errors for the likely intro group (individual) correlation[17]. This can be easily implemented through the cluster variance estimator (see STATA (2007)).

INSERT TABLE 5 "School enrolment model with selection into migration" HERE

Results are displayed in table 5. The model is estimated on repeated annual observations for 1860 young individuals between 13 and 25 years old, yielding an unbalanced panel of 12438 cases (one case corrisponding to the pair individual-year): of this total sample, 5624 observations refer to young resident in one of the five urban destinations. The main variable of interest, the origin village migrant network size, has a statistically significant negative impact on enrollment (column 1); this negative impact is confirmed when we control for origin village fixed effects (column 2). The magnitude of this impact is however small. Considering the specification with village fixed effects, the probability of enrollment decreases of 1.3% with one standard deviation change (3.3 in this sample) change in the network size.

As can be seen, there is no evidence of gender effects either on the propensity to pursue higher schooling or to migrate when young. Having educated parents raises the odds of being enrolled

[16] I opt for the second one estimating the model through the heckprob command in STATA. See the STATA reference manual (2007) for details on the command.

[17] The pooled probit estimator which considers the repeated observations for each individual as a large cross section is consistent, but inefficient. Higher efficiency can be achieved by a random effect probit estimator, but we need to be willing to assume that the individual random parameter is uncorrelated with the observed covariates (seeGreene (2003)).

in school after the age of 13, while having no apparent effect on the probability of migrating internally. Having an higher number of younger brothers and sisters does not affect schooling: educational choice does not seem to depend, in this context, on competition for scarse resources among siblings, as we would expect in a context where schooling costs are high relatively to household income and credit constraints exist. Coming from a larger family seems to positively affect the willingness to migrate. Speaking more than one language at home has no significant effects on schooling, while it is positively correlated with migration propensity: no univocal causal relation can be established here, since the fact of speaking more than one language at home can be associated with a family history of migration or interethnic marriages. In order to control for family wealth while reducing evident endogeneity concerns, I build a principal component measure of family long term income, condensing information on twelve productive and non productive assets: this wealth measure has a significant, positive explanatory power on the odds of enrollment, while being negatively associated with migration. With respect to those migrating to Bangkok, enrollment rates are higher for young who moved to Buriram (the regional capital), or to Khorat, while are sensibly lower for those choosing to migrate to the Eastern Seabord. As for the variables excluded for identification reasons, we observe that the overall migration investment of the village (obtained by counting together migrants from the village present in all the destinations with one year lag) is an highly relevant predictor of individual migration. The cumulated number of years of working experience outside Nang Rong of family members not in the sample (who are 40 years old or older), is another important explanatory factor of young migration. Rain shocks act as an important push factor for individual migration from a region where the colture of rice is still the main economic activity[18]. Villages with more difficult access to the Nang Rong town (as measured by distance to the major highway) tend to have a lower participation in migration.

4.2 Instrumental variable estimation of network effects

Instrumental variable estimation (IVE) is a powerful tool for dealing with three variations of the same statistical problem: measurement error, simultaneity bias and omitted variable bias. Essentially, all the three problems challenge the consistency of the model, pointing to the possibility that the network size regressor can be correlated with unobserved determinants of enrollment choice. Measurement error is an issue, since we dispose of an imperfect proxy of the true size of the network from the origin community: given that villages are not isolated one from the other, and that solidarity or mutual recognition can link individuals at an higher level than the village[19], it is possible that our measure is downwardly biased. Simultaneity bias can be a further

[18] This has been shown convincingly for Nang Rong data by Swee (2007), who build a measure of exposure to rain shock by interacting rain intensity with a estimated probability of being a net rice producer.

[19] Speaking the same dialect, or coming from the same geographical region (Buriram), might represent weaker ties among individuals, who can perceive themselves in the city as part of the same community even if coming from different villages. Given the data I dispose of, and the tradition of strong solidarity and identity traditionally

threat to the consistency of network effect estimation. The choice of youngster to be enrolled in school can be internalized by the community, which might then transfer information on job opportunities in the city to those who are still in the village; by this mechanism network size might not only cause but also be influenced by schooling behaviour. Finally, omitted variable bias is probably the greatest source of concern, since we are not able to control for the full set of unobserved factors affecting both movements from rural areas to the cities and the human capital accumulation of individuals who are part of this flow. The proposed instruments used in order to correct for the possible correlation between the network size regressor and the error term in the schooling equation are: 1) the village proportion of young males (at the ages of 20-21) who are balloted to serve in the military outside Nang Rong (MILITARY); 2) the exposure of the village to rain shocks (RAIN VOLUME); 3) the incidence of return migration to the village (RETURN). For what concerns the first instrument, recruitment for the military service in Thailand was by ballot until 1998, as the number of liable conscripts was far higher than the number needed by the armed forces[20]. Call-up took place once a year and each district was given a quota of the number of recruits needed by the armed forces. Liable males had to participate at the ballot and those who drew a red ticket had to perform military service, leaving the village of birth for up to two years. The random departures of the young imposed by the ballot system are likely to represent an external variation of information on living and working opportunities outside Nang Rong. A simple inspection of the data suggests that there is a significant correlation between the stock of migrants in a given year and the number of young balloted for the military the year before: migrant networks can be alimented directly by those young who choose not to go back to the village when the coscription period has expired, or indirectly by the information provided by conscripted movers to those considering the possibility to migrate (see figure 1). As well known, instrumental variables need to satisfy not only the requirement of relevance but also the one of instrumental exogeneity: restricting the relevant sample to individuals who are 19 years old or younger, we can believe that the randomized village-level participation of the young in the military has no independent effects on schooling choice of the young not liable for serving in the army.

For what concerns rainfall shocks, as already said there is only one weather station in Nang Rong, so we only dispose of rainfall variation over time, not villages: the village level measure is built by interacting the level of rainfall with the village-level average time investment in agriculture. The exogeneity of this second instrument can be reasonably assumed if we think that shocks at the origin village have no other impact on schooling choices of young migrants

existing in Thailand among people from the same village, I choose not to depart from the assumption that the village is the relevant bound for defining network size.

[20] Conscription was introduced in Thailand shortly after the First World War. In the 80s and 90s the recruitment system increasingly became subject to public debate. Obviously the system was likely to lead to favoritism on the hands of influential or rich people. I expect this favoritism to be less pronounced within the less developed rural areas of Nang Rong, and thus that young villagers had to face more or less the same risk of being recruited for the military. See the military recruitment dataset at: http://leav-www.army.mil/fmso/documents/mildat/RecruitmentCodebook.pdf

apart from altering the size of their relevant network[21].

The last instrument is a count of episodes of return from each destination to the origin village: the relevance of this instrument is straightforward, as substantial episodes of return are observable in the dataset once the village starts to have a relevant migration history. I use a two-year lagged value of this return variable to reduce the threat to validity represented by unobserved factors or shocks likely to affect both village level return behaviour and individual schooling decisions.

The system of equation to be estimated can be expressed as:

$$E_i = I(\beta' W_i + e_i > 0) \tag{10}$$

$$Net_i = \gamma' Z_i + u_i \tag{11}$$

where $W_i = (X_i, \widehat{Net}_i)$, with X_i being the usual set of controls and $\widehat{Net}_i$ being the fitted values from the network instrumented equation (11), $\widehat{Net}_i = \widehat{\gamma}' Z_i$.

The instruments in the vector Z_i are expected to satisfy the requirements of exogeneity and relevance.

The log likelihood for observation i is:

$$\ln L_i = E_i \ln \Phi(m_i) + (1 - E_i) \ln[1 - \Phi(m_i)] + \ln \phi(\frac{Net_{i-}\gamma' Z}{\sigma}) - \ln \sigma \tag{12}$$

with

$$m_i = \frac{\beta' W_i + \rho(Net_{i-}\gamma' Z)/\sigma}{(1 - \rho^2)^{\frac{1}{2}}} \tag{13}$$

The model is estimated using the IVProbit procedure in Stata which implements Amemiya's generalized least square estimator (Amemiya (1978); Newey (1987)), jointly estimating equations (10) and (11) via maximum likelihood; endogenous variables are treated as linear functions of their instruments as well as other exogenous variables. Standard error are corrected for the correlation induced by the fact that individuals are observed over multiple periods.

INSERT TABLE 6 "IV Estimation of network effects on schooling enrolment" HERE

Results are displayed in table 6. The models in column 1 to 3 are three different specifications of the Amemiya-Newey instrumental variable probit model, where the instruments used are enrolment in the military, rain intensity and return episodes. In the second column, we control for village fixed effects, while in the third we add, as an additional group level regressor, the average level of education of the network's population. In all the three specifications, network size enters as a negative and significant determinant of schooling enrolment: there is thus support for the hypothesis that young individuals, when migrating, have a lower propensity to acquire an higher level of education if they are part of a large enclave. The magnitude of the impact becomes

[21] The exogeneity of this second instrument can be challenged: if shocks at the origin alter demands of staying villagers for support by their family migrants, then the need to remit more can explain a change in labour supply of the young.

higher when network size is instrumented, a one standard deviation change in network size corrisponding to a decrease in the probability of enrollment of 6.6% (for specification in column 2). It is interesting to observe that the size of the effect increases when using instruments, as we suspected given that unobservable preferences for social interactions with one's origin group can magnify the externality and a measurement error in the network size could attenuate the non-IV estimates.

Regarding the instrumental variables, the lagged number of drafted soldiers is strongly correlated with the size of the network in all the specifications. Rain intensity, as expected, reduces out-migration from the village, even if its statistical significance is reduced when we add village fixed effects to the model specification. The number of episodes of return at period $t-1$ are an highly significant and positive predictor of the stock of migrants at any destination at period t.

Column 5 of table 6 is an attempt of estimating network effects on schooling choice making a simultaneous use of instrumental variables and of the Heckman type correction for selection into migration. This is done by simply adding, as an extra regressor in the IV model, the inverse Mill's ratio term λ obtained from the migration choice equation [22]. The selection term is statistically significant, and the instrumented network size effect shows no sizable variation.

Of course, the validity of the conclusions on the relevance of network effects depends on the validity of the model on which the IV estimates are based. A test of overidentifying restrictions needs to be undertaken in order to check the reliability of the identification strategy. Lung-Fei (1992) shows that the minimized distance for these estimators provides a test of overidentifying restrictions: the null of instrument exogeneity can not be rejected [23].

4.3 Controlling for unobserved heterogeneity

It is possible to control for unobserved heterogeneity exploiting the availability of repeated observations for each individual. Given that we dispose of data on both migrants and non migrants, a convenient way of extending the previous discussion to a panel data setting is to redefine the model as a multinomial one for the four possible status j, defined by the interaction of residential status (migrate to an urban destination or stay in Nang Rong) and enrollment choice (study or work):

1. Migrate, Study
2. Migrate, Work
3. Stay, Study

[22] It must be noted that this methodology is problematic as it requires a two-stage solution of the selection model instead of the full maximum likelihood approach: the two-stage estimation is biased when the outcome variable is binary, as in our case. (see STATA (2007)).

[23] The Lee (1992) test statistic is distributed as Chi-squared with (L-K) degrees of freedom under the null that the instruments are valid. The value of the statistic for the model with village fixed effects is 3.766, to which it corresponds a P-value of 0.1521.

4. Stay, Work

The fourth option (Stay and Work) is used as the reference category.

In this way, we are able to endogenize the location of residence decision with decisions in each period about going to school, controlling for unobserved heterogeneity across individuals.

I restrict the focus to a particularly convenient class of multinomial models, the mixed logit or logit Kernel model (Ben-Akiva et al. (2001), Prowse (2005)). In this setting, individual i rationally chooses one of the four options j to maximize his payoff V in each period t:

$$\begin{aligned} V_{i,1,t} &= \alpha' X_{i1t} + \beta' Net_{i1t-1} + \eta_{i,1} + \epsilon_{i,1,t} \\ V_{i,2,t} &= \alpha' X_{i2t} + \beta' Net_{i2t-1} + \eta_{i,2} + \epsilon_{i,2,t} \\ V_{i,3,t} &= \alpha' X_{i2t} + \beta' Net_{i2t-1} + \eta_{i,3} + \epsilon_{i,3,t} \\ V_{i,4,t} &= \epsilon_{i,4,t} \end{aligned} \tag{14}$$

Where X is the vector of controls, N_{t-1} is the network variable and the fourth option $V_{i,4,t}$ (Stay in Nang Rong and Work) is chosen as the reference category, according to which the parameters α and β in the first three equations are interpretable.

Individual i chooses alternative j at time t with the following probability:

$$P_{i,j,t} = P(\alpha' X_{i1t} + \beta' Net_{i1t-1} + \eta_{i,j} + \epsilon_{i,j,t}) > \max_{k=1,..,4,k\neq j} \left\{\alpha' X_{i1t} + \beta' Net_{i1t-1} + \eta_{i,j} + \epsilon_{i,j,t}\right\}) \tag{15}$$

The joint probability of the individual's observed sequence of choices is obtained by integrating w.r.t. to the distribution of the unobserved heterogeneity $F(\eta_i)$:

$$P_i = \int_S \prod_{t=1}^{T} \left(\frac{\exp(\alpha' X_{i1t} + \beta' Net_{i1t-1} + \eta_{i,j})}{\sum_{k=1,..,J} \exp(\alpha' X_{i1t} + \beta' Net_{i1t-1} + \eta_{i,j})} \right)^{Y_{i,j,t}} dF(\eta_i) \tag{16}$$

Estimation is performed by maximum simulated likelihood. In order to understand the intuition behind this method, observe that the unconditional log-likelihood can be interpreted as an expected value:

$$\ln L = \sum_{i=1}^{N} \log E_S \left[\left(\frac{\exp(\alpha' X_{i1t} + \beta' Net_{i1t-1} + \eta_{i,j})}{\sum_{k=1,..J} \exp(\alpha' X_{i1t} + \beta' Net_{i1t-1} + \eta_{i,j})} \right)^{Y_{i,j}} \right]$$

Simulation methods proceed by sampling R times from the distribution of η_i and constructing $P_i(\eta_i^r)$ for $r = 1....R$. The individual simulated likelihood are simply obtained by averaging $P_i(\eta_i^r)$ over the R draws from the distribution of η_i. By the SLLN, simulated maximum likelihood estimates converge almost surely to the true parameters as $R \to \infty$ and $N \to \infty$.

A variety of simulation methods exist for sampling from the distribution of η_i; we here use Halton draws, which have been shown to provide an high level of accuracy for a relatively low

computational time (Train (2000)) [24]. Results are displayed in table 7.

INSERT TABLE 7 "Mixed Logit estimation of Network effects" HERE

While no straightforward comparison is possible with previous models, new interesting insights emerge. The estimation is performed on the whole sample of migrants and non migrants, so to circumvent the problem of selection in migration choice. The first three columns (1) refer to a multinomial model with random intercepts, all the other variables having their cofficients kept fixed. The three columns on the right (2) refer to a second specification of the model with time invariant random intercepts and time invariant random coefficients for the network variables, all the other variables having fixed coefficients. The two specifications both allow correlations and/or heteroscedasticity in the within period and between period unobservables affecting the individual's payoffs from each of the alternatives.

Coefficients need to be interpreted with respect to the baseline alternative, staying in Nang Rong as a worker. For example, looking at the coefficients for family wealth (proxied by the asset score), one can see that those studying after the age of 13 come from relatively wealthier families; those migrating to internal destinations for working have a less wealthy background with respect to those who stay and find an employment in their origin region.

Concerning the network effect, there is evidence of a significant impact of the migrant network size on preferences for employment or education of those leaving their village. The young with a larger network tend to migrate and start working more at destination, while they exhibit a lower preference for moving and studying in the city. We can consider a discrete change in the network size to have an intuition on the magnitude of this effect: one standard deviation in the size of the network (an increase of 33,3 individuals present at destination) tends to increase the probability of choosing the option of migrating and working by 9,7%, while decreasing the probability of choosing to migrate and study by 15,5%, when these two options are evaluated with respect to the baseline choice of staying in the origin village to work.[25]

4.4 Robustness checks

In order to test the robustness of the results, we start by exploring whether the significance of our regressor of interest is affected by different specification of the selection and instrumental variable models. Enriching the models with other covariates which are at the group level and time varying is a reasonable way to inspect the sensitiveness of the estimated network size effect to different model specifications: we have thus tried to reestimate the models adding the mean education level of the network population as an additional regressor. As it can be seen in column 3 of table 5 for the selection model and in column 3 of table 6 for the instrumental variable model,

[24] The model is estimated on STATA using the mixlogit routine developed by Hole (2007). Another convenient estimator of the model is by Markov Chain Monte Carlo (MCMC). Similarities between the two methods are carefully explained by Train (2003). .

[25] This calculation is based on estimates from specification 1. There is a significant effect of network size on the choice of studying rather than working while staying in Nang Rong; however, this estimated effect is very small (one standard deviation change in network size affecting the relative probability of studying at origin by 0,6%).

the network size estimate is robust when controlling for this additional group characteristic.

One potential drawback to the application of the Heckman selection model is its sensitivity to the assumed parametric distribution of the unobservable error terms in the model. If the joint distribution of the error terms in equations (1) and (2) is misspecified, the second-step parameter estimator will be inconsistent in general. Newey et al. (1990) conditions on a polynomial of the estimated probability of participation in a regression framework, in order to avoid to impose parametric forms on error distributions. This flexible function of the estimated probability approximates the unknown conditional expectation of the error term, just as the Mills ratio terms, which are functions of the probability of participation, represent the conditional expectation of the unobservable under the normality assumption. In column 4 of table 5, the selection model is semi-parametrically estimated in two steps as suggested by Newey et al. (1990), using a quartic in the predicted probability of migration to approximate the true control function[26]: the coefficient on the network size decreases in size, while its statistical significance is not much altered[27].

Angrist (2001) argues that the common use of parametric models overly complicates inference when the statistic of interest is causal effects and suggests that linear approximation like standard 2SLS performs as well as parametric estimators in a labor-supply model.[28] In column 4 of table 5, we thus provide two stage least square estimates as an alternative to the instrumental variable probit model. The linear approximation yields results that are similar to the probit ones, confirming that the significance of network effect seems robust to alternative parametric assumptions.

INSERT TABLE 8 "Additional Robustness Checks for Selection Model" and TABLE 9 "Additional Robustness Checks for IV Model" HERE.

Other robustness checks performed on the two models are presented in table 7 and table 8. In column 2 of table 7, the selection model is re-estimated on a smaller sample, limiting the observation to the life histories of young individuals between the ages of 13 and 20. Again, the network size effect is stable. In the third column of table 7 and in column 2 of table 8, migrants to the regional capital of Buriram are excluded from the sample: those migrating to this city might face a different assimilation process, given closeness and lower ethnical diversity of this destination with respect to Nang Rong villages. In both models, the estimates of network effects seems only marginally affected by this change in the relevant sample. Finally, in column

[26] As stressed by Newey et al. (1990), this and other variants of semi-parametric estimators of the selection model would require some mechanism to choose the amount of "smoothing" imposed (here number of basis functions). We are not aware of any "rule of thumb" providing clear guidance on the lenght of the polynomial, and the quartic function has been simply chosen after testing for significance of alternative control functions.

[27] Results of the semi-parametric estimation of the selection model are only provided as a robustness check, since the two-step estimation of binary response models is problematic: the likelihood in the second step is biased, and the extent of the bias is proportional to the size of the correlation between the error terms of the two equations.

[28] Moffitt (2001) refines the conclusion of Angrist by arguing that the good performance of 2SLS depends crucially on where in the sample we look for causal effects. Due to the constant effect assumption of 2SLS, 2SLS and probit are bound to give different answers if we do not consider average effects or if we consider individuals with “extreme characteristics”, i.e. where the fitted value is not near zero.See Arendt & Holm (2007).

3 of table 8, we check whether the instrumental variable model results are sensitive to the choice of a specific set of instrument. The model is re-estimated with a weaker first stage regression, including only the military draft as instrumental variable. Statistically significance of the network size does not seem to be largely affected.

5 Conclusions

Economic integration and modernization benefit to a different extent individuals who are born in different regions: in developing countries, divides between enlocked rural regions and dynamic urban poles have tended to widen in the last decades, and so location of birth persists to be an important predictor of economic opportunities. Different economic models have illustrated how this opportunity gap is destined to shrink when mobility is not constrained by high migration costs. However, this is clearly true only if labor markets at destination do not segmentate in niches separating migrants from native, and if migrants are perceived and behave at destination labor market as the natives do. What we observe in reality is that migrants tend to concentrate in enclaves at destination, and networks as informal institutions have an important role in smoothing assimilation to urban life, either by promoting solidarity and risk sharing among their members, either by easing employment matching. In this paper, we looked at the dynamic implications of the strong ties linking young migrants to their own community of origin: if networks channel career choices so that individuals fail to take full advantage of the economic opportunities offered by the city, can we suspect dynamic inefficiencies and low upward mobility to arise as a consequence of networking?

We observe an high reliance on community support by Thai internal migrants; the data also show a very low propensity of young migrants to invest in acquiring higher education. Assuming that higher and technical education is needed for talented individuals to reap the full opportunities of booming cities, we have argued that a simple test of network effects on economic mobility can be undertaken by regressing network size (as a proxy for density of the community ties and for the probability that networks are active and effective) on the probability of the young to be enrolled as student at destination. In practice, empirical analysis of network effects is challenging, as potential selection and simultaneity bias need to be accounted for. We have proposed three different estimation frameworks in order to check the consistency of results, exploiting the uniqueness of a dataset which accurately describes migration and career choices of both stayers and movers[29]. We have first addressed the issue of selectivity of the migrants sample for which network effects are estimated, aggregating information on migrants and non migrants in order to correct for the non-random allocation of young people growing in a setting with dense migration history. Secondly, we have used instrumental variable estimation

[29] It is useful to repeat here that individuals are not defined as movers or stayers once and for all. Disposing of life history data, we can treat an individual as a migrant at a particular age if he is observed residing out of the village for most of the time at that particular age. Thus many individuals contribute to both the migrant sample and the non migrant sample at different periods of their observed life span (13 to 25 years old).

as an attempt to verify that the negative correlation between network size and young schooling choice is not spouriously determined by unobservable characteristics or shocks affecting either village migration and human capital accumulation at destination. Finally, we remarked that unobserved preferences matter in explaining migration behaviour and in determining perceived opportunity costs of higher schooling; exploiting the availability of repeated choices for migrants and non migrants, we have modelled the simultaneous schooling and migration decision using a multinomial mixed logit formulation which allows a straightforward accounting of individual heterogeneity.

The three models suggest that a statistically significant negative relation links young migrants' schooling decision with the size of the network of co-villagers present at destination. Even if the negative network externality is small in this sample, the paper provides evidence of a channel which can lead to persistence of low educational attainment for those born in rural areas. This finding has important implications for policies attempting to rise schooling participation: if the story presented in this paper is correct, young migrants integrated in sufficiently large enclaves perceive an higher opportunity cost of schooling and tend to join the labor force earlier. Since migrant networks are not likely to lose importance as rural urban integration proceeds, there is the need to study better why the clustering in solidal networks might undermine investment in education at destination.

Two main explanations should be considered. A first possible explanation link the migrant network size with the development of asymmetric or stratified communities within the city. Given an initial disadvantage with respect to natives, young migrants end up living in less dynamic neighborhood. Labor market institutions specific to these heterogenous communities - and in particular the word-of-mouth mode of dissemination of information about scarse job opportunities - also play a role, as individuals embedded in low skill networks develop lower expected returns to high education acquisition (Anderberg & Andersson (2007)). According to the second explanation, the community effectively discourages the acquisition of higher education by their members, as this behaviour is perceived as a deviation and sanctioned.

A social planner interested in raising educational achievements of the young migrants should sensibly think about measures affecting the community evaluation of higher schooling, such as reserving a preferential access of migrants to training and matching programs granting access to remurative, high skill jobs. It might also be worth experimenting policy measures reducing the dependency of young individuals on support of their own origin community (for example, scholarship schemes involving the phisical deplacement of the young to neighborhoods where concentration of their co-villagers is lower[30]).

What is valid for internal migration is potentially even more valid for international migration: consider that as the cost of migration and assimilation rise, so the value attached by the

[30] Evaluations of the Moving to Opportunity Program (for example Kling et al. (2005)), show significant improvement in young outcomes. In these programs, randomly selected families in disadvantaged neighborhoods receive a financial transfer to move to more affluent areas.

individual to partecipation in his community group rises.

Finally, further research is warranted to extend the analysis of migrant network effects on dimensions of upward mobility other than schooling, such as the relative participation of those with a migratory history to non traditional employment sectors.

References

Amemiya, T. (1978), 'The estimation of a simultaneous equation generalized probit model', *Econometrica* **46**(5), 1193–1205.

Anderberg, D. & Andersson, F. (2007), 'Stratification, social networks in the labour market, and intergenerational mobility', *Economic Journal* **117**(520), 782–812.

Angrist, J. D. (2001), 'Estimations of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice', *Journal of Business & Economic Statistics* **19**(1), 2–16.

Arendt, J. N. & Holm, A. L. (2007), Probit models with dummy endogenous regressors, Working paper series, University of Southern Denmark.

Austen-Smith, D. & Fryer, R. G. (2005), 'An economic analysis of "acting white"', *The Quarterly Journal of Economics* **120**(2), 551–583.

Banerjee, B. (1983), 'Social networks in the migration process: Empirical evidence on chain migration in india', *Journal of Developing Areas* **17**(2), 185–96.

Beckett, M., DaVanzo, J., Sastry, N., Panis, C. & Peterson, C. (1999), The quality of retrospective reports in the malaysian family life survey, Technical report.

Ben-Akiva, M., Bolduc, D. & Walker, J. (2001), Specification, estimation, identification of the logit kernel (or continuous mixed logit) model, Working paper, Massachussets Institute of Technology.

Benabou, R. (1996), 'Equity and efficiency in human capital investment: The local connection', *Review of Economic Studies* **63**(2), 237–64.

Borjas, G. J. (1995), 'Ethnicity, neighborhoods, and human-capital externalities', *American Economic Review* **85**(3), 365–90.

Bowles, S., Loury, G. & Sethi, R. (2008), Group inequality, Working papers, Unpublished manuscript.

Bowles, S. & Sethi, R. (2006), Social segregation and the dynamics of group inequality, Working Papers 2006-02, University of Massachusetts Amherst, Department of Economics.

Eckstein, Z. & Weiss, Y. (1998), The absorption of highly-skilled immigrants: Israel, 1990-1995, CEPR Discussion Papers 1853, C.E.P.R. Discussion Papers.

Friedberg, R. M. (2000), 'You can't take it with you? immigrant assimilation and the portability of human capital', *Journal of Labor Economics* **18**(2), 221–51.

Godley, J. (2001), 'Kinship networks and contraceptive choice in nang rong, thailand', *International Family Planning Perspectives* .

Greene, W. H. (2003), *Econometric Analysis*, Prentice-Hall International.

Heckman, J. J. (1979), 'Sample selection bias as a specification error', *Econometrica* **47**(1), 153–61.

Hoff, K. & Sen, A. (2005), The kin system as a poverty trap?, Policy Research Working Paper Series 3575, The World Bank.

Hole, A. R. (2007), 'Fitting mixed logit models by using maximum simulated likelihood', *Stata Journal* **7**(3), 388–401.

Jeong, H. (n.d.), Education and credit: Sources of growth with increasing inequality in thailand, Technical report.

Kling, J. R., Liebman, J. B. & Katz, L. F. (2005), Experimental analysis of neighborhood effects, NBER Working Papers 11577, National Bureau of Economic Research, Inc.

Lung-Fei, L. (1992), 'Amemiya's generalized least squares and tests of overidentification in simultaneous equation models with qualitative or limited dependent variables', *Econometric Reviews* **11**(3), 319–328.

Manski, C. F. (1993), 'Identification of endogenous social effects: The reflection problem', *Review of Economic Studies* **60**(3), 531–42.

Mills, M. B. (1997), 'Contesting the margins of modernity : women, migration and consumption in thailand', *The Quarterly Journal of Economics* **24**(1), 37–61.

Moffitt, R. A. (2001), 'Estimations of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice: Comment', *Journal of Business & Economic Statistics* **19**(1), 20–23.

Munshi, K. (2003), 'Networks in the modern economy: Mexican migrants in the u.s. labor market', *The Quarterly Journal of Economics* **118**(2), 549–599.

Munshi, K. D. & Rosenzweig, M. R. (2003), 'Traditional Institutions Meet the Modern World: Caste, Gender and Schooling Choice in a Globalizing Economy', *SSRN eLibrary* .

Newey, W. K. (1987), 'Efficient estimation of limited dependent variable models with endogenous explanatory variables', *Journal of Econometrics* **36**(3), 231–250.

Newey, W. K., Powell, J. L. & Walker, J. R. (1990), 'Semiparametric estimation of selection models: Some empirical results', *American Economic Review* **80**(2), 324–28.

Prowse, V. (2005), State dependence in a multi-state model of employment, Economics Papers 2005-W20, Economics Group, Nuffield College, University of Oxford.

Robert E. Lucas, J. (1989), On the mechanics of economic development, NBER Reprints 1176, National Bureau of Economic Research, Inc.

STATA (2007), *Base Reference Manual*, STATA PRESS.

Streufert, P. (2000), 'The effect of underclass social isolation on schooling choice', *Journal of Public Economic Theory* **2**(4), 461–82.

Swee, E. L. (2007), Network effects among migrants in the labour market: evidence from thailand, Working paper.

Tapananont, N. (2004), Socio-economic mixed in bangkok urban regeneration, Technical report, Technical Report at International Workshop on Asian Approach toward Sustainable Urban Regeneration,University of Tokyo.

Train, K. (2000), Halton sequences for mixed logit, Economics Working Papers E00-278, University of California at Berkeley.

Train, K. (2003), *Discrete choice methods with simulation*, Cambridge University Press.

Van de Ven, W. P. M. M. & Van Praag, B. M. S. (1981), 'The demand for deductibles in private health insurance : A probit model with sample selection', *Journal of Econometrics* **17**(2), 229–252.

Yamauchi, F. (2003), Are experience and schooling complementary?, Technical report.

6 Appendix

We here provide the proof for proposition 1 of section 2, following closely the original model of Bowles, Loury and Sethi (2008). Proposition 1 says that an initial inequality in the allocation of skills between groups can not be reversed in equilibrium. The proof is based on the definition of competitive equilibrium as an allocation of skill across groups satisfying equations (1)-(4), and the result follows from the assumed complementarity of low and high skill labor in the process of production. Assume that the proportion of skilled in two groups at the initial period is given by $(s_0^N, s_0^M) \in [0,1]^2$. Then, by equations (1) and (2), the proportion of skilled in the initial period in the overall economy, $\overline{s_o}$, and in each group network, $(L_0^N, L_0^M) \in [0,1]^2$, are uniquely defined. Define the function φ as follows:

$$\varphi(\overline{s}) = n(1 - G(\widetilde{a}_{t+1}(\delta(s_{t+1}), L_t^N))) + (1-n)(1 - G(\widetilde{a}_{t+1}(\delta(s_{t+1}), L_t^M))). \tag{17}$$

Note that $\varphi(\overline{s})$ is bounded between 0 and 1, and strictly decreasing, as $\varphi(0) = 1$ and $\varphi(1) = 0$. Then, given (L_0^N, L_0^M), there exists a unique value of $\overline{s}$ such that $\overline{s}$=$\varphi(\overline{s})$.From equation (1) and equation (4), $\overline{s}_1$ must satisfy $\overline{s}_1 = \varphi(\overline{s}_1)$ in equilibrium, so $\overline{s}_1$ is uniquely defined. This extends logically to the pair (s_1^N, s_1^M), which is also uniquely defined from (4). Having proven uniqueness, an initial situation of inequality $s_0^N > s_0^M$ will persist at any t given the construction of L^V in (2), the dynamics in (4) and the fact that the threshold $\widetilde{a}$ is decreasing in network quality L^V. This completes the proof for proposition 1.

Is this conclusion valid asymptotically? In other words, we wonder whether convergence can be achieved as $t \to \infty$ given an initial condition of disequality. Define a competitive equilibrium as a steady state if $(s_0^N, s_0^M) = (s_t^N, s_t^M)$ for all t. Steady states are symmetric if the two groups have a common skill share, $s_t^N = s_t^M$ At any symmetric steady state, the common skill share $s_t$must solve

$$s = 1 - G(\widetilde{a}(\delta(s), s)) \tag{18}$$

Since costs of training are bounded and benefits goes to infinite as $s \to 0$, everyone gets trained when $s \to 0$,i.e. $\lim_{s\to 0}\delta(s) = 0$. Samely, $\lim_{s\to 0}\delta(s) = 0$, as $\delta(1) = 0$.

Thus, there must exist at least one symmetric steady state. There will be only one symmetric steady state if $\widetilde{a}(\delta(s), s)$ is strictly increasing in s at any such state, or

$$\frac{d\widetilde{a}}{ds} = \widetilde{a}_1\delta' + \widetilde{a}_2 > 0 \tag{19}$$

where $\widetilde{a}_1$ and $\widetilde{a}_2$ are partial derivatives. The condition simply says that as s rises, the threshold raises too as the wage gap shrinks: this must be the case if the reduction in cost due to the externality is not high enough to compensate the effect of a reduction in the benefit of training (the wage gap) as s rises. Is this unique steady state stable? This depends on the level

of segregation, as it is shown now. Rewrite (4) as a recursive system:

$$s_t^V = f^V(s_{t-1}^N, s_{t-1}^M) \tag{20}$$

where f^V

$$f^V = 1 - G(\widetilde{a}(\delta(nf^N + (1-n)f^M)), \eta s_{t-1}^V + (1-\eta)(ns_{t-1}^N + (1-n)s_{t-1}^M)) \tag{21}$$

Assume that the externality effect is not too small:

$$G' \left|\widetilde{a}_2\right| > 1 \tag{22}$$

The stability of the symmetric steady state can be defined by examining the eigenvalues of the Jacobian:

$$J = \begin{bmatrix} f_1^N & f_2^N \\ f_1^M & f_2^M \end{bmatrix} \tag{23}$$

By some manipulations from (21), Bowles, Loury and Sethi (2008) show that these eigenvalues are equal to:

$$\begin{aligned} \lambda_1 &= -G'\widetilde{a}_2\eta, \\ \lambda_2 &= -G'\widetilde{a}_2(1-\gamma) \end{aligned} \tag{24}$$

where $\gamma = \frac{G'(\widetilde{a}_1\delta')}{1+G'(\widetilde{a}_1\delta')}$.

Both eigenvalues are positive and $\lambda_2 < 1$ for all parameter values. Thus the steady state is asymptotically stable if $\lambda_1 < 1$ and unstable if $\lambda_1 > 1$. From assumption (22), it follows that there is a level of segregation $\widehat{\eta} \in (0,1)$, such that the unique symmetric steady state is unstable if $\eta > \widehat{\eta}$, which is Bowles, Loury and Sethi (2008)'s main result.

Figure 1. Correlation between village migration and military conscriptions

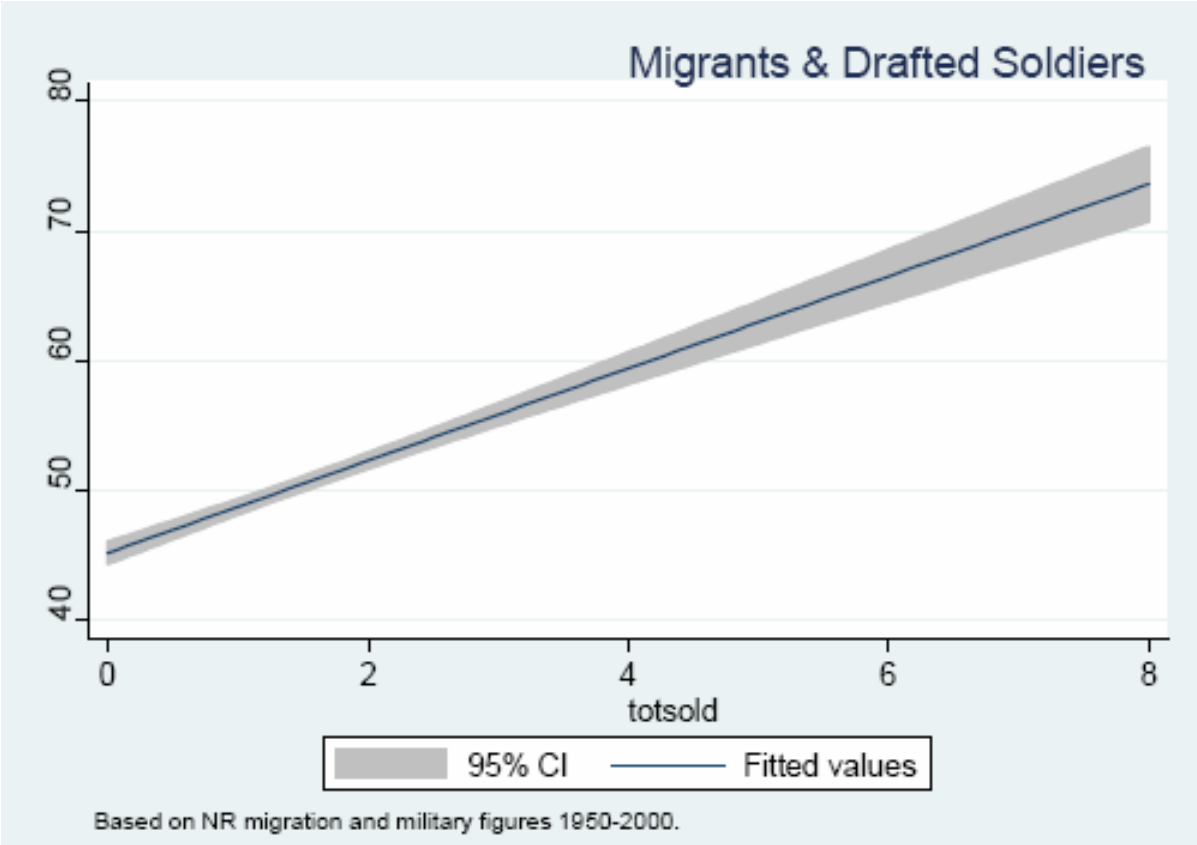

Table 1. List of variables

Variable	Description
	Dependent variables
Student	1=Enrolled as student, 0= Not Enrolled
Migrant	1=Living in a urban destination, 0= Living in a Nang Rong Village
	Main explanatory variables
Network Size	Number of year/destination/village specific established migrants, in multiples of ten. In all regressions it is computed with one year lag. For each young, migrants from his own family are excluded from the count
Total Network	Number of year/village specific established migrants, in multiples of ten. In all regressions it is computed with one year lag. For each young, migrants from his own family are excluded from the count
Family Migration	Family average of adult – 30 year-old of older - individual migration histories (years spent outside Nang Rong/years of adult life(14 years old or more))
Network's education	Destination/village specific average of years of education completed by established migrants. Computed with one year lag.
	Instrumental variables
Military	Number of drafted soldiers who left the village, inserted in all regressions with two years lag
Rain volume	Year average of monthly measures (in ml) of rain fallen in Nang Rong, interacted with averages (at village or household levels) of labor supply in agriculture (years spent working in agriculture /years of adult life (14 years old or more)).
Return	Number of Individuals returning to the village from one urban destination, inserted in all regressions with two years lag.
	Main control variables
Actual age	Age of the young at each period of the panel
Age in 2000	Age of the young when interviewed for the 2000 survey
Buriram	0= young lives in Buriram province, 1=young does not live in Buriram province
Eastern Seaboard	0= young lives in Rayong or Chanbury, 1=young does not live in Rayong or Chanbury
Khorat	0= young lives in Khorat, 1=young does not live in Khorat
Language	0= more than one language is spoken at home, 1= only one language is spoken at home
Male	0=female, 1=male
Other destination	0= young does not live in any Thai destination other than Bangkok, Khorat, Eastern Seaboard or Buriram at a given age; 1=young lives in another destination
Parental education	Average years of formal education completed by the parents. Years of education completed by the father (the mother), if only information on the father (the mother) is available
Road length	Distance in kilometres of the origin village from the main road to the town of Nang Rong
Siblings	Number of living younger brothers and sisters
School	1 = there is a secondary school in origin village, built in 1980 or earlier; 0= there is no such school in origin village
Single	0= married; 1= not married
Asset score	Principal component wealth score obtained by using 1984 data on origin family's ownership of the following assets: television, recorder, telephone, refrigerator, computer, washing machine, motorcycle, car and truck, extension of plots cultivated.
	Other descriptive variables
Salary	Deflated migrant's hourly wages (in bath) at first migration
Support	1= migrant received help to find a job or to set up an entrepreneurial activity by family or friends, at his first migration; 0= migrant did not receive this help
Education completed	Years of formal schooling completed
Family at destination	1= migrant has family members present at destination at the time of his first migration, 0=no family members present
Money received	Money (in bath) received from origin family during the first migration year
Money sent	Money (in bath) sent to the family in the village during the first migration year
Others came	1 = migrant moved with others at his first migration, 2=migrant moved alone
Others living	1 =migrant lived with family or friends at first migration, 0= migrant lived with no family or friends at his first migration
Job in 1 month	1 = migrant found a job within 1 month at first migration, 0=migrant did not find a job in 1 month
Education completed	Years of formal schooling completed
Neighbours from Isan	Percentage of neighbour coming from Isan for migrants in 2000. First refers to first migration and now to the year 2000

Table 2. Descriptive Statistics

		Nang Rong	All migrants	Bangkok	Khorat	East seabord	Buriram	Others
Age in 2000	mean	25.62	26.21	25.97	27.71	24.88	26.78	26.42
	sd	*5.21*	*4.55*	*4.30*	*5.71*	*4.58*	*5.32*	*4.54*
Male	mean	0.54	0.49	0.47	0.60	0.52	0.52	0.47
	sd	*0.50*	*0.50*	*0.50*	*0.49*	*0.50*	*0.50*	*0.50*
Single	mean	0.40	0.35	0.41	0.38	0.46	0.30	0.29
	sd	*0.49*	*0.48*	*0.49*	*0.49*	*0.50*	*0.46*	*0.45*
Parental Education	mean	8.79	8.88	8.80	9.18	8.92	9.74	8.75
	sd	*4.03*	*4.05*	*3.88*	*4.02*	*3.98*	*5.31*	*3.78*
Education completed	mean	6.33	6.39	5.79	6.67	5.13	8.48	7.02
	sd	*3.37*	*3.72*	*3.29*	*3.74*	*2.59*	*4.01*	*4.23*
Asset Score	mean	0.19	0.16	0.12	0.11	0.18	0.26	0.09
	sd	*1.34*	*1.44*	*1.31*	*1.39*	*1.43*	*1.93*	*1.44*
Siblings	mean	0.12	1.08	1.09	0.86	1.47	1.01	1.07
	sd	*0.67*	*2.06*	*2.14*	*1.95*	*2.29*	*1.80*	*1.98*
Road Length	mean	5.65	5.27	5.47	5.58	5.79	5.58	4.95
	sd	*3.30*	*3.11*	*3.23*	*3.08*	*3.21*	*3.10*	*2.96*
Network	mean		48.39	63.31	5.86	11.86	34.13	48.51
	sd		*30.83*	*30.22*	*3.78*	*10.17*	*19.46*	*25.47*
Family at destination	mean		1.36	1.32	0.54	1.08	0.83	1.26
	sd		*1.06*	*1.13*	*0.73*	*1.05*	*0.86*	*1.14*
Return	mean		4.51	5.29	3.22	2.82	3.49	4.65
	sd		*7.16*	*7.60*	*6.42*	*5.78*	*6.94*	*7.00*
Salary	mean		12.53	12.07	13.94	14.12	12.95	11.96
	sd		*12.53*	*11.91*	*8.20*	*9.49*	*10.17*	*14.48*
Money received	mean		3.88	3.29	3.44	4.09	4.45	4.22
	sd		*1.89*	*1.59*	*2.40*	*1.74*	*2.02*	*1.99*
Money sent	mean		4.54	4.62	4.56	4.51	4.87	4.17
	sd		*1.47*	*1.41*	*1.67*	*1.66*	*1.40*	*1.55*
Others came	mean		0.62	0.64	0.69	0.60	0.54	0.61
	sd		*0.48*	*0.48*	*0.47*	*0.49*	*0.50*	*0.49*
Others living	mean		0.73	0.75	0.68	0.68	0.60	0.76
	sd		*0.45*	*0.43*	*0.47*	*0.47*	*0.49*	*0.43*
Support	mean		0.61	0.67	0.60	0.56	0.34	0.66
	sd		*0.49*	*0.47*	*0.49*	*0.50*	*0.47*	*0.47*
Job in 1 month	mean		0.88	0.88	0.69	0.88	0.93	0.76
	sd		*0.33*	*0.33*	*0.47*	*0.33*	*0.25*	*0.43*
Isan Neighbours first	mean		62.17	60.78	64.72	70.97	54.54	63.61
	sd		*30.46*	*31.16*	*26.19*	*28.65*	*31.42*	*29.28*
Isan Neighbours now	mean		60.33	60.05	64.71	56.69	58.72	60.84
	sd		*31.73*	*31.92*	*32.00*	*34.14*	*30.62*	*31.34*
Isan friends now	mean		67.36	66.83	63.18	65.17	63.61	67.44
	sd		*39.48*	*40.12*	*41.69*	*38.15*	*42.14*	*38.66*

Table 3. Validating the network measure

Robust standard errors in brackets, clustered at the household level

Dependent variable: Ask someone from same village if experience money problems	Probit
Network size	0.11a
	[0.02]
Age in 2000	-0.04b
	[0.02]
Male	-0.04
	[0.12]
Single	-0.41a
	[0.13]
Education completed	0.001
	[0.02]
Siblings	0.04
	[0.03]
Year since first migration	-0.01
	[0.02]
Salary	-0.01c
	[0.01]
Others came	-0.17
	[0.12]
Others living	0.43a
	[0.15]
Owns a car	1.30c
	[0.75]
Asset Score (origin household)	0.0001
	[0.05]
Korat	0.82a
	[0.27]
Other province	0.2
	[0.13]
Buriram	0.71a
	[0.24]
Sector of occupation dummies	Yes
Constant	-0.35
	[0.61]
Observations	1296
Pseudo Log Likelihood	-758.582

c significant at 10%; b significant at 5%; a significant at 1%
Note: Estimated on Nang Rong Project 2000 data for migrant sample.

Table 4. Transition Matrix for joint mobility and study choices (13-25 years old / all cohorts)

	Stay and Work	Stay and Study	Migrate and Work	Migrate and Study
Stay and Work	23,474 (48.57)	75 (0.16)	2,117 (4.38)	22 (0.05)
Stay and Study	414 (0.86)	3,741 (7.74)	283 (0.59)	189 (0.39)
Migrate and Work	1,662 (3.44)	30 (0.06)	14,583 (30.17)	64 (0.13)
Migrate and Study	76 (0.16)	23 (0.05)	241 (0.5)	1,335 (2.76)

Note: Author's calculations on sample extracted from Nang Rong life history data for 2000; the numbers on the diagonal indicate the cases of no changes in status from one year to the following in the panel. Outside the diagonal are the cases of transitions from the state indicated in the first column at one year, to the state indicated in the first line at the following year (for example, in our sample there have been 414 cases of transitions from the state of residence in Nang Rong as a student (Stay and Study) to the state of residence in Nang Rong as a worker (Stay and Work)). In parenthesis percentages over all states observed are displayed.

Table 5. School enrolment model with selection into migration

	Probit with selection (1)		Probit with selection (2)		Probit with selection (3)		Semi-parametric (4)
Dependent Variable	student	migrant	student	migrant	student	migrant	student
Network Size	-0.05a		-0.04b		-0.04c		-0.04c
	[0.02]		[0.02]		[0.02]		[0.02]
Network's education					0.01	-0.02a	
					[0.01]	[0.01]	
Total network		1.16a		1.92a		0.62a	
		[0.12]		[0.15]		[0.15]	
Family migration		0.75a		0.73a		0.79a	
		[0.14]		[0.14]		[0.14]	
Rain volume (family)		-0.97a		-0.97a		-1.00a	
		[0.19]		[0.20]		[0.20]	
Age in 2000	0.02	-0.06c	0.03	-0.07c	0.02	-0.07c	0.03
	[0.07]	[0.04]	[0.07]	[0.04]	[0.07]	[0.04]	[0.09]
Male	-0.02	-0.05	-0.02	-0.05	-0.01	-0.06	-0.03
	[0.11]	[0.05]	[0.11]	[0.05]	[0.11]	[0.05]	[0.12]
Single	0.66a	-0.15a	0.67a	-0.16a	0.62a	-0.16a	0.71a
	[0.12]	[0.06]	[0.12]	[0.06]	[0.12]	[0.06]	[0.13]
Parental education	0.12a	0.01	0.11a	0.01	0.12a	0.01	0.12a
	[0.02]	[0.01]	[0.02]	[0.01]	[0.02]	[0.01]	[0.02]
Siblings	0.003	0.15a	0.003	0.15a	0.01	0.15a	0.06
	[0.03]	[0.02]	[0.02]	[0.02]	[0.02]	[0.02]	[0.06]
Language	0.13	0.23a	0.12	0.17a	0.12	0.17a	0.13
	[0.15]	[0.06]	[0.15]	[0.06]	[0.15]	[0.06]	[0.17]
Asset score	0.17a	-0.06a	0.16a	-0.05a	0.17a	-0.05a	0.18a
	[0.03]	[0.02]	[0.03]	[0.02]	[0.03]	[0.02]	[0.04]
Road length		-0.03a				-0.03a	0.2
		[0.01]				[0.01]	[0.36]
School	0.23	-0.09	0.22	-0.06	0.22	-0.06	-0.04
	[0.15]	[0.07]	[0.33]	[0.14]	[0.33]	[0.14]	[0.14]
Korat	0.46b		0.52b		0.52b		0.62b
	[0.23]		[0.23]		[0.22]		[0.24]
Eastern Seabord	-0.47b		-0.41c		-0.44c		-0.48c
	[0.21]		[0.23]		[0.21]		[0.26]
Buriram	0.81a		0.82a		0.85a		0.94a
	[0.18]		[0.17]		[0.17]		[0.16]
Other destination	0.1		0.12		0.11		0.15
	[0.14]		[0.14]		[0.14]		[0.15]
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effects	No	No	Yes	Yes	No	No	Yes
Constant	-2.59	1.22	-2.78	0.96	-2.6	0.97	1.72
	[2.59]	[1.39]	[2.50]	[1.45]	[2.48]	[1.37]	[1.59]
Observations	5624	12438	5624	12438	5317	11783	5366
Log Pseudolikelihood	-8467.88		-8344.54		-8119.854		-840.58

Robust standard errors in bracket, corrected for intra-cluster correlation at the individual level
c significant at 10%; b significant at 5%; a significant at 1%
Note: Semi-parametric in column 4 refers to the estimation method suggested in Newey et al.(1990), and use a qaurtic In the predicted probability of migration estimated as a first stage.

Table 6. IV Estimation of network effects on schooling enrolment

	IV Probit (1)		IV Probit (2)		IV Probit (3)		2SLS (4)	IV Probit (5)	
Dependent variable	student	network	student	network	student	network	student	student	network
Network size	-0.09c		-0.19a		-0.22a		-0.02b	-0.19a	
	[0.05]		[0.06]		[0.07]		[0.01]	[0.07]	
Network's education					-0.03	-0.12a			
					[0.02]	[0.03]			
Military		0.31a		0.13a		0.11a			0.13a
		[0.04]		[0.04]		[0.04]			[0.04]
Rain volume (village)	-3.71a		-1.52c		-1.57c			-0.59	
		[1.20]		[0.94]		[0.90]			[1.06]
Return		0.12a		0.08a		0.07a			0.08a
		[0.01]		[0.01]		[0.01]			[0.01]
Selection (IMR)							-1.94a	-1.38b	
								[0.60]	[0.56]
Age in 2000	0.001	-0.06	-0.01	0.04	0.05	0.06	0.001	0.09	0.12c
	[0.07]	[0.06]	[0.07]	[0.06]	[0.07]	[0.06]	[0.01]	[0.08]	[0.07]
Male	-0.07	0.14	0.1	-0.05	0.14	-0.04	-0.01	0.0002	0.14
	[0.14]	[0.16]	[0.13]	[0.13]	[0.13]	[0.12]	[0.02]	[0.14]	[0.13]
Single	-0.13	0.68a	-0.12	0.75a	-0.14	0.73a	0.09a	0.89a	-0.01
	[0.14]	[0.18]	[0.14]	[0.13]	[0.13]	[0.13]	[0.02]	[0.14]	[0.15]
Parental education	0.17a	0.004	0.17a	0.01	0.17a	0.001	0.03a	0.16a	0.002
	[0.02]	[0.02]	[0.02]	[0.03]	[0.03]	[0.02]	[0.00]	[0.03]	[0.02]
Siblings	0.03	0.01	0.03	0.02	0.03	0.01	0.004	-0.10b	-0.07c
	[0.02]	[0.03]	[0.02]	[0.02]	[0.02]	[0.02]	[0.003]	[0.05]	[0.04]
Language	-0.04	0.22	-0.26	-0.15	-0.24	-0.15	-0.04	-0.38c	-0.41b
	[0.18]	[0.21]	[0.19]	[0.19]	[0.19]	[0.19]	[0.02]	[0.21]	[0.20]
Asset score	0.19a	-0.05	0.17a	-0.14b	0.17a	-0.10c	0.04a	0.22a	-0.10c
	[0.05]	[0.07]	[0.05]	[0.06]	[0.05]	[0.06]	[0.01]	[0.05]	[0.06]
School	-0.04	-0.21	-0.29	-0.08	-0.1	-0.34	0.003	-0.1	-0.32
	[0.19]	[0.28]	[0.56]	[0.40]	[0.40]	[0.55]	[0.07]	[0.40]	[0.56]
Khorat	0.32	-5.91a	-0.14	-5.59a	-5.25a	-0.27	0.04	-0.13	-5.55a
	[0.37]	[0.18]	[0.29]	[0.45]	[0.49]	[0.29]	[0.08]	[0.46]	[0.29]
Eastern Seaboard	-0.49	-4.36a	-0.85b	-3.99a	-0.97b	-3.89a	-0.10b	-0.89b	-4.00a
	[0.32]	[0.21]	[0.37]	[0.23]	[0.39]	[0.20]	[0.05]	[0.38]	[0.22]
Buriram	0.07	-1.54a	-0.06	-1.43a	-1.52a	-0.13	0.004	-0.05	-1.43a
	[0.18]	[0.21]	[0.18]	[0.15]	[0.15]	[0.19]	[0.02]	[0.19]	[0.15]
Other destination	0.92a	-3.16a	-2.86a	0.70b	-3.02a	0.54c	0.21a	0.67b	-2.87a
	[0.24]	[0.25]	[0.27]	[0.28]	[0.33]	[0.27]	[0.05]	[0.29]	[0.28]
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Village fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-2.20	6.63a	-2.13	4.26	3.95	-2.05	0.04	2.86	3.45
	[2.80]	[2.52]	[0.14]	[2.45]	[2.69]	[0.02]	[0.36]	[3.05]	[2.39]
Pseudo Loglikelihood	-6791.00		-5979.58		-6673.57			-5947.49	
Adjusted R squared							0.34		

Robust standard errors in bracket, corrected for intra-cluster correlation at the individual level

c significant at 10%; b significant at 5%; a significant at 1%

Note: The estimation method is in all regression except column (4) is the Newey (1987) Amemiya (1978) instrumental variable probit.

Table 7. Mixed Logit estimation of Network effects

	Mixed logit by MSL (1)			Mixed logit by MSL (2)		
	Stay and Study	Migrate and Study	Migrate and Work	Stay and Study	Migrate and Study	Migrate and Work
Network Size	-0.03a	-0.02a	0.01a	-0.02a	-0.01a	0.01a
	[0.003]	[0.002]	[0.001]	[0.003]	[0.003]	[0.001]
Age in 2000	0.61a	-0.48a	0.31a	0.91a	-0.06a	0.30a
	[0.18]	[0.15]	[0.06]	[0.19]	[0.15]	[2.02]
Actual Age	-0.79	1.96a	1.29a	-0.83	1.169a	1.14a
	[0.49]	[0.36]	[0.13]	[0.45]	[0.37]	[0.12]
Actual Age squared	-0.04a	-0.05a	-0.04a	-0.05a	-0.04a	-0.03a
	[0.01]	[0.01]	[0.00]	[0.01]	[0.01]	[0.002]
Male	-0.29	-0.38	-0.37b	0.038	-0.781	-0.45b
	[0.35]	[0.28]	[0.17]	[0.38]	[0.33]	[0.18]
Parental education	0.75a	0.10a	0.60a	0.91a	0.22a	0.87a
	[0.05]	[0.02]	[0.05]	[0.05]	[0.02]	[0.04]
Siblings	0.60a	0.67a	0.46a	0.84a	0.68a	0.82a
	[0.07]	[0.04]	[0.08]	[0.07]	[0.05]	[0.07]
Asset score	0.12	0.41a	-0.30a	0.14a	0.30a	
	[0.11]	[0.08]	[0.06]	[0.09]	[0.08]	[0.33a]
School	1.08c	0.64c	0.16	1.03c	0.64c	-0.1
	[0.61]	[0.35]	[0.22]	[0.37]	[0.38]	[0.23]
Road Lenght	0.05	-0.09a	0.19a	-0.05	-0.14a	-0.05a
	[0.06]	[0.03]	[0.05]	[0.04]	[0.02]	[0.04]
Village dev. score	0.29b	-0.30a	0.15a	0.12b	0.20a	0.15a
	[0.12]	[0.09]	[0.06]	[0.1]	[0.09]	[0.05]
	-21.84a	-25.71°	-10.19b	31.04	24.08	-17.85a
	[5.94]	[2.04]	[4.82]	[6.4]	[2.12]	[4.74]
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	47504			47504		

Standard errors in brackets, c significant at 10%; b significant at 5%; a significant at 1%
Note: Estimation is by Maximum Simulated Likelihood, using 200 Halton draws.

Table 8. Additional Robustness Checks for Selection Model

	Probit with selection *(1)* *Only 13-20 years old*		Probit with selection *(2)* *Migrants to Buriram dropped*	
	student	migrant	student	migrant
Network Size	-0.05b		-0.06b	
	[0.02]		[0.02]	
Total network		1.97a		1.87a
		[0.19]		[0.15]
Family migration		0.88a		0.71a
		[0.20]		[0.15]
Rain volume		-1.17a		-0.94a
		[0.25]		[0.20]
Age in 2000	0.01	-0.01	0.04	-0.06
	[0.07]	[0.04]	[0.08]	[0.04]
Male	-0.02	-0.11c	-0.14	-0.07
	[0.13]	[0.06]	[0.12]	[0.05]
Single	0.79a	-0.23a	0.52a	-0.13b
	[0.14]	[0.07]	[0.12]	[0.06]
Parental education	0.16a	-0.01	0.11a	0.01
	[0.03]	[0.01]	[0.02]	[0.01]
Siblings	0.01	0.09a	0.01	0.15a
	[0.03]	[0.02]	[0.02]	[0.02]
Language	-0.16	0.18b	0.11	0.19a
	[0.18]	[0.08]	[0.16]	[0.07]
Asset score	0.19a	-0.05c	0.15a	-0.08a
	[0.05]	[0.03]	[0.04]	[0.02]
School	-0.03	-0.11	0.25	-0.05
	[0.37]	[0.19]	[0.30]	[0.15]
Khorat	0.58b		0.44c	
	[0.27]		[0.23]	
Eastern Seaboard	-0.3		-0.48b	
	[0.28]		[0.23]	
Other destination	0.12		0.06	
	[0.15]		[0.13]	
Buriram	0.99a			
	[0.21]			
Time fixed effects	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes
Village fixed effects	Yes	Yes	Yes	Yes
Constant	-2.25	-0.8	-3.31	0.23
	[2.67]	[1.42]	[2.95]	[1.57]
Observations	6984	6984	11795	11795

Robust standard errors in brackets, corrected for intra-cluster correlation at the individual level
c significant at 10%; b significant at 5%; a significant at 1%

Table 9. Additional Robustness Checks for IV Model

	IV Probit (1) Migrants to Buriram dropped		IV Probit (2) Military as unique instrument	
	student	network	student	network
Network Size	-0.19a		-0.46a	
	[0.07]		[0.07]	
Rain volume (village)		-1.01		
		[1.01]		
Return		0.08a		
		[0.01]		
Military		0.14a		0.07
		[0.05]		[0.04]
Age in 2000	0.04	0.002	0.04	0.02
	[0.07]	[0.06]	[0.05]	[0.07]
Male	-0.23	0.07	0.01	0.09
	[0.14]	[0.12]	[0.13]	[0.09]
Single	0.60a	-0.13	-0.06	0.35
	[0.14]	[0.13]	[0.14]	[0.27]
Parental education	0.17a	0.002	0.09	0.01
	[0.03]	[0.02]	[0.02]	[0.06]
Siblings	0.03	0.004	0.02	0.02
	[0.03]	[0.02]	[0.02]	[0.02]
Language	-0.15	-0.26	-0.17	-0.27
	[0.21]	[0.18]	[0.13]	[0.20]
Asset score	0.16a	-0.13c	0.04	-0.13b
	[0.05]	[0.07]	[0.08]	[0.06]
School	-0.04	-0.51	-0.29	-0.14
	[0.40]	[0.59]	[0.31]	[0.57]
Khorat	-0.15	-5.54a	-2.16a	-5.64a
	[0.51]	[0.32]	[0.68]	[0.28]
Eastern Seaboard	-0.93b	-4.22a	-1.91a	-4.00a
	[0.41]	[0.24]	[0.30]	[0.23]
Other destinations	-0.11	-1.42a	-1.42a	-0.56a
	[0.18]	[0.16]	[0.19]	[0.16]
Buriram			-0.71	-2.86a
			[0.58]	[0.29]
Time fixed effects	Yes	Yes	Yes	Yes
Age fixed effects	Yes	Yes	Yes	Yes
Village fixed effects	Yes	Yes	Yes	Yes
Constant	-3.81	6.46a	0.57	4.66c
	[2.98]	[2.33]	[2.19]	[0.71]
Observations	2225		2627	
Pseudo Loglikelihood	-4560.7		-8378.4	

Robust standard errors in brackets, corrected for intra-cluster correlation at the individual level
c significant at 10%; b significant at 5%; a significant at 1%