Determinants of the Choice of Migration Destination*

Marcel Fafchamps† Forhad Shilpi‡

July 2011

Abstract

This paper examines migrants' choice of destination conditional on migration. The study uses data from two rounds of Nepal Living Standard Surveys and a Population Census and examine how the choice of a migration destination is influenced by various covariates, including income differentials across possible destinations. We find that migrants move primarily to nearby, high population density areas where many people share their language and ethnic background. Better access to amenities is significant as well. Differentials in average income across destination districts are significant in univariate comparisons but not once we control for other covariates. Differentials in consumption expenditures are statistically significant but smaller in magnitude than other determinants. It is differentials in absolute, not relative, consumption between destination districts that are correlated with the destination of work migrants. Except for the latter, results are robust to different specifications and datasets.

*We thank for their excellent comments Todd Sorensen, Måns Söderbom, and the participants to seminars at the University of Gothenburg and UC Riverside. We are very grateful to Prem Sangraula and the Central Bureau of Statistics of the Nepal whose assistance with the data was essential for the success of this endeavor. Financial support for this research was provided by the World Bank.

†Department of Economics, University of Oxford. Email: marcel.fafchamps@economics.ox.ac.uk.

‡DECRG, The World Bank

1 Introduction

There has been a long tradition of research on migration issues in development literature (Greenwood 1975, Borjas 1994). Recent research has highlighted the methodological issues in estimating returns to migration, in assessing the role of migration networks in actual migration flows and in evaluating the effect of migration on economic well-being. This literature has contributed significantly to the understanding of the migration process. There is a large descriptive literature, dating back to Ravenstein's original work in the 1880s of what drives migration, along with whole literatures in urban economics, sociology and demography. But with the exception of Stark and Taylor on Mexico and Lokshin et al on Nepal, there is little work by economists on how migrants choose their destination in the context of poorer developing countries. This paper seeks to fill this gap in the economics literature. By focusing on the choice of destination, this research seeks to shed light on the respective role of various location attributes in the choice of migration destination.

The literature on migration maintains that differences in income and infrastructure – suitably corrected for price differentials – play a dominant role in the choice of a place to live. To investigate this issue, we develop an original empirical strategy focusing on the choice of destination conditional on the migration decision. The econometric analysis seeks to identify the main factors influencing the choice of migration destination. We limit our analysis to adult males who have migrated outside their birth district for work reasons. We begin by constructing a measure of expected income differentials between the place of origin and all domestic migration destinations. These differentials are allowed to vary depending on observable migrant characteristics believed to affect labor market outcomes, such as education and language. We also construct measures of social proximity between a migrant's place of birth and each possible destination, using detailed available data on ethnicity, caste, language, and religion. The empirical analysis is conducted by combining LSMS survey data with the 2001 population Census from Nepal.

We also investigate a number of factors that may influence the choice of migration destination but have not received much attention in the existing literature. Fafchamps and Shilpi (2009) have shown that the subjective welfare cost of geographical isolation is high. To investigate this issue, we include regressors controlling for population density and for the average distance to various amenities. Fafchamps and Shilpi (2008) have further shown that migrants are concerned with their welfare relative to that of their birth district as well as to that in their destination location. We examine whether relative welfare considerations influence the choice of migration destination. Additional regressors include distance and prices.

It has long been observed that migrants often are better educated than non-migrants.[1] Migrants may differ from non-migrants in terms of unobservables as well. A number of recent studies have sought to estimate returns to migration that are immune to selection on unobservables (Gabriel and Schmitz,1995; Akee, 2006; and Mckenzie, Gibson and Stillman, 2006). Their results suggest that simply comparing the earnings of migrants and non-migrants overestimates the return to migration. For instance, Mckenzie, Gibson and Stillman (2006) use an experimental design to show that ignoring selection bias leads to an overestimation of the gains from migration by 9 to 82 percent. Similar evidence is reported by researchers investigating the relationship between education and migration (Dahl, 2002).[2] Our empirical strategy sidesteps the isse of selection into migration by focusing on the choice of destination conditional on migrating, rather than on the decision to migrate itself.

Results show that migrants move primarily to high population density areas that are nearby,

[1] A related strand of work points out that migration prospects raise investment in education (de Brauw and Giles, 2006; Batista and Vicente, 2008).

[2] The view that it is the better educated and more able who migrate has not gone unchallenged, however (Borjas, 1994). According to Borjas' negative selection hypothesis, the less skilled are those most likely to migrate from countries/locations with a high skill premia and earnings inequality to countries/locations with a low skill premia and earnings inequality. Chiquiar and Hanson (2005) test and reject this hypothesis for Mexican immigrants in the US and conclude instead for intermediate selection.

have good access to amenities, and where many people share their language and ethnic background. These results confirm earlier work on the factors affecting the subjective welfare cost of isolation (Fafchamps and Shilpi, 2008). Differentials in consumption expenditures are significantly correlated with migrants' destination but the magnitude of the relationship is less important than anticipated. Moreover, it is differentials in absolute, not relative, consumption that are correlated with the destination of work migrants.

The paper is organized as follows. The conceptual framework and testing strategy are presented in Section 2. The data is discussed in Section 3, together with the main characteristics of the studied population. Econometric results are presented in Section 4. Conclusions follow.

2 Conceptual framework

We are interested in factors that are correlated with migrants' likelihood of moving to one of N possible destinations.[3] Let utility of individual h in location $i = \{1, .., N\}$ be denoted U_i^h. The probability of migrating from i to s is expected to increase in $U_s^h - U_i^h$. Our empirical strategy is to use one, anterior dataset to construct estimates of U_s^h for *all* locations to which a migrant h might relocate within the study country, and to use a subsequent dataset to test whether migrants' choice of destination is predicted by $U_s^h - U_i^h$.

Following the literature, let us assume that utility U_s^h in location s is a function of the consumption (or income) level y_s^h that the migrant is likely to achieve, of the prices p_s he will

[3] Others who have studied migration decisions with respect to the place of destination (e.g., Sorensen et al.) have included non-migrants in their analysis. We decided against this approach because it would require controlling for push factors that influence the decision to migrate but not the choice of destination. For instance, some individuals may have access to plenty of land in their place of origin, or they may have relatives they wish to stay close to. Since we do not observe these factors, we would have to control for them by adding an individual-specific place-of-origin fixed effect ϕ_i^h. But then including the place of origin in the analysis of the choice of destination adds no information. The reason is that, for those who do not migrate, there is always a value of ϕ_i^h that accounts for their not moving. For this reason, we chose not to include the place of origin in the analysis and to drop all non-migrants. This means that we are estimating the preferences of migrants. But, ultimately it is the migrants who migrate, so it is their preferences that help us understand where migrants go.

face, and a vector of location-specific amenities A_s (Bayoh, Irwin and Haab, 2006):

$$\begin{aligned} U_s^h &= U^h(y_s^h, p_s, A_s) \\ &\approx y_s^h - \alpha p_s + \beta A_s \end{aligned}$$

Income y_s^h in turn depends on observable z^h and unobservable μ^h characteristics of migrant h:

$$y_s^h = \delta_s + \eta_s z^h + \gamma_s \mu^h + \varepsilon_s^h \tag{1}$$

where ε_s^h is a disturbance independent of z^h and μ^h. Parameters η_s and γ_s vary across locations to capture the idea that returns to talent differs with the mix of activities undertaken in that location (Fafchamps and Shilpi, 2005).

The relative gain a migrant achieves by moving from i to s also depends on the physical and social distance d_{is}^h between i and s (e.g., including differences in religion, language, or caste). As recent papers by Munshi (2003) and Beaman (2006) have shown, social networks play a role in finding employment. Migrants may also value social interaction with neighbors and friends in the place of destination (for entertainment, mutual support, marriage market, etc.). Let d_{is}^h denote a vector of physical and social distances for individual h. We assume that the probability of moving to location s falls with d_{is}^h.

Let M_{is}^h describe h's choice of destinations: $M_{is}^h = 1$ if individual h migrates from location i to location s, and 0 otherwise. By construction, each individual in the sample is a migrant, and each migrant only migrates to a single location. Since we condition on migrating (i.e., $M_{ii}^h = 0$), we can only identify the effect of differences between destinations on the choice of destination. We do not seek to estimate the likelihood of migrating itself. We seek to estimate a model of the form:

$$\begin{aligned}\Pr(M_{is}^h = 1 | M_{ii}^h = 0) &= \lambda\left(\rho E(U_s^h - U_i^h | z^h, \mu^h) - \omega d_{is}^h\right) \\ &= \lambda(\rho(\delta_s - \delta_i + (\eta_s - \eta_i)\, z^h + (\gamma_s - \gamma_i)\, \mu^h \\ &\quad -\alpha(p_s - p_i) + \beta(A_s - A_i)) - \omega d_{is}^h)) \qquad (2)\end{aligned}$$

where $\lambda(.)$ is a logit function. Given the symmetry of the underlying migration choice, we have assumed that coefficient vectors ρ and ω are the same across locations. Estimation is achieved by generating, for each migrant, N observations on M_{is}^h and the regressors and by estimating (2) using logit.[4] Interdependence across observations arises from the fact that, by construction, migrants can only go to a single destination. This generates a pattern of positive and negative correlation between the error terms relative to individual h.[5] We correct for this interdependence by clustering standard errors. A similar approach is used by Fafchamps and Gubert (2003) to estimate dyadic regressions.[6] Here we cluster standard errors by district of origin. This takes care not only of interdependence across observations for each migrant h, but also of possible correlation in the choice of destination by all migrants originating from the same district. We also include individual fixed effects to correct for unobserved differences in U_i^h across migrants; the migrant fixed effect absorbs any effect due to U_i^h, such as differences across migrants in terms of $\delta_i, \eta_i, \gamma_i, p_i$, or A_i. Since non-migrants are omitted from the regression, this means that coefficients ρ,α and β are identified solely from variation across possible destination districts.

[4] The dropped observation corresponds to the location of origin M_{ii}^h which, as explained above, we do not include in the analysis since including M_{ii}^h would mean de facto including the decision of whether to migrate or not.

[5] To see why, consider the simple case when all destinations are equally likely.

[6] Train (2003) discusses other possible estimation methods, such as joint maximum likelihood estimation using multiple integration, or Bayesian methods using Gibbs sampling. With a choice of over 70 possible destinations, multiple integration is out of the question. Gibbs sampling remains a possibility but would require extensive programming. We choose instead to keep the logit approach but to correct the standard errors for possible correlation in errors across choices. The possible efficiency gain achieved by Bayesian methods does not appear to justify the programming cost.

In terms of implementation, we begin by estimating equation (1) using data from an household survey. This yields an estimate of:

$$E[\widehat{y_s^h - y_i^h}|z^h] = \widehat{\delta}_s - \widehat{\delta}_i + (\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$$

for each possible destination. We also use the survey data to obtain information on prices p_s and amenities A_s. We then use these and $\widehat{\delta}_s - \widehat{\delta}_i$ and $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ to estimate equation (2) using census data.

How adequately does this approach take care of unobserved heterogeneity? We begin by noting that, in general $E[z^h \mu^h] \neq 0$: observable and unobservable talents are correlated. For those who wish to estimate the return to a specific individual characteristic z^h, this correlation is problematic. For our purpose, this correlation is good news. To see this, consider the extreme case in which μ^h is a deterministic function of z^h:

$$\mu^h = \lambda z^h$$

Inserting in (1), we get:

$$y_i^h = \delta_i + (\eta_i + \gamma_i \lambda) z^h + \varepsilon_i^h$$

In this case the estimated coefficient of z^h also captures the effect of unobserved heterogeneity on income:

$$E[\widehat{\eta}_i] = \eta_i + \gamma_i \lambda$$

and $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ in equation (2) controls for *both* observed and unobserved heterogeneity.

What happens if z^h and μ^h are only imperfectly correlated? Say we have:

$$\mu^h = \lambda z^h + v^h$$

with $E[v^h] = 0$ and $E[z^h v^h] = 0$. Inserting in (1), we get:

$$y_i^h = \delta_i + (\eta_i + \gamma_i \lambda) z^h + \gamma_i v^h + \varepsilon_i^h$$

It follows that:

$$p \lim[\widehat{\delta}_i] = \delta_i + \gamma_i p \lim[v^h] = \delta_i$$

In v^h there probably remains variation in returns to unobserved individual characteristics. This variation may affect the choice of migration destination. It should not, however, affect the coefficient of $\widehat{\delta}_s - \widehat{\delta}_i$ and $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ in the migration regression since, by construction, they are orthogonal to v^h.

What of equation (2)? It can be rewritten:

$$\begin{aligned} \Pr(M_{is}^h &= 1) = f_+[\rho(\delta_s - \delta_i + (\eta_s - \eta_i + \lambda(\gamma_s - \gamma_i))\, z^h \\ &\quad -\alpha(p_s - p_i) + \beta(A_s - A_i)) - \omega d_{is}^h + u_{is}^h] \qquad (3) \\ u_{is}^h &\equiv (\gamma_s - \gamma_i)\, v^h \end{aligned}$$

which shows that, since v^h is uncorrelated with z^h by construction, $(\widehat{\eta}_s - \widehat{\eta}_i) z^h$ is uncorrelated with the disturbance term u_{is}^h.

We have discussed unobserved heterogeneity in income generation. There can also be unobserved heterogeneity in migration costs. We are particularly concerned about the large proportion of surveyed households who still live in their birth district. This population includes households who chose not to migrate, but also many households for whom the cost – or the risk – of migrating were probably too high. Munshi and Rosenzweig (2005) have shown that mutual insurance within castes in India provides a strong disincentive to migrate. The same probably applies to our study country, which is neighboring India. It follows that the decision not to migrate at all – $M^h_{ii} = 1$ – is distinct from the choice of a destination, conditional on migrating. This is why, to minimize the bias that self-selection into migration may generate, we drop M^h_{ii} and estimate (3) with migrants only. Since we have no data on individuals who have left the country, our analysis is only pertinent to internal migrants.

We worry about possible circularity resulting from general equilibrium effects (Dahl, 2002; Hojvat-Gallin, 2004; Borjas, 2006; Bayer, Khan and Timmins, 2008). If many people migrate to a specific location, such as the capital city, this is likely to affect wages, incomes, and access to amenities in that location.[7] This would generate a potential endogeneity bias due to the fact that incomes and amenities in that location result in part from the decision of many migrants to locate there.

To minimize this bias, we estimate income regressions (1) using anterior data. More precisely, let T be the period for which we have income information and $T + t$ the period at which we observe migrants. The income regression is estimated using data for period T. Migrants are defined as those who migrated between T and $T + t$. Migration decision are thus assumed to be taken based on income differentials at time T, that is, prior to migration. Assuming that migrations depend on income levels at T is a reasonable assumption given that most migrants in our dataset come from rural areas of Nepal and are unlikely to be particularly good at forecasting differential income trends in multiple locations.

We also examine whether migrants consider relative incomes – rather than absolute incomes

[7]The effect could be negative – e.g., congestion – or positive – e.g., agglomeration externalities.

– when deciding where to migrate. This point was already touched upon by Stark and Taylor (1991), who do not consider relative deprivation in the migrant's destination but show that households' relative deprivation in their village of origin is significant in explaining migration to destinations where a reference group substitution is unlikely and the returns to migration are high. More recent work in economics and psychology has shown that subjective well-being depends on relative achievement, of which one dimension is income (see Fafchamps and Shilpi, 2008 and 2009 for brief surveys of the literature). This raises the question of whether people choose the migration destination that, on the basis of their individual characteristics, promises them a high income relative to that of others in that location. To investigate this idea, we reestimate the model by replacing y_i^h with $y_i^h/\overline{y}_i$ in equation (1) and proceeding as outlined above. If migration decisions are based on relative rather than absolute income, then the coefficients of $\widehat{\delta}_s - \widehat{\delta}_i$ and $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ should be positive and significant only when they are computed using $y_i^h/\overline{y}_i$.

3 The data

Having described the conceptual framework and estimation strategy, we now present the data. The data used in this paper come from two sources: living standard household surveys, and population census.

The living standard data come from two rounds of Nepal Living Standards Survey (NLSS). The first round was conducted in 1995/96 while the second took place in 2002/3. The NLSS surveys collected detailed information on households and individuals using nationally representative samples. The 1995/96 NLSS survey is used as source of detailed information about locally available amenities. It is also used to estimate the income regression (1).

Survey data are complemented with information from the 2001 population census. The short population census questionnaire was administered to the whole population. It contains information about ethnicity and language. For a randomly selected 11% of the census population, additional information was collected using a second, longer questionnaire. This questionnaire collected information on district of current residence, district of residence 5 years prior to the census, and district of origin. Detailed information is also available on gender, age, education, unemployment, occupation, and motive for migration, if any. The Nepalese Central Bureau of Statistics was kind enough to merge the short and long questionnaire datasets for the 11% of the population covered by the long questionnaire. This provides a very large data set on which we estimate the migration regression (3).

Nepal is divided into 75 districts and further subdivided into 3915 VDCs and 35235 wards. The 11% population census covers approximately 2.5 million individuals in 520624 households. 345349 of these individuals are living in a district other than their district of residence and 119475 have moved in the five years preceding the census, that is, in the period between the 1995/96 NLSS and the 2001 census. Most of these individuals have moved for reasons other than work. Marriage is the dominant reason for moving among women; study is the dominant reason for moving among children and youths. In contrast, of the adult males who migrated during last 5 years, 69% moved for work reasons.

Because our focus is on work migration, we restrict our attention to adult males. Among those, 16850 are recorded as having moved in the five years preceding the census specifically for work reasons. These individuals are the focus of our analysis. We note that, by construction, this approach excludes those who have migrated outside Nepal. Our focus is thus on internal migrants. We do not have data on India but since there is no big Indian city within 200 Km of the Nepalese border, commuting to India for work while residing in a Nepalese district is rare, making it unlikely that economic opportunities in neighboring India affected the choice of migration destination within Nepal.

Figures 1 and 2 show the geographical distribution of work migrants in terms of district of residence and origin. We see that a small number of destination districts have a high proportion of work migrants. In contrast, districts of origin are distributed widely across the country. This reflects the fact that much work migration is from remote rural areas to towns and cities.

The main characteristics of work migrants are reported in Table 1, together with those of non-migrant adult males. We see that work migrants are on average younger and better educated. The census contains detailed information about ethnicity, language, and religion. In the Nepal census, the term 'ethnicity' is used to capture a hodgepodge of caste and tribal distinctions. The census distinguishes up to 103 ethno-caste categories. Most of these categories only account for a tiny proportion of the total population. In terms of the total adult population, the most common ethno-caste categories are Chhetri, Brahmin, and Newar who, together, account for 35% of adult males in the 11% census. All three categories are regarded as upper castes. As we see from Table 1, migrants are much more likely to be upper caste than non-migrants.

The census distinguishes 84 different languages. The main ones are Nepali and Maithili, spoken by 58% of the population. In Table 1 we see that work migrants are much more likely to speak Nepali, the main language in the country. While the Nepalese population is heterogeneous in terms of ethnicity and language, it is relatively homogeneous in terms of religion: 81% of adult males are Hindu and 11% are Buddhist. We see in Table 1 that work migrants are predominantly Hindu.

The dependent variable M_{is}^h in our main regression of interest, regression (3), is constructed as follows. We begin by creating, for each of the 16850 work migrants h identified in the 11% census, 75 M_{is}^h observations corresponding to each of the possible 75 district destinations s. We set $M_{is}^h = 1$ if migrant h moved from district i to district s in the 5 years preceding the census, and 0 otherwise. We then drop M_{ii}^h since we focus on migrants. By construction a migrant reside in one district. For each migrant, variable M_{is}^h thus takes value 1 once and value 0 73 times.

Since the migrant can only move to a single destination, the 74 M_{is}^h observations are not independent and residuals in (3) are correlated. Dependence across M_{is}^h observations combines negative and positive correlation. To illustrate this point, imagine for a moment that all destinations are equivalently attractive to the migrant. The probability $\Pr(M_{is}^h = 1)$ of selecting one of them is thus 1/74. Further assume that one of them is selected at random; for this observation, we have $u_{is}^h = 1 - \Pr(M_{is}^h = 1) = 73/74$. For all other observations, the residual $u_{is}^h = -1/74$. We see that, for individual h, the observation in which $M_{is}^h = 1$ is negative correlated with observations in which $M_{is}^h = 0$. We also see that observations in which $M_{is}^h = 0$ are positively correlated with each other. This combination of positive and negative correlation means that a standard fixed or random effect approach is not sufficient to ensure correct inference; clustering standard errors by individual is necessary. This is what we do.[8]

Having described how the dependent variable is constructed, we turn to regressors. We begin by describing how we construct an estimate of $\widehat{E[y_s^h|z^h]}$, the level of consumption (or income) y_s^h that a migrant with characteristics z^h can expect to earn in district s.[9] To construct such an estimate, we use the 1995/96 NLSS data. The reason for using the 1995/96 data instead of the 2002/3 NLSS survey is to avoid reverse causation, i.e., migration causing a change in income patterns. Migrants are unlikely to be able to accurately predict the evolution of incomes in each district over time. Income and consumption levels observable before they migrated are thus a

[8] To be more precise, we cluster by district of origin, and this encompasses clustering by individual.

[9] Districts are divided into wards. Ideally we would have wanted to estimate $\widehat{E[y_s^h|z^h]}$ for each ward, as this would yield a more accurate expected income proxy. But we do not have NLSS data for all wards. Furthermore, NLSS sample size within each ward (12 households) is too small to permit estimation of the slope coefficients $\widehat{\eta}_s$ in each ward. We also do not have many of the other regressors at the ward level.

reasonable starting point. While income is in principle a better choice of regressor to explain migration patterns, it is also subject to more measurement error. For this reason, we also use consumption expenditures.

Using the NLSS data we begin by estimating a regression of the form:

$$y_s^h = \delta_s + \alpha(a_s^h - \overline{a}) + \beta_s(E_s^h - \overline{E}_s) + \chi_s(H_s^h - \overline{H}_s) + \zeta x_s^h + v_s^h \tag{4}$$

where y_s^h is the log of consumption (or income) of household h residing in district s, coefficients δ_s, β_s and χ_s vary by district, a_s^h stands for the age and age squared of the household head, E_s^h is the education level of the head measured in years of completed education, $H_s^h = 1$ if the head's mother tongue is other than Nepali, the national language, and x_s^h is a vector of household composition variables. Since income or consumption are expressed in logs, β_s and χ_s can be thought of as education and language income premia, respectively. Female headed households are excluded from the regression since the focus is on migrant males. Vector $\overline{a}$ denotes the average age and age squared of observations across the sample. Variables $\overline{E}$ and $\overline{H}_s$ denote the district-specific averages of E_s^h and H_s^h. By demeaning regressors, we ensure that $\widehat{\delta}_s$ measures the unconditional, district-specific average of y_s^h. Household size and the share of adult males and females are included as controls because larger households with more adults should earn more income and consume more; omitting them would overestimate incomes in districts where households are larger, e.g., rural districts, and this may bias results.[10] Other household characteristics are not included because they are possibly affected by migration.

Equation (4) is estimated with correct sampling weights using data on all individuals, mi-

[10] We revisit this assumption when we present robustness checks without household size and composition in regression (4). The literature has often emphasized that migrations can serve an important role in household formation. For migrants, the prospect of forming a large, successful household may be one of the purposes of migration.

grants and non-migrants.[11] In the 1996/6 LSMS the overwhelming majority of household heads (i.e., more than 80%) still resided in their birth village, probably because the economic and psychological costs of migrating were high. This means that the distribution of unobserved talent μ^h among 1995 district residents corresponds roughly to the distribution of talent in the population at large. This implies that the bias in estimating δ_i is probably small when we estimate (1) using data on all district residents. In the robustness section, we examine whether our results differ when we only use non-migrants and correct for selection correction. We cannot estimate (4) using migrants only because there are not enough observations, especially for rural districts.

Regression estimates for equation (4) are summarized in Table 2 where we show the coefficients α and ζ of the control variables as well as the average and standard error of $\widehat{\delta}_s, \widehat{\beta}_s$ and $\widehat{\chi}_s$. The coefficients $\widehat{\delta}_s, \widehat{\beta}_s$ and $\widehat{\chi}_s$ are large and jointly significant. There is considerable variation across districts not only in average log income and consumption but also in the income or consumption premia associated with education and language. These results are used to construct, for each of the 16,000 or so work migrants in the census, a measure of the income or consumption they can expect to achieve in each of the possible destination districts. Formally, this measure is calculated as:

$$\widehat{E[y_s^h|z^h]} = \widehat{\delta}_s + \widehat{\beta}_s(E_s^h - \overline{E}_s) + \widehat{\chi}_s(H_s^h - \overline{H}_s) \tag{5}$$

where E_s^h and H_s^h are the education and language dummy for migrant h. Age is ignored from the calculation since work migrants typically migrate around the same age, i.e., in early adulthood.

Formula (5) can be decomposed into two parts: $\widehat{\delta}_s$, which measures the average income level in district s, and $\widehat{\eta}_s z^h \equiv \widehat{\beta}_s(E_s^h - \overline{E}_s) + \widehat{\chi}_s(H_s^h - \overline{H}_s)$ which captures individual-specific variation

[11] The 1995/96 NLSS survey adopted the following sampling strategy. Within each district a small number of wards were selected at random. Within each ward, 12 randomly selected households were interviewed. Because the wards differ widely in terms of population, applying sampling weights is essential in order to obtain consistent estimates of δ_s.

in income. Migration models predict that, other things being equal, the choice of migration destination should depend on $E[\widehat{y_s^h|z^h}]$. This means that if we regress the choice of destination separately on $\widehat{\delta}_s$ and $\widehat{\eta}_s z^h$, they should have the same coefficient.

A similar methodology is used to construct other variables that may affect the choice of destination. Building on a growing literature documenting the relationship between subjective welfare and relative income, Fafchamps and Shilpi (2008) show that Nepalese households care about their consumption level relative to that of others in the same location. If this is the case, it is conceivable that migrants choose their destination not so much for the absolute gain in income it may provide but for the gain in relative status that would ensue. For instance, if returns to education and ability are higher in an urban setting, an educated individual may improve his relative position in society by moving from a rural to an urban setting. To investigate this possibility, we estimate equation (4) using the log of relative income (or relative consumption) as dependent variable and construct a predicted relative income measure using the same formula (5). These are shown in columns 3 and 4 of Table 2.

Theories of work migration predict that individuals move to increase their utility or welfare. The 1995/96 NLSS asked respondents a number of questions regarding their subjective satisfaction level with various dimensions of consumption – namely, food, clothing, housing, health care, and child schooling. They were also asked their subjective satisfaction with their level of total income. We apply the same methodology to these data – i.e., we estimate a regression of the same form as (4) and apply formula (5) to construct an expected subjective satisfaction index. Estimation results are shown in columns 5 to 10 in Table 2. If migrants correctly anticipate the subjective satisfaction they will enjoy from moving to different destinations, these subjective satisfaction measures may offer a better way of controlling for expected welfare differences across destinations.

To control for migration costs, we construct variables proxying for geographical and social distance. For geographical distance between districts, we use the arc distance between the district of origin and each possible district of destination, computed from the average longitude and latitude of each districts.[12] We expect the cost and risk of migration to increase with physical distance.

Social distance is proxied by the proportion of individuals in the district who share the same language, religion, and ethno-caste group. This is implemented as follows. From the census we have information on ethnic, religious, and language diversity in all districts of the country. From these we construct an index of similarity between individual h and the population of each district. Let m denote a specific trait – e.g., ethnicity, religion or language – and let p_s^m be the proportion of the population of district s that has trait m. Consider the trait m_h of individual h. We expect h's chances of finding a job, etc, to increase in the proportion of individuals in the district of destination who share the same trait. We construct, for each destination and each migrant, a variable $p_s^{m_h}$ equal to the proportion of members of h's with trait m_h. For this migrant, the social distance between two locations i and s is $p_s^{m_h} - p_i^{m_h}$. The idea behind this measure is that individual h 'fits' better in district s if the proportion of like individuals is higher than in his district of origin. We construct similar indices for language and religion. Note the similarity between $p_s^{m_h}$ and the commonly used index of ethno-linguistic fractionalization (ELF). The ELF index measures the probability that two individuals taken at random belong to the same ethnic or linguistic group. Variable $p_s^{m_h}$ measures the probability that an individual taken at random in the population of district s belongs to the same ethno-caste or linguistic group as the migrant, and is thus the individual-equivalent of the ELF index for groups.

We seek to control for price differences p_s across locations. This is difficult because we do not

[12] The average longitude and latitude of a district are obtained as a weighted average of the longitude and latitude of all the VDC's in the district, where the population of each VDC serves as weight.

have detailed price data. We use the price of rice as a proxy for the price of common household goods. This is not entirely satisfactory but, in the absence of a district-level consumer price index, this is the best we can do. Given the mountainous nature of Nepal, rice cannot be grown in many parts of the country. The price of rice thus tends to rise with altitude and geographical isolation, as we expect the prices of many manufactures to do as well. The 1995/96 NLSS collected information on the quantity and price paid for rice by individual households. From this we compute a unit price per Kg. The log of the district median is used as our price index proxy.

To capture amenities A_s and other location effects, we construct a district-specific housing rental premium. To the extent that people are mobile, differentials in housing costs capture, in a reduced form, the effect of location attributes such as proximity to jobs and access to public amenities. To construct a proxy for location attributes, we take advantage of a section of the 1995/96 NLSS survey focusing on housing. The survey collected information on hypothetical and actual house rental values of each household together with house characteristics such as square footage, number and type of rooms, quality of materials, and the availability of various utilities. We use these data to construct an hedonic index of housing premium for each district. Let r_s^k be the house rental price paid (or estimated) by household h in district s and let x_s^h denote a vector of house characteristics. We estimate a regression of the form:

$$\log r_s^k = a_s + bx_s^h + e_s^k$$

to obtain estimates of $\widehat{a}_s$, the housing premium in each district s. Since the dependent variable is in log form, $\widehat{a}_s$ measures the percentage housing premium in each district. Regression results are shown in Table A1 in appendix. Many house characteristics are significant with the expected sign, e.g., larger, better built houses with better in-house amenities get a higher rent. District differentials in housing premia are large and jointly significant. To the extent that the housing premium captures differences in amenities, we expect migrants to be attracted by districts with a high $\widehat{a}_s$. To further control for access to amenities, we include travel time to the nearest road (a measure of market access) and to the nearest bank (a measure of financial and commercial development).

Finally, we include a number of regressors to control for geographical isolation. Fafchamps and Shilpi (2009) have shown that, in Nepal, subjective welfare is negatively associated with geographical isolation. Census data on total population and population density in each district are used as proxies for urbanization and geographical proximity: the denser the population, the less geographically isolated individuals are likely to be. We also include data on the average elevation in each district. Nepal being a mountainous country, the higher the average elevation of a district, the more costly it is to build roads, raising transport and delivery costs to the district. Ceteris paribus, we expect migrants to seek out districts with a higher population density and a lower elevation.

4 Econometric Results

We now investigate the choice of migration destination. We begin with descriptive statistics before presenting the econometric results.

4.1 Descriptive analysis

Descriptive statistics for all variables used in the analysis are presented in Table 3. All variables in the Table are of the form $\Delta_{is}^h = x_s^h - x_i^h$ where i is the district of origin of migrant h and s is each of 74 possible districts of destination. We examine the average value of Δ_{is}^h for the destination district and compare it to the value of Δ_{is}^h for alternative destinations. For instance, let x_s^h be population density in district s. The average value of Δ_{is}^h for the actual destination of the migrant tells us whether the destination district is more densely populated than the district of origin. The comparison between Δ_{is}^h for actual and hypothetical destinations tells us whether the actual district of destination is more densely populated than alternative destinations.

Estimated district averages $\widehat{\delta}_s$ appear at the top of the Table. We have two estimates of $\widehat{\delta}_s$, one obtained using reported income data, and the other based on reported consumption data. Given that most respondents to the NLSS survey are self-employed, measurement error is typically larger for income than for consumption. We see that our estimates of log income and consumption $\widehat{\delta}_s$ are on average 23% and 12% higher in the district of destination than in the district of origin, respectively. Migrating to one of the 73 alternative destinations would, on average, have reduced income and consumption relative to the district of origin. The difference in anticipated income and consumption between actual and hypothetical destinations is statistically significant. Migrants thus tend to move to districts where consumption and income are unconditionally higher.

Next we examine whether there are significant differences in returns to individual characteristics $\widehat{\eta}_s z^h$. For income, $\widehat{\eta}_s z^h$ is on average lower in the district of destination than in the district of origin. The difference is large enough to be statistically significant at the 10% level. This implies that better educated, Nepali-speaking migrants gain relatively less from migrating to actual destination districts than less educated, non-Nepali speaking migrants. In contrast, $\widehat{\eta}_s z^h$ estimates based on consumption data show an increase relative to the district of origin. But the difference with alternative destinations is not significant.

Differences in relative log income and consumption are displayed next. Predicted relative log income and consumption are generated using the same formula $\widehat{\delta}_s + \widehat{\beta}_s(E_s^h - \overline{E}_s) + \widehat{\chi}_s(H_s^h - \overline{H}_s)$ used for log income, except that, relative income (or consumption) is used as dependent variable.

By construction, $\widehat{\delta}_s = 0$. We see that relative income falls between the district of origin and the district of destination while it would have risen in alternative destinations. The difference is statistically significant. In contrast, relative consumption is higher in the destination district than in the district of origin but the difference between actual and hypothetical destinations is not significant.

We then turn to differences in subjective welfare. The equivalent of $\widehat{\delta}_s$ is used as for log income. We begin with subjective perceptions regarding the adequacy of total income. Relative to their district of origin, the average subjective satisfaction with total income is found to rise between the district of origin and the district of destination. Whether this is fully anticipated by migrants is unclear. Fafchamps and Shilpi (2008) show that in assessing their subjective satisfaction migrants still compare themselves to those in their district of origin.

Results regarding subjective satisfaction from the consumption of food, clothing, housing, health care, and schooling are shown next. We see that in all cases the district of destination has a much larger level of subjective satisfaction, both relative to the district of origin and relative to other possible destinations. We also compute the equivalent of $\widehat{\eta}_s z^h$ and find it to be positive in five out of six cases. All migrants improve their consumption adequacy relative to their district of origin – and alternative destinations – but better educated, Nepali-speaking migrants improve it more. The only exception is income, a finding that is consistent with the fall in $\widehat{\eta}_s z^h$ found for income between the districts of origin and destination.

We then turn to prices and amenities. We observe on average an 9% fall in the median price of rice between the districts of origin and destination. Migrating to alternative destinations would have raised the price of rice instead of reducing it. This is consistent with our interpretation that the price of rice captures differences in delivery costs driven by isolation. In contrast, we find a 38% average increase in housing premium between the districts of origin and destination.

Moving to an alternative destination would also have raised average housing costs but by less than that in the actual destination district. Travel time to various facilities and infrastructures falls uniformly between the district of origin and that of destination. Since these differences are strongly correlated with each other, we only report two: travel time to the nearest road, and travel time to the nearest bank. Both fall massively between district of origin and destination, and both would have risen had the migrant moved to an alternative destination.

We observe a strong negative difference in elevation between the district of origin and district of destination. Moving to an alternative destination would, on average, have resulted in a higher elevation than the district of origin. This implies that migrants on average move down from the mountains. They also tend to go to districts with a larger and more dense population than the district of origin and alternative destinations. Migration is thus primarily from rural to urban areas.

In terms of social proximity, we see that migrants on average face a population that is more different from them in terms of both language and ethno-caste than in their district of origin. This is true for the actual destination district but also for alternative districts. We do not observe the same pattern for religion; if anything, migrants are more likely to face someone of their religion in their district of destination. The difference is small, however.

Finally, the geographical distance between the district of origin and the actual destination is on average much smaller than that between the district of origin and alternative destinations: migrants tend to go to a district that is much closer to their district of origin than alternative migration destinations. The difference is strongly statistically significant and large in magnitude.

To summarize, simple bivariate analysis shows that migrants tend to move to a district with: a larger population and population density; a lower elevation; a higher average income and consumption; higher subjective consumption adequacy; lower rice prices and a higher housing premium; better access to public amenities; and close to the district of origin. In contrast, migrants move to districts where they have a lower relative income compared to their district of origin. They also tend to move to districts where fewer people speak their language and belong to their caste or ethnic group, but more share their religion.

4.2 Multivariate analysis

We have seen that there are strong differences between actual and alternative migration destinations. Many of these characteristics are correlated with each other, however. To disentangle them we turn to multivariate analysis and estimate the migration regression (3). As explained in the previous section, regressors include: prices as described above; geographical and social distance; and access to amenities. We also include the log of total population, population density, and average elevation as additional controls.

We begin by estimating (3) with $\widehat{\delta}_s - \widehat{\delta}_i$ computed from the log income data.[13] Results are shown in the first column of Table 4. As discussed earlier, reported results include individual fixed effects and standard errors clustered by district of origin.[14] The univariate analysis showed that income was significant on its own. Once we control for distance, population, prices and amenities, the difference in expected income is no longer significant.[15] Most of other variables remain significant, though. Distance has the expected negative sign and is strongly significant – on average the migration destination is closer to the district of origin than alternative des-

[13] The issue of correcting standard errors for the use of predicted regressors is discussed in detail in appendix.

[14] Omitting individual fixed effects does not affect results much, but standard errors are very different without clustering, confirming that observations are indeed not independent.

[15] It should be noted that the regression is comparing income at destination with that in other possible destinations – not with income in the district of origin, the effect of which is nullified by the migrant fixed effect. What the results show is that, after controlling for population density, migration costs and amenities, income at destination is not significantly higher than possible alternative destinations. Compared with income at origin, the first column in Table 3 shows clearly that income at the destination is higher. This is consistent with the prediction of the migration literature that income prospect is an important determinant of migration flow between origin and destination, but it is not the focus of our analysis.

tinations. The destination district also has a significantly larger population and population density, a lower elevation, and a lower rice price. The housing premium in contrast is higher in the destination district than in alternative destinations, probably because they control for the availability of amenities and other public goods.[16] We also see that the destination district has a significantly shorter average travel time to the nearest road. Once we control for road distance, travel time to the nearest bank is no longer significant.[17]

The univariate analysis showed that migrants on average move to destinations where they are *less* likely to find people like them in terms of language or ethnicity. The results presented in Table 4 present a different picture. Conditional on the other regressors, the ethno-caste and language proximity indices are significant with the anticipated positive sign: social proximity between the migrant and the population of the destination district is higher than in alternative destinations. The religion proximity index is not significant. Taken together, these results suggest that, conditional on material benefits from migration, migrants prefer to move to a destination where they integrate more easily – and possibly enjoy network benefits in terms of access to jobs and housing (Munshi 2003, Beaman 2006).

It is surprising that income differences are not significant once we control for geography, population, prices and amenities. This may be because we have not included individual-specific income differentials across districts. We therefore reestimate (3) with $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ as well as $\widehat{\delta}_s - \widehat{\delta}_i$. Results are shown in column 2 of Table 4. We now find a significantly positive association between $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ and the choice of destination. In column 3 we replace absolute differences in log income with relative differences. The constructed regressor, which by construction depends only on $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$, is again positive and statistically significant. Finally in

[16] Since the housing premium capitalizes both observed and unobserved location characteristics, it controls for amenities not directly included in the regression.

[17] Lall, Timmins and Yu (2009) also find that access to amenities and services (health, education, electricity) is a major determinant of migrant's destination choice in the case of Brazil.

column 4 we compute $\widehat{\delta}_s - \widehat{\delta}_i$ and $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ using answers to the question regarding the subjective adequacy of total income. Estimate coefficients are significant, but with opposite signs: only the $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ part as the anticipated positive sign.

The results presented in Table 4 indicate that, once we control for other district characteristics, individual-specific income differentials play a role in the choice of destination. But differentials in *average* income across districts are not significant in columns 1 and 2, suggesting that they do not plays a clear role in the choice of destination once we control for other factors. It is conceivable that this is due to measurement error: income is notoriously difficult to measure in a poor, primarily self-employed population. In such environment, consumption is often regarded as a more accurate measure of standards of living. To investigate this possibility, we reestimate (3) using NLSS consumption data to construct $\widehat{\delta}_s - \widehat{\delta}_i$ and $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$.

Results, shown in Table 5, are more in line with expectations. Average log consumption in the district is now significant (columns 1 and 2), albeit only at the 10% level. The coefficient of the consumption differential due to education and language $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ is strongly significant (column 2). So is the coefficient of the combined $\widehat{\delta}_s - \widehat{\delta}_i +\, (\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ variable (column 3). We also find a significant positive coefficient when the combined $\widehat{\delta}_s - \widehat{\delta}_i +\, (\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ variable is constructed using relative rather than absolute log consumption (column 4). If we include $\widehat{\delta}_s - \widehat{\delta}_i +\, (\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ computed both from absolute and relative income, only absolute consumption is significant, suggesting that it is an increase in absolute – not relative – standards of living that affects the choice of migration destination. The coefficients of other regressors are essentially unaffected.

We also estimate similar regressions using subjective consumption adequacy questions to construct $\widehat{\delta}_s - \widehat{\delta}_i$ and $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$. Results, not shown here to save space, are generally less significant. The only exception is food consumption but, as we found in column 4 of Table 4, estimated coefficients have opposite signs so the results are difficult to interpret.

4.3 Robustness checks

We conduct numerous robustness checks. We first try to better understand the difference between the univariate and multivariate results for average district income. To this effect, we estimate a series of simple regressions that include $\widehat{E[y_s^h|z^h]}$ (measured in terms of income) together with one of the additional regressors appearing in Tables 4 and 5. We find that $\widehat{E[y_s^h|z^h]}$ remains highly significant with all regressors with a single exception: as soon as the average travel time to the nearest road is included in the regression, $\widehat{E[y_s^h|z^h]}$ loses all significance. We already know from Fafchamps and Shilpi (2008) that income is strongly negatively correlated with geographical isolation. What this suggests is that once we control for geographical isolation, average district income no longer matters. Similar findings are reported for Brazil and Mexico by Timmins (2008), using a different methodology.

Next, we investigate in different ways whether our failure to find a significant coefficient for average district income in Tables 4 and 5 is due to income mis-measurement. The income regression (4) controls for household size and composition. The rationale for doing so is that larger households have more manpower and potentially more income. Household size and composition may be endogenous to the migration decision, however, – e.g., individuals who migrate to the city may opt to have a smaller household. Furthermore, migrants may derive satisfaction from the total income jointly earned by the household they head. To investigate whether this is responsible for the low income coefficients, we reestimate the income regression (4) without the log of household size and the share of adult males and females, and we replicate the analysis using the revised $\widehat{E[y_s^h|z^h]}$. The results, which are not shown here to save space, are very similar from those reported in Tables 4 and 5. Whatever the reason for non-significant average income coefficients in Table 4, it is not the inclusion household size and composition as controls.

Estimates of income and consumption levels enjoyed by households in various districts play a central role in our estimation of the relationship between income or consumption and the choice of migration destination. To check the robustness of our results, we reestimate all income and consumption regressions (4) using non-migrants only. The reason for doing so is that non-migrants represent the bulk of the population and thus $E[v^h|\text{do not migrate}] \approx E[v^h]$. The generates a loss of observations, however. As a result, $\widehat{\delta}_i$ and $\widehat{\eta}_s$ may be estimated less precisely. Regression results, not shown here to save space, are fairly similar for income. For consumption, coefficient estimates are slightly smaller in magnitude, suggesting attenuation bias. The main difference is that, in the last column of Table 5, it is relative income that is now significant, not absolute income. This suggests that this particular result is not robust, so we should refrain from drawing inference as to whether migrants move to districts where their relative or absolute income is higher.

Dropping migrants does not control for possible self-selection: if more talented individuals migrate, remaining households may be less productive. As a result, they may earn less than migrants in the same location. To correct for the self-selection of non-migrants we need variables that affect the decision to migrate but are unlikely to affect income. Family background variables such as the education and occupation of the father may serve this purpose because they affect the ability of the migrant's father to help finance the cost of migration. Given that most migrants migrate early in their adult life, it is reasonable to expect that parental influences play a role in the decision to migrate – and in the financing of migration costs. We use the education and occupation of the father to construct two selection correction terms for the income regressions

– one selection term for migrants, and one for non-migrants (Wooldridge, p. 631):

$$\begin{aligned} y_s^k &= \delta_s + \alpha(a_s^k - \overline{a}) + \beta_s(E_s^k - \overline{E}_s) + \chi_s(H_s^k - \overline{H}_s) \\ &\quad + \rho_1 m \frac{\phi(z\widehat{\theta})}{\Phi(z\widehat{\theta})} + \rho_2(1-m)\frac{\phi(z\widehat{\theta})}{1-\Phi(z\widehat{\theta})} + v_s^k \end{aligned} \tag{6}$$

where $\phi(z\widehat{\theta})$ and $\Phi(z\widehat{\theta})$ are the normal density function and cumulative distribution from the selection regression of migrant status m on determinants z.

The selection regression is shown in Table A2 in Appendix. Other variables are the same as those appearing in the income and consumption regressions (4). We see that family background variables are significant. Using this selection regression we construct the two Mills ratio shown in equation (6), one for migrants and one for non-migrants, and we reestimate the income and consumption regressions with these additional regressors, obtain corrected $\widehat{\delta}_s$ and $\widehat{\eta}_s$ estimates, and reestimate the destination choice regressions. Results are very similar to those reported in Tables 4 and 5. They are omitted here to save space. With the selection correction, we again find that it is absolute differences in consumption that matter, not relative differences.

We also experimented with an alternative selection correction for migration suggested by Dahl (2002). In this approach, observed propensities to migrate from and to each location are used as selection correction terms in the income regressions, in lieu of the Mills ratios in (6). Because these propensities vary only by district, they drop out of the income regressions given the inclusion of district fixed effects. To circumvent this problem, we interact migration propensities with the education level of each worker so that we obtain variation within districts. The idea is that educated workers are more likely to migrate and therefore more subject to self-selection. Using this approach we reestimate all $\widehat{\delta}_s$ and $\widehat{\eta}_s$ and reestimate the destination choice regressions. Results, not shown here to save space, are again very similar to those reported in Tables 4 and

5. In this case, when we include both absolute and relative consumption differentials, neither is statistically significant, again confirming that in this particular respect our results are not robust.

When constructing $\widehat{E[y_s^h|z^h]}$ we implicitly assume that migrants are well informed about incomes in all potential destinations. But it is possible that they are better informed about certain destinations, for instance, destinations chosen by migrants from their district in the past. Failing to control for this possibility may lead to an attenuation bias in the income coefficient, as suggested by the work of McKenzie et al. (2004) for Tonga. To investigate this possibility, we interact the income variable with a proxy for the availability of income information. If migrants only respond to income differences for those districts on which they have more accurate information, the coefficient of the interacted term should be significant even if the uninteracted term is not.

As proxy for the availability of information, we use the proportion P_{is} of adult males who migrated more than 5 years ago (that is *before* the migrants themselves) from the district of origin i to each of the districts of destination s. To avoid spurious inference, P_{is} is also included as a separate regressor. We find that P_{is} is strongly significant, suggesting persistence in migration patterns. But the coefficient of the interacted term is either not significant, or significant with the wrong (negative) sign. The inclusion of P_{is} as separate regressor reduces the magnitude of nearly all coefficients, and drives the coefficient of population density below standard levels of significance. The coefficients of income and absolute consumption become non significant; only the coefficient of relative consumption remains significant at the 10% level. From this we conclude that the small or non-significant coefficients we have found on income and consumption are probably not due to insufficient information about potential destinations.

So far we have estimated the migration regression using all male heads of household who migrated for work reasons. There often are multiple reasons for migration, so that the one reason listed as response to a census question may not capture all those who migrated at least in part for work reasons. Furthermore, it is of interest to know whether other male migrations follow a similar pattern. To investigate this, we reestimate the migration regression using all male adults who migrated, for any reason, in the five years preceding the census.[18] Results, shown in Tables 6 and 7, are to be compared with Tables 4 and 5. in all cases, the coefficients of income and consumption variables are larger in magnitude and, in a couple cases, more statistically significant. Other coefficients are also similar.

As a final robustness check, we reestimate the model using migrant data from the NLSS 2002/03. The number of migrants is significantly smaller, so results may be less precise. The advantage of this approach is that it serves as cross-validation. Results are presented in Tables 8 and 9. Table 8 should be compared with Table 4, and Table 9 with Table 5.

Comparing Tables 8 and 4, we again find that anticipated income, whether absolute or relative, is either non-significant or negative. Most of our other results obtain. Exceptions include the rice price – which now appears with the wrong sign – and elevation and population density – which are no longer significant. Comparing Tables 9 and 5, we find that in the smaller NLSS 2002/3 dataset none of the anticipated consumption variables is statistically significant. Other results are as before.

There are other robustness checks we wished we could undertake, but are not possible due to data limitations. Expected income $\widehat{E[y_s^h|z^h]}$ may evolve over time, for instance because of aggregate shocks or inflow of migrants. Our $\widehat{\delta}_s - \widehat{\delta}_i$ and $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ estimates are based in 1995 data and are therefore reasonably accurate for those who migrated in, say, 1996. But they are less accurate estimates of $\widehat{E[y_s^h|z^h]}$ for those who migrated later, causing a possible attenuation

[18]In addition to work, the census records three reasons for migrating: 'study', 'marriage' and 'other'. For male adults, these account for 11%, 0%, and 20% of responses, respectively.

bias. A possible solution would have been to reestimate the model with only those who migrated immediately after 1995. Unfortunately, the census data do not record the date of migration, so this cannot be done.

4.4 Magnitude

Our discussion so far has focused on statistical significance. We are also interested in the magnitude of income coefficients relative to other regressors. It is for instance conceivable that $\widehat{\delta}_s - \widehat{\delta}_i$ and $(\widehat{\eta}_s - \widehat{\eta}_i)\, z^h$ have a coefficient that is large in magnitude, even though it is not always strongly significant.

To investigate this possibility, we multiply the coefficients estimated in Tables 4 and 5 by the standard deviation of their respective regressors. This gives an idea of the relative magnitude of a one standard deviation change in each regressor, keeping other regressors unchanged. For regressors other than income and consumption, we average over the various coefficient values reported in Tables 4 and 5. This is crude, but given that coefficient estimates are very similar across regressions and that we are only interested in orders of magnitude, it is sufficient for our purpose.

Calculations are summarized in Table 10, using the standard deviations reported in column 1. The larger the value reported in the last column, the more influence the regressor has on the choice of a destination district.

The most important regressors in terms of magnitude are distance (by far the strongest), the price of rice, travel time to the nearest road, the housing premium, and language similarity. Consumption variables have an effect on migration destination that is smaller in magnitude: a one standard deviation increase in anticipated relative consumption, for instance, has an effect on destination that corresponds to a third of the effect of a one standard deviation in the housing premium – and one sixth of a one standard deviation in the log of rice price. The magnitude of the coefficients of income variables is negligible in comparison. These calculations confirm our earlier assessment.

5 Conclusion

Combining data from a household survey and an 11% census of the population, we have estimated destination choice regressions for Nepalese internal migrants. Results show that distance, population density, social proximity, and access to amenities are strongly correlated with migrants' choice of destination. These results confirm earlier work on the factors affecting the subjective welfare cost of isolation (Fafchamps and Shilpi, 2008).

Differentials in income and consumption expenditures across potential destination districts are significant in univariate comparisons but are found to be less important than expected once we control for covariates. Average district income is not statistically significant in any of the regression. Average district consumption is only marginally significant. What matter are income and consumption differentials across possible destinations that are driven by differences in education and language. These results are robust to different specifications and datasets. We also find that, in line with what is assumed in most economic models, migrants appear to respond to gains in absolute – not relative – consumption. This latter result, however, is not robust: depending on what sample we use to construct income predictions, the finding is either reversed or disappears.

The analysis reported here is based on one critical maintained assumption, namely, that average income and consumption levels obtained by district residents in the recent past are reasonable proxies for the anticipations of subsequent migrants. There are several reasons why this need not be the case. For instance, migrants may select their destination based on differentials in returns to unobservable characteristics. This would naturally lead to a downward bias in the income coefficient. Undoubtedly it would be better to have direct measurements of what migrants actually anticipate to earn and consume in different districts upon migration, along the lines of the data that McKenzie et al. (2006) collected for migrants from Tonga to New Zealand. Unfortunately such data is difficult to collect for all possible destinations. Our results ultimately rest on the assumption that the biases caused by measurement error are not so large as to invalidate all inference.

With this important caveat, our results suggest that an urban environment and access to amenities are key considerations when internal migrants choose a migration destination. Anticipated income and consumption expenditures relative to others in the district of destination play a significant but relatively secondary role. This does not imply that income differentials do not affect the decision to migrate, an issue that we have sidestepped by focusing on the choice of destination conditional on migrating.

It is difficult to draw causal inference from observational data. This study is no exception. The results presented here are nevertheless sufficiently suggestive to cast doubt on the theory that the choice of migration destination is driven primarily or exclusively by income differentials. Other factors seem to play a strong – and probably more important – role.

References

1. Adams, Richard, *Remittances, Investment, and Rural Asset Accumulation in Pakistan*, International Food Policy Research Institute, Washington, D.C., 1997

2. Akee, Randall K.Q., "Deciphering Immigrant Self-Selection: New Evidence from a Developing Country," Kennedy School of Government, Harvard University, Cambridge Mass., 2006 (mimeograph)

3. Bayer, Patrick J., Shakeeb Khan, and Christopher Timmins, "Nonparametric Identification and Estimation in a Generalized Roy Model," NBER Working Paper No. W13949. March 2008

4. Bayoh Isaac, Elena G. Irwin and Timothy Haab, "Determinants of Residential Location Choice: How Important Are Local Public Goods in Attracting Homeowners to Central City Locations?," *Journal of Regional Science*, 46(1), 97-120, February 2006.

5. Beaman, Lori A.,"Social Networks and the Dynamics of Labor Market Outcomes: Evidence from Refugees Resettled in the US," Department of Economics, Yale University, New Haven, 2006 (mimeograph)

6. Borjas, George J., "The Economics of Immigration," *Journal of Economic Literature*, 32(4), 1667-1717, December 1994

7. Borjas, George J., Native Internal Migration and the Labor Market Impact of Immigration, *Journal of Human Resources*, 41(2), 221-58, Spring 2006

8. Borjas, George J., Stephen G. Bronars and Stephen J. Trejo, "Assimilation and the Earnings of Young Internal Migrants," *Review of Economics and Statistics*, 74(1), 170-75, February 1992

9. Carrington William J., Enrica Detragiache and Tara Vishwanath, "Migration with Endogenous Moving Costs," *American Economic Review*, 86(4), 909-30, September 1996

10. Chiquiar, Daniel and Gordon H. Hanson, "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 113(2), 239-81, April 2005

11. Dahl, Gordon B., "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets," *Econometrica*, 70(6), 2367-2420, November 2002

12. de Brauw, Alan and John Giles, "Migrant Opportunity and the Educational Attainment of Youth in Rural China," IFPRI and Department of Economics, Michigan State University, Washington DC, September 2006 (mimeograph)

13. de la Briere, Benedicte, "Household Behavior toward Soil Conservation and Remittances in the Dominican Republic," University of California, Berkeley, 1996 (unpublished PhD thesis)

14. Fafchamps, Marcel and Forhad Shilpi, "The Spatial Division of Labor in Nepal," *Journal of Development Studies*, 39(6), 23-66, 2003.

15. Fafchamps, Marcel and Forhad Shilpi, "Cities and Specialization: Evidence from South Asia," *Economic Journal*, 115(503), 477-504, April 2005.

16. Fafchamps, Marcel and Forhad Shilpi, "Subjective Welfare, Isolation, and Relative Consumption," *Journal of Development Economics*, 2008 (forthcoming)

17. Fafchamps, Marcel, and Forhad Shilpi, "Isolation and Subjective Welfare: Evidence from South Asia," *Economic Development and Cultural Change*, 2009 (forthcoming)

18. Fujita, Masahisa, Paul Krugman and Anthony J. Venables, *The Spatial Economy: Cities, Regions, and International Trade*, MIT Press, Cambridge and London, 1999

19. Gabriel, Paul E. and Suzanne Schmitz, "Favorable Self-Selection and the Internal Migration of Young White Males in the United States," *Journal of Human Resources*, 30(3), 460-71, Summer 1995

20. Greenwood, Michael J., "Research on Internal Migration in the United States: A Survey," *Journal of Economic Literature*, 13(2), 397-433, June 1975

21. Harris, John and M. Todaro, "Migration, Unemployment and Development: A Two-Sector Analysis," *Amer. Econ. Rev.*, 60, 126-142, 1970

22. Henderson, J. Vernon, *Urban Development: Theory*, Fact, and Illusion, Oxford University Press, New York, 1988

23. Hojvat-Gallin, Joshua, "Net Migration and State Labor Market Dynamics," *Journal of Labor Economics*, 22(1), 1-21, January 2004

24. Hole, Arne Risa (2006), "Calculating Murphy–Topel variance estimates in Stata: A simplified procedure", *The Stata Journal*, 6, Number 4, pp. 521–529

25. Jackman, Richard and Savvas Savouri, "Regional Migration in Britain: An Analysis of Gross Flows Using NHS Central Register Data," *Economic Journal*, 102(415), 1433-50, November 1992.

26. Lall, S., C. Timmins and S. Yu, "Connecting Lagging and Leading Regions: The Role of Labor Mobility," *Brookings-Wharton Papers on Urban Affairs*, pp.151-174, The Brookings Institution, Washington, D.C., 2009.

27. Lokshin, Misha, Mikhail Bontch-Osmolovski and Elena Glinskaya, "Work Migration and Poverty in Nepal," World Bank, mimeograph, 2007.

28. Mansuri, Ghazala, "Migration, School Attainment and Child Labor: Evidence from Rural Pakistan," DECRG, The World Bank, Washington DC, April 2006 (mimeograph)

29. McCall, B.P and J.J. McCall, A Sequential Study of Migration and Job Search, *Journal of Labor Economics*, Oct. 1987, 5(4), 452-76

30. McCormick, Barry and Jacqueline Wahba, "An Econometric Model of Temporary International Migration and Entrepreneurship," Department of Economics, University of Southampton, Southampton, 2001 (mimeograph)

31. McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior", in P. Zarembka (ed.), *Frontiers in Econometrics*, Academic Press, New York, pp.105-42, 1974

32. McKenzie, David, John Gibson and Steven Stillman, "How Important Is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration," *IZA Discussion Papers 2087*, Institute for the Study of Labor, 2006

33. Mesnard, Alice and Martin Ravallion, "Is Inequality Bad for Business? A Nonlinear Empirical Model of Entrepreneurship," ARQADQ, University of Toulouse and DECRG, The World Bank, Toulouse and Washington D.C., 2001 (mimeograph)

34. Munshi, Kaivan and Mark Rosenzweig, "Why Is Mobility in India so Low? Social Insurance, Inequality and Growth," Economic Growth Center, Yale University, New Haven, July 2005 (mimeograph)

35. Munshi, Kaivan, "Networks in the Modern Economy: Mexican Migrants in the US Labor Market," *Quarterly Journal of Economics*, 118(2), 549-99, May 2003

36. Roy, Andrew D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3: 135-46, 1951

37. Sorensen, Todd, Price Fishback, Samuel Allen, and Shawn Kantor, "Migration Creation, Diversion, and Retention: New Deal Grants and Migration: 1935-1940", IZA DP No. 3060, September 2007

38. Stark,Oded and Robert E. Lucas, “Migration, Remittances, and the Family,” *Economic Development and Cultural Change*, 36, no.3, 465-481, Apr. 1988

39. Stark, Oded and J. Edward Taylor, ”Migration Incentives, Migration Types: The Role of Relative Deprivation,” *Economic Journal*, 101(408), 1163-78, September 1991

40. Timmins, Christopher, “Estimating Spatial Differences in the Brazilian Cost of Living with Household Location Choices,” *Journal of Development Economics*, 80(1): 59-83, June 2006

41. Train, Kenneth E., *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, 2003

42. Uhlig, Harald, “Regional Labor Markets, Network Externalities and Migration: The Case of German Reunification,” *American Economic Review*, 96(2), 383-87, May 2006.

43. Winters, Paul, Alain de Janvry and Elisabeth Sadoulet, “Family and Community Networks in Mexico-U.S. Migration,” *Journal of Human Resources*, 36(1), 159-84, Winter 2001

44. Woodruff, Christopher, “Remittances and Microenterprises in Mexico,” Graduate School of International Relations and Pacific Studies, UCSD, La Jolla, August 2001 (mimeograph)

45. Wooldridge, Jeffrey M., *Econometric Analysis of Cross-Section and Panel Data*, MIT Press, Boston, 2002

Appendix

In the migration regression income and consumption are predicted regressors. The reader may worry that this affects standard errors. To correct standard errors for the use of predicted regressors, Hole (2006) offers an implementation of the Murphy-Topel (hereafter MT) correction for maximum likelihood estimators in Stata. The Murphy–Topel estimate of variance for a two-step model is given by:

$$\widehat{V}_2 + \widehat{V}_2 \left(\widehat{C}\widehat{V}_1\widehat{C}' - \widehat{R}\widehat{V}_1\widehat{C}' - \widehat{C}\widehat{V}_1\widehat{R}' \right) \widehat{V}_2$$

where $\widehat{V}_1$ and $\widehat{V}_2$ are the estimated covariance matrices from model 1 (the predicting regression) and model 2 (the regression using predicted regressors from model 1). The assumed data generating process for Models 1 and 2 is of the following form:

$$\begin{aligned} y_1 &= \lambda_1(X_1\theta_1) \\ y_2 &= \lambda_2(X_2\theta_2 + y_1\beta) \end{aligned}$$

where $\lambda_1(.)$ and $\lambda_2(.)$ are suitable maximum likelihood models, e.g., logit, probit, or linear regression. Further let:

$$\begin{aligned} \widehat{C} &= (p \times q) \text{ matrix given by } \left\{ \sum_{i=1}^{n} \left(\frac{\partial \ln f_{i2}}{\partial \widehat{\theta}_2} \right) \left(\frac{\partial \ln f_{i2}}{\partial \widehat{\theta}_1'} \right) \right\} \\ \widehat{R} &= (p \times q) \text{ matrix given by } \left\{ \sum_{i=1}^{n} \left(\frac{\partial \ln f_{i2}}{\partial \widehat{\theta}_2} \right) \left(\frac{\partial \ln f_{i1}}{\partial \widehat{\theta}_1'} \right) \right\} \end{aligned}$$

where q is the number of regressors in model 1, p is the number of regressors in model 2, and f_{i1} and f_{i2} are observation i's contribution to the likelihood function of models 1 and 2, respectively. – i.e., the scores. In our case, $\widehat{R}$ is zero by construction since the two samples are distinct. The correction term thus boils down to $\widehat{V}_2 \left(\widehat{C}\widehat{V}_1\widehat{C}' \right) \widehat{V}_2$.

Implementing this correction, however, raises many difficuties. The first difficulty is that we need to merge the migrant data from the census with the income data from the NLSS in order to calculate $\widehat{C}$. This results in a very large dataset (over 1 gigabite) which not only makes estimation time-consuming but also precludes any attempt at bootstrapping standard errors.

Next, the likelihood model we estimate in the paper is a conditional (fixed effect) logit model. This model does not nicely fit the data generating process assumed above because the likelihood function does not take single observations one by one, but rather as a group. Hence derivatives $\frac{\partial \ln f_{i2}}{\partial \widehat{\theta}_2}$ and $\frac{\partial \ln f_{i2}}{\partial \widehat{\theta}_1'}$, which are necessary to compute $\widehat{C}$, are not well defined since f_{i2} is not well defined.

We try overcoming this difficulty in several ways. We first estimate a standard logit model with migrant-specific dummies. Given that we have 68 observations[19] per migrant, the incidental parameter bias is likely very small. The problem is that the number of migrant dummies is very large (i.e., more than 16,000). Since the coefficients of migrant dummies are part of the parameter vector θ_2, they must enter the construction of $\widehat{C}$. Unfortunately this exceeds our matrix manipulation capacity. To sidestep the difficulty we turn to a linear probability model that we estimate in deviation to the migrant-specific mean.[20] Doing so keeps the number of regressors to a minimum, in which case calculating $\widehat{C}$ is straightforward. But by construction the average of the dependent variable is 1/68 – a migrant moves to one of 68 districts. A linear probability model is hardly suited to such a small average value for the dependent variable. Logit would be a definite improvement, if only the MT correction could be calculated.

The second difficulty has to do with the clustered standard errors. Throughout our analysis, we cluster standard errors by the district of origin of the migrant. This means clustering in excess of one million observations into 68 clusters. Unfortunately, the MT correction method applies to maximum likelihood covariance matrices, not, to the best of our knowledge, to clustered covariance matrices. Hence the MT correction and clustering correction are not compatible. Not clustering by district of origin results in dramatically smaller standard errors, which is the

[19] Or more, dependent on the regressor set.

[20] In regressions of fixed effect models in demeaned form, a degree of freedom correction should be applied to standard errors. In our case, because the number of observation per migrant is large, this correction is very small, e.g., of the order of 1.02 or less.

opposite of the desired effect. Keeping clustered standard errors is definitely better.

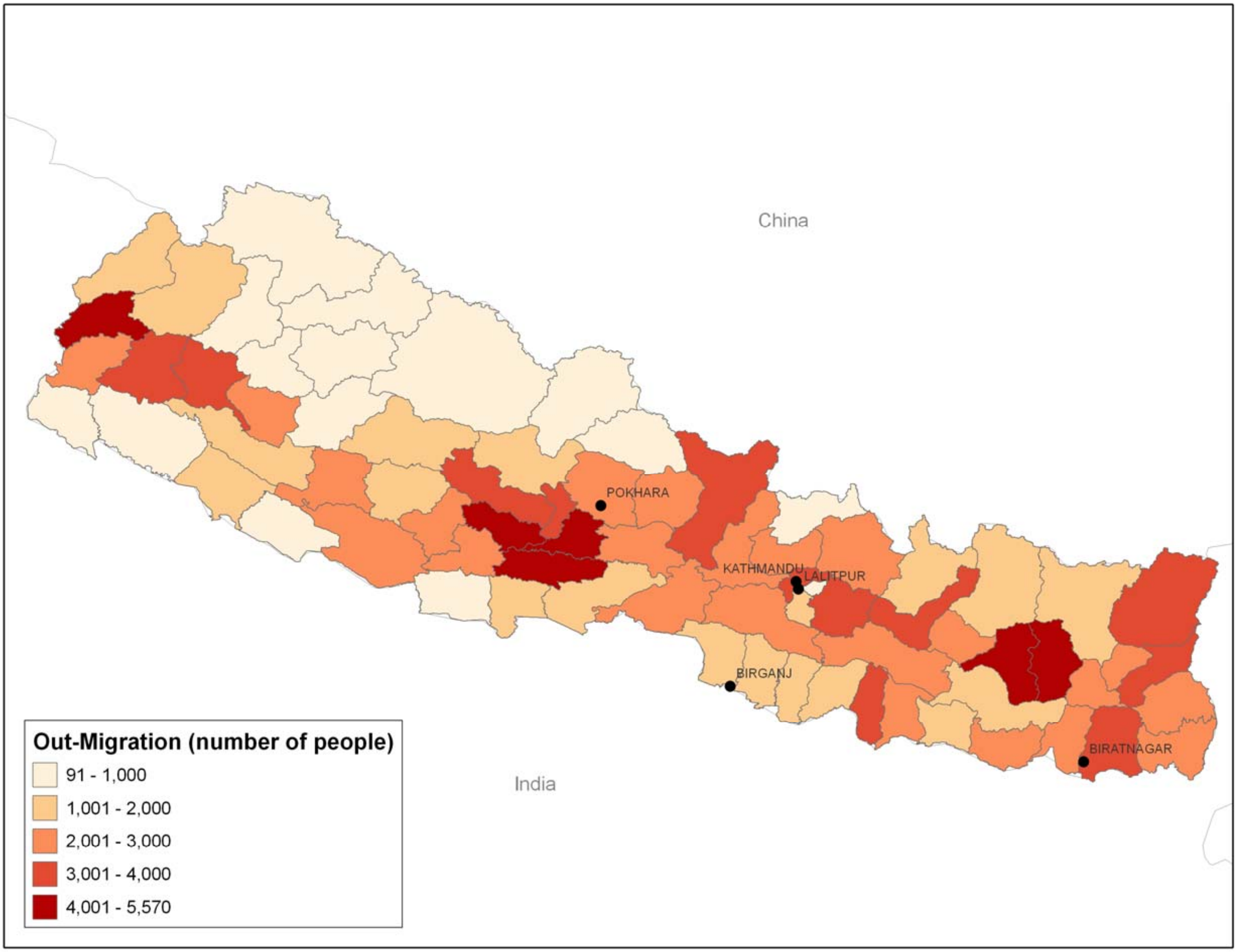
A third difficulty has to do with the fact that our predicted regressors do not fit Model 1 above. This is particularly true for $E[\widehat{y_s^h - y_i^h}|z^h]$, which includes two predicted y_1's and thus cannot be obtained from a single regression of the form $y_1 = \lambda_1(X_1\theta_1)$. To bypass this problem, we estimate a modified migration model in which the income variable takes the form $E[\widehat{y_s^h|}z^h]$. This is, however, not the model we believe drives the data generation process, and this creates a specification error. Moreover, in our estimation we divide $E[\widehat{y_s^h - y_i^h}|z^h]$ into two parts, $\widehat{\delta}_s$ and $\widehat{\eta}_s z^h \equiv \widehat{\beta}_s(E_s^h - \overline{E}_s) + \widehat{\chi}_s(H_s^h - \overline{H}_s)$. This cannot be accommodated by the MT correction.

A fourth difficulty has to do with the fact that the MT correction assumes a data generating process that is different from the one we envisage. To illustrate the problem, in the MT correction the data generating process, in linear form, is assumed to follow:

$$\begin{aligned} y_1 &= X_1\theta_1 + e_1 \\ y_2 &= X_2\theta_2 + y_1\beta + e_2 \end{aligned}$$

We see that y_1 enters the second equation directly. This is different from our model where y_2, the decision to migrate, is affected by $E[y_1|X_1] = X_1\theta_1$, the expected – or average – y_1, but never by the realized value of y_1 for a particular individual. This is because migrants do not know the realized value of y_1 before they migrate. In the migration equation, $X_1\widehat{\theta}_1$ is used in lieu of $X_1\theta_1$, not in lieu of y_1, so there is some sample variation in the predicted regressor. Clearly, the prediction variance of sample average $X_1\widehat{\theta}_1$ is much smaller than the variance of individual realizations y_1. Yet the covariance matrix that enters the correction term $\widehat{V}_2\left(\widehat{C}\widehat{V}_1\widehat{C}'\right)\widehat{V}_2$ is $\widehat{V}_1$, which is computed on the basis of the (conditional) variation of y_1 and thus overstates the variance of $X_1\widehat{\theta}_1$. It is in principle possible to correct $\widehat{V}_1$ for this, e.g., by dividing by the sample size in the y_1 regression. When we do so, the magnitude of $\widehat{V}_2\left(\widehat{C}\widehat{V}_1\widehat{C}'\right)\widehat{V}_2$ shrinks, so much so that the correction is no longer noticeable.

In order to derive an upper bound on the MT correction, we apply the MT correction to a linear probability model in demeaned form, to net out fixed effects. Predicted regressor $\widehat{y}_s^h|z^h$ is obtained by OLS. Standard errors are not clustered by birth district, which of course means that they are much smaller than those reported in the paper. The $\widehat{V}_2\left(\widehat{C}\widehat{V}_1\widehat{C}'\right)\widehat{V}_2$ correction increases the magnitude of standard errors, as expected. But the increase is negligible for all regressors except $\widehat{y}_s^h|z^h$, for which it is small. Since estimates of expected income and consumption effects reported in the paper are, if anything, small in magnitude and not always significant, correcting standard errors for the use of predicted regressors – if it were possible – would not affect our qualitative conclusions.


China
India
POKHARA
KATHMANDU
LALITPUR
BIRGANJ
BIRATNAGAR
Out-Migration (number of people)
91 - 1,000
1,001 - 2,000
2,001 - 3,000
3,001 - 4,000
4,001 - 5,570

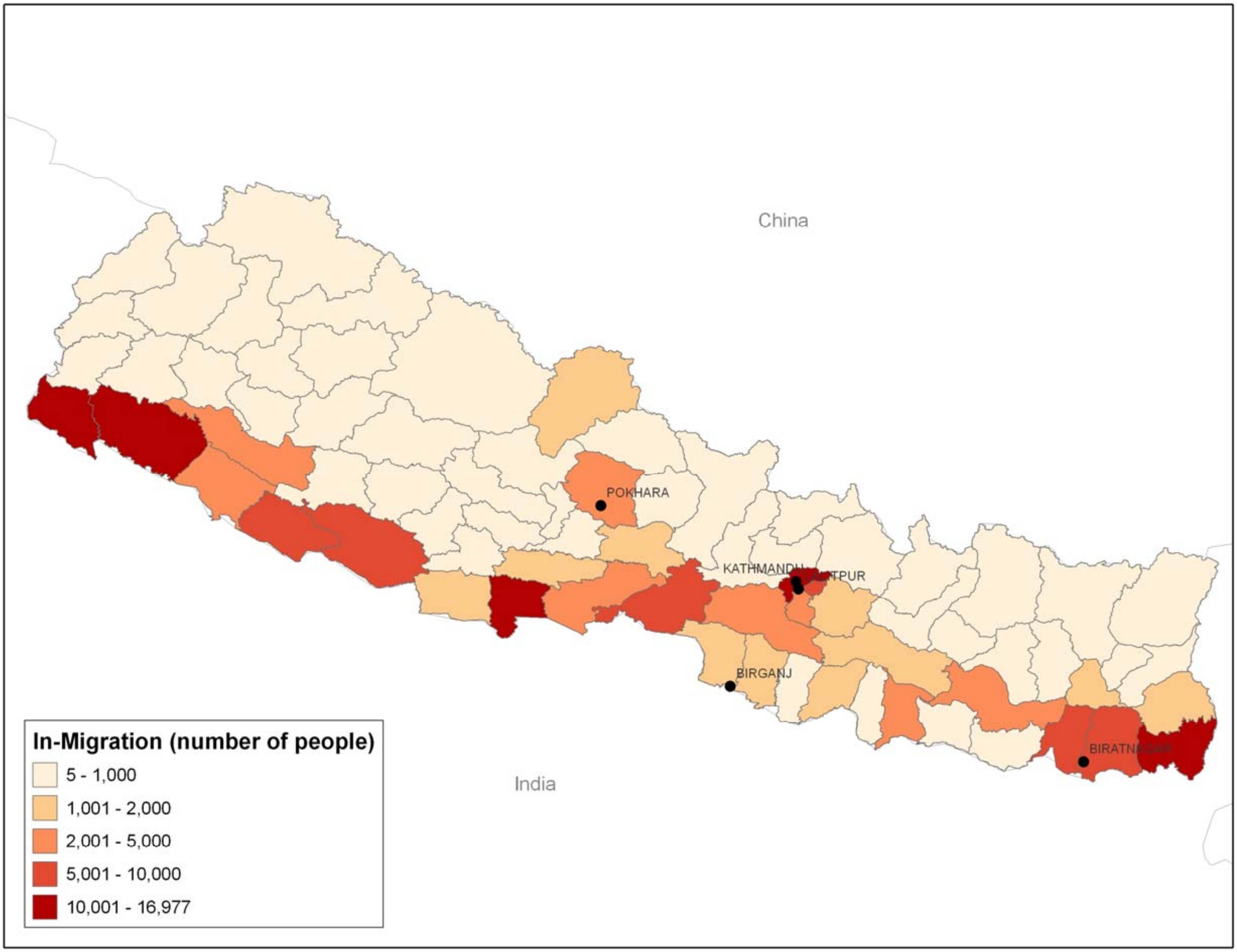

China
POKHARA
KATHMANDU
BIRGANJ
India
In-Migration (number of people)
5 - 1,000
1,001 - 2,000
2,001 - 5,000
5,001 - 10,000
10,001 - 16,977

Table 1: Summary Statistics

	Work Migrant	Adult Male
Age		
Mean	35.3	43.9
Standard Deviation	10.6	13.9
Education (years)		
Mean	8.0	3.0
Standard Deviation	5.0	4.3
Ethnicity (Percentage)		
Brahmin	34.5	11.7
Chhetri	21.5	15.6
Newar	7.4	7.9
Tharu	3.1	6.7
Magar	6.1	6.0
Tamang	4.2	5.9
Other	23.2	46.2
Language (Percentage)		
Nepali	73.9	45.3
Maithili	6.2	13.2
Bhojpuri	1.3	7.3
Newar	4.4	6.1
Tharu	2.0	5.8
Tamang	3.7	5.5
Other	8.5	16.8
Religion (Percentage)		
Hindu	89.6	81.0
Buddheism	7.2	11.7
Muslim	0.9	3.7
Kirat	1.5	2.9
Christian	0.6	0.3
Others	0.2	0.4

Table 2. Income and Consumption regressions using NLSS 95/96

		Absolute		Relative		Consumption adequacy index					
Dummies and human capital:		**ln(income)**	**ln(cons.)**	**ln(income)**	**ln(cons.)**	**Food**	**Clothing**	**Housing**	**Healthcare**	**Schooling**	**Income**
District dummies:	mean	10.289	10.325	n.a.	n.a.	1.496	1.357	1.404	1.412	1.446	1.251
	st.dev.	*0.340*	*0.340*			*0.213*	*0.196*	*0.184*	*0.198*	*0.200*	*0.156*
Education coefficients:	mean	0.201	0.184	0.019	0.018	0.134	0.057	0.114	0.079	0.052	0.067
	st.dev.	*0.200*	*0.133*	*0.020*	*0.013*	*0.123*	*0.101*	*0.106*	*0.115*	*0.120*	*0.094*
Language coefficients:	mean	-0.053	-0.030	-0.005	-0.003	0.067	0.050	0.032	-0.008	0.014	-0.020
	st.dev.	*0.624*	*0.308*	*0.060*	*0.029*	*0.319*	*0.258*	*0.301*	*0.252*	*0.332*	*0.240*
Controls		coef	coef	coef	coef	coef	coef	coef	coef	coef	coef
Age of household head	coef.	0.010	0.011	0.001	0.001	0.001	0.005	0.004	0.002	-0.005	0.003
	t-stat.	*1.810*	*2.770*	*1.711*	*2.656*	*0.260*	*1.350*	*0.879*	*0.511*	*-1.207*	*0.986*
Age squared/10000	coef.	-0.007	-0.368	0.004	-0.031	0.382	-0.194	0.134	0.128	0.833	-0.001
	t-stat.	*-0.011*	*-0.896*	*0.070*	*-0.793*	*0.886*	*-0.483*	*0.325*	*0.321*	*1.807*	*-0.002*
Log(household size)	coef.	0.982	0.883	0.095	0.086	0.148	0.018	0.062	0.041	0.023	0.070
	t-stat.	*21.998*	*24.501*	*21.847*	*24.414*	*4.641*	*0.620*	*2.041*	*1.362*	*0.668*	*2.318*
Share of adult males	coef.	0.655	0.359	0.064	0.035	0.247	0.123	0.163	0.168	0.038	0.248
	t-stat.	*5.254*	*4.202*	*5.268*	*4.237*	*3.031*	*1.675*	*2.155*	*2.236*	*0.401*	*3.322*
Share of adult females	coef.	0.632	0.411	0.060	0.039	0.296	0.158	0.144	0.119	0.259	0.202
	t-stat.	*5.384*	*4.726*	*5.316*	*4.674*	*3.289*	*1.975*	*1.771*	*1.526*	*2.585*	*2.669*

Each column corresponds to a different regression. The estimator is weighted least squares, using sampling population weights.

Table 3. Comparing the actual destination to alternative destinations

All figures are relative to the district of origin	Actual Destination	Mean in Alt. Destin.	Diff. in mean t-stat	
Income and consumption				
Average income (log)	0.195	-0.037	-61.774	***
Differential in log income due to education and language	-0.004	0.004	1.940	*
Average consumption expenditures (log)	0.075	-0.047	-33.587	***
Differential in log consumption due to education and language	0.009	0.012	0.907	
Relative log income	-0.001	0.000	2.352	**
Relative log consumption	0.001	0.001	1.325	
Subjective consumption adequacy				
Average consumption adequacy index: total income	0.094	-0.010	-44.058	***
Differential due to education and language: total income	-0.024	0.010	13.100	***
Average consumption adequacy index: food	0.076	-0.019	-42.957	***
Differential due to education and language: food	0.010	-0.020	-15.012	***
Average consumption adequacy index: clothing	0.070	-0.028	-47.461	***
Differential due to education and language: clothing	0.011	0.000	-5.412	***
Average consumption adequacy index: housing	0.081	-0.022	-46.650	***
Differential due to education and language: housing	0.008	0.003	-2.229	**
Average consumption adequacy index: health care	0.094	-0.011	-46.011	***
Differential due to education and language: health care	0.016	-0.012	-11.462	***
Average consumption adequacy index: children schooling	0.054	-0.008	-35.556	***
Differential due to education and language: children schooling	0.010	-0.014	-13.540	***
Prices and amenities				
Log of rice price	-0.088	0.021	47.577	***
Housing price premium (log)	0.377	0.210	-12.188	***
Travel time to nearest paved road	-0.745	0.104	79.624	***
Travel time to nearest bank	-0.373	0.092	71.245	***
Population and distance				
Population density	0.281	-0.033	-86.063	***
Log(population)	0.329	-0.208	-73.980	***
Elevation in meters	-0.316	0.167	56.973	***
Ethno-caste similarity index	-0.042	-0.060	-13.660	***
Language similarity index	-0.123	-0.101	7.429	***
Religion similarity index	0.008	-0.016	-13.706	***
Distance in '000 Km	0.116	0.278	115.060	***

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

Table 4. Income and the choice of migration destination -- using all male work migrants

District difference in:	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat
Income								
Average log income	0.074	0.447	0.095	0.570				
Differential in log income due to education and language			0.167***	2.652				
Relative log income controlling for education and language					1.668**	2.537		
Average consumption adequacy index: total income							-1.207***	-5.254
Differential due to education and language total income							0.510***	4.523
Prices and amenities								
Log of rice price	-4.265***	-6.262	-4.187***	-6.378	-4.184***	-6.368	-4.300***	-6.346
Housing price premium (log)	0.414***	6.729	0.408***	6.552	0.412***	6.675	0.431***	7.179
Travel time to nearest paved road	-0.840***	-9.427	-0.832***	-9.406	-0.846***	-9.954	-0.850***	-10.195
Travel time to nearest bank	0.322*	1.749	0.308*	1.650	0.293	1.588	0.029	0.138
Elevation in '000 meters	-0.389**	-2.209	-0.419**	-2.462	-0.395**	-2.370	-0.168	-0.920
Population								
Population density	0.883***	6.835	0.901***	6.800	0.915***	6.962	0.866***	6.911
Log(population)	0.344**	2.369	0.319**	2.240	0.334**	2.482	0.429***	3.219
Ethno-caste similarity index	0.853***	3.082	0.820***	2.918	0.818***	2.914	0.780***	2.854
Language similarity index	1.388***	7.145	1.419***	7.120	1.428***	7.162	1.347***	7.323
Religion similarity index	-0.324	-1.042	-0.325	-1.050	-0.315	-1.024	-0.097	-0.314
Distance								
Distance in '00 Km	-9.951***	-13.531	-9.945***	-13.504	-9.932***	-13.491	-10.007***	-14.111
Log-Likelihood	-47,987.92		-47,970.50		-47,973.28		-47,839.05	
Number of observations	1,072,804		1,072,804		1,072,804		1,072,804	
Pseudo R2	0.287		0.287		0.287		0.289	

The estimator is Fixed Effect Conditional Logit. Standard errors are corrected for clustering across district of origin. *** p<0.01, ** p<0.05, * p<0.1

Table 5. Consumption and the choice of migration destination -- using all male work migrants

Consumption	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat
Average consumption expenditures (log)	0.318*	1.687	0.331*	1.753						
Log consumption differential due to education and language			0.533***	6.330						
Combined average and differential					0.450***	4.136			0.333*	1.791
Relative log consumption controlling for education and language							5.439***	6.376	2.082	1.201
Prices and amenities										
Log of rice price	-4.486***	-6.555	-4.492***	-6.809	-4.568***	-6.878	-4.265***	-6.443	-4.493***	-6.817
Housing price premium (log)	0.424***	7.153	0.428***	7.314	0.430***	7.467	0.419***	7.011	0.427***	7.307
Travel time to nearest paved road	-0.767***	-8.122	-0.770***	-8.297	-0.739***	-8.587	-0.861***	-9.958	-0.771***	-8.299
Travel time to nearest bank	0.291*	1.649	0.276	1.569	0.272	1.552	0.296	1.628	0.276	1.567
Elevation in '000 meters	-0.454***	-2.670	-0.512***	-3.149	-0.539***	-3.291	-0.418***	-2.585	-0.512***	-3.142
Population										
Population density	0.876***	6.614	0.947***	7.316	0.929***	7.472	0.963***	7.641	0.947***	7.319
Log(population)	0.322**	2.361	0.268**	2.099	0.262**	2.008	0.306**	2.380	0.267**	2.093
Ethno-caste similarity index	0.881***	3.204	0.749***	2.607	0.782***	2.728	0.724**	2.508	0.750***	2.611
Language similarity index	1.362***	7.008	1.486***	8.037	1.455***	7.853	1.520***	8.152	1.487***	8.039
Religion similarity index	-0.282	-0.908	-0.443	-1.416	-0.403	-1.310	-0.473	-1.519	-0.442	-1.415
Distance										
Distance in '00 Km	-9.974***	-13.558	-9.932***	-13.425	-9.950***	-13.355	-9.895***	-13.380	-9.931***	-13.426
Log-Likelihood	-47,966.66		-47,880.35		-47,885.72		-47,905.87		-47,880.27	
Number of observations	1,072,804		1,072,804		1,072,804		1,072,804		1,072,804	
Pseudo R2	0.288		0.289		0.289		0.288		0.289	

The estimator is Fixed Effect Conditional Logit. Standard errors are corrected for clustering across district of origin. *** p<0.01, ** p<0.05, * p<0.1

Table 6. Income and the choice of migration destination -- using all male migrants

District difference in:	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat
Income								
Average log income	0.143	0.785	0.172	0.939				
Differential in log income due to education and language			0.217***	3.033				
Relative log income controlling for education and language					2.069***	2.735		
Average consumption adequacy index: total income							-1.145***	-4.439
Differential due to education and language total income							0.498***	4.151
Prices and amenities								
Log of rice price	-3.908***	-5.327	-3.800***	-5.438	-3.798***	-5.433	-3.939***	-5.395
Housing price premium (log)	0.390***	5.326	0.383***	5.162	0.388***	5.236	0.408***	5.625
Travel time to nearest paved road	-0.918***	-9.685	-0.907***	-9.647	-0.934***	-9.987	-0.937***	-10.102
Travel time to nearest bank	0.400*	1.757	0.383*	1.667	0.355	1.571	0.105	0.407
Elevation in '000 meters	-0.402**	-2.050	-0.445**	-2.356	-0.397**	-2.126	-0.171	-0.833
Population								
Population density	0.938***	6.883	0.963***	6.831	0.986***	6.952	0.933***	6.912
Log(population)	0.444***	2.875	0.410***	2.724	0.440***	3.085	0.538***	3.773
Ethno-caste similarity index	0.900***	3.425	0.855***	3.225	0.851***	3.216	0.824***	3.195
Language similarity index	1.506***	7.826	1.549***	7.805	1.564***	7.851	1.471***	8.032
Religion similarity index	-0.612*	-1.880	-0.623*	-1.910	-0.601*	-1.866	-0.382	-1.178
Distance								
Distance in '00 Km	-10.632***	-13.548	-10.627***	-13.518	-10.599***	-13.516	-10.651***	-14.107
Log-Likelihood	-67,188.57		-67,146.99		-67,161.69		-67,007.60	
Number of observations	1,555,740		1,555,740		1,555,740		1,555,740	
Pseudo R2	0.312		0.312		0.312		0.314	

The estimator is Fixed Effect Conditional Logit. Standard errors are corrected for clustering across district of origin. *** p<0.01, ** p<0.05, * p<0.1

Table 7. Consumption and the choice of migration destination -- using all male migrants

Consumption	**coef**	**t-stat**	**coef**	**t-stat**	**coef**	**t-stat**	**coef**	**t-stat**	**coef**	**t-stat**
Average consumption expenditures (log)	0.418**	2.036	0.433**	2.105						
Log consumption differential due to education a			0.614***	7.197						
Combined average and differential					0.541***	4.992			0.446**	2.202
Relative log consumption controlling for educat							6.181***	7.126	1.681	0.825
Prices and amenities										
Log of rice price	-4.181***	-5.768	-4.173***	-6.016	-4.239***	-6.020	-3.895***	-5.509	-4.181***	-6.026
Housing price premium (log)	0.405***	5.744	0.409***	5.905	0.412***	6.023	0.397***	5.554	0.409***	5.907
Travel time to nearest paved road	-0.826***	-8.245	-0.831***	-8.431	-0.802***	-8.877	-0.952***	-9.925	-0.828***	-8.415
Travel time to nearest bank	0.350*	1.646	0.335	1.589	0.331	1.576	0.362	1.628	0.335	1.587
Elevation in '000 meters	-0.485**	-2.496	-0.553***	-2.978	-0.581***	-3.155	-0.418**	-2.308	-0.555***	-2.983
Population										
Population density	0.937***	6.636	1.019***	7.356	1.003***	7.604	1.037***	7.665	1.017***	7.353
Log(population)	0.419***	2.892	0.357***	2.669	0.350**	2.560	0.412***	3.046	0.356***	2.654
Ethno-caste similarity index	0.934***	3.593	0.794***	2.921	0.821***	3.027	0.762***	2.786	0.797***	2.934
Language similarity index	1.476***	7.717	1.622***	8.812	1.594***	8.599	1.665***	8.881	1.620***	8.797
Religion similarity index	-0.552*	-1.702	-0.744**	-2.265	-0.707**	-2.183	-0.780**	-2.403	-0.739**	-2.255
Distance										
Distance in '00 Km	-10.660***	-13.587	-10.619***	-13.481	-10.636***	-13.393	-10.564***	-13.446	-10.620***	-13.483
Log-Likelihood	-67,142.70		-66,980.53		-66,986.51		-67,044.31		-66,981.61	
Number of observations	1,555,740		1,555,740		1,555,740		1,555,740		1,555,740	
Pseudo R2	0.312		0.314		0.314		0.313		0.314	

The estimator is Fixed Effect Conditional Logit. Standard errors are corrected for clustering across district of origin. *** p<0.01, ** p<0.05, * p<0.1

Table 8. Income and the choice of migration destination -- using migrants from the NLSS 2002/3

District difference in:	coef	t-stat	coef	t-stat	coef	t-stat	coef	t-stat
Average log income	-0.881*	-1.925	-0.883*	-1.926				
Differential in log income due to education and language			-0.060	-0.179				
Relative log income controlling for education and language					-0.698	-0.201		
Average consumption adequacy index: total income							-2.425***	-3.480
Differential due to education and language: total income							-1.021	-1.488
Prices and amenities								
Log of rice price	3.061***	2.939	3.046***	2.941	2.658**	2.541	2.455***	2.579
Housing price premium (log)	0.320***	2.754	0.323***	2.769	0.313***	2.899	0.368***	3.270
Travel time to nearest paved road	-1.741***	-4.374	-1.745***	-4.387	-1.525***	-3.305	-1.501***	-3.534
Travel time to nearest bank	0.986	1.634	1.001*	1.660	1.127**	2.018	0.800	1.352
Elevation in '000 meters	0.284	0.740	0.286	0.742	0.097	0.245	0.406	1.082
Population								
Population density	-0.479	-1.188	-0.484	-1.187	-0.607	-1.368	-0.661	-1.578
Log(population)	2.628***	4.400	2.633***	4.390	2.574***	4.073	2.885***	4.821
Ethno-caste similarity index	0.910	0.900	0.915	0.907	0.867	0.896	0.966	0.954
Language similarity index	3.432***	3.481	3.420***	3.498	3.436***	3.438	3.147***	3.290
Religion similarity index	-0.666	-0.470	-0.622	-0.437	-0.577	-0.402	-0.767	-0.542
Distance								
Distance in '00 Km	-11.141***	-6.430	-11.144***	-6.448	-11.186***	-6.534	-11.571***	-6.953
Log-Likelihood	-620.08		-620.05		-622.72		-612.82	
Number of observations	16,214		16,214		16,214		16,214	
Pseudo R2	0.391		0.391		0.388		0.398	

The estimator is Fixed Effect Conditional Logit. Standard errors are corrected for clustering across district of origin. *** p<0.01, ** p<0.05, * p<0.1

Table 9. Consumption and the choice of migration destination -- using migrants from the NLSS 2002/3

Consumption	**coef**	**t-stat**	**coef**	**t-stat**	**coef**	**t-stat**	**coef**	**t-stat**	**coef**	**t-stat**
Average consumption expenditures (log)	0.373	0.453	0.375	0.455						
Log consumption differential due to education and language			-0.060	-0.154						
Combined average and differential					0.126	0.309			0.417	0.506
Relative log consumption controlling for education and language							-0.989	-0.247	-5.343	-0.565
Prices and amenities										
Log of rice price	2.304*	1.645	2.300	1.640	2.553**	2.237	2.671**	2.547	2.256	1.619
Housing price premium (log)	0.312***	2.997	0.314***	3.039	0.305***	2.901	0.314***	2.933	0.316***	3.071
Travel time to nearest paved road	-1.351***	-2.858	-1.349***	-2.819	-1.467***	-3.664	-1.520***	-3.266	-1.326***	-2.774
Travel time to nearest bank	1.045*	1.937	1.050**	2.002	1.079**	2.101	1.122**	2.064	1.044**	1.989
Elevation in '000 meters	0.023	0.046	0.024	0.046	0.071	0.158	0.094	0.239	0.014	0.026
Population										
Population density	-0.636	-1.505	-0.644	-1.471	-0.595	-1.313	-0.614	-1.362	-0.653	-1.491
Log(population)	2.592***	4.085	2.602***	3.973	2.556***	3.876	2.581***	4.001	2.612***	3.968
Ethno-caste similarity index	0.888	0.934	0.896	0.926	0.855	0.871	0.876	0.890	0.904	0.936
Language similarity index	3.413***	3.474	3.409***	3.496	3.444***	3.425	3.441***	3.442	3.401***	3.495
Religion similarity index	-0.750	-0.476	-0.740	-0.465	-0.692	-0.447	-0.605	-0.414	-0.740	-0.465
Distance										
Distance in '00 Km	-11.204***	-6.558	-11.208***	-6.605	-11.184***	-6.512	-11.188***	-6.568	-11.213***	-6.614
Log-Likelihood	-622.38		-622.37		-622.65		-622.71		-622.24	
Number of observations	16,214		16,214		16,214		16,214		16,214	
Pseudo R2	0.388		0.388		0.388		0.388		0.388	

The estimator is Fixed Effect Conditional Logit. Standard errors are corrected for clustering across district of origin. *** p<0.01, ** p<0.05, * p<0.1

Table 10. Relative magnitude of effect of regressors on choice of migration destination

	Standard deviation	Relative effect
Income and consumption		
Combined income effect	0.70	0.12
Relative log income controlling for education and language	0.06	0.09
Combined consumption effect	0.54	0.24
Relative log consumption controlling for education and language	0.03	0.17
Prices and amenities (*)		
Log of rice price	0.29	-1.28
Housing price premium (log)	1.74	0.73
Travel time to nearest paved road	1.34	-1.09
Travel time to nearest bank	0.83	0.22
Elevation in meters	1.08	-0.46
Population (*)		
Population density	0.47	0.43
Log(population)	0.92	0.29
Ethno-caste similarity index	0.17	0.13
Language similarity index	0.38	0.54
Religion similarity index	0.23	-0.08
Distance (*)		
Distance above 100 Km	0.19	-1.84

Relative effect of a one standard deviation calculated as coefficient x standard deviation,
(*) averaged over the different regressions reported in Tables 4 and 5.

Table A1. Hedonistic regression of house rental value

	Coef.	t-stat
Area of dwelling		
Log(sq.ft of the dwelling)	0.179	(3.08)**
Log(sq.ft of the plot)	-0.093	(1.91)
Kitchen garden (yes=1)	-0.202	(2.72)**
Number of rooms and room composition		
Log(number of rooms)	0.553	(6.37)**
Share of Kitchen	-1.467	(0.69)
Share of toilet/bathroom	-2.619	(1.21)
Share of bedrooms	-2.113	(1.00)
Share of living/dinning room]	-1.517	(0.72)
Share of office	-1.185	(0.55)
Share of mixed use room	-2.256	(1.07)
Share of other rooms	-2.358	(1.11)
Construction material of outside wall		
Mud Bricks/stone (yes=1)	-0.197	(1.66)
Wood/branches (yes=1)	-0.369	(2.36)*
Other (yes=1)	-1.455	(7.90)**
Floor material		
Wood, Stone,Cement/tile or other (yes=1)	0.461	(3.66)**
Roof material		
Galvanized Iron (yes=1)	0.823	(6.75)**
Concrete, Cemnet(yes=1)	0.882	(4.90)**
Tiles/slate(yes=1)	0.44	(4.79)**
Characteristics of windows		
Shutters (yes=1)	0.379	(4.43)**
Screen/glass(yes=1)	0.496	(2.64)**
Other (yes=1)	-0.602	(2.32)*
Drinking water source		
Covered Well/Hand Pump	-0.25	(1.99)*
Open Well	-0.309	(1.80)
Other (yes=1)	-0.474	(3.27)**
Amenities		
Sanitary System (yes=1)	0.115	(0.88)
Garbage Disposal (yes=1)	0.121	(0.78)
Non-Flush/Communal Toilet (yes=1)	-0.48	(2.90)**
No toilet (yes=1)	-0.596	(3.47)**
Electric Light (yes=1)	-0.003	(0.08)
District dummies	Yes	

The dependent variable is the log of the rental value of the dwelling.
Rental value is either actual or estimated in case of owner occupation.
Based on NLSS 1995/96.

Table A2. Migration Selection Equation

	Coef/z-stat
Age	0.011
	(0.78)
Age squared	-0.000
	(0.41)
Father's education level	0.036
	(2.60)**
Father's employment in non-farm sector	0.344
	(3.61)**
High caste dummy	0.253
	(3.78)**
Education	0.033
	(0.87)
Constant	-1.532
	(4.58)**
Observations	2762

The dependent variable is 1 if head was born outside district of residence
Robust z statistics in parentheses
* significant at 5%; ** significant at 1%