A Neglected Phenomenon in Migration Research – Return Migration: Explorations Based on Household Surveys in China

*Key words: Labor migration; Household surveys; Return-migration; China; Questionnaire design; Population-environment nexus.*

## 1. Introduction

### 1.1. Context

Rural labor migration is a major aspect of population change in processes of socio-economic development, especially in developing countries (Bilsborrow et al., 1984), and both affected by and affecting environmental conditions and dynamics across the globe (DeSherbinin et al., 2012; Liang, 2016). Many rural regions have experienced substantial transformations due to shifts in land use and environmental changes which are often linked to and/or affecting climate change and altering economic opportunities in rural areas, as a push factor prompting out-migration as people seek better livelihoods by migrating to urban centers (Fotso, 2007). This migration has profound implications not only for the populations and economies of both rural and urban areas but also for their environments (Gray & Bilsborrow, 2014).

While rural-urban migration has been extensively studied, return migration—in which people who out-migrated to urban or other areas later return to their origin rural households—has not. Indeed, most existing work on return migration is on its effects in origin households and areas not on its determinants, and furthermore in recent decades is mostly on international migration (refs. to be added). Return migration adds a layer of complexity to migration by introducing dynamic, non-linear migration trajectories (Bijak, 2022). Many migrants do not follow a straightforward path of moving from rural to urban areas; instead, their migration can involve multiple moves and different periods of return. For example, those who initially migrate to urban centers for better work opportunities may later return to their rural origins due to changing economic conditions in the origin or destination, achieving a target amount of savings, getting married or divorced, to help care for older persons or children in the origin household, or retirement. Return migration can have significant implications for rural development and community dynamics, as return migrants often bring back new skills, knowledge, and financial resources that can reduce poverty in their origin household and contribute to local development.

Motives for return migration can be diverse and multifaceted, including job loss, poor pay or working conditions, declining economic opportunities, health problems, or dissatisfaction with urban life (noise and air pollution). Also people want to go back to rural areas due to family ties, help care for family members or work on the family farm, or desires for a simpler life. Some may return to reconnect with their roots, pursue agricultural or non-agricultural ventures, or capitalize on opportunities created by technological advancements and digital connectivity. Exploring these motivations provides valuable insights into the evolving nature of migration decisions and the potential for return migration to support sustainable rural development.

Despite its potential importance, return migration is rarely studied, and almost never are the factors influencing it studied (refs)—the focus of this paper. We shall see what kinds of data are needed for this and how existing questionnaire designs fall short.

### 1.2. Theoretical frameworks

The theoretical focus of this study is multidisciplinary, drawing upon economics, sociology, geography and psychology. Decision-making processes underlying both rural out-migration and return migration are complex and influenced by a multitude of factors. Similar theories underly out-migration and return migration, as both are similar in terms of their driving factors. Among these are human capital theories of economists, which focus on the human capital/individual characteristics of individuals seeking to improve their incomes in alternative locations (Sjaastad, 1966; Taylor & Martin, 2001). Push-Pull theory explains migration as also due to contextual factors involving push factors in the origin (e.g., environmental degradation, economic stagnation) and pull factors in destinations (e.g., better job prospects, improved living conditions) (Lee, 1966). areas. Social Networks highlight the role of personal connections and social capital in migration (Munshi, 2020), as people often migrate based on information and support from family and friends in other locations, which can ease the migration process. Resilience and Adaptation explores how communities adapt to environmental and socio-economic changes, including migration (Crate and Nuttall, 2016). Recent theories of Carling (2017) and deHaas (2010, 2023) focus on migrant capabilities and aspirations--such as access to information and financial resources, influence migration decisions.

## 2. Materials and Methods

### 3.1. Description of study sites

Data were collected to study out-migration and other topics in rural areas of two provinces of China, with very different environmental and socioeconomic conditions. Tiantangzhai Township in Anhui Province, nestled in the Dabie Mountain Range, has a subtropical monsoon climate, with fairly mild temperatures and abundant precipitation. It has a rugged terrain and is classified as impoverished, with many households engaged in subsistence farming but increasingly diversifying into other agricultural and non-agricultural activities. Out-migration occurs to both local and distant destinations for better job opportunities, reflecting its status as a migrant-sending area. In contrast, the Checheng and Jichang townships in Ji County, Shanxi Province, are in the semi-arid Loess Plateau region with a dry climate and lower temperatures, with agriculture focused on dryland farming of corn and cultivation of fruit and nut orchards. The area supports a larger population with significant livestock raising and off-farm activities.

### 3.2. Data

Household surveys: During June-August 2014 in Anhui and July-August 2015 in Shanxi, similar household surveys were conducted using a comprehensive questionnaire addressing demographics, socioeconomic conditions, cropland use, economic activities, and participation in two government programs to stimulate replacement of cropland by trees (CCFP or GFG) and retention of existing forest (NFCP). University graduate students were recruited and trained to carry out the surveys under our supervision. Due to low CCFP participation rates (17% in Anhui and 15% in Shanxi), stratified disproportionate sampling (Bilsborrow, 2016; Bilsborrow et al., 1984) was used to ensure approximately equal numbers of household were recruited participating and not participating in CCFP (Zhang et al., 2018, 2020). GPS units were used to record household locations, leading to final samples of 481 households in Anhui and 251 in Shanxi. Each household has 1 to 10 members, including previous ones who had migrated out to other places. The surveys also gathered community-level data on infrastructure and services. To address migration complexities, questionnaires were developed to carefully identify appropriate migrants and return migrants,

Out-migrants, return migrants and non-return migrants: *Out-migrants* were defined as persons who, when they left their household, since 2000, and were aged between 16 and 59 at the time, and who lived away from the household for at least six consecutive months. *Return migrants* are out-migrants who have returned to their origin household by the time of the survey, while *non-return migrants* are out-migrants who have not returned to their origin household and continue to reside elsewhere.

Data preparation and pre-processing: We are interested in modeling return migration decisions for out-migrants during the past 10 years up to the survey time. Thus, we needed to reconstruct, to the extent possible from our questionnaire, time-varying characteristics, particularly age, to reflect the statuses of individuals at the time of their migration decision of returning or not. Once we labeled out-migrants as return or non-return migrants, we reconstructed an age-specific panel dataset for all out-migrants, identifying the age at return for return migrants. The composition of the origin survey by age and sex was also reconstructed for every year based on the data on when migrants left and returned. The ages of non-return migrants in households where someone returned needed to be calculated to include in the model, to establish the appropriate comparison population (Bilsborrow et al., 1984, 1997). For this purpose, we divided non-return migrants into two groups including those whose households have return migrants and those whose households do not. For the group with households having return migrants, we reconstructed ages of non-return migrants to the year when their corresponding return migrant returned; for the group with their households having no return migrants, we reconstructed ages for non-return migrants to be 4 years before the survey date (2010 in Anhui and 2011 in Shanxi) to represent a middle point of return year for all return migrants. We also excluded those aged below 16 at the year of return (or not), considering them as dependents and not decision-makers (e.g., children returning home following parents). The final sample for the analysis includes a total of 997 individuals from 523 households,

### 3.3. Model specification

We fit multilevel mixed-effects logistic regression models to explore the driving factors hypothesized to influence return migration decisions. The multilevel method is used because the dataset is hierarchical or nested, featuring multiple levels that include individual persons, farm households, and contextual factors. Fixed Effects represent the average effects of predictors (independent variables) on the outcome (dependent variable) across all groups, such as age and education among all individuals. Random Effects account for variability at different levels of the hierarchy. In the two-level model, random effects capture the group level (e.g., farm households), allowing each group to have its own intercept and/or slope. This means that the effect of a predictor can vary from one household to another. Logistic transformation is applied considering that the outcome variable is binary indicating whether an individual out-migrant decides to return or not.

## 4. Preliminary Results

Statistical summary: Among the 997 individual out-migrants, 227 are return migrants, 22.8%. Most households just have 1-2 members participating in the study, with the maximum number 6. The log likelihood observed for the table is -402. The results are most interesting, with major variables affecting return migration were time away, marital status, relationship to head, education (with more, stay in city), elevation (more remote), hh has someone working off-farm, travel time, and city size (more likely to return from large mega cities over 10 million than smaller cities, interestingly). Lack of space limits further interpretation. Other runs with different variables and time frames yield methodologically interesting findings. Data limitations resulting from questionnaire design will be discussed, yielding recommendations for improvements in questionnaire design.

**Table 1.** Estimated effects of explanatory variables on return migration based on mixed-effects logistic regression model.

| **Variable** | **Coef.** |
|---|---|
| Age at departure | -0.116 |
| Years since departure | 0.460*** |
| Gender (female) | -0.491 |
| Married or not (1/0) | 0.905** |
| Education (years completed) | -0.112* |
| Son or daughter of household head | 2.392*** |
| Household size | 0.100 |
| Total household land | 0.038 |
| Whether household has animals | -0.528 |
| Household wellness/assets score | 0.049* |
| Whether household receives GFG subsidy | -0.001 |
| Whether household receives NFCP subsidy | -0.001* |
| Household elevation in meters | -0.006*** |
| Time to get to paved road walking | 0.008 |
| Household has member working off-farm | 0.892** |
| City size score | 0.496*** |
| Travel time to destination city | -0.002** |
| Province (Anhui=0; Shanxi=1) | 1.967** |
| Constant | -1.507 |
| Variance (Constant) | 3.159 |
| chibar2(01) | 11.68*** |
| Number of Obs. | 656 |