Popul Res Policy Rev (2010) 29:659–685
DOI 10.1007/s11113-009-9165-2

# Increasing Migration, Diverging Communities: Changing Character of Migrant Streams in Rural Thailand

**Filiz Garip · Sara Curran**

Received: 26 February 2008 / Accepted: 18 September 2009 / Published online: 2 October 2009


**Abstract** This paper studies how increasing migration changes the character of migrant streams in sending communities. Cumulative causation theory posits that past migration patterns determine future flows, as prior migrants provide resources, influence, or normative pressures that make individuals more likely to migrate. The theory implies exponentially increasing migration flows that are decreasingly selective. Recent research identifies heterogeneity in the cumulative patterns and selectivity of migration in communities. We propose that this heterogeneity may be explained by individuals' differential access to previously accumulated migration experience. Multi-level, longitudinal data from 22 rural Thai communities allow us to measure the distribution of past experience as a proxy for its accessibility to community members. We find that migration becomes a less-selective process as migration experience accumulates, and migrants become increasingly diverse in socio-demographic characteristics. Yet, selectivity within migrant streams persists if migration experience is not uniformly distributed among, and hence not equally accessible to, all community members. The results confirm that the accumulation and distribution of prior migrants' experiences distinctly shape future migration flows, and may lead to diverging cumulative patterns in communities over time.

**Keywords** Internal migration · Cumulative causation · Selectivity · Thailand

F. Garip (✉)
Department of Sociology, Harvard University, Cambridge, USA
e-mail: fgarip@wjh.harvard.edu

S. Curran
Henry M. Jackson School of International Studies and Daniel J. Evans School of Public Affairs, University of Washington, Seattle, USA
e-mail: scurran@u.washington.edu

## Introduction

This paper studies how social, economic, and demographic characteristics of migrant streams change as migration gains prevalence in sending communities. Prior research suggests that accumulation of past migration experience initiates a process of 'cumulative causation' through which migration flows become self-sustaining (Massey 1990). With each migrant from a community, social networks that connect migrants in destination to individuals in origin expand across community members. These networks alter the context of future migration by providing access to resources, influence, or normative pressures from prior migrants. The theory posits that access to migrant networks makes future migration more likely, eventually dampening the effect of other social, economic, or demographic factors on migration, and lowering the selectivity within migrant streams, predictions that are confirmed by empirical evidence (Dunlevy 1991; Massey and Espinosa 1997; Massey et al. 1994).

Recent research also points to heterogeneities in migration outcomes that cannot be completely explained through the current cumulative causation framework. Studies find that the effect of social networks on migration is not necessarily uniform across settings, and may be shaped by the structure of community networks (Garip 2008), gender relations (Curran et al. 2005; Kanaiaupuni 2000), and sending or receiving community contexts (Curran and Rivero-Fuentes 2003; Fussell and Massey 2004). These findings imply that social networks can create differential migration outcomes for different groups of individuals or in different settings. Because social networks feed a cumulative migration mechanism, these differences in their effects are likely to grow and create divergent migration patterns across communities over time. Despite their critical theoretical and practical implications, these ideas have not been incorporated into mainstream cumulative causation theory.

This paper seeks to fill this gap. Prior work focuses on the level of past migration in explaining future migration flows and effectively assumes that all individuals in a community can easily and equally access resources or influence provided by prior migrants (Massey et al. 1994). We extend prior work by considering how the level of accumulated migration experience, as well as individuals' extent of access to it, influence future migration flows. We use the distribution of past migration experience across individuals in a community as a proxy measure for observing individuals' potential for accessing prior migrants. We suggest that the level and distribution of past migration experience can distinctly shape migration flows out of a community, and lead to divergent patterns of cumulative causation across communities.

## Cumulative Migration Patterns in Thailand

We study 22 villages in Northeastern Thailand to understand the relationship between cumulative migration patterns and the changes in the character of migrant streams. These villages are located in the historically poor, rural district of Nang Rong, and due to poverty, past high fertility and limited arable land for future development, provide an important source of migrants to urban centers, primarily Bangkok. We observe the period from 1984 to 2000, when Thailand's shift from an agriculture-based export

economy to a manufacture-based export economy occurred and migration took on added significance in Thai livelihoods (Bello et al. 1998; Phongpaichit 1980; Phongpaichit and Baker 1998; Warr 1993; Warr and Nidhiprabha 1996).

During the period from the mid-1980s to mid-1990s, Thailand's economy grew on average 10% per year (Bello et al. 1998; Warr and Nidhiprabha 1996). This growth was fueled by production in export manufacturing, which was a result of rising wages in nearby, newly industrialized countries and changes in textile import quotas to the US, as well as increases in foreign direct investment, primarily from Japan (Nidhiprabha 1994; Phongpaichit and Baker 1998). With the rise in manufacturing export came an increased demand for labor. Rural migrants, mostly from the Northeastern part of the country, provided much of this labor (Chamratrithirong et al. 1995; Mills 1997; Phongpaichit and Baker 1998). This period of expansive growth began to slow in the mid-1990s. In 1996 export growth slumped from over 20% to zero, partly due to increasing competition from China and India. In 1997 the Asian financial crisis hit Thailand, leading to a devaluation of the Thai currency, baht, and precipitating a brief recession. Unemployment rates increased as a consequence, and migration flows from rural to urban regions slowed.

Employing social survey data from the Nang Rong Project, we study this period of economic growth and its downturn in Thai history from 1984 to 2000. The data contain information on all individuals (between the ages 13 and 41) in the study villages, not just a random sample, and allow for the only longitudinal analysis of the accumulation and distribution of migration experiences over a 16-year period.

Figure 1 shows a pattern of dramatic growth in the migration prevalence of villages over 16 years. Migration prevalence is defined as the percentage of people who have ever migrated in a village up to a given year. There is also considerable variation across villages, which is maintained over time. In 1984, at the low end, in one village only 7% had ever migrated and at the higher end 31% had ever migrated. By 2000 all villages had increased their prevalence rates, albeit at different rates, but the wide range between the high and low prevalence villages is still apparent.

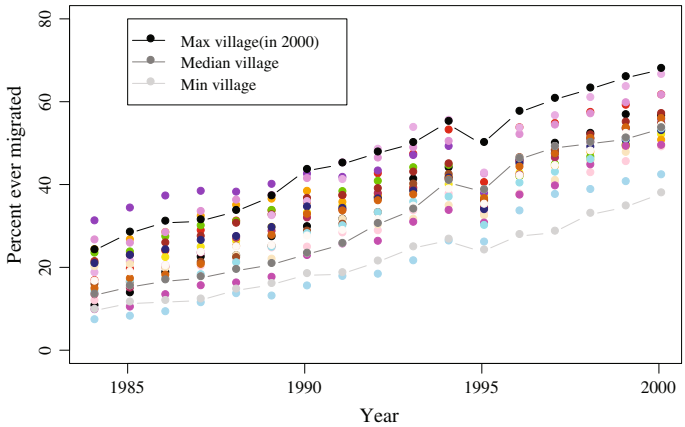


**Fig. 1** Migration prevalence in 22 villages in Nang Rong, Thailand. *Lines* for three villages with the maximum, median and minimum prevalence in 2000 are shown separately

These descriptive data lend preliminary support to cumulative causation theory, which suggests that accumulated migration experience in a community reduces the risks of migrating for new migrants and makes future migration more likely (Massey 1990). The theory also predicts declining migrant selectivity with increasing community migration experience. This hypothesis has only been tested in one context, Mexico-US migration (Massey et al. 1994), where patterns of accumulated migration experience were summarized into a measure of a community's migration prevalence. We can also evaluate this hypothesis in the Thai context. However, our evaluation elaborates the measure of migration prevalence in two ways by creating an index that is the product of the number of *migration trips*, rather than a simple count of migrants, and the inequality in the distribution of that experience across individuals in a community. We claim and show that the level and distribution of past migration experience differentially affects future patterns and selectivity of migration in Thai villages.

There are several differences between the Thai and Mexico-US migration cases, which may have implications for the generalizability of our results. While Mexico-US flows began almost a century ago in many Mexican communities, migration in Thailand is a rather recent phenomenon, which has gained prevalence from the mid-1980s onward. Our data allow us to observe the initiation and perpetuation of cumulative migration flows, which are unlikely to have reached saturation by 2000. Despite the short tenure of migration, migration prevalence rates we observe are higher than the US prevalence rates in Mexican communities (Massey et al. 1994). This difference may be explained by the type of migration (internal versus international) or types of communities studied (small villages in Thailand versus large localities in Mexico).

Rural to urban moves within Thailand carry significant costs and risks for migrants, albeit not commensurate with those faced by international migrants. During our fieldwork in Nang Rong, villagers told stories of migrants being mistreated, not being paid in full, or losing all their belongings to drinking and gambling.[1] Social ties to other migrants, according to our informants, helped alleviate concerns of exploitation in destination, and mitigated the risks of migrating for potential migrants.

These costs and risks are minor compared to the grave dangers faced by international, mostly undocumented, migrants from Mexico to the US. To give one example, a significant number of Mexican migrants lose their life to hypothermia, dehydration, heat stroke, drowning, traffic accidents or homicide while trying to cross the US border (Cornelius 2001; Durand and Massey 2003; Eschbach et al. 1999). Given the higher risks of migrating, Mexican migrants are more likely to rely on the prior experience of trusted family or community members compared to rural-urban migrants in Thailand. Moreover, in the Thai case, most migrants visit their communities twice a year, at which time they take along new migrants to urban destinations. The frequency of contact with origin is likely to be lower for Mexican

[1] Both authors have spent significant time in the field including in 1992, 1994, 1997, 2000, 2005, 2007. In 2005, over a 3-week period we conducted in-depth and focus group interviews with current and return migrants in eight of the 22 villages, with the participation of 158 individuals. Although these qualitative data are not used directly in this study, they inform our theoretical insights and empirical analysis.

migrants in the US due to the cost and difficulty of travel across the border. As a result, in the Mexican case, while it is potentially more important to have contact with prior migrants, it is less easy to do so compared to the Thai case. Given these differences, we expect our results from Thailand to provide a lower bound for the strength of the mechanisms we study. Specifically, if the level and accessibility of resources from prior migrants affect migrant selectivity in the Thai case, then we would expect a similar, but stronger, relationship in the Mexico-US migration case, all else equal.

The remainder of this article is organized as follows: We first provide an overview of cumulative causation theory and its predictions regarding selectivity of migration. We review empirical studies that suggest heterogeneity in cumulative migration patterns and provide an extension to cumulative causation theory that leads to novel predictions. We then describe the data, operational measures of concepts and devise a novel method to capture the cumulative migration patterns and changing character of migrant streams. We present our results from the empirical analysis and conclude with remarks on the implications for future research.

## Theoretical Background

Prior research identifies migration as a dynamic and cumulative process by demonstrating how past migration patterns influence future levels and directions of movement. We can distinguish two components of this process conceptually. First, the objective conditions creating a migration stream in the first place are likely to continue and cause further migration (Fuller et al. 1985). Second, social networks to prior migrants are likely to expand and facilitate the attraction of new migrants over time, propelling more migration. Focusing on this second aspect, in his seminal article, Massey (1990) coins this dynamic and self-feeding mechanism of migration as 'cumulative causation.'

The idea underlying cumulative causation is that, by connecting prior migrants to community members, social networks alter the context within which future migration decisions are made. First, networks act as hubs of information or help from prior migrants, and reduce the costs and risks of migrating for individuals (Massey and Garcia-Espana 1987). The visible signs of increased earnings of prior migrants, in some cases, cause other community members to migrate as well (Stark and Taylor 1991). Ties to prior migrants also foster the diffusion of migration norms, and migrating becomes a 'rite of passage' for young adults (Kandel and Massey 2002; Piore 1979).

The theory predicts that, through resource exchange, influence, or normative pressures, social networks to prior migrants increase future migration flows, eventually dampening the effect of other social, economic, or demographic factors on migration and lowering the selectivity within migrant streams. Prior work confirms these empirical predictions. Dunlevy (1991) and Massey and Espinosa (1997) demonstrate exponentially increasing migration patterns between Mexico and the US, while Massey et al. (1994) note declining selectivity and increasing diversity of migrant streams over time. These findings hold while controlling for

conditions in origin communities or despite changes that raise or lower barriers to migration. Similar to prior work, *we expect to observe a declining migrant selectivity in terms of socio-demographic characteristics due to the accumulation of past experience in Thai villages*.

Recent work, however, shows that the process of growing migration momentum is not uniform, but exhibits significant heterogeneity across origin communities, destinations, and among migrants. Fussell and Massey (2004) demonstrate that cumulative causation varies in its influence across rural and urban origin communities. Similarly, Taylor (1986) and Curran and Rivero-Fuentes (2003) find that the influence of social networks varies across destinations. Others show how the gendered nature of migrant networks in both origin and destination yield varied impacts on migrant outcomes (Curran et al. 2005; Curran and Rivero-Fuentes 2003; Hagan 1998). Research also suggests that social categories of kinship, ethnicity, or community differentially structure social networks and impact migration outcomes (Bauer and Zimmerman 1997; Curran et al. 2005; Winters et al. 2001). These findings provide evidence that supports the theory of cumulative causation, but also suggest that there are critical mechanisms creating divergent outcomes that have yet to be systematically explored.

This study seeks to identify a network-based mechanism that may lead to divergent patterns and selectivity of migration across communities. Rather than observing specific destination conditions or gender relations as in prior work, we focus on a general characteristic of social networks: their differential accessibility to individuals in a community. Cumulative causation theory assumes that all individuals in a community are equally likely to have access to and be influenced by migrant networks. However, based on gender differences, ethnic cleavages, or kinship structures, access to migrant networks may be constrained for certain groups of individuals in a community. Differences in network access may differentiate migration outcomes for individuals and significantly alter cumulative migration trajectories across communities.

We propose that individuals' differential access to migrant networks in a community is a critical mechanism moderating the cumulative causation of migration. The ideal way to measure access is to collect social network data in communities, and construct a measure of ties to prior migrants for each individual. In the absence of network information, we suggest that the degree to which migration experiences are concentrated among a few individuals or distributed more broadly across many individuals within a community provides an alternative measure for individuals' access to migrant networks. To clarify, consider two villages of equal size with equal amounts of past migration experience, say, accumulated trips. In village A, there are only 5 migrants making those trips, while in village B there are 20 migrants. Although the total amount of resources available to potential migrants is the same in both villages, the resources are clearly more widespread and accessible in village B since more people contribute to their accumulation. We assume that the distribution of migration experience among community members serves as a reasonable proxy for individuals' degree of access to migratory resources. In communities where the distribution of past experience is not uniform, individuals will be less likely to access resources (or be exposed to influence) from prior

migrants, in which case their own characteristics will be important in determining migration outcomes. In the Thai villages, *we expect to observe a persistent migrant selectivity if the accumulated migration experience is not uniformly distributed among, and hence not equally accessible to, community members*.

A competing explanation for declining selectivity of migration proposes exogenous factors that are independent of past migration paths. In *Men in a Developing Society*, Balan et al. (1973) attribute the declining selectivity of migration to static and diminishing reservoir of potential migrants. The authors argue, "a combination of increasing demands put upon the rural reservoir by the continuous high rate of urbanization, and in particular by the rapid growth of [destination], with either an absolute decline in the size of the reservoir or at least a decline in its relative rate of growth produces a situation whereby the migratory flow must be supplied from an increasingly representative sample of the origin population." (p. 168) As an alternative to cumulative explanations of migration, *we can expect to observe decreasing migrant selectivity due to an increasing demand for migrants in the destination communities, under the condition that the rate of demand grows faster than the replenishment rate of individuals*. Given the rapid economic growth from mid-1980s to mid-1990s in Thailand, and the resulting increase in the labor demand in urban centers, this alternative hypothesis may be especially relevant in our case.

## Methods

### Data

We use data from the Nang Rong survey, a longitudinal data collection effort conducted by the University of North Carolina and Mahidol University in Thailand.[2] Nang Rong is a relatively poor, rural district in Northeastern Thailand, and a major provider of migrants to urban centers, primarily Bangkok. The survey captures the period from 1984 to 2000, when Thailand's economy shifted from agriculture to manufacturing, propelling the migration from rural to urban regions.

We employ the first three waves of data for our analysis. The 1984 survey was a census of 51 villages and included information on individual demographic data, household assets and village characteristics. The 1994 data collection not only replicated the 1984 survey, but also included a retrospective life history about education, work, and migration, as well as key social and demographic events from all individuals aged 13 to 35. The 2000 survey followed all of the 1984 and 1994 respondents as well as new residents, and collected additional life history information.

The 1994 and 2000 surveys included a migrant follow-up component in 22 of the original 51 villages, which identified the migrants who were absent from their origin households during the time of the survey and found them in four major destinations: the provincial capital, Buriram; the regional capital, Korat or Nakhon Ratchasima;

[2] More information can be found at http://www.cpc.unc.edu/projects/nangron.

Bangkok and the Bangkok Metropolitan Area; or Eastern Seaboard provinces.[3] On average, about 44% of the migrants were successfully interviewed at some point in the 6 months following the village surveys. The success at finding migrants is considered remarkably high for this kind of follow-up (Rindfuss et al. 2007). (The potential bias due to missing migrants is analyzed in Appendix 2.)

The Nang Rong survey data provide unique advantages for testing cumulative causation of migration. Most surveys collect data from randomized samples of respondents. Our data cover *almost all individuals* between the ages of 13 and 41 who had ever resided in the study villages between 1983 and 2000.[4] We are able to measure the distribution of migration experiences in villages, which are central to our account of cumulative causation. The data used in this analysis are based on the life history survey, which begins with 13–35 year olds in 1994 and follows them retrospectively from 1984 to 1994. Adding the 6-year retrospective life history collected from 18 to 41 year olds in 2000, the resulting data set covers a 16-year period. To obtain household and village level indicators, life history data are merged with four cross-sections of household censuses (1984, 1988, 1994, 2000) and three cross-sections of village-level surveys (1984, 1994, 2000).

An important shortcoming of the data is the skewed age distribution due to retrospective data collection. We observe 13–25 year old individuals in 1984, 13–35 year olds in 1994 and 18–41 year olds in 2000. The non-uniform age composition over time could bias our results on migrant selectivity. We address this issue by conducting additional analysis on a sample of 18–25 year-old individuals (the age group present in each year) and present the results in Appendix 2.

## Operational Measures

We define a migrant as a person who has been out of Nang Rong for more than 2 months in a year.[5] To measure community experience, we build on the 'migration prevalence ratio' proposed by Massey et al. (1994) and defined as the percentage of people who have ever been a migrant in a community-year. We improve this measure on two accounts. First, prevalence ratio treats migration as a binary outcome, and equates a community's migration history with a count of its migrants. The implicit assumption is that each migrant contributes equally to the stock of migration information. Rather than a count of migrants, the accumulated number of trips to destination better captures the extent of available resources, while allowing

[3] We restrict our analysis to 22 villages (out of the 51 original villages) where migrants were followed-up in destination.

[4] As with any household registration-based data collection, the Nang Rong survey may have missed individuals residing in household that were not registered during survey years, but were present in non-survey years. Based on our regular visits to the site in between survey years, and given the small size of villages, this number is very minor. Other households and individuals lost to follow-up include entire households who moved out of the village to destinations not pursued through the study in 1994 and 2000. However, these instances were relatively rare, since, more often than not, at least one member of the origin household remained in the village living in another house.

[5] This definition is from the survey and reasonable in the Thai setting, since the majority of migrants make one trip of long duration to their destinations.

for variability across migrants (Curran et al. 2005; Massey and Zenteno 1999). Second, prevalence ratio treats community as a homogeneous entity and implicitly assumes that migration experiences in a community are uniformly accessible to all individuals. However, past migration experiences can be concentrated in few individuals in a community, limiting individuals' degree of access to the resources embodied in them. An ideal measure should incorporate the distribution of past experiences as well as their amount.

We measure the amount of past experiences in a community using the accumulated number of prior migrant trips by community members up through the previous year. A migrant trip is defined as a move to a destination followed by return to the origin village. This definition captures migrants' frequency of contact with origin villages, which is important for migrants to transmit resources or influence to potential migrants.[6] We combine the level and distribution of past experiences in an indicator as follows:

$$\text{Migration History Index}_{vt} = \mu_{vt}(1 - I_{vt}) \tag{1}$$

where $\mu_{vt}$ is the mean number of trips to a destination by members of village $v$ up to time $t - 1$ and $I_{vt}$ is the inequality in the distribution of trips among village members in time $t - 1$. Inequality implies reduced access to social networks for some individuals, and the mean experience, $\mu_{vt}$ has a lower contribution to the migration history the higher this inequality. The measure is decomposable into mean experience and inequality components, which allows us to observe how each component differentially alters migrant selectivity in the Thai villages.

To measure inequality, we compute the commonly used coefficient of variation, $\text{CV}_{vt}$ of village $v$ up through time $t - 1$, by dividing the standard deviation of accumulated individual trips, ($\sigma_{vt}$), by its mean ($\mu_{vt}$). The coefficient of variation is rescaled to vary between 0 and 10. If every village member has equal number of trips, standard deviation equals zero and inequality index reaches its minimum value of 0.

We identify age, sex, marital status, education (measured by years of schooling) and wealth (proxied by total land owned in rai, about 0.5 acres) as important characteristics that migrants are likely to be selected on. A binary indicator of village remoteness to urban centers[7] is included in our models to capture development level. Migrant follow-up rate accounts for differences among villages in survey success. Several annual statistics capture the changing labor market conditions in Thailand from 1984 to 2000.[8] *Unemployment rate* and *annual growth in GDP* are included to

[6] During our fieldwork, participants indicated that only few households had telephones in Nang Rong villages during the study period, and migrants typically contacted their households through return visits, and rarely via letters. Participants repeatedly noted that it was during the return visits that migrants helped potential migrants by giving them information, or taking them along to their places of destination.

[7] A village is considered remotely located if there are three or more obstacles to traveling to the district town. The obstacles are the presence of a portion of the route to the district town that is a cart path (unpaved, rutted, and narrow), the lack of public transportation to the district town, travel to the district town takes an hour or more, that during the year there are four or months of difficult travel to leave the village, and it is 20 or more kilometers to the district town.

[8] These statistics are compiled from various resources, such as International Labor Organization Database, Thai National Statistics Office, and reports prepared by the World Bank and Thailand Development Research Institute, and are available from the authors upon request.

measure the changes in labor supply and demand, respectively. *Productivity-wage gap in agriculture* measures the trend in wages for farm jobs relative to average productivity in those jobs. A negative value would indicate wages lagging behind productivity and suggest a potential 'push' factor to migrate to urban areas to seek manufacturing or service jobs. *Relative average wage in Bangkok versus Northeast* provides a proxy for destination-origin wage differentials. *Percent employed in manufacturing* captures the trend in labor demand in manufacturing. (Descriptive statistics are provided in the Appendix Table 6.)

### Analytic Strategy

We begin with a descriptive analysis similar to Massey et al. (1994). We group 22 villages from 1984 to 2000 (374 village-years) with respect to migration history index as opposed to their migration prevalence ratio. We then observe how migrant selectivity in age, education, sex, marital status and wealth changes as villages move from low to high levels of migration history. Next, we employ regression analysis to adjudicate between different explanations for the variation in migrant selectivity across villages or time. We construct a model of whether a person migrates in any given year with the socio-demographic characteristics (age, education, sex, marital status and wealth) as explanatory variables. To capture the social context of migration, we include indicators of the level and inequality of community migration experience. Given our theoretical argument, we expect community migration experience to increase an individual's likelihood of migration, and the inequality in the distribution of experience to decrease it. Among the macro-economic indicators included, we expect increasing unemployment rate and productivity-wage gap to decrease individuals' probability of migration, while increasing annual GDP growth, relative average wage in Bangkok versus Northeast, and percent employed in manufacturing should increase it.

We build on this baseline model to test two hypotheses, which suggest that migrant selectivity (i) declines with increasing level of community migration experience, and (ii) rises with increasing inequality of that experience. The first hypothesis implies that the effect of an individual's socio-demographic characteristics on migration should decrease with increasing migration experience. Assume education has a positive effect on migration. An interaction term between years of education and village trips should obtain a negative coefficient. The second hypothesis implies that an interaction term between individual characteristics and inequality in migration experience should amplify the effect of these characteristics on the probability of migration. A third, alternative, hypothesis suggests that migrant selectivity decreases with increasing demand for migrants in destination. An interaction term between individual characteristics and indicators of demand should diminish the effect of these characteristics on migration.

Our data set contains multiple observations for the same individual over time. Let $y_{it}$ represent our dependent variable, such that $y_{it} = 1$ if person $i$ ($i = 1,\ldots,n$) is a migrant at time $t$ ($t = 1984,\ldots,2000$) and $y_{it} = 0$ otherwise. Let $x_{it}$ contain a set of time-invariant (e.g. sex) and time-varying (e.g. age, education, marital status) explanatory variables. To allow for a correlation between the observations of an

individual across time, we introduce an individual-level effect, $u_i$, in the model. The probability of being a migrant, that is $p_{it} = Pr(y_{it} = 1)$, is modeled as a function of $x_{it}$ and $u_i$, using the following *logit* regression model:

$$\log\left(\frac{p_{it}}{1 - p_{it}}\right) = \alpha + \beta' x_{it} + u_i \tag{2}$$

The individual effects, $u_i$, can be treated as either fixed or random. In either case they represent unobserved time-invariant individual-level variables, but a random-effects specification is more appropriate in our case because it allows us to estimate the effects of time-invariant variables (e.g., sex) that cannot be estimated by a fixed-effects model.

## Results

Table 1 presents a descriptive analysis of the migration patterns in Nang Rong villages. The first two columns display the village size and mean number of migration trips in 2000. The fourth and fifth columns document the migration history index in 1994 and 2000, the mid- and end-period of data collection respectively. Migration history index is the average number of accumulated migrant trips in the village weighted negatively by its variation. Villages are ranked by their migration history levels in 2000 and listed in ascending order with respect to their rank (column 3). Given that these villages are in the same district and subject to similar economic conditions, there is striking variation in their migration experiences. Mean number of cumulative trips per person in 2000 reaches very high levels for some villages (1.02 in the highest village), while it lingers at moderate levels for others (0.44 in the lowest village). Villages also differ in their growth of migration history from 1994 to 2000 (column 6). Some villages more than double their migration experience in a period of 6 years, while others show modest growth. Migration prevalence ratio, defined as the percentage of individuals who have ever migrated (column 7), provides a ranking consistent with the migration history index for only five villages. This discrepancy suggests that migration history index captures a different aspect of the migration process than the prevalence ratio.

To track the changes in migrant characteristics, we classify village-years into five progressive categories based on the quintiles of the migration history index. The implicit assumption is that each village-year is 'similar' in terms of migration context to others in its category. We then average the characteristics of individuals in each category, and analyze how these averages vary across levels of migration history.[9]

[9] Massey et al. (1994) use migration prevalence ratio to categorize villages. We performed the descriptive analysis in Tables 2 and 3 using this ratio and the results were similar. This is expected as both indices provide roughly similar, albeit not identical, classifications of villages into quintiles. Although the individual rankings of villages may differ significantly across the two indices, as shown in Table 1, the categorical assignments to quintiles are roughly consistent across the prevalence ratio and the migration history index. However, the similarity of the two indices end there. Prevalence ratio is a simple count of migrants, and cannot be decomposed into mean experience and inequality components as the migration history index. This decomposition is crucial to test our first and second hypotheses in the analysis presented in Tables 4 and 5.

**Table 1** History of internal migration (data collected from 13 to 41 year olds in 22 villages in Nang Rong, Thailand in 1984–2000)

| (1) Village size 2000 | (2) Mean trips per person 2000 | (3) Migration history rank 2000 | (4) Migration history 2000 | (5) Migration history 1994 | (6) Growth in history 1994–2000 (%) | (7) Migration prevalence rank 2000 | (8) Migration prevalence 2000 |
|---|---|---|---|---|---|---|---|
| 315 | 0.44 | 1 | 34 | 20 | 76 | 1 | 38 |
| 259 | 0.57 | 2 | 47 | 20 | 136 | 2 | 42 |
| 293 | 0.60 | 3 | 49 | 29 | 72 | 4 | 49 |
| 191 | 0.65 | 4 | 56 | 26 | 115 | 5 | 49 |
| 349 | 0.66 | 5 | 56 | 35 | 59 | 3 | 49 |
| 293 | 0.68 | 6 | 59 | 37 | 59 | 7 | 52 |
| 328 | 0.68 | 7 | 61 | 39 | 55 | 13 | 54 |
| 323 | 0.71 | 8 | 61 | 41 | 47 | 6 | 50 |
| 441 | 0.68 | 9 | 61 | 37 | 63 | 11 | 53 |
| 362 | 0.72 | 10 | 63 | 33 | 92 | 10 | 53 |
| 409 | 0.72 | 11 | 63 | 40 | 58 | 9 | 53 |
| 468 | 0.71 | 12 | 63 | 36 | 75 | 12 | 54 |
| 312 | 0.70 | 13 | 64 | 42 | 54 | 17 | 56 |
| 254 | 0.75 | 14 | 64 | 44 | 46 | 8 | 53 |
| 302 | 0.73 | 15 | 64 | 39 | 65 | 14 | 54 |
| 233 | 0.74 | 16 | 67 | 47 | 43 | 15 | 55 |
| 254 | 0.76 | 17 | 68 | 39 | 75 | 16 | 56 |
| 322 | 0.78 | 18 | 70 | 46 | 52 | 18 | 57 |
| 273 | 0.78 | 19 | 72 | 55 | 31 | 20 | 61 |
| 310 | 0.82 | 20 | 77 | 52 | 48 | 19 | 61 |
| 312 | 0.90 | 21 | 86 | 59 | 45 | 22 | 68 |
| 122 | 1.02 | 22 | 100 | 66 | 52 | 21 | 66 |

Migration history index is the average cumulative trips of migration in the village negatively weighted by its variation, scaled to 0–100 range. Migration prevalence ratio is the percentage of individuals who have ever migrated

The first panel in Table 2 traces the shifts in migration prevalence as communities pass through different phases of migration history. At the first stage, few people have any migration experience: only 17% of men and 12% of women. In later stages, migration spreads increasingly throughout the population. The prevalence of migration among women slightly lags behind that of men at all stages of the migration process, but the differential grows increasingly smaller. The second panel shows how migration trips accumulate as migratory behavior becomes more diffused in communities. At the first stage, the mean number of trips in a community is 32, a figure that considerably increases in each of the four subsequent categories and reaches 120 trips at the highest level of migration history. Inequality in the distribution of trips follows a reverse trend: as migration history moves from the first to the last stage, inequality (measured by the coefficient of variation scaled to 0–10) declines from 5.7 to 0.9.

**Table 2** Cumulative migration experience (data collected from 13 to 41 year olds in 22 villages in Nang Rong, Thailand in 1984–2000)

| Quintiles | Migration history in community | | | | |
|---|---|---|---|---|---|
| | I | II | III | IV | V |
| Prevalence ratio (%) | | | | | |
| Males | 17 | 27 | 35 | 45 | 56 |
| Females | 12 | 22 | 32 | 40 | 50 |
| All | 15 | 24 | 34 | 43 | 53 |
| Cumulative migrant trips in village | | | | | |
| Males | 20 | 35 | 48 | 58 | 61 |
| Females | 12 | 27 | 44 | 55 | 59 |
| All | 32 | 63 | 92 | 113 | 120 |
| Inequality in distribution of cumulative trips (0–10) | | | | | |
| Males | 4.9 | 3.1 | 2.2 | 1.5 | 0.8 |
| Females | 7.4 | 3.9 | 2.5 | 1.8 | 1.1 |
| All | 5.7 | 3.4 | 2.3 | 1.6 | 0.9 |
| Destination for male migrants | | | | | |
| Bangkok (%) | 30 | 41 | 46 | 45 | 38 |
| Bangkok Metropolitan Area (%) | 6 | 9 | 11 | 14 | 20 |
| Eastern Seaboard (%) | 6 | 9 | 10 | 12 | 15 |
| North East (%) | 42 | 30 | 23 | 20 | 19 |
| Other (%) | 15 | 12 | 10 | 9 | 8 |
| Diversity (above groups, $n = 5$) | 84 | 87 | 87 | 88 | 93 |
| Destination for female migrants | | | | | |
| Bangkok (%) | 44 | 51 | 51 | 47 | 40 |
| Bangkok Metropolitan Area (%) | 6 | 11 | 16 | 20 | 29 |
| Eastern Seaboard (%) | 4 | 6 | 7 | 9 | 11 |
| Northeast (%) | 37 | 24 | 20 | 20 | 17 |
| Other (%) | 9 | 8 | 6 | 4 | 3 |
| Diversity (above groups, $n = 5$) | 77 | 81 | 81 | 84 | 85 |
| Community-years ($N$) | 75 | 75 | 75 | 75 | 74 |
| Years covered | 1984–1993 | 1984–1997 | 1986–2000 | 1991–2000 | 1993–2000 |

In the lower two panels of Table 2, we observe the dispersion of migrant trips across five major destinations (Bangkok, Bangkok Metropolitan Area, Eastern Seaboard, North Eastern Region and Other) using Shannon's entropy index.[10] For

[10] Diversity index is defined as:

$$\text{Diversity} = \frac{-\sum_{i=1}^{n} p_i \times \log(\mathrm{p_i})}{\log(n)} \times 100 \tag{3}$$

where $n$ is the number of possible destinations and $p$ is the proportion of trips to destination $i$. The index varies between 0 and 100. Minimum diversity occurs when all trips are concentrated in one destination and the index equals zero. Maximum diversity occurs when each destination category contains the same proportion of trips, yielding an index of 100.

both men and women, diversity of migrant destinations increases as villages reach high levels of migration history. While majority of migrants in early stages move to the North East, the closest destination to Nang Rong, migrants in later stages spread out to farther and riskier destinations, such as Bangkok or Eastern Seaboard.

Cumulative causation theory posits that the socio-demographic base of migrants should broaden as migration history accumulates in a community due to declining risks of migrating for new migrants. This hypothesis receives preliminary support from the data presented in Table 3, which compares migrants and the overall sample[11] in terms of sex composition, marital status, years of education, and land owned. Each characteristic is summarized in a panel, where its average value for migrants and the overall sample is reported along with their ratio across quintiles of the migration history index. If there is no selectivity and migrants are a random draw from the overall sample, then the migrants-to-overall ratio of means should equal approximately unity.

The first panel shows that migrant stream becomes increasingly female, rising from 37% at the lowest level of migration history to 49% at the highest level. The migrants-to-overall means ratio increases from 0.81 in the earliest stage to 0.98 in the latest stage of migration, suggesting that migrants become increasingly representative of the overall population. In the second panel, we observe an older group of individuals in later stages of migration history. Average age increases from 20.3 to 26.3 for male migrants, and from 19.6 to 25.6 for female migrants. This pattern is an artifact of the retrospective data, where older individuals appear in later years. Compared to migrants, the overall sample is slightly older, yielding a ratio of means that is smaller than 1. This ratio declines from 0.98 to 0.94 for males, and from 0.98 to 0.94 for females, suggesting that migrants become increasingly younger compared to the overall population as migration history grows. Migrants are less likely to be married than the rest of the population as the ratio of the married migrants to the married in the overall sample is always less than 1. This ratio becomes larger at higher levels of migration history, suggesting a declining selectivity of migrants on marital status.

The subsequent two panels in Table 3 examine migrant selectivity on education. Moving from the lowest level of migration history to the highest, the mean years of education increases from 6.3 to 7.5 for male migrants and from 5.9 to 7.2 years for female migrants, possibly reflecting a time trend in the data. Migrants are more educated than the overall population at each stage of the migration history. The selectivity of migrants on education first declines, only to rise again to reach or surpass its former level, as suggested by the trend in the migrants-to-overall ratio of mean education. In terms of wealth, both male and female migrants become less likely to come from poor families as migration history evolves. The ratio of average

[11] Our data was collected retrospectively in 1994 and 2000, and the age distribution of the sample is not uniform across years. The 1994 data set begins with 13–35 year old individuals in 1994, and contains retrospective information on their migration patterns from 1984 to 1994. Similarly, the 2000 wave begins with those aged 18–41 in 2000 and gathers retrospective information from the period 1994–2000. The changing age distribution over time makes it difficult to evaluate the trends in migrant selectivity. We circumvent this problem by comparing migrants to the overall sample when assessing the changes in selectivity across phases of community migration history.

**Table 3** Education, demographics, and wealth of migrants and the overall sample (data collected from 13 to 41 year olds in 22 villages in Nang Rong, Thailand in 1984–2000)

| Quintiles | Migration history in community | | | | |
|---|---|---|---|---|---|
| | I | II | III | IV | V |
| Female (%) | | | | | |
| Migrants | 37 | 43 | 45 | 47 | 49 |
| Overall | 46 | 48 | 48 | 50 | 51 |
| Migrants/overall | 0.81 | 0.91 | 0.93 | 0.96 | 0.98 |
| Mean age of males | | | | | |
| Migrants | 20.3 | 21.5 | 22.9 | 24.7 | 26.3 |
| Overall | 20.8 | 22.3 | 23.8 | 25.8 | 27.9 |
| Migrants/overall | 0.98 | 0.96 | 0.96 | 0.96 | 0.94 |
| Mean age of females | | | | | |
| Migrants | 19.6 | 20.8 | 21.9 | 23.8 | 25.6 |
| Overall | 20.1 | 21.4 | 23.1 | 25.3 | 27.7 |
| Migrants/overall | 0.98 | 0.97 | 0.95 | 0.94 | 0.92 |
| Married (%) among males | | | | | |
| Migrants | 20 | 25 | 35 | 46 | 53 |
| Overall | 30 | 38 | 45 | 53 | 61 |
| Migrants/overall | 0.65 | 0.67 | 0.78 | 0.86 | 0.87 |
| Married (%) among females | | | | | |
| Migrants | 26 | 30 | 37 | 50 | 58 |
| Overall | 37 | 43 | 52 | 63 | 73 |
| Migrants/overall | 0.71 | 0.70 | 0.72 | 0.80 | 0.79 |
| Mean education of males (in years) | | | | | |
| Migrants | 6.3 | 6.6 | 6.8 | 7.0 | 7.5 |
| Overall | 5.7 | 6.0 | 6.2 | 6.6 | 6.7 |
| Migrants/overall | 1.10 | 1.09 | 1.09 | 1.07 | 1.11 |
| Mean education of females (in years) | | | | | |
| Migrants | 5.9 | 6.4 | 6.5 | 6.7 | 7.2 |
| Overall | 5.3 | 5.7 | 6.0 | 6.3 | 6.5 |
| Migrants/overall | 1.11 | 1.11 | 1.10 | 1.07 | 1.11 |
| Land owned by males' households (in rai) | | | | | |
| Migrants | 24.1 | 22.5 | 22.3 | 22.0 | 20.3 |
| Overall | 26.9 | 24.5 | 23.3 | 22.4 | 19.7 |
| Migrants/overall | 0.90 | 0.92 | 0.96 | 0.98 | 1.03 |
| Land owned by females' households (in rai) | | | | | |
| Migrants | 21.5 | 21.4 | 21.1 | 21.6 | 20.5 |
| Overall | 27.5 | 24.5 | 23.3 | 22.1 | 20.0 |
| Migrants/overall | 0.78 | 0.87 | 0.91 | 0.98 | 1.02 |
| Sample size | | | | | |
| Migrants | 4,119 | 6,176 | 8,205 | 8,458 | 7,715 |
| Overall | 17,806 | 21,799 | 24,443 | 25,192 | 22,428 |
| Years covered | 1984–1993 | 1984–1997 | 1986–2000 | 1991–2000 | 1993–2000 |

land of a migrant to that of a village resident increases from 0.90 to 1.03 for men and from 0.78 to 1.02 for women. Migrants become less likely to be negatively selected on wealth as migration history grows.

These results suggest that, as Nang Rong villages move through different stages of migration, migrants become more likely to be selected on age and education, but less likely to be selected on sex, marital status, and land. These patterns may partly reflect time trends, as stages of migration history tend to occur at different historical periods. While the earliest stage of migration history includes communities observed from 1984 to 1993, the final stage covers communities from 1993 to 2000. Another concern is our strategy to group villages by their migration history, which disregards potential heterogeneity across villages. To address both issues, we include a village-specific analysis and compare the trends in migrant selectivity over time.

Box plots in Fig. 2 show the distribution of migrants' average characteristics *relative* to the overall population in the twenty-two villages from 1984 to 2000. Some common trends are: (i) declining age of migrants relative to the population (i.e., migrants-to-overall mean ratio drops below 1 around 1990), (ii) migrants' increasing level of education in comparison to the population, and (iii) first increasing, and then decreasing, relative wealth of migrants. These trends over time are similar to those observed across migration history quintiles in Table 3. To parse out migration-stage effects from year effects, we include trend lines for the two villages with the highest and lowest migration history in 2000. If migration history affects selectivity independently from time, then we should observe differences in selectivity between the two villages. We find that migrants tend to be relatively older and less educated in the village with the highest migration history compared to the village with the lowest one. Imagine a horizontal line crossing the *y*-axis at 1, which would represent the case where migrants are identical to the overall population in terms of key characteristics. The trend line for age and education in the highest migration history village is closer to this imaginary 'no selection' line compared to the line for the lowest village. The pattern for land is more complicated: compared to the overall sample, migrants in the high migration village are poorer in the early years, and richer after 1993. Migrants in the low migration village are similar to the overall population in wealth over time. This simple comparison suggests that accumulation of migration experience affects migrant selectivity independently from time, albeit providing mixed evidence on the direction of this effect. We now turn to more rigorous regression analysis to adjudicate how community migration history and time-specific economic conditions differentially shape the selectivity of migrants.

Table 4 displays model estimates for each of the five migration history categories to observe changes in the importance of key characteristics across migration stages. Included in all models are indicators of age, sex, years of education, marital status, land owned, the level and inequality of migrant trips in village. Controls for remoteness to urban centers and migrant follow-up rate in surveys capture village-specific conditions. Macro-economic indicators of unemployment rate, productivity-wage gap in agriculture, destination-to-origin wage ratio (Bangkok/Northeast) and percent employed in manufacturing capture year-specific changes in the internal migration context of Thailand. The findings mirror our conclusions from the prior descriptive

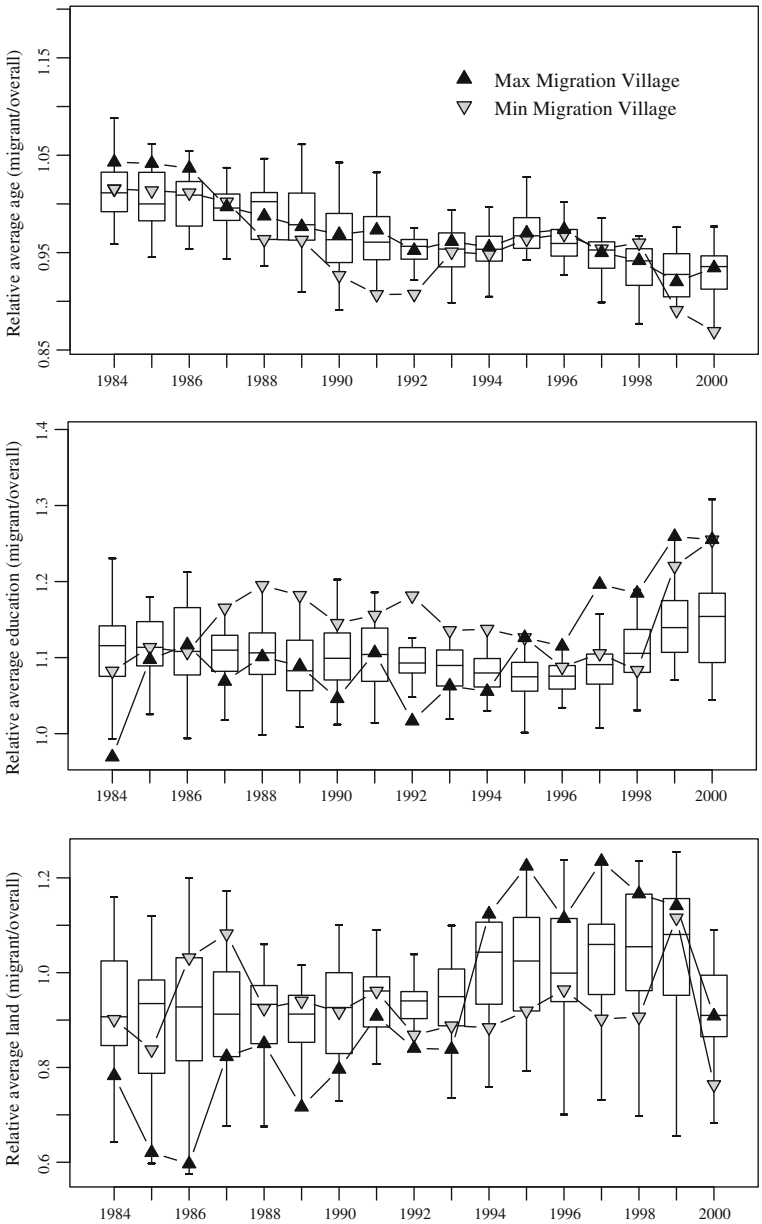


**Fig. 2** Trends in migrant selectivity in villages over time. *Box plot* shows the distribution of migrants' characteristics compared to the overall population across villages. The *lines* for the minimum and maximum migration history villages are shown separately

analysis. The odds ratio of age drops from 1.15 in the first stage of migration history to 0.9 in the fifth stage. The higher propensity of migration for males compared to females declines from 174 to 23% (not significant) as a community moves from the

**Table 4** Random effects logistic estimation of odds of being a migrant in a year models by quintiles of the migration index (data collected from 13 to 41 year olds in 22 villages in Nang Rong, Thailand in 1984–2000)

| Quintiles | Migration history in community | | | | |
|---|---|---|---|---|---|
| | I | II | III | IV | V |
| Individual characteristics | | | | | |
| Age | 1.15** | 1.09#, *** | 1.03#, *** | 0.98#, ** | 0.90*** |
| Sex (male = 1) | 2.74*** | 1.95#, *** | 1.60*** | 1.34#, ** | 1.23# |
| Education (in years) | 1.48*** | 1.50*** | 1.44#, *** | 1.43#, *** | 1.50#, *** |
| Married | 0.10*** | 0.13#, *** | 0.18#, *** | 0.24#, *** | 0.46#, *** |
| Land (in rai) | 0.98*** | 0.98*** | 0.99#, *** | 0.99#, ** | 1.00# |
| Village characteristics | | | | | |
| Village remote? | 0.63** | 1.05 | 1.06 | 1.10 | 1.39#, ** |
| Migrant follow-up rate | 1.06*** | 1.03*** | 1.05*** | 1.05*** | 1.03* |
| Migration context of the community | | | | | |
| Migrant trips in the community | 1.89 | 2.82#, * | 2.69** | 1.89 | 2.58#, *** |
| Inequality of migrant trips | 0.76 | 0.40** | 0.75 | 0.30** | 0.19*** |
| Thai Economic Context | | | | | |
| Unemployment rate (%) | 1.18 | 0.95 | 1.05 | 0.73*** | 0.61*** |
| Annual GDP growth (%) | 1.01 | 1.00 | 1.03** | 1.03 | 1.12*** |
| Productivity-wage gap in agriculture (%) | 1.00 | 1.00 | 1.01** | 1.03*** | 1.05*** |
| Wage ratio (Bangkok/Northeast) | 1.53 | 1.20 | 2.78*** | 5.63*** | 45.56*** |
| Employment in manufacturing (%) | 1.08 | 0.90 | 1.53*** | 1.38*** | 1.62** |
| Wald $\chi^2$ | 6,600*** | 7,951*** | 9,098*** | 10,593*** | 9,663*** |
| *N* (person-years) | 17,806 | 21,799 | 24,443 | 25,192 | 22,428 |

Year and constant are included in all models. Results are presented in odds ratios. Migration experience and inequality are standardized to mean 0, SD 1

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

# $p < 0.1$ (test of difference between coefficients of subsequent quintile models)

first to the final stage of migration history. The detrimental effect of marriage on migration also wanes as migration history grows. The odds of migrating is 90% lower for individuals in the first stage of migration compared to 54% in the final stage. The effect of educational attainment on migration changes in a nonlinear fashion, first increasing then decreasing only to increase again through higher levels of migration history. By contrast, the selectivity of migrants on land declines as migration history grows. In the first stage of migration each rai of land reduces the odds of migrating by 2%, which in later phases drops to 1% and eventually to nil.

The results show how the level and distribution of migrant trips (standardized to zero mean and unit standard deviation for comparability) differentially influence migration propensities. In the first stage of migration history, neither variable has a significant effect. In the second stage, a standard deviation increase above average community trips more than doubles individuals' likelihood of migrating, while a

similar change in the inequality of trips reduces the odds by more than half. The positive effect of migration trips drops only slightly at later stages. By contrast, the negative effect of the inequality of migrant trips increases as migration history grows. The relative importance of the distribution of migration experience, compared to its extent, seems to increase at later phases of a community's migration history.

Village-specific follow-up rate has a positive and consistent effect on migration across the phases of migration history. Not surprisingly, migration seems more likely in villages where surveyors were more successful in identifying migrants in destination. Remote villages, which are initially less likely to send migrants, become more likely to do so as migration history accumulates. Indicators of the economic context, unemployment rate, productivity-wage gap in agriculture, Bangkok-to-Northeast wage ratio and percent employed in manufacturing significantly affect migration patterns in later stages of migration history, possibly capturing historical period effects.

Overall, the results in Table 4 show that the effects of various determinants on migration vary considerably depending on a community's migration stage. As villages move from low to high levels of migration, migrants become more likely to be selected on age and education, and less likely to be selected on sex, marital status and land. Migrant streams tend to be younger and more educated, include increasingly women and the married, and become representative of the overall population in terms of wealth.

The analysis so far has shown, first, that migrant selectivity changes as a community moves from low to high levels of migration, and second, that the level and inequality of migration experience, as well as changing economic conditions, affect migration patterns in communities. We now combine these two points and and ask: How does migrant selectivity change as a result of changes in (i) the level and (ii) inequality of community migration experience, and (iii) the economic context? Based on our hypotheses, the level of migration experience is expected to decrease selectivity; inequality of experience is expected to increasing it. Increasing demand in destination should also decrease migrant selectivity. These hypotheses are tested in models in Table 5.

The first column shows the baseline model for migration estimated on the whole sample. The propensity to migrate increases with age (by 3%), decreases with marriage (by 67%), is higher for men than women (by 32%). Each additional year of education increases the likelihood of migrating by 35%, while each rai of land owned decreases it by 0.1%. Living in a remote village increases migration odds by 6%. Increasing community migration trips by one standard deviation above its mean more than doubles the odds of migrating, while a commensurate increase in inequality decreases the likelihood of migration by 17%. Given the small size of these villages, where most people know each other, the large negative effect of unequal access is striking.

The economic indicators significantly shape migration flows. Migration rates decrease with increasing unemployment rate, and increase with the growth in the GDP. The growing gap between productivity and wages in agriculture, which provides a 'push' factor to migrate for rural farmers, also increases the odds of migrating. Individuals become more likely to migrate as the ratio of average wages

**Table 5** Random effects logistic estimation of odds of being a migrant in a year—interaction models (data collected from 13 to 41 year olds in 22 villages in Nang Rong, Thailand in 1984–2000)

| | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| Individual characteristics | | | | |
| Age | 1.03*** | 1.06*** | 1.06*** | 1.32*** |
| Sex (male = 1) | 1.32*** | 1.33*** | 1.34*** | 1.75* |
| Education (in years) | 1.35*** | 1.27*** | 1.27*** | 1.82*** |
| Married | 0.33*** | 0.25*** | 0.25*** | 0.49** |
| Land (in rai) | 0.999* | 0.996*** | 0.996*** | 0.973*** |
| Village characteristics | | | | |
| Village remote? | 1.06** | 1.02 | 1.01 | 1.03 |
| Migrant follow-up rate | 1.02*** | 1.02*** | 1.03*** | 1.02*** |
| Migration context of the community | | | | |
| Migrant trips in the community | 2.11*** | 27.11*** | 19.90*** | 5.60*** |
| Inequality of migrant trips | 0.83*** | 1.29*** | 0.90 | 0.94 |
| Thai Economic Context | | | | |
| Unemployment rate (%) | 0.89*** | 0.93*** | 0.93*** | 0.93*** |
| Annual GDP growth (%) | 1.01*** | 1.00 | 1.00 | 1.00 |
| Productivity-wage gap in agriculture (%) | 1.02*** | 1.01*** | 1.01*** | 1.01*** |
| Wage ratio (Bangkok/Northeast) | 2.38*** | 1.58*** | 1.54*** | 1.21*** |
| Employment in manufacturing (%) | 1.38*** | 1.20*** | 1.19*** | 2.07*** |
| Interactions b/w Ind char and migration experience | | | | |
| Age × trips | | 0.92*** | 0.92*** | 0.95*** |
| Sex × trips | | 0.80*** | 0.89** | 0.94 |
| Educ × trips | | 1.00 | 1.02** | 1.08*** |
| Married × trips | | 0.91*** | 0.79*** | 0.87 |
| Land × trips | | 1.00*** | 1.01*** | 1.01*** |
| Interactions b/w Ind char and inequality of mig experience | | | | |
| Age × inequality | | | 1.00 | 0.99 |
| Sex × inequality | | | 1.14** | 1.15** |
| Educ × inequality | | | 1.03*** | 1.02* |
| Married × inequality | | | 0.85** | 0.84** |
| Land × inequality | | | 1.01*** | 1.01*** |
| Interactions b/w Ind char and economic context | | | | |
| Age × employment in manuf. | | | | 0.98*** |
| Sex × employment in manuf. | | | | |
| Educ × employment in manuf. | | | | 0.97*** |
| Married × employment in manuf. | | | | 0.94** |
| Land × employment in manuf. | | | | 1.00*** |
| Wald $\chi^2$ | 40,435*** | 41,847*** | 41,846*** | 41,909*** |
| *N* (person-years) | 111,668 | 111,668 | 111,668 | 111,668 |

Year and constant are included in all models. Results are presented in odds ratios. Migration experience and inequality are standardized to mean 0, SD 1

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

in destination (Bangkok) and origin (Northeast) increases. The availability of manufacturing jobs, measured by percent employed in manufacturing in the country, also increases the odds of migrating.

The second model introduces interactions between community migration trips and socio-demographic characteristics to test the hypothesis of declining selectivity with increasing community migration experience. This hypothesis is supported for age and sex, but not for marital status, education or land owned. The effect of age on migration decreases by 8% for a standard deviation increase above average community trips, while the effect of sex decreases by a remarkable 20%. By contrast, the negative effect of being married on migration becomes more pronounced as community migrant trips increase. An increase of a standard deviation in community trips decreases the odds of migrating for married individuals by 9%. The positive effects of years of education and land owned do not change with community migration experience, as the interaction-term coefficients for both education (not significant) and land are close to unity.

Interactions between the inequality of community trips and socio-demographic characteristics are introduced in the third model. Given the hypothesis of increasing migrant selectivity with increasing inequality of trips in community, we expect the interaction term coefficients to be positive for age, sex, education and land, and negative for marital status. The estimates confirm our expectations for all variables but age. The higher likelihood of men's migration compared to women becomes more pronounced in communities with an unequal distribution of migration experience. A standard deviation increase in inequality of trips increases the effect of sex on migration by 14%. Similarly, the effect of each year of education on migration increases by 3%, while the effect of each rai of land owned increases by 1%. The negative effect of being married becomes larger as inequality increases, and married individuals become 15% less likely to migrate with a standard deviation increase above the mean inequality.

In sum, the hypothesis of declining selectivity with increasing community migration experience holds for age and sex, but not for marital status, education or land. The opposing hypothesis of increasing selectivity with increasing inequality of migration experience holds for sex, marital status, education and land, but not for age. We now consider the alternative hypothesis that decreasing migrant selectivity results from an increasing demand in destination. This hypothesis is especially viable in the rapidly changing economic context of Thailand during the study period. We introduce interaction terms between socio-demographic characteristics and an indicator of labor demand (percent employed in manufacturing) in the fourth model of Table 5 to test this idea. The estimates show that as demand for manufacturing workers increases, migrants become less likely to be selected on age, sex (not significant), education and wealth, but more likely to be selected on marital status. Reassuringly, the inclusion of these interactions does not alter the results related to the first two hypotheses, which are the main focus of our analysis. Even controlling for the changing economic context of the country, the level and inequality of migration experience in communities uniquely alter migrant selectivity as the theory of cumulative causation suggests. (A number of robustness checks with alternative samples are presented in Appendix 2.)

## Conclusion

In an age of increasing migration, anticipating and directing migration flows is a major concern for policy makers around the world. A critical research finding in the migration literature shows that migration flows can develop a self-sustaining momentum that is difficult to control or redirect. This phenomenon, called the cumulative causation of migration, occurs because prior migrants provide resources of information or assistance, influence or normative pressures that make individuals in origin communities more likely to migrate. Cumulative causation explains how past migration patterns determine future magnitudes and directions of movement and diminish the importance of other social, economic, or demographic factors that influence migration.

Empirical studies evaluating the cumulative causation theory are substantial, especially for the Mexican-US migration flows, demonstrating exponentially increasing patterns of migration that are decreasingly selective on individual characteristics, as the theory predicts. However, recent research also shows significant heterogeneity in patterns and selectivity of migration across communities. In this article, we proposed that this heterogeneity in migration outcomes can be explained by further theorizing the mechanisms underlying cumulative causation. We argued that the differential accessibility of previously accumulated migration experience to individuals is one such mechanism that may disrupt the cumulative migration dynamic and trends in selectivity, and lead to divergent migration patterns in communities.

We built on the analytical approach of Massey et al. (1994) to study the patterns of migration and migrant selectivity out of 22 rural communities in Northeastern Thailand. This approach categorizes communities by their count of migrants or 'migration prevalence ratios' and observes patterns of change in migrant characteristics. Differently, we proposed a 'migration history index,' which combines the extent of past migration experience and its distribution among individuals, which signifies its accessibility to community members. We used this index to categorize the 22 villages over a 16-year time period (1984–2000) into five progressive stages of migration history. We found that, as a community moves from initial to later stages of migration history, migrants become less likely to be selected on sex, marital status and land. Hence, migrant streams tend to include increasingly women and the married, and become representative of the overall population in terms of wealth.

To better evaluate the sources of the declining migrant selectivity, we decomposed the migration history index into its two constitutive components, the level and distribution of migration experience. Focusing on the level of migration experience, similar to prior studies and in line with the expectations of the cumulative causation theory, we found that migration becomes a less-selective process as migration experience accumulates, and migrants become increasingly diverse in terms of age and sex in the context of Thai internal migration. Contrarily, we also found that the selectivity within migrants by education and wealth remains constant, while selectivity by marital status increases as migration gains prevalence. Different than any prior study, we also considered how the distribution of migration

experience in a community, signifying its accessibility to individuals, shapes the selectivity of migration. We found that migrants' likelihood of being selected on age, sex, marital status, education and wealth persists or increases with increasing inequality in the distribution of migration experience in a community. These findings remained robust when we controlled for the potential time trends in migrant selectivity due to the rapidly changing economic context of Thailand during the study period.

Hence, we provide additional evidence from Thailand to a dynamic relationship between community migration experience and migrant selectivity first identified in the Mexico-US context; we show that similar mechanisms govern internal migration as international migration; and demonstrate the usefulness of a new methodological tool to disentangle these dynamics. Theoretically, our findings help qualify the predictions of cumulative causation theory regarding migrant selectivity, and point to the importance of considering the distribution, as well as the level, of community migration experience. We thus extend the reach of cumulative causation theory to explain the heterogeneity in patterns and selectivity of migration observed in reality.

Future research should systematically analyze how the distribution of migration experience shapes migration flows or selectivity in other settings to generalize our findings. Another vein of research should focus on identifying and measuring the nature of social mechanisms underlying cumulative causation. The literature on cumulative causation has assumed, rather than shown, that individuals are influenced by others in their social networks. These networks are never empirically measured, but rather supposed to exist between individuals sharing the same household or community. Furthermore, while connected individuals are believed to affect each others' behavior, the underlying mechanisms (imitation, learning or social influence) have not been identified. In future work, migration scholars should collect data on social networks along with longitudinal data on individuals' migration moves. These data should further be connected to qualitative information illuminating the social mechanisms underlying the correlated behavior of connected individuals. This effort presents the next frontier in identifying the social determinants of migration behavior.

**Acknowledgements** This research was funded by research grants from Center for Migration and Development at Princeton University and NSF (SES-0525942). The authors thank the research team from the Carolina Population Center at the University of North Carolina and the Institute for Population and Social Research at Mahidol University for their data collection efforts and the villagers of Nang Rong district, Buriram province, Thailand for their cooperation.

## Appendix 1

Descriptive Statistics

(See Table 6).

**Table 6** Descriptive statistics (data collected from 13 to 41 year olds in 22 villages in Nang Rong, Thailand in 1984–2000)

| | Mean | SD | Minimum | Maximum |
|---|---|---|---|---|
| Individual characteristics | | | | |
| Age | 23.99 | 6.99 | 13.00 | 41.00 |
| Sex (male = 1) | 0.51 | 0.50 | 0.00 | 1.00 |
| Education (in years) | 6.14 | 2.77 | 0.00 | 19.00 |
| Married | 0.50 | 0.50 | 0.00 | 1.00 |
| Land (in rai) | 23.22 | 24.60 | 0.00 | 912.00 |
| Village characteristics | | | | |
| Village remote? | 0.37 | 0.48 | 0.00 | 1.00 |
| Migrant follow-up rate | 55.36 | 7.95 | 39.71 | 69.51 |
| Migration context of the community | | | | |
| Migrant trips in the community | 0.43 | 0.19 | 0.08 | 1.02 |
| Inequality of migrant trips | 2.61 | 1.61 | 0.00 | 10.00 |
| Thai Economic Context | | | | |
| Unemployment rate (%) | 2.98 | 0.84 | 1.51 | 4.37 |
| Annual GDP growth (%) | 6.39 | 5.49 | −10.51 | 13.29 |
| Productivity-wage gap in agriculture (%) | −3.66 | 10.51 | −22.48 | 11.77 |
| Wage ratio (Bangkok/Northeast) | 2.20 | 0.27 | 1.62 | 2.54 |
| Employment in manufacturing (%) | 11.36 | 2.03 | 7.80 | 14.50 |
| *N* (person-years) | 111,668 | | | |

## Appendix 2

### Robustness Checks with Alternative Samples

We perform robustness checks to address two data-related issues that might bias our results. First issue is related to the migrant follow-up rate in survey data. In the 22 Nang Rong villages, migrants who were absent at the time of the survey were followed up in four major migrant destinations. On average, 44% of migrants were successfully located. To see how the exclusion of the remainder of migrants biases our results, we use the variability among villages in migrant follow-up rates, which range from 40 to 70%. We repeat our most comprehensive analysis (Model 4 of Table 5) on a restricted sample of four villages with the highest follow-up rates (all above 65%). Comparing the coefficient estimates for the whole and restricted samples in Appendix Table 7, we find that, despite the drastic change in sample size, the estimates are mostly similar. The only major change is in the coefficient of sex, which is much higher in the restricted sample. Accordingly, the coefficients for the interaction terms including sex differ remarkably in the two samples. Other coefficients remain consistent in direction, and differ negligibly in magnitude, across samples. This evidence increases our confidence that low follow-up rates in some villages do not bias our results.

**Table 7** Random effects logistic estimation of odds of being a migrant in a year—interaction models with alternative samples (data collected from 13 to 41 year olds in 22 villages in Nang Rong, Thailand in 1984–2000)

| | All sample | Follow-up >65% | 18–25 year olds |
|---|---|---|---|
| Individual characteristics | | | |
| Age | 1.32*** | 1.16** | 0.97 |
| Sex (male = 1) | 1.75* | 30.41*** | 6.20*** |
| Education (in years) | 1.82*** | 1.72*** | 1.28*** |
| Married | 0.49** | 2.86 | 0.44** |
| Land (in rai) | 0.97*** | 0.99 | 0.97*** |
| Village characteristics | | | |
| Village remote? | 1.03 | 0.81*** | 0.83*** |
| Migrant follow-up rate | 1.02*** | 0.89* | 1.04*** |
| Migration context of the community | | | |
| Migrant trips in the community | 5.60*** | 2.99** | 0.91 |
| Inequality of migrant trips | 0.94 | 0.42** | 0.91 |
| Thai Economic Context | | | |
| Unemployment rate (%) | 0.93*** | 0.94** | 0.92*** |
| Annual GDP growth (%) | 1.00 | 1.00 | 1.02*** |
| Productivity-wage gap in agriculture (%) | 1.01*** | 1.00 | 1.02*** |
| Wage ratio (Bangkok/Northeast) | 1.21*** | 1.26 | 2.76*** |
| Employment in manufacturing (%) | 2.07*** | 1.87*** | 1.29* |
| Interactions b/w Ind char and migration experience | | | |
| Age × trips | 0.95*** | 0.94*** | 0.99 |
| Sex × trips | 0.94 | 1.69*** | 1.02 |
| Educ × trips | 1.08*** | 1.13*** | 1.05** |
| Married × trips | 0.87 | 1.25 | 1.40** |
| Land × trips | 1.01*** | 1.01*** | 1.00 |
| Interactions b/w Ind char and inequality of mig experience | | | |
| Age × inequality | 0.99 | 1.03* | 0.97 |
| Sex × inequality | 1.15** | 1.76*** | 1.22 |
| Educ × inequality | 1.02* | 1.01 | 1.03 |
| Married × inequality | 0.84** | 0.88 | 1.38** |
| Land × inequality | 1.01*** | 1.01*** | 1.00 |
| Interactions b/w Ind char and economic context | | | |
| Age × employment in manuf. | 0.98*** | 0.99 | 1.01 |
| Sex × employment in manuf. | 0.98 | 0.76*** | 0.88*** |
| Educ × employment in manuf. | 0.97*** | 0.97*** | 1.00 |
| Married × employment in manuf. | 0.94** | 0.81*** | 0.96 |
| Land × employment in manuf. | 1.00*** | 1.00 | 1.00** |
| Wald $\chi^2$ | 41,909*** | 8,566*** | 15,470*** |
| *N* (person-years) | 111,668 | 21,779 | 44,142 |

Year and constant are included in all models. Results are presented in odds ratios. Migration experience and inequality are standardized to mean 0, SD 1

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The second issue is related to the age structure in the data. The retrospective life history survey was administered to 13–35 year olds in 1994, and 18–41 year olds in 2000. Thus, we observe 13–25 year olds in 1984, 13–35 year olds in 1994 and 18–41 year olds in 2000. The changing age distribution over time may bias the results. To address this issue, we restrict our sample to 18–25 year olds (the age group present in each year), and estimate Model 4 of Table 5. The results presented in Appendix Table 7 show a number of differences from those for the overall sample. First, the coefficient of sex is higher in the age-restricted sample, and the coefficient for age is insignificant. The latter is expected due to the narrow age range of the restricted sample. Different than the overall sample, the selectivity in marital status in the age-restricted sample declines with increasing migration experience. Other coefficients remain similar in direction, but differ slightly in magnitude or significance, across the samples. Despite these minor differences, our main conclusions (regarding the differential effect of the level and inequality of migration experience on selectivity) remain unaltered.

## References

Balan, J., Browning, H. L., & Jelin, E. (1973). *Men in a developing society: Geographic and social mobility in Monterrey, Mexico*. Austin: University of Texas Press.

Bauer, T., & Zimmerman, K. (1997). Network migration of ethnic Germans. *International Migration Review, 31*, 143–149.

Bello, W., Cunningham, S., & Poh, L. K. (1998). *A Siamese tragedy: Development and disintegration in Modern Thailand*. London: Zed Books, Ltd.

Chamratrithirong, A., Archavanitkul, K., Richter, K., Guest, P., Varachai, T., Boonchalaksi, W., et al. (1995). *National Migration Survey of Thailand*. Bangkok, Thailand: Institute for Population and Social Research, Mahidol University.

Cornelius, W. A. (2001). Death at the border: Efficacy and unintended consequences of US Immigration Control Policy. *Population and Development Review, 27*, 661–685.

Curran, S., & Rivero-Fuentes, E. (2003). Engendering migrant networks: The case of Mexican Migration. *Demography, 40*, 289–307.

Curran, S. R., Garip, F., Chung, C., & Tangchonlatip, K. (2005). Gendered migrant social capital: Evidence from Thailand. *Social Forces, 84*, 225–255.

Dunlevy, J. A. (1991). On the settlement patterns of recent Caribbean and Latin immigrants to the United States. *Growth Change, 22*, 54–67.

Durand, J., & Massey, D. S. (2003). The costs of contradiction: US Border Policy 1986–2000. *Latino Studies, 1*, 233–252.

Eschbach, K., Hagan, J., Rodriguez, N., Hernandez-Leon, R., & Bailey, S. (1999). Death at the border. *International Migration Review, 33*, 430–454.

Fuller, T. D., Lightfoot, P., & Kamnuansilpa, P. (1985). Rural-urban mobility in Thailand: A decision-making approach. *Demography, 22*, 565–579.

Fussell, E., & Massey, D. S. (2004). The limits to cumulative causation: International migration from Mexican urban areas. *Demography, 41*, 151–171.

Garip, F. (2008). Social capital and migration: How do similar resources lead to divergent outcomes? *Demography, 45*, 591–617.

Hagan, J. M. (1998). Social networks, gender, and immigrant incorporation: Resources and constraints. *American Sociological Review, 63*, 55–67.

Kanaiaupuni, S. (2000). Reframing the migration question: An analysis of men, women, and gender in Mexico. *Social Forces, 78*, 1311–1347.

Kandel, W., & Massey, D. S. (2002). The culture of Mexican migration: A theoretical and empirical analysis. *Social Forces, 80*, 981–1004.

Massey, D. S. (1990). Social structure, household strategies, and the cumulative causation of migration. *Population Index, 56*, 3–26.

Massey, D. S., & Espinosa, K. (1997). What's driving Mexico-U.S. migration? A theoretical, empirical, and policy analysis. *American Journal of Sociology, 102*, 939–999.

Massey, D. S., & Garcia-Espana, F. (1987). The social process of international migration. *Science, 237*, 733–738.

Massey, D. S., Goldring, L., & Durand, J. (1994). Continuities in transnational migration: An analysis of nineteen Mexican communities. *American Journal of Sociology, 99*, 1492–1533.

Massey, D. S., & Zenteno, R. M. (1999). The dynamics of mass migration. *Proceedings of the National Academy of Sciences, 96*, 5328–5335.

Mills, M. B. (1997). *Thai women in the global labor force: Consuming desires, contested selves*. New Brunswick, NJ: Rutgers University.

Nidhiprabha, B. (1994). *determinants of private investment expenditures and driect foreign investment in Thailand*. Bangkok: Thailand Development Research Institute.

Phongpaichit, P. (1980). The open economy and its friends: The "development" of Thailand. *Pacific Affairs, 53*, 440–460.

Phongpaichit, P., & Baker, C. (1998). *Thailand's boom and bust*. Chiang Mai, Thailand: Silkworm Press.

Piore, M. J. (1979). *Birds of passage: Migrant labor and industrial societies*. New York: Cambridge University Press.

Rindfuss, R. R., Kaneda, T., Chattopadhyay, A., & Sethaput, C. (2007). Panel studies and migration. *Social Science Research, 36*, 374–403.

Stark, O., & Taylor, J. E. (1991). Migration incentives, migration types: The role of relative deprivation. *The Economic Journal, 101*, 1163–1178.

Taylor, J. E. (1986). *Differential migration, networks, information and risk* (Vol. 4). Greenwich, CT: JAI Press.

Warr, P. (1993). *The Thai economy in transition*. New York: Cambridge University Press.

Warr, P., & Nidhiprabha, B. (1996). *Thailand's macroeconomic miracle: Stable adjustment and sustained growth*. Washington, DC: The World Bank.

Winters, P., de Janvry, A., & Sadoulet, E. (2001). Family and community networks in Mexico-U.S. migration. *Journal of Human Resources, 36*, 159–184.