# Is Learning by Migrating to a Megalopolis Really Important? Evidence from Thailand

| | |
|---|---|
| | Machikita Tomohiro |
| | Copyrights / Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO) http://www.ide.go.jp |
| journal or publication title | IDE Discussion Paper |
| volume | 82 |
| year | 2006-12-01 |
| URL | http://hdl.handle.net/2344/334 |

INSTITUTE OF DEVELOPING ECONOMIES

DISCUSSION PAPER No. 82

# Is Learning by Migrating to a Megalopolis Really Important? Evidence from Thailand

Tomohiro MACHIKITA*

* Researcher, Regional Integration Studies Group, Inter-disciplinary Studies Center
IDE (machi@ide.go.jp)

**Abstract**

We examine the effects of learning by migrating on the productivity of migrants who move to a "megalopolis" from rural areas using the Thailand Labor Force Survey. The main contribution is to the development a simple framework to test for self-selection on migration decisions and learning by migrating into the urban labor market, focusing on experimental evidence in the observational data. The role of the urban labor market is examined. In conclusion, we find significant evidence for sorting: the self-selection effects test (1) is positive among new entrants from rural areas to the urban labor market; and (2) is negative among new exits that move to rural areas from the urban labor market. Further, estimated effects of learning by migrating into a "megalopolis" have a less significant impact. These results suggest the existence of a natural selection (i.e. survival of the fittest) mechanism in the urban labor market in a developing economy.

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

Institute of Developing Economies (IDE), JETRO
3-2-2, Wakaba, Mihama-ku, Chiba-shi
Chiba 261-8545, Japan

# Is Learning by Migrating to a Megalopolis Really Important? Evidence from Thailand[*]

Tomohiro Machikita[†]

Version Prepared in December 2006 - First Version in February 2005

## Abstract

We examine the effects of learning by migrating on the productivity of migrants who move to a "megalopolis" from rural areas using the Thailand Labor Force Survey. The main contribution is to the development a simple framework to test for self-selection on migration decisions and learning by migrating into the urban labor market, focusing on experimental evidence in the observational data. The role of the urban labor market is examined. In conclusion, we find significant evidence for sorting: the self-selection effects test (1) is positive among new entrants from rural areas to the urban labor market; and (2) is negative among new exits that move to rural areas from the urban labor market. Further, estimated effects of learning by migrating into a "megalopolis" have a less significant impact. These results suggest the existence of a natural selection (i.e. survival of the fittest) mechanism in the urban labor market in a developing economy.

**JEL Classification Numbers** : D83, J61, R23

**Keywords** : Self-selection; Learning by migrating; Survival of the fittest; Exits; Thailand

[†]Corresponding address: Institute of Developing Economies (IDE-JETRO), 3-2-2 Wakaba, Mihama Chiba 261-8545, JAPAN. Email: machi@ide.go.jp. Phone: +81-43-299-9758. Fax: +81-43-299-9763.

# 1   Introduction

The aim of this paper is to examine the effects of learning by migrating on the productivity of migrants who move to a "megalopolis" from rural areas using the Thailand Labor Force Survey Data, 1994 to 1996. The main contributions of this paper are: (1) providing a simple empirical framework to identify between self-selection effects in migration decisions and learning by migrating effects; (2) discussing the role of the urban labor market (i.e. natural selection or location of human capital accumulation). This study is one contribution to understanding the unobserved characteristics of migrants in a concentrated area. Policy makers, economists, and researchers of other fields such as sociology and anthropology are also interested in sorting through urban immigration and unobserved heterogeneity of urban workers. In particular, policy makers, macroeconomists, and labor economists are interested in the role of such a concentrated area to provide job matching and their aggregate properties. Due to the causal effects of migration decisions and individual characteristics, it has been difficult to identify the true impact of concentrated areas on migrant's wages and job matching. Thanks to exogenous sources of variation (or natural experiments) in the available empirical data, the impact of migration on wages is consistently estimated. Based on these estimates, we can begin to discuss the role of the urban labor market, active labor market policies in urban areas, and their aggregate implications.

Trying to identify self-selection effects and learning effects among migrants is a growing field. Using observational data, whether we use repeated cross-section or panel data become a serious concern. Using cross-section data, Borjas, Bronars and Trejo (1992) find positive self-selection among inter-state migrants in the United States. Tunali (2000) finds evidence of a lottery in the outcomes from migration decisions in Turkey. His result suggests that a substantial portion of migrants realize negative gains and a minority realize very high gains from migration. Using panel-data, Clerides, Lach and Tybout (1998) study a similar question on the export market using panel-data. No empirical evidence of learning effects (i.e. improving productivity) by exporting are observed from their analysis. Self-selection in the domestic market is the main explanation for becoming exporters. Glaeser and Maré (2001) finds learning by migrating effects in cities after movements from small cities to large cities. On the other hand, studying internal learning by migrating effects is also a growing field: Yamauchi (2003) finds that complementarities between schooling and experience are reinforced as a migrant's experience increases in the destination market (Bangkok Metropolitan Area) using Thailand Labor Force Survey, 1994 to 1996. Using the same data as Yamauchi (2003), Yamauchi and Tanabe (2006) conclude that the employment probability of recent migrants is negatively affected by a large population size of previous migrants originating from the same region and positively af-

fected by success of the previous migrants in getting work. Kimura (2004) examines the main explanations for the urban wage premium: learning skills and learning job opportunities in the urban labor market using household-block-level data in the Thailand Labor Force Survey. Like previous studies on the aggregate labor market and urban immigration, this study uses pooled cross-sections. Munshi (2003) focuses on employment at the destination among Mexican migrants in the U.S. labor market. He uses rainfall in the origin (Mexico) as an instrumental variable to identify origin-communities network effects on employment opportunities at the destination. Combes, Duranton and Gobillon (2003) estimate a model wage determination across the French local labor market using a large panel of workers. They control for worker characteristics, worker-fixed effects, industry-fixed effects, and the conditions of the local labor market using a competitive equilibrium model. They find that ability sorting is the main explanation for spatial wage disparities.

Two unique characteristics in the available data are useful for our analysis: the "reason for migration" and "duration of stay" of migrants. A controlled experiment is constructed with these variables in the Thailand Labor Force Survey. This experiment enables us to identify self-selection effects and learning effects and to estimate the effects of concentrated area using observational data. The variable the "reason for migration" includes two types of migrants; job-seekers and migrants who move along with the household head. The location choice for job-seekers is self-selective based on their observed and unobserved characteristics. On the other hand, the location choice for migrants who move with the household head seems to be independent of their characteristics. Location choice is exogenous for these household migrants. We can observe true location specific returns for migrants who move with the household head. To see the degree of self-selection bias for job-seekers, this paper compares the returns to location between job-seekers and household relations migrants. Clear results are drawn from our identification strategy. This is very similar to the study estimating the heterogeneity of reason of displacement (plant-closing versus lay-off) on re-employment outcomes by Gibbons and Katz (1991) and Gibbons and Katz (1992). Our paper is the first attempt to identify self-selection and learning effects related to migration using exogenous sources of variation in migration decisions. The variable "duration of stay" for migrants suggests the possibility to examine learning effects of migration. There is a large wage difference between short-staying and long-staying migrants in each location. This pattern is quite different for migration streams; rural-rural, urban-rural, rural-urban, and urban-urban. The difference between cohorts provides evidence of improving average productivity. The cohort difference between the reasons for migration (i.e. difference in differences) also provides the solution to the difference of learning by migrating effects between two types of migrants.

The innovative feature of this work is that we utilize experimental evidence from observational data to

identify the self-selection of unobserved characteristics and examine the effects of learning by migrating in urban and rural areas, respectively. The main results are: first, positive self-selection among new migrants to urban from rural areas (i.e. new entrants into the urban labor market); secondly, negative self-selection among new migrants in rural areas (i.e. "new exits" from the urban labor market). These results suggest the existence of a natural selection (survival of the fittest) mechanism in the urban labor market. "New exits" from the urban area seem to have the potential to take their acquired skill. These results show learning by migrating effects in urban area over time. This effect is larger in urban than rural areas.

This paper proceeds as follows. Section 2 describes a simple model for understanding some empirical hypotheses. Section 3 shows the structure of the dataset of the Thailand Labor Force Survey. Section 4 contains a simple identification framework for studying self-selection effects on unobserved abilities and learning by migrating effects. While Section 5 deals with the estimation of self-selection bias on individual characteristics, section 6 focuses on learning by migrating effects. In the final section, we conclude the paper and discuss the remaining issues.

## 2    The Attraction of Cities: The Bangkok Megalopolis

Let us start by focusing on the geography of Thailand. Table 1 shows the patterns of urbanization of Thailand. We observe a unique position for Thailand in the world from three indices of urbanization patterns. An urban area is defined as place with over one hundred thousand inhabitants. This criteria is the lower bound of the definition of urban area.[1] *Primacy* means the level of urban primacy and in the case of Thailanld is the ratio of urban population residing in Bangkok compared to the second largest city. The usrban primacy of Thailand (about 25) is the highest in the world. *Megalopolitan* is the people living in Bangkok to the total urban population in Thailand. Almost 77% of urban residents are concentrated in Bangkok. These two indices show an agglomeration of economic activity in Bangkok and also demonstrate that there is only one megalopolis in Thailand. Finally, *Urbanization* means the ratio of the number of urban residents to the whole domestic population. Only 15% of the whole population is located in urban areas (i.e. almost 85% of the population is located in rural areas.). These indices show a clear contrast between urban and rural areas in Thailand.

These indices of economic geography in Thailand can simplify our empirical analysis. The same is not

---

[1]We should notice that this criteria is correlated with domestic population. For example, the USA seems to easily satisfy this criteria because it has a population of about two hundred million. On the other hand, Thailand (60 million population $\approx 3/10$ of the USA) does not satisfy this.

true, however, of the USA data. Large cities seem to be distributed discretely in the USA. If we see the USA data, we have to set up multiple discrete choice models. We refer to the Greater Bangkok Area as the urban area. We assume that workers can commute to the center of Bangkok as long as they locate in GBA. GBA seems to be a kind of basin of attraction. We also define all rural areas as non-GBA.

# 3  A Model of Migrating and Learning in Megalopolis

This section presents a simple model of decisions to migrate to the Bangkok labor market among workers in rural areas. The model is constructed by three stages: migration, production, and reshuffling (exits). Stage 1 is the migration decision stage. Production and reshuffling are stage 2 and 3, respectively. The model is based on the study of migration decisions to the U.S. labor market among Mexican migrants by Munshi (2003) and the schooling (English or local language) choice in the Bombay labor market among castes by Munshi and Rosenzweig (2006) (stage 1). The searching and production in the Bangkok labor market after migration is based on the model of imperfect information, learning, and worker mobility by Gibbons and Katz (1992) (stage 2 and 3).

## 3.1  Setup

The model provides an explanation and overview of the causal relationship between abilities, decisions to migrate to Bangkok, and natural selection through exits from Bangkok labor market. The geography is simply divided into rural areas and the Bangkok labor market. We assume that the Bangkok labor market is ruled by the geographic boundaries of the Greater Bangkok Area (hereafter GBA). The timing is divided into three periods. In the first stage, potential migrants (in rural areas) decide whether they will move to GBA or stay in the rural area. The migration decision is based on the wage that the potential migrants will receive in an urban job and rural job. Wages in the urban job ($\gamma_j \omega_i$) are assumed to be based on the migrants' ability or productivity. The returns to ability in the Bangkok labor market is described by $\gamma_j$. Wages in rural jobs such as profits from an agricultural area $R$ are not contingent on the worker's ability or productivity. This assumption can be relaxed in the empirical specification using an industry specific shock ($\sigma$) in the rural area. This shock is considered to involve the agricultural sector or export-oriented industries in the rural area. The returns to working in the rural area can be summarized as $R(\sigma)$.[2] On the other hand, individual's abilities $\omega_i$ are distributed uniformly. The ability distribution of premigration (born) level of ability does not between rural and GBA.

The second stage starts from the beginning of work in the urban job. Each firm (or sector) $j$ is endowed with a type of technology $\gamma_j$. This is also returns to ability in the Bangkok labor market. Each migrant produces a commodity and receives wages $W_{ij} = \gamma_j \omega_i$ under his/her own abilities and firm technologies. Following the setting of Gibbons and Katz (1992), technology is restrict to two types; ability-sensitive

---

[2]In July 1997, the returns to agricultural sector and export-oriented industries rose due to the baht devaluation. These sectors were booming because of exogenous currency advantages. The causal effects of the financial shock in 1997 on rural and urban labor market is task for the future.

technology ($\gamma_A$) and ability-insensitive technology ($\gamma_B$). We assume that the information on worker's abilities is imperfect at this stage but improves over time as long as migrants stay in the urban labor market. Job matching with urban jobs is random at the second stage. We also apply the story of Gibbons and Katz (1992) to our framework. Endogenous mobility improves the job matching between workers and industries. Highly able workers employed in firm $B$ can switch to firm $A$ while workers with low ability employed in firm $A$ can switch to firm $B$ at stage 3. This process can be explained by the learning effects in the urban area. The final stage starts from the time when there is perfect observation of worker's ability. The matching process is not random in the stage 3. The search markets are segmented, and positively associated with the worker's abilities and the firm's technologies at this stage. There is endogenous mobility of workers among two types of firms in the final stage.[3]

## 3.2 Abilities and the Migration Equilibrium

Migration decisions are based on the returns to urban job. The returns to firm (or sector) in urban area is determined by a lottery. Each worker knows the level of his/her own abilities. Potential migrants observe the expected value of returns to abilities in urban area $E(\gamma)$ in the second and third stage for individual $i$. For simplicity, we assume that $E(\gamma)$ summarizes the returns to ability at both the second and third stage. The expected returns to the urban job for individual $i$ with ability $\omega_i$ in the rural area is $E(W_i) = E(\gamma) * \omega_i$. Following Munshi and Rosenzweig (2006), ability is assumed to have a level of $\omega_i \in \{0, \frac{1}{2}, 1\}$. The expected returns to the urban job is zero, $E(\gamma)/2$, and $E(\gamma)$ respectively. Three types of migration equilibria arise: (1) only low type workers remain in rural areas, (2) low and medium types remain in rural areas, (3) everyone remains in rural areas. Each condition can be sustained within each rural area.

**Condition 1**   $R < E(\gamma)/2$

**Condition 2**   $E(\gamma)/2 < R < E(\gamma)$

**Condition 3**   $E(\gamma) < R$

The statistical inference of the conditions is our target. We are able to test for self-selection on the migration decision using the observational data including characteristics of migrants and employment outcome at the destination.

---

[3]Before turning to observe the migration decision, we add some mention of learning in the urban labor market. Instead of deriving the learning process (job shopping/sectoral mobility/and beween-job mobility) following Jovanovic (1979), Gibbons and Katz (1992), Topel and Ward (1993), Farber and Gibbons (1996), Neal (1999), and providing direct evidence from NLSY by Yankow (2003), this paper assumes that the effects of learning by migrating on wages appears in sectoral movers and between-job movers in urban area.

Our theoretical framework predicts that (1) the probability of staying in a rural area is negatively related to the individual's abilities and (2) the returns to the urban job $E(W_i)$ are positively related to the abilities $\omega_i$ and assignment to the urban migrants sample $D$.

$$Pr(D_i = 1) = 1 - \omega_i \tag{1}$$

$$\log W_{ij} = \alpha + \beta D(\omega_i) + \omega_i + \gamma_j \tag{2}$$

These specification need to be verified empirically to get a consistent estimate of the effects of urban migrants on the returns to the urban job. We can estimate the returns to the urban job using observational data. The main point of this expression is the causal effects of abilities $\omega_i$ on the incidence of urban migrants $D(\omega_i)$. The empirical methodology for estimating the returns to the urban job is shown in section 5.

## 3.3   Testable Hypotheses

We summarize our simple model for describing some testable hypotheses here. The model has the following theoretical implications: (1) the probability of staying in rural area is negatively related to the individual's abilities and (2) the returns to the urban job are positively correlated to the abilities and incidence of the urban migrants. It is time now to formulate some empirical questions or some testable hypotheses regarding the rural and urban labor markets: do individuals become more productive after moving to the megalopolis? Three testable hypotheses are described simply here. The first hypothesis supports self-selection in the migration decision: relatively efficient individuals become migrants and these individuals also have a good job-match in the new location. We test it for each migration streams. The second hypothesis support the improvement of average productivity over time: the average performance of migrants with long experience is better than those with short experience due to sorting or learning by migrating.

**Hypothesis 1** *There is a positive self-selection on individual abilities for "new entrants" into the urban labor market. There is a negative self-selection on individual abilities for "new exits" from the urban labor market.*

This hypothesis suggests that (1) a young worker with high abilities moves from the rural to the urban labor market and (2) a young worker with low abilities moves from the urban to the rural labor market. This implies the existence of natural selection mechanism exists in the urban labor market.

**Hypothesis 2** *Average productivity for a long stayer is higher than that of a short stayer in the rural and urban labor market through the two-sided learning process between firm and worker (i.e. selection mechanism over time) or learning by migrating.*

This hypothesis proposes that (1) sorting matters among the local labor market and (2) learning by migrating exist in the local labor market.

# 4  Data

In this section, we examine the "reason for migration" and "duration of migration" provide statistical evidence based on the wages of migrants, using data from the Thailand Labor Force Survey, 1994 to 1996. The data set presents three issues related to (1) geography; (2) the reason for migration and duration of migration; (3) evidence on wages. In fact, the proper treatment of these issues provides the key to understanding the self-selection mechanism and learning by migrating effects for migrants.

## 4.1  The Thailand Labor Force Survey

The data source used in this paper is The Thailand Labor Force Survey (hereafter LFS), 1994-1996 by the The National Statistical Office (NSO) of Thailand. This individual-level data provides the information on many individual characteristics: gender, structure of family, years of schooling, years of labor market experience, wages (or profit for self-employed household and profit for agricultural household), labor force status, migration status, hours and days of weekly work, occupation, industry, region, marital status; and employer characteristics: firm size, industry, and fringe benefits.

LFS is implemented four times per year. The first round of the survey is done in February, the dry season in Thailand. The third round is done in August, during the monsoon (agricultural) season. We use only the third round survey because we can neglect seasonal labor migration at the dry season. The second and third rounds are carried out in May and November, respectively. Because LFS does not follow individuals from year to year, this study cannot provide the information on labor mobility from the pre-crisis period to post crisis period.

The sample used in this paper comes from not only the "Greater Bangkok Area" and rural areas: we use the whole sample of the Kingdom of Thailand, year 1994 to year 1996. We would like to mention some of the geographic characteristics. This paper constructs a GBA (Greater Bangkok Area) dummy variable equals to 1 if the province is included in the Bangkok metropolitan area. Almost all industry and occupation tend to agglomerate in GBA.

## 4.2 Data on Migrants

Here, we shall examine our pooled sample of the Thailand Labor Force Survey (hereafter LFS). LFS is a random sampling of all households in Thailand taken during two survey rounds (February and August) every year. This paper pools the annual sample of LFS in 1994 to 1996. We do not add our sample to the LFS from 1997 to present. But the impacts were uneven for each occupation and industry. The number of pooled observations adds up to approximately one million. This paper leaves out agricultural and self-employed workers, housewives, and students. The result is a sample of 250,000 wage workers in which 53,529 are migrants. We concentrate on this sample in this paper to do a statistical inference with aggregate patterns of migration and its microeconomic consequences. We do not examine the 200,000 residents (i.e. non-migrants) among the sample of wage workers. The main contents of LFS are the following: gender, age, years of education, weekly wage, bonus, occupation, industry, firm size, unemployment spells, reason for migration, and length of stay when the survey was taken.

For the moment, let us look closely at our key variables for empirical analysis. Variables such as the "reason for migration" and "length of stay when the survey was taken" (or duration and exposure to destination) have unique characteristics. The variable the "reason for migration" can be grouped into seven categories: (1) Job search; (2) Job transfer; (3) Education/Training; (4) Medical treatment; (5) Back to former place of residence; (6) Move with household head; and (7) Other reasons. These reasons are automatically recorded for migrants who have 0 to 4 years of experience in each destination. Next, duration or exposure to each destination is recorded for migrants who also have 0 to 4 years of experience in each destination.[4] Based on the length of stay in migrant primacy destination, we shall classify migrants into two main groups: (A) Job search, job transfer, and back to former place of residence; (B) Move with the household head. We call the former group "Job-seeking" or "Job related". This distinct classification is useful for our identification and estimation in the next section. We exclude the following three categories: (3) Education/Training; (4) Medical treatment; and (7) Other reasons. This paper focuses on the following categories: (1) Job search; (2) Job transfer; (5) Back to former place of residence; (6) Move with household head. Secondly, migration streams are also divided into four types: (*1*) from rural to rural areas; (*2*) from urban to rural areas; (*3*) from rural to urban areas; and (*4*) from urban to urban areas. Thirdly, the duration of migrant status is divided into five categories: less than 1 year; 1 year; 2 years; 3 years; and 4 years of duration. All migrants are classified by reason of migration, migration stream, and length of stay.

---

[4]Migrants who have 5 to 9 years of experience are also recorded. However, there is no record of the original area, making it impossible for us to specify their migration streams from original area to destination area. Thus, we do not include these migrants who have 5 to 9 years of experience in our empirical analysis.

We tabulate the relationship between the reason for migration and duration of destination in Table 2. The main group is rural to rural movers who accounts for almost 40,000 migrants among the total sample of 53,000 Thai migrants. Our focus is on migrants who have experience in the Greater Bangkok Area: urban to rural movers and rural to urban movers. Migration attributed for the purpose of study also constitutes about 20% of all migrants. First, the tendency of exit rises to a peak within the first year of the move and after 3 years for migrants due to looking job searches. On the other hand, the tendency of exit also rises to a peak after 3 years for migrants due for migrants who move with the household head.

## 4.3   Evidence on Wages by Migration Status

We are now ready to look at the evidence of wages on migration status. We already assumed two labor markets: rural and urban. First, we look at the sample of new entrants into the rural labor market from rural areas. Secondly, we also look at the sample of new entrants into the rural labor market from urban areas. The mean, standard deviations, and number of observations are shown in each cell by migration status. The descriptive statistics for new entrants into the rural labor market is shown in Table 3, using the whole sample. Comparing the wage differentials between less than 1 year and 4 years of duration in the sample of rural-rural migrants, we find a gradual growth (from 6.656 to 6.959) in the long-staying migrants. Comparing the wage differentials between less than 1 year and 4 years of duration in the sample of urban-rural migrants, there is a sharp growth (from 6.705 to 7.114) among the long staying migrants. This sample experienced steeper wage differentials than rural-rural migrants. Next, we discuss the difference between origin of migration. The level of wages in rural areas is higher for migrants with some experience of urban area than migrants moving from rural areas. This difference seems to represent age. Urban to rural migrants are concentrated in their late 20s or 30s. They are usually older than rural to rural migrants or rural to urban migrants.

The descriptive statistics for new entrants into the urban labor market are shown in Table 4. Comparing the wage differentials between less than 1 year and 4 years of duration in the sample of rural-urban migrants, we also find a gradual growth (from 6.705 to 7.114) among the long staying migrants. We also see a slow growth (from 6.652 to 6.909) in the long staying migrants from rural areas. The level of urban wages is also lower than for rural to rural migrants or urban to rural migrants. This is because of a differential of age and labor market experience among migrants. Looking at the urban-urban category, we compare the wage differentials between less than 1 year and 4 years of duration in the sample of urban-urban migrants. There is also a gradual growth (from 6.963 to 7.266) in the long staying migrants. Next, we compare the difference between origin of migration. The level of wages in urban area is higher for migrants who have

more experience in urban areas than new migrants from rural areas. This difference seems to be represented by age and urban experience. Finally, we look at the wage differentials between migration streams: from rural to urban and from urban to urban. As expected, the level of wages is considerably higher for new entrants from urban areas than new entrants from rural areas. This is most likely due to differentials with respect to age, urban experience (i.e. benefits from searches within urban areas or improving productivity), and sectors (i.e. formal and informal) between the migration streams.

# 5    Identification Strategy

This section provides a simple framework for empirically testing our hypotheses. We assume that a young worker has all the information to evaluate his or her own ability as well as the returns to this ability. It is easy for us to imagine a correlation among own ability, migration decision, destination, and returns to ability. If there is a self-selection bias in the migration decision and destination, it is not easy to evaluate the true learning effects in urban areas. This paper proposes a novel method for evaluating self-selection and learning by migrating effects using the unique characteristics of the Thailand Labor Force Survey: the "reason for migration" and "length of stay" for migrants. Our identification approach is quite different from that of Clerides et al. (1998) who study learning by exporting effects among Columbian, Moroccan, and Chilean exporters using establishment level panel-data. Our identification strategy is also quite different from Glaeser and Maré (2001), who examine learning effects in cities using US panel-data; PSID and NLSY.

First, every worker has to make a migration decision in every period: stay in the current labor market or move to another area. We call this decision variable $M \in \{0, 1\}$. Staying is captured by $M = 0$ and moving is captured by $M = 1$. In this paper we restrict our analysis to movers (i.e. $M = 1$). Secondly, every worker chooses his/her labor market every period: rural or urban labor market. We call this location choice variable $K \in \{R, U\}$. A rural worker is captured by $K = R$ and urban worker by $K = U$. Thirdly, we define the variable $Z \in \{0, 1\}$ which is the "reason for migration" for movers. We have already classified migrants into two main groups: Group (A) Job-Seeking (which includes job search, job transfer, and back to former place of residence) and Group (B) Move with the household head. With respect to household-related reasons of migration, we assume that the member moving with the household head is captured by $Z = 1$ and the member actively seeking employment or another reason is captured by $Z = 0$. These bring us to the second point: the description of decision and state space for each individual.

These assumptions on migration decision and migration streams lead us further into an empirical investigation. We assume that the household head decides whether his family will migrate or not. Therefore we use heterogeneity in the reason for migration as an exogenous source of variation in endogenous variables. This paper constructs a *household related* dummy variable $D_{1H}$. If a young worker follows the household head in moving, then the econometrician will treat the member as $D_{1H} = 1$. If the job-seekers move, $D_{1H} = 0$. The indicator variable $D_H^K$ is generated as follows:

$$
D_H^K = \begin{cases} 1 & \text{if } Z = 1, K \text{ for } M = 1 \\ 0 & \text{if } Z = 0, K \text{ for } M = 1 \end{cases}
$$

The outcome variable for individual worker $i$ in location $K$ and at survey week $s$ is defined as $Y_{is}^K$. The

cross-section outcome function is formalized as an additive separable form.

$$Y_{is}^K = f(X_{is}) + \Gamma \cdot D_H^K + g\left(\omega_{is},\ \xi_{is}\right)$$

where $f(X)$ is a function of a vector of observed individual characteristics, $D_H^K$ is a dummy variable equal to one if individual $i$ follows the household head to move to location $K$, and $g\left(\omega_{is},\ \xi_{is}\right)$ is a function of unobserved characteristics for an individual worker and firm: $\omega_{is}$ is an error term for unobserved abilities for individual $i$ at survey week $s$, and unobserved firm specific characteristics $\xi_{is}$.

The choice of location $K$ correlates to pecuniary returns to individual characteristics: an observed component $X_i$ and component of unobserved ability $\phi_i$. The high frequency of movement to urban area for young and more educated workers is a known and observed fact. Young and highly educated workers know the urban area to be a thick labor market (with varieties in occupation, industry, and technology). The return to schooling is also higher in urban areas. Recent literature Borjas et al. (1992), Wheeler (2001), Dahl (2002), Moretti (2004a), and Moretti (2004b) also show this using USA data. This is the logic of self-selection. The location choice $K$ of the migrant in the household related subgroup ($Z = 1$) is assumed to be orthogonal to his/her ability $\omega$, and the location choice $K$ of the migrant job-related subgroup ($Z = 0$) is assumed to be non-orthogonal to ability $\omega$ by definition.

$$\begin{aligned} \omega_i &\propto & D_{1H}^R = 0,\ D_{1H}^U = 0 \\ \omega_i &\perp & D_{1H}^R = 1,\ D_{1H}^U = 1 \end{aligned}$$

Undoubtedly, a worker in the local labor market $K$ has a location specific premium; however, the econometrician cannot distinguish between the true premiums in location and the self-selection bias in the migration decision. Our identification method suggests that movement with the household head is exogenous on migration decision. We call the member of $D_{1H} = 1$ the treatment group. We can examine the true effects of moving to local labor market $K$ on individual outcome by looking at the coefficient $\Gamma$. This coefficient $\Gamma$ signifies the premium differentials between the household related ($D_{1H} = 1$) and job related ($D_{1H} = 0$) subgroup in location $K$. This paper develops a new and a simple identification method for distinguishing between the true premium in $K$ and the impact of the self-selection bias on the migration decision to $K$.

Next, we try to identify learning by migrating effects (i.e. productivity increasing based on migration) in location $K$. For example, learning by migrating effects in urban areas include formal training, learning by doing, knowledge spillovers by communication, reduction in mismatches by turnover, and R&D investments by firms. We assume that firms can offer a wage after removing the returns to investment in technology.

If this assumption is valid, then we can exclude the latest possibility of learning by migrating effects: investment in technology by firms. Now we may restrict our discussion to learning effects due to individual efforts (i.e. learning and job turnover) and spillovers.

To study learning effects, we use the variable $\tau$ for duration of stay after migration and relate this to experience in $K$. The data "length of migrant stay" $\tau \in \{0, 1, 2, 3, 4\}$ years is useful for the identification of learning by migrating effects. This paper divides years $T \in \{S, L\}$ according to the duration of stay $\tau$. The short-staying migrant worker is captured by $T = S$. The longer-staying migrant worker is captured by $T = L$. We define (1) a short experience as $S \in \{0, 1, 2\}$ years and (2) a long experience as $L \in \{3, 4\}$ years of moving to current location $K$ based on the survey week. We are also now able to expand the individual decision space from $D$ to $D_T$ with duration of stay.

# 6  Testing for Self-selection on Migration

## 6.1  Specification and Estimation

First, we test whether there is a self-selection bias on ability when a young worker moves to location $K$. The outcome variable is the wage level. The wage represents self-selection about employment and productivity. To test for self-selection in observed characteristics and unobserved abilities, we estimate the reason differentials from the cross-section wage function following Gibbons and Katz (1991) and Gibbons and Katz (1992) who studied the impacts of reason for separation from the last job on the wage level of new employment using an establishment closure sample and layoff sample. We run following linear regression formulations:

$$\log W_{is}^K \;\; = \;\; X_{is}\beta + \Gamma \cdot D_H^K + \omega_{is} + \xi_{is}^K \tag{3}$$

where the dependent variable $\log W_{is}^K$ is the log of weekly earnings for individual $i$ at survey week $s$ in location $K$, $X$ is a vector of observed individual characteristics (gender, age, education, square of years of education, and marital status), the dummy variable $D_H^K = 1$ if individual $i$ moves to location $K$ due to household related reasons, and otherwise if individual $i$ moves to location $K$ to actively seek employment. $\omega_{is}$ is unobserved abilities for individual $i$, $\xi_{is}^K$ is unobserved firm technology if individual $i$ is employed in this firm. Finally, $u_{is}^K$ is a mixed error component of $\omega_{is}$ and $\xi_{is}^K$.

## 6.2 Results

The expected signs of the estimates on self-selection are as follows. First, the coefficient $\Gamma$ of urban-rural migrant is expected to be positive (i.e. job-seeking migrants receive lower rural premiums than migrants based on household relations because of the negative self-selection on exit decisions from the urban labor market).

$$H_0 \quad : \quad \Gamma^{UR} < 0$$

$$H_1 \quad : \quad \Gamma^{UR} > 0 \quad \text{(negative self-selection for job-seekers)}$$

Secondly, the coefficient $\Gamma$ of rural-urban migrant is expected to be negative (i.e. job-seeking migrants receive higher urban premiums than migrants based on household relations because of the positive self-selection on entry decisions into the urban labor market).

$$H_0 \quad : \quad \Gamma^{RU} \geq 0$$

$$H_1 \quad : \quad \Gamma^{RU} < 0 \quad \text{(positive self-selection for job-seekers)}$$

The estimates for the four subsamples urban (rural) migrants to urban (rural) areas in Table 5 provide evidence that there are highly significant on self-selection effects of unobserved abilities: (1) there is positive self-selection for new entries into both the urban and rural labor markets from rural areas; (2) there is negative self-selection for new exits from the urban labor market and movers within urban areas. The estimates show that the household related subgroup has approximately 4.8% lower wages than migrants in the job related subgroup in the rural to rural subsample. Alternatively, the estimates show that job-seeking migrants earn 4.8% higher wages than migrants in the household related subgroup. Job turnovers and established residences in the rural area have positive self-selection effects. For the urban to rural subsample, there is a negative and highly significant self-selection bias of unobserved abilities for job-seeking migrants, who have 5.3% lower wages than movers with household head.

On average, job-seeking migrants in the urban-rural subsample have a 5.3% lower level of rural true premium than migrants in the household related subgroup. In other words, if the job-seeking subgroup sample had the same level of ability as the household related subgroup, the job search sample will have a true rural premium. Therefore, the "low ability" of the job search sample reduces their own rural premium from the true rural premium to a lower level. Nearly all urban to rural movers were former migrants to urban areas from rural areas.

For the rural-urban and urban-urban subgroups, migrants for household related reasons have a true

16

urban premium independent of self-selection effects with respect to location choice and ability based on identification strategy in the previous section. The coefficient of this dummy variable suggests that there is a differential of unobserved ability from the urban premium. Estimates for the rural-urban subsample show a positive and highly significant self-selection bias of unobserved abilities for job-seekers, who have 8.8% higher wages than the household related control group. Estimates for the job-seeking migrants in the rural-urban subsample show a positive self-selection effect of abilities. Estimates for the urban-urban subsample also show negative and significant self-selection effects of unobserved ability for migrant job-seekers who have 8.0% lower wages than those in the household related control group. Job turnovers and established residence in the urban area do not seem to have positive self-selection effects.

We are now ready to say that there are positive and quite significant (1% level) self-selection effects of unobserved abilities for migrants from rural areas for both the rural to rural and rural to urban subsamples. But on average, there is a significant negative self-selection bias of unobserved abilities for migrants from urban areas for both urban to rural and urban to urban subsamples.[5] In summary, the high ability of young migrants from rural areas is clearly reflected in wages in rural and urban areas on average. Our model predicts that a migrant with high ability can keep a job as long as he/she obtains a job requiring skill/training/know-how early in the migration process.

Migrants who find jobs choose to stay. But migrants who cannot find jobs return to their place of origin or move to another location to seek jobs. Both rural movers from urban areas and urban movers from urban areas seem to have a bad match in urban area. However, the results using the urban to rural and urban to urban subsamples (see Table 5) show that there are smaller learning effects in urban area for the job-seeking subgroup than for the household related subgroup. If there are positive spillovers into the urban area, the gaps between the two types of migrants mean that there is heterogeneity in learning through migrating effects in urban areas.

# 7 Testing for Learning by Migrating

## 7.1 Specification and Estimation

We test our next hypothesis on learning through migrating effects to improve productivity after a young worker moves to location $K$. Learning through migrating effects are mentioned by Glaeser and Maré (2001)

---

[5]These results support the "lemon" effect for migrants from urban area. On the other hand, they also support the no lemon effect for migrants from rural area.

using panel regressions. It is a comprehensive work. However, there are shortcomings concerning the self-selection bias on migration decision and destinations.

The outcome variable is also the wage level. We divide the explanatory variable $D_H^K$ into the *short* $D_H^K(S)$ and *long* $D_H^K(L)$ cohort. To test for learning through migrating effects, we compare the coefficients $\Gamma(S)$ and $\Gamma(L)$ of the household related subgroup dummy variables. The coefficient $\Gamma(L)$ includes various sources of improved productivity due to individual efforts in location $K$ (for example, formal training or learning by doing), due to knowledge spillovers in location $K$, and due to reallocation effects by sorting and two-sided learning between individuals and firms. The occurrence of reallocation represents self-selection on ability. Information on matching quality is accumulated by firms and individuals after production.

If we assume that ability does not change over time, we can say that reallocation effects are self-selective innate (or natural) ability. If we assume that individual ability changes over time through learning by migrating effects, we can say that reallocation effects are self-selective in terms of acquired ability after migration. If any doubt remains about identification between learning through migrating effects (i.e. individual efforts and spillovers) and reallocation effects, it is clear that average productivity can be higher for a *long* cohort $L$ than a *short* cohort $S$.[6] To see these effects, we also specify and estimate cohort differentials in estimates of reason differentials from the cross-section wage function by the following linear regressions.

$$\log W_{is}^K = X_{is}\beta + \Gamma(S) \cdot D_H^K(S) + \Gamma(L) \cdot D_H^K(L) + \omega_{is} + \xi_{is}^K \tag{4}$$

where $\log W_{is}^K$ is the log of weekly earnings for individual $i$ in location $K$, $X_{is}$ a vector of observable individual characteristics (gender, age, years of education, square of years of education, marital status), the dummy variables $D_H^K(S)$ and $D_H^K(L)$ are equal to 1 if individual $i$ moved with the household head to location $K$ and has short (long, respectively) experience in location $K$, and $u_{is}^K$ is an error term of individual unobserved abilities $\phi_i$ and firm's technology, $\xi_{is}$.

## 7.2 Results

We also estimate the effects of dummy variables $D_H^K(0)$, $D_H^K(1)$, $D_H^K(2)$, $D_H^K(3)$, and $D_H^K(4)$ in the four migration flow subsamples: (*1*) rural to rural; (*2*) urban to rural; (*3*) rural to urban; and (*4*) urban to urban. The coefficient of dummy variable $\Gamma$ signifies the differentials of the location specific premium between

---

[6]The literature on TFP shows in detail that the higher level of productivity among long-lived firms than short-lived firms can be explained by the exiting of non-productive firms. This reallocation effect can also be drawn from the literature on the export market and two-sided learning in the labor market. Our theoretical model also predicts natural selection or survival of the fittest through job mobility.

the two migrant subgroups of job-seekers and household relations in each location $K$. The coefficient $\Gamma(0)$ means the difference of the location premium for staying for the two migrant groups with less than 1 year of duration. The same is true for the coefficient $\Gamma(1)$, $\Gamma(2)$, $\Gamma(3)$, and $\Gamma(4)$ for long-staying workers in the two groups for 1 year, 2 years, 3 years, and 4 years of duration respectively.

The estimates for each sample in Table 6 and Table 7 show that there are smaller reallocation effects in the urban labor market and there are no learning effects in urban areas for "new exits for job-seeking" from urban areas: (1) the wage gap between job-seekers and household related migrants in the rural-rural subsample increases due to the improvement of job-matching effects (or reallocation effects through self-selection) for job-seeking migrants looking for jobs in rural area; (2) the wage gap between job seekers and household related migrants in the urban-rural subsample due to learning through migrating effects for household related migrants; (3) the wage gap between job seekers and household relations in the rural-urban subsample increases due to the improvement job-matching effects (or reallocation effects through self-selection) for migrants of looking for jobs in urban areas; (4) the wage gap between job-seekers and household related migrants in the urban-urban subsample due to learning through migrating effects for household related migrants.

It is worthwhile to mention the relationship between the short- and the long-duration migrants in each migration stream. Table 2 presents some interesting facts regarding the similarity and differences between the reasons for migration. Here, we look more carefully into our empirical results. The estimate for the rural-to-rural subsample with short-staying (1 year and 2 years of duration) in rural areas shows that job-seekers have 3.4% higher wages and 7.2% higher wages, respectively, than household related migrants. On the other hand, for long-staying (3 years and 4 years of duration) migrants in rural areas after migration from another rural area, job-seekers have 10.9% and 9.1% higher wages, respectively, than household related migrants. There is a steep rise from short to long-experienced migrants between the two types of migrants. It seems reasonable to suppose that there is learning through migrating effects in rural areas or reallocation effects gained from the sorting process through job-matching. Increased average productivity of job-seekers can be explained by the differences of the survival rate between the two types of migrants: job-seekers and household relations migrants. Migrants who cannot find a good job match seem to change their location at an early stage, before the second year. However, migrants who find a good match will stay in rural areas and become a long-duration migrant. It is likely that an improvement in average productivity for a long cohort of job-seekers can be explained by the survival of the fittest in the rural to rural subsample.[7]

---

[7]A possible explanation is: if learning by migrating effects exist, then the wage gap between short and long durations can

The estimate for the urban-to-rural subsample with the short-staying migrants in Table 6 shows that migrants looking for jobs have 11.2% and 15.5% lower wages than migrants who move with the household head when they reach 2 and 3 years of duration in rural areas. On the other hand, for migrants staying less than 1 year and 1 year, there is no significant difference between job-seeking and household relation migrants. We can construct two hypotheses: (1) self-selection based on abilities; (2) learning through migrating effects in each location. The wage gap between the short and long cohort among the two types of migrants can be explained by the difference of learning speeds in rural areas.[8] From Table 2, we see that the exit patterns of the two types of migrants are similar: job-searching versus moving with the household head. We may say that the difference between exit patterns is not the main reason. The result for migrants with 4 years of duration in rural area suggests that there is no significant difference between job-seeking and household relation migrants. The household relations migrant who can find a good match stay in rural area and enter long duration status, with 4 years of duration.

Let us, for the moment, examine urban migrants. The estimate for the rural to urban subsample with short-experienced migrants in urban areas in Table 7 shows that job-seeking migrants with 1 and 2 years of duration in urban areas experience 17% higher wages and 8.6% higher wages, respectively, than household relation migrants. On the other hand, especially for long-staying migrants with 4 years of duration job-seekers experience 20.6% higher wages than household relation migrants. The wage gap between the two types of migrants decreases due to convergence of abilities between the two groups or due to their remaining in urban job-seeking. Exit patterns are also quite similar for job-seekers versus household relation migrants, as seen from Table 2. If a young migrant has a bad match in an urban area, then he/she will change his/her location to look for another job match, the average productivity of migrants in urban area can be improved over time. We observe that this reallocation effect is common for the two types of migrants. We can present two explanations: first, a decline in the reallocation effects among job-seeking migrants is the main explanation, it is likely that less able workers will remain in the urban area due to the thickness of the market. This leads to our finding of a declining gap between the two types of migrants. Secondly, if learning by migrating effects is the main explanation, we may say that household relation migrants catch

---

be explained by the difference in learning speeds between the two types of migrants. Another explanation is: if reallocation effects (through self-selection) are the main reason, then the wage gap between short and long durations can be explained by the difference of exit speeds from the rural labor market between job-seekers and household relation migrants.

[8]In the previous section, we discussed the fact that there is negative self-selectivity in ability for urban to rural movers. There are complementarities between ability and learning speed. We cannot identify whether ability is main explanation or not here. Even if any complication remains about complementarities, it is clear that there is gap between the two types of migrants from urban to rural areas.

up with job-seeking migrants: there is an observed convergence among urban migrants.

Our results on urban migrants show that in the urban-to-urban subsample, short-staying job-seeking migrants with less than 1 year of duration experience 5.1% higher wages than household relation migrants. But this estimate is not significant. This result suggests there are no significant differences between the two types of migrants among the new migrants from urban to urban areas. On the other hand, long-staying (3 years and 4 years of duration) job-seeking migrants experience 22% and 19.6% lower wages than household relation migrants. We shall now look more carefully into the two explanations. First, there is a substantial difference in the total number of movers between those looking for a job versus those moving with the head of family. If reallocation effects are the main explanation of the wage gap, then we will see a decrease in this gap: there are strong reallocation effects among migrants looking for jobs, and sorting effects force them to improve their average productivity. But this hypothesis contradicts the gap of the two estimates. Average productivity among migrants looking for jobs decreases. Secondly, we can refer to the possibility of learning by migrating effects in the urban area. It seems reasonable that the differentials in ability or differentials of learning speed for the two types of migrants are quite significant on average. The initial gap between migrants looking for jobs and migrants moving with the household head seems to wide over time.

## 7.3    Robustness Check

In earlier parts of the paper, we discussed the differences in the learning by migrating effects due to various reasons for the four migration streams. Next, we test for learning effects related to the length of stay of migrants by the reason for migration. Two dummy variables are used to estimate the impacts of the length of stay on productivity for two types of migrants: job-seekers and those based on household relations. For job-seeking migrants (i.e. $Z = 0$), the dummy variable $D_H^K$ is equal to 1 if individual $i$ moved to location $K$.

To see the learning by migrating effects of the length of stay on productivity, we specify and estimate reason differentials based on cohort differentials from the cross-section wage function. We run the following regression equation for the reason of migration: job-seeking migrants ($D_H^K = 0$) and those moving with household head ($D_H^K = 1$).

$$\log W_{is}^K = X_{is}\beta + \Gamma_S(1) \cdot d_1 + \Gamma_S(2) \cdot d_2 + \Gamma_S(3) \cdot d_3 + \Gamma_S(4) \cdot d_4 + \omega_{is} + \xi_{is}^K \tag{5}$$

where where $\log W_{is}^K$ is the log of weekly earnings for individual $i$ in location $K$, $X_{is}$ is a vector of observable individual characteristics (gender, age, years of education, square of years of education, marital status), coefficient $\Gamma_S(1)$ captures the difference in productivity between movers who have a duration of less than 1 year and a duration of 1 year for the job-seeking migrants. On the other hand, coefficienct $\Gamma_S(2)$ also

captures the difference of productivity between movers who have a duration of less than 1 year and of 2 years among migrants based on household relations, and so on.

These coefficients provide us with an empirical understanding of the impact of duration on abilities. There is an "endogeneity problem" : each individual chooses the length of stay in location $K$ after the migration, and the pitfall is related to comparing two coefficients due to sample truncation. However, we can focus on the results of the share of long-experienced migrants for the two types of migrants because Table 2 suggests same patterns of sample-attrition for job-seekers and people with household relations among urban-rural migrants and among rural-urban migrants.

Empirical results are given in Table 8 for the rural labor market and Table 9 for the urban labor market. There is sharp empirical observation of learning by migrating effects for job-seekers in rural areas: we have several results to show that the coefficient is positively significant at the 1% level. Job-seeking migrants from rural to rural areas experienced a gradual wage increase between 2.3% (1 year of duration), 5.5% (2 years), 3.5% (3 years), and 8.3% (4 years), respectively. On the other hand, we have the following observation regarding the learning by migrating effects for rural-rural migrants based on household relations: the coefficient is also positively significant at the 1% level. Household-related migrants from rural to rural areas also experienced a slight wage increase of 4.2% (1 year of duration), 5.3% (2 years), 3.8% (3 years), and 4.3% (4 years) respectively. For the learning by migrating effects of urban-rural migrants based on household relations, We also obtained significant coefficients. In conclusion: (1) there is a steep wage increase for job-seekers with *long* experienced (3 years of duration and more) in the rural labor market and (2) however, there is sudden wage declining for migrants with *short* experience (1 year and 2 years of duration) based on household relations in the rural labor market. The wage level for migrants with household relations is higher when they newly enter a in rural area from an urban area. The main difference in the reason for migration seems to be derived from the difference of search intensity in the early stages after migration.

The empirical results in Table 9 also suggest that the advantage of migrants with *long* experience exists. "New entrants" into the urban labor market experience learning by migrating effects for job-seekers: the coefficients are positively significant at the 1% level. While "new entrants" into the urban labor market experience learning by migrating effects for those with household relations: the coefficients are negatively significant at the 1% level. Job-seeking migrants from rural to rural areas experienced a gradual wage increase of 6.8% (1 year of duration), 9.9% (2 years), 14.9% (3 years), and 22.4% (4 years), respectively. Comparing the impacts of duration on wages between rural and urban areas for job-seeking migrants, there

is an advantage of staying because of the thick market externalities in the urban area. Job-seeking migrants within the urban labor market also benefit from search activities. On the other hand, migrants based on household relations do not experience any benefit from staying in the urban labor market. There is no advantage for household relation migrants to stay in the urban labor market: the coefficient is negative and significant. Job-seeking migrants from urban to urban areas experienced a gradual wage decrease of -13.3% (3 years of duration) and -7.9% (4 years), respectively. Household related migrants from rural to rural areas also experienced a slight wage decrease between -15.1% (1 year of duration) and -10.7% (3 years), respectively. We therefore conclude that both *short* (1 year and 2 years of duration) and *long* (3 years and 4 years of duration) experience job-seekers of rural origin with gains in the urban labor market. These advantages of duration of stay come from search activities in the urban labor market (thick market externalities).

# 8    Discussion and Conclusion

Some issues remain. The first is, the validity of our instrumental variable used in the reason for migration. Recent works by Rosenzweig and Wolpin (2000) and Angrist and Krueger (2001) argue for general shortcomings of instrumental variables due to natural experiment. Our identification strategy depends on whether moving with the household head to a new location is orthogonal to individual abilities: following the head is exogenous. We have to consider that the location choice for migrants who move with the household head seems to be independent of their characteristics. The choice of location seems to be exogenous for these household relation migrants. But we can conjecture the existence of a co-location problem for husbands and wives, following Costa and Kahn (2001). If there are non-random strategic complementarities to co-location between wives and husbands, our identification strategy fails. It is an unsettled question. There is a room for a further structural estimation of a model of collective behavior and household bargaining process in migration decisions. Secondly, to study the self-selection effects on abilities and the learning through migrating, we used wage worker migrants. This leads to: sample selection bias. We have to examine transitions by migrants to agricultural, self-employed, wage workers, and household workers such as housewives. Thirdly, we do not control for the categories of occupation, industry, and firm size. Self-selection effects on abilities and learning by migrating effects can be quite different for these categories.

In this paper we developed a simple framework for identifying the learning by migrating effects and self-selection effects on abilities with an instrumental variable. This is the first attempt to identify learning by migrating effects from self-selection effects using an exogenous source of variation, i.e., the reason for

migration. This paper is useful for understanding the role of the urban labor market: natural selection or the location of learning. Our empirical results for self-selection on migrating are summarized below: (1) there is a significant positive self-selection based on abilities for rural-urban job-seekers; and (2) there is a significant negative self-selection based on abilities for urban-rural job-seekers significantly. Evidence of a positive selection for "new entrants" to urban areas from rural areas also supports findings of Borjas et al. (1992). However, evidence of negative self-selection for "new exits" to the rural area from urban area does not support for findings of Glaeser and Maré (2001) using PSID and NLSY. The main difference between the empirical results of Glaeser and Maré (2001) and our results using the Thailand and Greater Bangkok dataset is attributed to the number, size, and matching externalities in cities.

In short, these rigorous inferences suggest the existence of the survival of the fittest in the urban labor market. Highly able job-seekers tend to move to the urban labor market from rural areas, while less skilled job-seekers tend to exit from the urban labor market to the rural labor market. The origin of migrants from the urban labor market to rural areas can be rural areas. These migrants (i.e. returnees) moved from the rural to the urban labor market and then back to the rural area because of bad job-matching experiences in the urban labor market. Our empirical results for learning by migrating effects are summarized below: (1) natural selection or survival of the fittest plays a significant role in the urban labor market; and (2) there are learning effects (through job-matching) for job-seeking migrants after migrating to urban areas. This effect is larger in urban than in rural areas.

It seems reasonable to state that better job-matching can be found by rural-urban migrants based on household relations after *learning where better job opportunities are located* in the thick market. It is due to the difference in the necessity of getting a job between job-seekers and household relation migrants. Household relation migrants have more search or waiting premiums than *impatient* job-seekers. The results for both rural-urban migrants can apply to those results for both urban-rural migrants and urban-urban migrants but not for rural-rural migrants. These results lead to the conclusion that there is a search option in the thick market.

# References

**Angrist, Joshua D. and Alan B. Krueger**, "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiment," *Journal of Economic Perspectives*, 2001, *15*, 69–85.

**Borjas, George J., Stephen G. Bronars, and Stephen J. Trejo**, "Self-Selection and Internal Migration in the United States," *Journal of Urban Economics*, 1992, *32*, 159–185.

**Clerides, Sofronis, Saul Lach, and James R. Tybout**, "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco," *Quarterly Journal of Economics*, 1998, *113*, 903–947.

**Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon**, "Spatial Wage Disparities: Sorting Matters!," Mimeograph 2003.

**Costa, Dora and Matthew E. Kahn**, "Power Couples: Changes in the Locational Choice of the College Educated, 1940-1990," *Quarterly Journal of Economics*, 2001, *115*, 1287–1315.

**Dahl, Gordon**, "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets," *Econometrica*, 2002, *70*, 2367–2420.

**Farber, Henry S. and Robert Gibbons**, "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 1996, *111*, 1007–1047.

**Gibbons, Robert and Lawrence F. Katz**, "Layoffs and Lemons," *Journal of Labor Economics*, 1991, *9*, 351–380.

_ **and** _ , "Does Unmeasured Ability Explain Inter-Industry Wage Differentials?," *Review of Economic Studies*, 1992, *59*, 515–535.

**Glaeser, Edward and David Maré**, "Cities and Skills," *Journal of Labor Economics*, 2001, *19*, 316–342.

**Jovanovic, Boyan**, "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 1979, *87*, 972–990.

**Kimura, Yuichi**, "The Role of Big Cities in Human Capital Accumulation: Evidences from Thailand," *Ph.D. Dissertation, Kyoto University*, 2004.

**Moretti, Enrico**, "Estimating the Social Return to Higher Education: Evidence From Longitudinal and Repeated Cross-Sectional Data," *Journal of Econometrics*, 2004, *121*, 175–212.

— , "Human Capital Externalities in Cities," in "Handbook of Regional and Urban Economics," North Holland-Elsevier, 2004.

**Munshi, Kaivan**, "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics*, 2003, *118*, 549–597.

— **and Mark Rosenzweig**, "Traditional Institutions Meet the Modern World: Caste, Gender and Schooling Choice in a Globalizing Economy," *American Economic Review 96(4)*, 2006, *96(4)*, 1225–1252.

**Neal, Derek**, "The Complexity of Job Mobility Among Young Men," *Journal of Labor Economics*, 1999, *17*, 237–261.

**Rosenzweig, Mark R. and Kenneth I. Wolpin**, "Natural "Natural Experiments" in Economics," *Journal of Economic Literature*, 2000, *38*, 827–874.

**Topel, Robert and Michael Ward**, "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 1993, *107*, 441–479.

**Tunali, Insan**, "Rationality of Migration," *International Economic Review*, 2000, *41(4)*, 893–920.

**Wheeler, Cristopher H.**, "Search, Sorting, and Urban Agglomeration," *Journal of Labor Economics*, 2001, *19*, 879–899.

**Yamauchi, Futoshi**, "Are Experience and Schooling Complementary? Evidence from Migrants' Assimilation in the Bangkok Labor Market," *Journal of Development Economics*, 2003, *74(2)*, 489–513.

— **and Sekiko Tanabe**, "Nonmarket Networks Among Migrants: Evidence from Metropolitan Bangkok, Thailand," *Forthcoming Journal of Population Economics*, 2006.

**Yankow, Jeffrey J.**, "Migration, Job Changes, and Wage Growth: A New Perspective on the Pecuniary Return to Geographic Mobility," *Journal of Regional Science*, 2003, *43(3)*, 483–516.

Table 1: **Position of Urbanization Patterns for Thailand in the World**

| Rank | Country | Primacy | Country | Megalopolis | Country | Urbanization |
|------|---------|---------|---------|-------------|---------|--------------|
| 1 | **Thailand** | 24.994 | Angola | 0.857 | UAE | 0.997 |
| 2 | Angola | 15.520 | Azerbaijan | 0.846 | South Korea | 0.736 |
| 3 | Chile | 14.847 | Ireland | 0.841 | Dominican | 0.727 |
| 4 | Peru | 10.332 | Paraguay | 0.832 | Lebanon | 0.722 |
| 5 | Lebanon | 9.847 | Sierra Leone | 0.828 | Japan | 0.696 |
| 6 | Sierra Leone | 9.221 | Lebanon | 0.814 | USA | 0.689 |
| 7 | Madagascar | 9.078 | Kyrgyzstan | 0.785 | Australia | 0.659 |
| 8 | Argentina | 8.787 | Tajikistan | 0.7712 | Venezuela | 0.644 |
| 9 | Hungary | 8.707 | **Thailand** | 0.7706 | Mexico | 0.633 |
| 10 | Mali | 8.542 | El Salvador | 0.750 | Chile | 0.623 |
| | | | | | | |
| 91 | Ecuador | 1.307 | Venezuela | 0.203 | India | 0.159 |
| 92 | South Africa | 1.282 | South Africa | 0.197 | Mali | 0.154 |
| 93 | USA | 1.267 | Poland | 0.173 | **Thailand** | 0.150 |
| 94 | Vietnam | 1.236 | Germany | 0.167 | Tajikistan | 0.147 |
| 95 | China | 1.194 | Netherlands | 0.159 | Kenya | 0.141 |
| 96 | Cameroon | 1.157 | Ukraine | 0.138 | Madagascar | 0.131 |
| 97 | Australia | 1.150 | Russia | 0.135 | Sri Lanka | 0.092 |
| 98 | UAE | 1.103 | India | 0.113 | Niger | 0.091 |
| 99 | Syria | 1.059 | USA | 0.085 | Ethiopia | 0.062 |
| 100 | Netherlands | 1.035 | China | 0.047 | Nepal | 0.056 |

*Notes*: Primacy means the level of urban primacy: ratio of urban population residing in the largest city, to the second largest city. The urban primacy of Thailand (about 25) is the highest in the world. Megalopolitan population means the ratio of the Greater Bangkok Area to total urban population in Thailand. Almost 77% of urban residents are concentrated in Bangkok. These two indices show an agglomeration of economic activity in Bangkok and also show that there is only one megalopolis in Thailand. Finally, Urbanization means the ratio of the number of urban residents to whole domestic population. Only 15% of the whole population is located in urban areas (i.e. almost 85% of the population is located in rural areas.).

*Source*: Author's calculation from *United Nations World Urban Prospects 2000*.

Table 2: **Sample Size by Reason, Migration Streams, and Duration**

| Subsample | Less than 1 year | 1 year | 2 years | 3 years | 4 years | Total |
|---|---|---|---|---|---|---|
| Rural-to-rural | 11,530 | 9,377 | 8,141 | 6,412 | 4,061 | 39,521 |
|     Job-seeking | 9,495 | 7,364 | 6,028 | 4,527 | 2,822 | 30,236 |
|     Household relations | 2,035 | 2,013 | 2,113 | 1,885 | 1,239 | 9,285 |
| Urban-to-rural | 2,451 | 1,588 | 1,275 | 854 | 533 | 6,701 |
|     Job-seeking | 2,094 | 1,339 | 1,056 | 711 | 435 | 5,635 |
|     Household relations | 357 | 249 | 219 | 143 | 98 | 1,066 |
| Rural-to-urban | 1,185 | 1,318 | 1,282 | 1,113 | 696 | 5,594 |
|     Job-seeking | 1,083 | 1,200 | 1,170 | 1,032 | 625 | 5,110 |
|     Household relations | 102 | 118 | 112 | 81 | 71 | 484 |
| Urban-to-urban | 473 | 405 | 383 | 305 | 147 | 1,713 |
|     Job-seeking | 248 | 218 | 216 | 147 | 57 | 886 |
|     Household relations | 225 | 187 | 167 | 158 | 90 | 827 |

*Notes*: The first column shows migration streams and the reason for migration by migration stream. Each column shows years of stay from time of migration until the survey week. We exclude students, self-employed, housewives, and farmers in the analysis. We focus on migrant wages worker only. We also drop the sample of Education/Training, Medical treatment, and Other reasons. We classify migrants related to job search, job transfer, and back to former place of residence into "Job-seeking" migrants.

*Source*: Thailand Labor Force Survey, 1994-1996.

Table 3: **Log of Weekly Wages for New Entrants into the Rural Labor Market by Duration**

| Subsample | Mean | S.D | Obs. |
|---|---|---|---|
| Rural-to-rural migrants | 6.807 | .660 | 39521 |
| Less than 1 year | 6.656 | .668 | 11530 |
| 1 year | 6.796 | .657 | 9377 |
| 2 years | 6.870 | .634 | 8141 |
| 3 years | 6.915 | .641 | 6412 |
| 4 years | 6.959 | .647 | 4061 |
| | | | |
| Urban-to-rural migrants | 6.859 | .694 | 6701 |
| Less than 1 year | 6.705 | .670 | 2451 |
| 1 year | 6.853 | .690 | 1588 |
| 2 years | 6.944 | .672 | 1275 |
| 3 years | 7.025 | .674 | 854 |
| 4 years | 7.114 | .729 | 533 |

*Notes*: All new entrants into rural labor market are classified by reason for migration, migration streams, and years of stay. All new entrants is restricted to wage workers aged 40 or below.

*Source*: Thailand Labor Force Survey, 1994-1996.

Table 4: **Log of Weekly Wages for New Entrants into the Urban Labor Market by Duration**

| Subsample | | Mean | S.D | Obs. |
|---|---|---|---|---|
| Rural-to-urban migrants | | 6.771 | .474 | 5594 |
| | Less than 1 year | 6.652 | .460 | 1185 |
| | 1 year | 6.740 | .466 | 1318 |
| | 2 years | 6.777 | .443 | 1282 |
| | 3 years | 6.840 | .490 | 1113 |
| | 4 years | 6.909 | .490 | 696 |
| | | | | |
| Urban-to-urban migrants | | 7.075 | .507 | 1713 |
| | Less than 1 year | 6.963 | .480 | 473 |
| | 1 year | 7.070 | .496 | 405 |
| | 2 years | 7.093 | .516 | 383 |
| | 3 years | 7.142 | .494 | 305 |
| | 4 years | 7.266 | .548 | 147 |

*Notes*: All new entrants into the urban labor market are classified by reason for migration, migration stream, and duration of stay. All new entrants are restricted to wage workers aged 40 or below.

*Source*: Thailand Labor Force Survey, 1994-1996.

Table 5: **Job-Seeking Migrants from Rural to Urban Had Higher Wage Level while Job-Seeking Migrants from Urban to Rural Experienced Lower Wage Level**

| Subsample | $\Gamma$ | Adjusted $R^2$ | Obs. |
|---|---|---|---|
| Rural-to-rural | $-.048^{***}$ | .428 | 39521 |
| | (.005) | | |
| Urban-to-rural | $.053^{***}$ | .342 | 6701 |
| | (.018) | | |
| Rural-to-urban | $-.088^{***}$ | .329 | 5594 |
| | (.022) | | |
| Urban-to-urban | $.080^{***}$ | .474 | 1713 |
| | (.020) | | |

*Notes*: The dependent variable is the log of weekly wages. The household related subgroup dummy variable equals 1 if individual $i$ moves to location $K$ for household related reasons, and otherwise if individual $i$ moves to location $K$ to actively seek employment. The explanatory variables are gender, age, years of education, square of years of education, marital status, and a household related subgroup dummy. All individual explanatory variables are highly significant at the 1% level. This table focuses on each coefficient of the exogenous variable on the migration decision: household related subgroup dummy. Numbers in parentheses are standard errors.

*Source*: Thailand Labor Force Survey, 1994-1996.

*** significant at the 1 % level.

Table 6: **Household Related Subgroup Dummies in Wage Equations of Each Duration in the Rural**

| Subsample | $\Gamma_0$ | $\Gamma_1$ | $\Gamma_2$ | $\Gamma_3$ | $\Gamma_4$ |
|---|---|---|---|---|---|
| Rural-to-rural | .013 | $-.034^{***}$ | $-.072^{***}$ | $-.109^{***}$ | $-.091^{***}$ |
| | (.011) | (.012) | (.012) | (.012) | (.016) |
| Adjusted $R^2$ | .359 | .429 | .455 | .482 | .438 |
| Obs. | 11530 | 9377 | 8141 | 6412 | 4061 |
| | | | | | |
| Urban-to-rural | .004 | .037 | $.112^{***}$ | $.155^{***}$ | .050 |
| | (.032) | (.038) | (.035) | (.052) | (.055) |
| Adjusted $R^2$ | .230 | .340 | .406 | .437 | .516 |
| Obs. | 2451 | 1588 | 1275 | 854 | 533 |

*Notes*: The dependent variable is the log of weekly wages. The explanatory variables are gender, age, years of education, square of years of education, marital status, and a household relation migrants dummy. All individual explanatory variables are highly significant at the 1% level. This table also focuses on each coefficient of the exogenous variable on the migration decision: household relations migrants dummy. Numbers in parentheses are standard errors.

*Source*: Thailand Labor Force Survey, 1994-1996.

*** significant at the 1% level.

Table 7: **Household Related Subgroup Dummies in Wage Equations of Each Duration in the Urban**

| Subsample | $\Gamma_0$ | $\Gamma_1$ | $\Gamma_2$ | $\Gamma_3$ | $\Gamma_4$ |
|---|---|---|---|---|---|
| Rural-to-urban | .069 | $-.170^{***}$ | $-.086^{**}$ | $-.003$ | $-.206^{***}$ |
| | (.048) | (.047) | (.044) | (.063) | (.049) |
| Adjusted $R^2$ | .356 | .363 | .349 | .275 | .292 |
| Obs. | 1185 | 1318 | 1282 | 1113 | 696 |
| | | | | | |
| Urban-to-urban | $-.051$ | $.115^{***}$ | .036 | $.220^{***}$ | $.196^{***}$ |
| | (.040) | (.040) | (.049) | (.041) | (.076) |
| Adjusted $R^2$ | .418 | .585 | .427 | .556 | .421 |
| Obs. | 473 | 405 | 383 | 305 | 147 |

*Notes*: The dependent variable is the log of weekly wages. The explanatory variables are gender, age, years of education, square of years of education, marital status, and a household relations migrants dummy. All individual explanatory variables are highly significant at the 1% level. This table also focuses on each coefficient of the exogenous variable on the migration decision: household relations migrants dummy. Numbers in parentheses are standard errors.

*Source*: Thailand Labor Force Survey, 1994-1996.

*** significant at the 1% level.

** significant at the 5% level.

Table 8: **Each Duration Dummy in Wage Equations by Reason for Migration in the Rural**

| Subsample | 1 year | 2 years | 3 years | 4 years | Adjusted $R^2$ | Obs. |
|---|---|---|---|---|---|---|
| Rural-to-rural | | | | | | |
| Job-seeking | .023*** | .055*** | .035*** | .083*** | .464 | 30236 |
| | (.007) | (.008) | (.008) | (.010) | | |
| Household relations | .042*** | .053*** | .038*** | .043*** | .338 | 9285 |
| | (.015) | (.015) | (.015) | (.017) | | |
| Urban-to-rural | | | | | | |
| Job-seeking | $-.010$ | .050*** | .069*** | .102*** | .347 | 5635 |
| | (.018) | (.020) | (.025) | (.032) | | |
| Household relations | $-.188$*** | $-.168$*** | $-.064$*** | $-.049$*** | .338 | 1066 |
| | (.054) | (.058) | (.058) | (.071) | | |

*Notes*: The dependent variable is the log of weekly wages. The explanatory variables are gender, age, years of education, square of years of education, marital status, and durations of experience in rural area dummies. The benchmark duration is less than 1 year. All individual explanatory variables are highly significant at the 1% level. This table also focuses on each coefficient of *Short* experience dummies. Numbers in parentheses are standard errors.

*Source*: Thailand Labor Force Survey, 1994-1996

\*\*\* significant at the 1% level.

Table 9: **Each Duration Dummy in Wage Equations by Reason for Migration in the Urban**

| Subsample | 1 year | 2 years | 3 years | 4 years | Adjusted $R^2$ | Obs. |
|---|---|---|---|---|---|---|
| Rural-to-urban | | | | | | |
| Job-seeking | .068*** | .099*** | .149*** | .224*** | .357 | 5110 |
| | (.016) | (.016) | (.016) | (.018) | | |
| Household relations | −.087*** | −.190*** | −.155** | −.173*** | .277 | 484 |
| | (.071) | (.073) | (.073) | (.070) | | |
| Urban-to-urban | | | | | | |
| Job-seeking | −.039 | −.044 | −.133*** | −.079*** | .434 | 886 |
| | (.035) | (.032) | (.037) | (.052) | | |
| Household relations | −.151*** | .003 | −.107*** | −.058 | .520 | 827 |
| | (.042) | (.040) | (.041) | (.054) | | |

*Notes*: The dependent variable is the log of weekly wages. The explanatory variables are gender, age, years of education, square of years of education, marital status, and durations of experience in rural area dummies. The benchmark duration is less than 1 year. All individual explanatory variables are highly significant at the 1% level. This table also focuses on each coefficient of *Short* experience dummies. Numbers in parentheses are standard errors.

*Source*: Thailand Labor Force Survey, 1994-1996

*** significant at the 1% level.

** significant at the 5% level.