



UNU-MERIT Working Paper Series

#2011-047

# International migration and local employment: analysis of self-selection and earnings in Tajikistan

**By Aziz Atamanov and Marrit van den Berg**

**Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)**
email: info@merit.unu.edu | website: http://www.merit.unu.edu

**Maastricht Graduate School of Governance (MGSoG)**
email: info-governance@maastrichtuniversity.nl | website: http://mgsog.merit.unu.edu

Keizer Karelplein 19, 6211 TC Maastricht, The Netherlands
Tel: (31) (43) 388 4400, Fax: (31) (43) 388 4499

**UNU-MERIT Working Papers**

**ISSN 1871-9872**

**Maastricht Economic and social Research Institute on Innovation and Technology, UNU-MERIT**

**Maastricht Graduate School of Governance MGSoG**

***UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at UNU-MERIT and MGSoG to stimulate discussion on the issues raised.***

# International migration and local employment: analysis of self-selection and earnings in Tajikistan

August 2011

Working paper


**Abstract**

This paper addresses the issue of self-selection of individuals in international labour migration, non-agricultural and agricultural employment in Tajikistan and its link to earnings from these activities. Unlike most empirical studies, we could attribute selection bias on unobservable characteristics to the allocation of individuals to alternative employment sectors and analyse its impact on earnings abroad and at home. We have found positive selection in migration against local non-agricultural activities and positive selection in local non-agricultural activities against local agricultural activities. This indicates that the most capable individuals with regards to unobservable characteristics choose to migrate, while the somewhat less able choose non-agricultural activities, and individuals with the worst capabilities stay in poorly-paid agricultural activities. Controlling for self-selection, labour income returns to education of migrants and individuals in non-agricultural activities are slightly lower than those from Ordinary Least Squares (OLS).



Keywords: international migration, self-selection, earnings, Tajikistan
JEL classification: J24, J31, F22, O15



**Authors:**

Aziz Atamanov[1*] and Marrit van den Berg[2]

[1] Maastricht Graduate School of Governance, Maastricht University, Maastricht, The Netherlands
[2] Development Economics, Wageningen University, Wageningen, The Netherlands

* Contact author:
[1]Email: azizatamanov@hotmail.com

[2]Email: Marrit.vandenBerg@wur.nl

## 1. Introduction

International migration is often viewed as an important alternative to economic difficulties at home in developing countries related to the lack of employment opportunities, risky and seasonal agriculture, and market imperfections (Stark, 1991). Migration has a potential to absorb excessive labour, to overcome seasonal income variability of agricultural activities and to finance self-employment activities through remittances (Taylor, 1999). However, since migration is a selective process, often attracting better educated and more productive local workers, it can constrain local income-generating activities as well. Therefore, understanding self-selection into international migration and its relationship with local activities is important for policy making in developing countries (Mora and Taylor, 2005).

Human capital is one of the most important determinants of international migration and local non-agricultural activities. For instance, there are numerous studies showing that education determines access to rural non-agricultural activities and has a positive relationship with income from them (Reardon et al., 2001). At the same time, individuals that have higher expected earnings in other countries due to characteristics such as education, abilities, and experience are more likely to migrate (Taylor and Martin, 2001). This skill bias in migration may have a negative impact on local activities absorbing the most productive workers, as discussed in the brain drain literature (De Haas, 2007).[1] Moreover, self-selection in migration causes a bias in estimates of returns to human capital in different income activities. This can lead to the misleading evaluation of government programs and may blur the picture on successful allocation of human capital across the most productive activities (Rosenzweig and Wolpin, 1988; Dimova and Gang, 2007).

The empirical analysis of the impact of human capital characteristics on participation and earnings from migration and local income activities has been limited due to methodological problems. Traditionally used selectivity correction methods (Heckman, 1979; Lee, 1983) could not explain the selection bias in earnings in a specific activity from re-allocation of individuals from other alternatives and therefore provided an incomplete picture. For instance, traditional sample selection models do not clarify whether a positive bias in migrants' earnings is related to reallocation of people with better unobserved characteristics from the non-agricultural sector to migration or from agriculture to migration, which may have different policy implications. The selectivity correction methodology developed by Dubin and McFadden (1984) and recently modified by Bourguignon et al. (2007) makes such a distinction possible. This enables new

[1] Due to the focus of this work on self-selection and earnings we are not discussing potential positive effects of brain drain on stimulating investment in education in home countries, the positive impact of remittances and obtained skills on productivity at local communities.

empirical studies that contribute to the migration, human capital, and nonfarm literature and promote better understanding of their complex interrelationships.

Still, the number of relevant studies is limited due to a lack of comprehensive data comprising both earnings of migrants abroad and non-migrants at home. Most relevant are Lanzona (1998) and Wu (2010). Lanzona (1998) focused on how self-selection in migration influenced local wages because information on migrants' earnings was not available. He showed that positive selection in migration led to the underestimation of returns to education and experience in local wages. In contrast, Wu (2010) analysed the impact of human capital on migrants' earnings controlling for self-selectivity and participation in local agricultural and non-agricultural employment, but due to lack of data he could not analyse the returns to human capital in local employment. The author found positive selection in migration on unobserved characteristics relative to working in agriculture, and negative selection relative to participation in the local non-agricultural sector.

This paper contributes to the existing literature by identifying the selection bias on observable and unobservable characteristics of human capital and its effect on returns to human capital across three all-encompassing employment choices: international migration, local non-agricultural and local agricultural activities. We use data for Tajikistan from the national household budget survey in 2007, which offers detailed information on both migrants and non-migrants, and on their earnings. Tajikistan is a poor, predominantly rural Central Asian country that became a prominent supplier of labour migrants to the Russian Federation. It is one of the largest recipients of remittances in the world in percentage to GDP (ILO, 2010; Mohapatra et al., 2010). According to Mughal (2006), about 370,000 Tajiks work temporarily and permanently abroad (mostly in the Russian Federation), which is equivalent to 28% of the economically active male population and 17% of the total economically active population. Still, to the best of our knowledge, the determinants of international migration in Tajikistan have not been studied quantitatively yet.[2]

Tajikistan lacks land resources, and agriculture alone cannot provide its young population sufficient jobs. In these circumstances both non-agricultural activities and international migration can absorb the excessive labour force, but there is growing anecdotal evidence that the brain drain associated with international migration leads to shortages of skilled manpower in the Republic (ILO, 2010:54). Therefore, it is important for policymakers to understand how the reallocation of human capital affects participation and income from migration and local income activities. This study shows how self-selection affects estimates of returns to education in

[2] Asian Development Bank (2008b) and World Bank (2009) present comprehensive descriptive analysis of migration, remittances and welfare implications, but do not analyze the impact of self-selection on migrants' and local income.

different income generating activities, which brings more accurate information on the most productive employment choices and effectiveness of policy programs, such as public education and trainings.

## 2. Estimation strategy

The theoretical and empirical framework in this section is mostly based on Lanzona (1998), while detailed descriptions of the models discussed are presented in Lee (1983), Dubin and McFadden (1984) and Bourguignon et al. (2007).

A wage framework can be presented by the following equations:

$$\log w_{si} = X_{si}\beta_{si} + u_{si} \quad (1)$$

$$Y_{si}^{*} = Z_{si}\gamma_{si} + \eta_{si} \quad (2)$$

Where $X_{si}$ refers to the vector of exogenous variables that determine wages at home and abroad ($\log w_{si}$). $Z_{si}$ represents the set of explanatory variables affecting the probability of choosing sth option ($Y_{si}^{*}$) which is participation in job market at home and abroad.

Imagine that there are J number of job choices (j=1,…,J). In the data the individual has thee choices: i) international migration; ii) work in the agricultural sector; and iii) work in the non-agricultural sector. Disturbance $u_k$ satisfies $E(u_k \mid X, Z) = 0 \quad and \quad v(u_k \mid X, Z) = \delta_k^2$. We assume that the same factors that determine the choice of workers to migrate abroad can also explain the choice between local agricultural and non-agricultural sectors. An individual considers all existing market opportunities at home and abroad and choose one that maximizes his utility (Lanzona, 1998).

If there are unobservable characteristics, for instance abilities, that affect the participation in an employment activity and the earnings from it, disturbances in equations (1) and (2) will be correlated, leading to biased estimates of $\beta_k$ when (1) is estimated using OLS (Heckman, 1979). Unbiased estimation requires the use of selection correction methods for multiple choices as developed by Lee (1983) and Dubin and McFadden (1984). The difference between these two approaches is the assumption behind the direction of correlations between the unobserved determinants of the choice of alternative j against any other alternative and the unobservable determinants of the outcome ($\log w_{si}$). Lee (1983) assumes that all correlations have the same sign and that one correction term in the model explains the selection bias. Dubin and McFadden (1984) consider this assumption to be very strong and propose to use multiple correction terms summing up to zero to control for self-selection in the $k^{th}$ alternative as related to each other alternative. Another advantage of the Dubin and McFadden approach is that it identifies not only

the direction of the selection bias, but also links the selection bias to the allocation of individuals to each other alternative (Wu, 2010).

In this paper we use a modified version of the Dubin and McFadden approach developed by Bourguignon et al. (2007, hereafter BFG(2)) with normalized residuals that relaxes the restriction that the correlation coefficients sum up to zero. They show in a Monte Carlo experiment that this model outperforms the Dubin and Mc Fadden model if the restriction is violated and the independence of irrelevance alternative (IIA) hypothesis is questionable. Moreover, the BFG(2) provides similar results to the Dubin and McFadden approach when the restriction holds. Estimation of this model is done in two steps. In the first step, a multinomial logit model of activity choice is estimated. In the second step, the predicted probabilities are included in the wage equations, which are estimated by OLS. In order to correct inefficient standard errors due to two-step procedure, a bootstrap method is used.

The BFG(2) produces the same number of bias correction terms as the number of multinomial logit choices. Each term shows the direction of the bias related to the allocation of individuals to a specific sector and also explains from which choice among other alternatives this bias stems from. For instance, a positive bias-correction coefficient related to the non-agricultural sector selection equation in the migration earnings equation demonstrates higher earnings of migrants compared to earnings of random individuals due to the allocation of people with better unobserved characteristics from the non-agricultural sector to international migration.

## 3. Data and model specification

### 3.1 Data

For our empirical analysis we use data from the Tajikistan Living Standard Survey (TLSS) implemented by the National Statistical Committee of Tajikistan with support from the World Bank and the United Nations Children's Fund in 2007. The data is available through the Living Standards Measurement Study project of the World Bank. This survey provides detailed information about employment, labour income, and demographic characteristics of all household members. It also includes detailed information on demographic characteristics, employment and income of migrants, who were absent from the household during the survey. For the purpose of this study, we have chosen the subsample of working adults (15-65 years) who earned labour income. After accounting for individuals with missing income and education information, we end up with a subsample of 7184 individuals. We classify all labourers into three categories: international migrants working abroad at the time of survey (N=874), individuals with primary employment in non-agricultural activities (N=4364) and individuals with primary employment in agricultural activities (N=1946). Non-agricultural and agricultural activities include both wage

and self-employment. Monthly income of local labourers includes net monthly wages, bonuses and payments in kind. As we know how many hours an individual spends on this work during the month, we were able to compute hourly earnings of local labourers. Unfortunately, we only know monthly income and not hours worked for migrants.

Table 1. Descriptive statistics across occupational status, 2007

| | Migration | Non-agricultural employment | Agricultural employment |
|---|---|---|---|
| Monthly income, somoni | 1066.7 | 344.5 | 140.7 |
| Age, years | 28.0 | 38.1 | 33.3 |
| Individual is male, % | 93.0 | 72.0 | 48.7 |
| Individual has higher education, % | 10.4 | 27.2 | 3.8 |
| Individual has vocational education, % | 11.1 | 23.2 | 10.1 |
| Individual has secondary education, % | 65.8 | 39.8 | 59.7 |
| Individual does not have secondary education, % | 12.7 | 9.8 | 26.5 |
| Individual lives in urban area, % | 21.4 | 44.8 | 4.3 |
| No ownership of land, % | 24.3 | 42.5 | 5.6 |
| Size of owned land per capita between 0.01 and 0.1 ha, % | 29.3 | 29.9 | 25.4 |
| Size of owned land per capita between 0.1 and 0.2 ha, % | 21.5 | 13.2 | 32.3 |
| Size of owned land per capita more than 0.2 ha, % | 24.9 | 14.3 | 36.7 |
| Altitude below 400 meters, % | 12.0 | 11.3 | 25.2 |
| Altitude between 400 and 1000 meters, % | 45.4 | 64.8 | 62.9 |
| Altitude between 1000 and 1500 meters, % | 10.3 | 11.2 | 7.4 |
| Altitude above 1500 meters, % | 32.3 | 12.7 | 4.5 |
| Number of observations | 874 | 4364 | 1946 |

Source: Tajik National Statistical Committee, authors' calculation.

As a result, we will use logarithms of hourly agricultural, non-agricultural labour income and logarithm of migrant monthly income as dependent variables.

The data do not contain information on years spent on education, only on the highest degree an individual obtained. Based on this information, we grouped individuals into four main categories: higher education, vocational education, secondary education, and the reference group of individuals with education lower than secondary.[3] With regards to working experience, there is not information how long an individual has worked after studying. The only available information is about the current tenure. For migrants, we also know the time when they left home last time, which can be used as an imperfect measure of their migration experience. Descriptive statists across chosen employment categories are presented in table 1.

The average monthly income of international migrants is three times higher than the average non-agricultural income and almost nine times higher than the average agricultural income. Migration attracts mostly young males. Older people are more likely to be involved in non-agricultural activities. More than half of those employed in agriculture are females. With regards to education, every tenth migrant has higher education, which is less than in non-agricultural activities (every forth has higher education), but much higher than among those employed in agricultural activities. Ownership of land seems to be an important factor explaining

[3] Vocational education does not belong to higher education in Tajikistan and is a post-secondary professional education.

participation in different employment categories. There are more landless among migrants and individuals employed in non-agricultural activities than among those employed in agriculture. Poor agricultural conditions also seem to explain why people engage in non-agricultural activities or migrate abroad. For example, 32% of international migrants originate from areas higher than 1500 meters above the sea level, in comparison to 4.5% of individuals engaged in agriculture.

It is interesting to check the main sectors that migrants and non-agricultural workers were employed in. The most important primary non-agricultural sectors were: trade (19%), education (16.5%), construction (15%), transport (8.7%) and public administration (8.7%). The majority of migrants were employed in construction (52.8%) and trade (9.6%).[4] More than 60% of migrants were unemployed before migration. With regards to local agricultural activities, 66% of individuals were employed by non-household members, 26.6% worked on rented farms or farms owned by relatives and 7.5% were self-employed on their own farms.

In Tajikistan, the agricultural sector has undergone serious changes, but reforms are still far from accomplished. According to USAID (2004), farmers' control over land resources is still limited and many rural individuals are employed at large and inefficient collective farms. Moreover, in cotton growing areas many farmers do not have the freedom to choose which crops to grow but have to supply cotton to the state.

### 3.2 Model specification

In order to explain wages and the choice of employment sector estimate reduced form equations for individual participation in local employment and migration, along with equations for logarithmic income from each employment strategy. The choice of explanatory variables is based on the preceding empirical studies of selection bias in earnings, determinants of migration, and nonfarm activities (Lanzona, 1998; Mora and Taylor, 2006; Shi et al., 2007; Dimova and Gang, 2007; Wu, 2010).[5]

Our empirical specification is as follows:

$Sector = \mathrm{F}(X_s)$,

$\ln Y_s = \mathrm{G}_s(X_y)$, $s$ = migration, farm, nonfarm

where sector indicates the sector of employment (international migration, agriculture, local nonfarm) and $Y_s$ is the hourly wages in the case of farm or and local nonfarm employment and monthly earnings in the case of migration. As mentioned before, due to the limitation of data we

[4] Direct comparison between migrants and nonagricultural workers is not possible since about 20% of migrants worked as unskilled laborers without indicating the sector they are employed in. Therefore, indicated percentages most probably underestimate shares of migrants in construction and trade.

[5] Full description of variables and descriptive statistics are provided in the appendix.

cannot estimate hourly earnings of migrants. $X_s$ and $X_y$ are vectors of explanatory variables for the selection and wage/income equations, respectively. The choice of variables for each equation and their detailed explanation is provided below.

We utilise an extended version of classical Mincer (1974) equation to explore returns to observable characteristics of migrants, and individuals employed in agricultural and non-agricultural sectors in earnings equations. Observable characteristics include education, age and its squared term, tenure in the present occupation, gender, and ethnicity. We also included regional and urban/rural dummies to control for regional differences. For the agricultural and non-agricultural earnings' equations we have included the distance to regional centre measured at the community level. Wages in remote areas are expected to be lower than in more developed areas closer to the centre.

Education is measured as a range of binary variables for higher, vocational and secondary education. Education less than secondary is chosen as a base category. We proxy working experience by age and dummies for the number of years an individual has been working at the current job/ year of last migration visit. Working less than seven month at the current job and migrating in 2007 are the base categories.

For the selection equation we need exogenous variables explaining the choice of employment activity that are correlated with the sector, but not affect market wages (Dimova and Gang, 2007: 619). Household assets and family background are often used in selection equations (Lanzona, 1998; Wu, 2010). We use a set of household, individual and local characteristics to explain the choice between sectors based on existing studies of self-selection bias, nonfarm literature and migration studies (Reardon et al., 2006; Martin and Taylor, 2001).

Possession of land seems to be an important factor affecting incentives to undertake non-agricultural activities or to migrate in rural areas. In particular, we expect individuals from areas with insufficient amount of land to choose migration or non-agricultural activities. As individual access to land can be endogenous, because wealthier people or individuals with better unobserved abilities may have connections to get better and large land holdings, we use total land available per capita at the community level.

Household size and shares of children and old people in the household measure the human capital of the family and its vulnerability. For example, as shown in Shi et al. (2007), old people may both promote and constrain migration depending on their health. In particular, they can offer care for children which enables migration or in contrast require care for themselves and constrain migration. The same logic applies in the case of children.

Another variable used to explain the choice between agricultural, non-agricultural activities and migration is the education of the head of household's father. Contrary to individual

education attainment, we expect that this variable does not affect earnings in a given sector. We anticipate that individuals from households where the head's father is better educated have more chances to engage into non-agricultural activities or to migrate due to better knowledge and connections rather than staying in the poor paying agricultural sector (human capital plays crucial role both in international migration and nonfarm activities, see Reardon et al., 2006 and Martin and Taylor, 2001). Education less than secondary is chosen as a base category. We also included a proxy for migration networks in the selection equation. Migration networks are frequently used in the literature to instrument for migration as helping new migrants to decrease migration costs through providing access to information, assistance with housing and work abroad (McKenzie and Rapoport, 2007). The proxy is measured as the percentage of people who has lived abroad at least for three month during 1998-2003 to adult local population (older than 14 years) at the district level.

We used the distance in kilometres to regional centre and the availability of a central water system at the local level to proxy costs which may affect the decision to migrate and the capacity to undertake non-agricultural activities. Nonfarm studies often indicate importance of the access to infrastructure for development of and access to nonfarm activities (Reardon et al., 2006). Local economic potential is also captured by the inclusion of several dummies for different altitudes which may affect incentives and capability of individuals to choose income generating activities. An altitude above 1500 metres is chosen as a base category. Living in mountainous areas can stimulate nonfarm activities and international migration due to adverse agricultural conditions, but can also constrain them due to worse access to infrastructure, remoteness and so forth.

Individual's age and gender are included in the selection equation as well. As the majority of migrants work in construction and trade sectors, it is expected that younger males are more likely to migrate. Older people are more likely to choose non-agricultural activities which may require experience and connections, while migration can be a less attractive strategy for them because they have families to care for and have less time to repay investments (Wu, 2010). Finally, in order to control for regional effects, we included dummies for regions and urban/rural residence.

## 4. Results of the econometric analysis

### 4.1. Results from the multinomial logit

Results from the first step multinomial regression are presented in table 1. We explain employment choice, distinguishing between participation in international migration, agricultural and non-agricultural activities. Agricultural and non-agricultural employment is compared to the

base category of international migration. The Small-Hsiao test did not reject the independence of irrelevant alternatives (IIA), while the Hausman test provides mixed results (Hausman & McFadden, 1984; Small & Hsiao, 1985).[6] Fortunately, the method we use provides a fairly good selection correction for the outcome equation even if the IIA hypothesis is violated (Bourguignon et al., 2007). We also use Wald and log-likelihood tests to check whether the outcomes can be combined, but the tests reject these hypotheses.

Table 2 reports risk ratios, which are the coefficients in exponential form and which indicate how the risk of the outcome falling in the comparison group compares to the risk of the outcome falling in the reference group changes with the variable in question. A ratio greater than (less than) one indicates a higher (lower) probability of choosing non-agricultural activities or agricultural activities over international migration.

The empirical results reveal several factors which significantly affect the choice between employment opportunities. First of all, as expected, in areas with larger share of migration stock to local population during 1998-2003 years there is less likelihood to choose agricultural or non-agricultural activities over international migration. Migrants are also found to be different based on their observable individual characteristics. Males are less likely to undertake local agricultural or non-agricultural activities than females. This is in line with theoretical and previous empirical studies on Central Asia which show that international labour migrants are predominantly young males (Asian Development Bank, 2008ab).

Family background plays an important role in employment choice. Thus, individuals from households where the head's father had a higher education are more likely to choose local non-agricultural activities over migration. This suggests that returns to education may be higher in local non-agricultural activities in comparison to international migration. Having more children and old people in the household is positively related to participation in international migration versus non-agricultural activities which may indicate a "push" nature of this process when individuals from more vulnerable households are prone to migrate.

Table 2. Risk ratios after multinomial logit for the employment choice of individuals (N=1784)

| Variables | Non-agricultural employment | Agricultural employment |
|---|---|---|
| Proxy for migration network, % | 0.922*** | 0.988 |
| | [0.0140] | [0.0168] |
| Age, years | 1.008 | 0.853*** |
| | [0.0402] | [0.0344] |
| Age squared | 1.002*** | 1.004*** |
| | [0.000624] | [0.000631] |
| Total land per capita at the local level | 1.008 | 1.042** |
| | [0.0158] | [0.0189] |
| Father of the head of household had higher education | 1.791*** | 0.851 |
| | [0.379] | [0.225] |
| Father of the head of household had vocational education | 1.352 | 1.269 |

[6] Results of tests are available in the appendix.

| Variables | Non-agricultural employment | Agricultural employment |
|---|---|---|
| | [0.277] | [0.306] |
| Father of the head of household had secondary education | 1.216 | 0.941 |
| | [0.176] | [0.165] |
| Dummy, altitude less than 400 meters | 2.458*** | 1.957** |
| | [0.688] | [0.625] |
| Dummy, altitude between 400 and 1000 metres | 3.592*** | 1.515 |
| | [0.917] | [0.436] |
| Dummy, altitude between 1000 and 1500 metres | 4.185*** | 0.869 |
| | [1.082] | [0.279] |
| Dummy for gender, male | 0.168*** | 0.0451*** |
| | [0.0239] | [0.00679] |
| Distance to the regional centre in kilometres | 1.002*** | 1.002*** |
| | [0.000614] | [0.000803] |
| Dummy for central water supply at the local level | 1.629*** | 1.078 |
| | [0.205] | [0.152] |
| Share of children in the household size,% | 0.378*** | 0.703 |
| | [0.107] | [0.230] |
| Share of old people in the household size, % | 0.315** | 0.669 |
| | [0.177] | [0.484] |
| Household size, number of people | 1.193*** | 1.224*** |
| | [0.0288] | [0.0324] |
| Dummy , Tajik | 1.135 | 0.768* |
| | [0.148] | [0.111] |
| Dummy for urban areas | 2.077*** | 0.326*** |
| | [0.328] | [0.0749] |

Notes: *** p<0.01, ** p<0.05, * p<0.1. For brevity we do not report coefficients on regional dummies. International migration is the base outcome. Robust standard errors clustered at the household level are in parentheses.

Location characteristics measuring economic capacity of local areas also provide interesting information. As expected the likelihood to choose agricultural activities over migration depends positively on land availability, while worse agricultural conditions due to high altitude stimulate individuals to choose non-agricultural activities over international migration. This may happen because migration can become too costly due to the remoteness of the place of residence, making non-agricultural activities important in the remote depressed regions with poor agricultural potential. Finally, better access to infrastructure increases chances to choose non-agricultural activities versus international migration.

### 4.2. Estimation of earnings

In the second step of our empirical model we estimate regressions modelling earnings of migrants and local labourers in agricultural and non-agricultural activities. We report results from OLS and the selection corrected income estimates based on the approach described in Bourguignon et al. (2007). Selection correction coefficients (m1, m2, m3) in BFG(2) are based on the results from multinomial logit discussed above. A positive (negative) selectivity coefficient related to any of the alternative employment choices implies higher (lower) wages than those of randomly chosen individuals due to the allocation of people with worse (better)

unobserved characteristics from this sector to the respective alternative sector. For example, the positive significant non-agricultural selection coefficient we obtained in the equation for migrant earnings means higher than random rewards to human capital of migrants due to the allocation of people with worse unobserved characteristics from migration to the non-agricultural sector or alternatively allocation of people with better unobserved characteristics from the non-agricultural sector to migration.

There is also strong positive self-selection in the non-agricultural sector stemming from the fact that people with better unobservable characteristics leave the poorly paying agricultural sector and undertake non-agricultural activities. The migration selection coefficient is found to be negative in the agricultural earnings regression, which implies allocation of people with worse unobserved characteristics from migration to farm activities.

With regards to returns to observable characteristics, neither secondary nor vocational education increase migrants' earnings. However, higher education has a significant positive association with migrants' earnings. We have not found any positive relationship between age and migrants' earnings, but those migrants who left Tajikistan earlier tend to earn more in comparison to those who left in 2007. There is a positive bias in returns to higher education and migration experience in OLS regression due to positive self-selection when people with better abilities choose migration over non-agricultural activities.

Positive returns to higher and secondary education are found in non-agricultural activities as well. The results of the OLS regression are slightly overestimated, since self-selection on unobservable characteristics is not taken into account. This is related to positive selection when individuals with better unobservable characteristics leave the agricultural sector for non-agricultural activities. Experience in the non-agricultural sector pays most during the first seven month of work and returns to experience decreases afterwards which seem to be counterintuitive, but can just mean that casual labour pays best and that experience has not effect. Non-agricultural earnings have also an inverted U-shape relationship with age with a turning point at 37 years.

In contrast to non-agricultural income, we have not found a significant relationship between the tenure and agricultural earnings. Moreover, there are no positive returns to education in the agricultural sector.

Table 3. Individual income estimates from OLS and BFG(2)

| Variables | Non-agricultural income | | Agricultural income | | Income of migrants | |
|---|---|---|---|---|---|---|
| | BFG(2) | OLS | BFG(2) | OLS | BFG(2) | OLS |
| Age, years | 0.0157*** | 0.0213*** | 0.0209*** | 0.0169*** | 0.0105 | 0.000869 |
| | [0.00342] | [0.00357] | [0.00629] | [0.00486] | [0.00640] | [0.00531] |
| Age squared | -0.000222*** | -0.000279*** | -0.000206** | -0.000189*** | -9.95E-05 | 3.94E-05 |
| | [4.10e-05] | [4.50e-05] | [8.35e-05] | [6.60e-05] | [9.44e-05] | [7.57e-05] |
| Dummy for secondary education | 0.0602*** | 0.0609*** | -0.0129 | -0.0132 | 0.031 | 0.035 |
| | [0.0226] | [0.0231] | [0.0242] | [0.0217] | [0.0232] | [0.0234] |
| Dummy for vocational education | 0.0422* | 0.0458* | -0.0401 | -0.0543 | 0.0123 | 0.0118 |
| | [0.0235] | [0.0252] | [0.0398] | [0.0389] | [0.0352] | [0.0340] |
| Dummy for higher education | 0.0977*** | 0.0994*** | 0.0784 | 0.0688 | 0.0666* | 0.0772** |
| | [0.0241] | [0.0240] | [0.0658] | [0.0674] | [0.0357] | [0.0350] |
| Dummy for gender, male | 0.224*** | 0.251*** | 0.0915** | 0.136*** | 0.106** | 0.102*** |
| | [0.0196] | [0.0142] | [0.0443] | [0.0197] | [0.0438] | [0.0313] |
| Dummy , Tajik | -0.0109 | 0.00152 | 0.0237 | 0.0195 | -0.0423** | -0.0443** |
| | [0.0167] | [0.0160] | [0.0237] | [0.0210] | [0.0201] | [0.0197] |
| Distance to the regional centre in kilometres | -0.000297*** | -0.000337*** | 0.000737*** | 0.000747*** | | |
| | [8.11e-05] | [7.44e-05] | [0.000153] | [0.000160] | | |
| Worked at this job longer than 6 years | -0.152*** | -0.149*** | -0.0545 | | -0.0443 | |
| | [0.0214] | [0.0205] | [0.0565] | | [0.0575] | |
| Worked at this job during 3-5 years | -0.146*** | -0.142*** | -0.00549 | | 0.0036 | |
| | [0.0224] | [0.0212] | [0.0559] | | [0.0558] | |
| Worked at this job during 1-2 years | -0.100*** | -0.0977*** | -0.0594 | | -0.0529 | |
| | [0.0231] | [0.0216] | [0.0557] | | [0.0556] | |
| Worked at this job during 7-12 month | -0.0884*** | -0.0844*** | -0.0516 | | -0.0386 | |
| | [0.0259] | [0.0246] | [0.0732] | | [0.0695] | |
| Migrated earlier than 2005 | | | | | 0.0466 | 0.0560* |
| | | | | | [0.0316] | [0.0301] |
| Migrated in 2005 | | | | | 0.105*** | 0.113*** |
| | | | | | [0.0324] | [0.0304] |
| Migrated in 2006 | | | | | 0.0742*** | 0.0822*** |
| | | | | | [0.0281] | [0.0282] |
| m1 (related to migration) | 0.109 | | -0.409** | | 0.039 | |
| | [0.129] | | [0.161] | | [0.0339] | |
| m2 (related to non-agricultural work) | -0.00719 | | 0.0733 | | 0.256*** | |
| | [0.0835] | | [0.136] | | [0.0879] | |
| m3 (related to agricultural work) | 0.331** | | -0.0404 | | 0.0444 | |
| | [0.153] | | [0.0814] | | [0.114] | |
| N | 4364 | | 1946 | | 874 | |
| $R^2$ | 0.15 | | 0.19 | | 0.06 | |

Notes: ***p<0.01, **p<0.05, *p<0.1. For brevity we do not report coefficients for regional dummies. Bootstrapped standard errors, based on 200 replications, are in parentheses.

This can be a result of the unfinished restructuring of this sector, which suffers from state intervention in agricultural decisions. Lack of returns to education and low earnings in general may also explain why this sector looses the most productive individuals to non-agricultural activities. Among other interesting findings one may notice a persistent gender wage gap across all employment choices and lower migration income for ethnic Tajik people abroad.

In sum, migration and local nonfarm activities seem to enhance efficiency by reallocating people with better unobserved and observed characteristics to migrate or undertake non-

agricultural activities and to avoid the agricultural sector with low returns. Moreover, the empirical results confirm the usefulness of the BFG methodology as opposed to OLS. Ignoring self-selection of migrants and individuals employed in non-agricultural activities leads to a slight overestimation of returns to education in these sectors.

## 5. Conclusion

This paper addressed the question of self-selection based on observable and unobservable characteristics of individuals in international labour migration, non-agricultural and agricultural employment in Tajikistan and its link to earnings from these activities. Unlike most empirical literature, we could attribute selection bias to the allocation of individuals to alternative employment sectors. Tajikistan was chosen as one of most prominent suppliers of international migrants among Former Soviet Union countries with migration and remittances playing a crucial role in social-economic development of the country.

The employment choice between international migration, agricultural and non-agricultural activities is found to be strongly affected by observable individual, family and locational characteristics. Males originating from households with larger share of children and old people, living closer to regional centres in areas with larger migration networks are more likely to migrate rather than to choose local agricultural or non-agricultural employment. The difference between migrants and individuals in local activities also stems from their family background. In particular, individuals from households where the father of the head of household had higher education are more likely to choose non-agricultural activities over migration. Descriptive statistics show that local non-agricultural activities attract the most educated people, followed by migration, which leaves the least educated in the agricultural sector.

With regards to returns to education across different employment options, we found positive returns to higher education for migrants and even more so for non-agricultural labourers. In the nonfarm sector also the returns to secondary and vocational education are positive. We did not find any positive returns to education in agriculture, which can be the result of unfinished structural reforms in this sector. Dimova and Gang (2007) did not find positive returns to education in self-employment in Bulgaria, but obtained strong positive results after completion of structural reforms.

We have found positive selection in migration against local non-agricultural activities and positive selection in local non-agricultural activities against local agricultural activities. This indicates that the most capable individuals with regards to unobservable characteristics choose to migrate, the somewhat less able choose to undertake local non-agricultural activities, while individuals with the worst capabilities stay in poorly paid agricultural activities. These results are

different from the results in China obtained by Wu (2001), where individuals with best unobserved characteristics prefer local nonfarm work. The identified self-selection bias in unobservable characteristics slightly distorts estimates for earnings. Thus, returns to education in migrants' and non-agricultural earnings are overestimated in OLS equations which do not take into account the self-selection bias.

Our findings lead to several important policy implications. First of all, local non-agricultural activities seem to enhance efficiency by reallocating better educated and more productive individuals from badly paying agricultural activities. Secondly, a potential income gain, lack of experience and connections with migration networks seem to stimulate migration of young males with better unobservable characteristics from the non-agricultural sector. This may have a negative effect on the development of non-agricultural activities in Tajikistan, which could be (partly) compensated by remittances and obtained skills of return migrants though. This latter issue has not been addressed in the current study and can be explored in future. Finally, ignoring self-selection on unobservable characteristics in migration and non-agricultural activities can lead to slight overestimation of returns to education.

## Acknowledgments

Authors are thankful to the Living Standards Measurement Study project of the World Bank for providing access to 2003 and 2007 Tajikistan Living Standards Surveys used in this paper.

## References

Asian Development Bank. 2008a. A study on international migrants' remittances in Central Asia and South Caucasus, Country Report on Remittances of International Migrants and Poverty in the Kyrgyz Republic.

Asian Development Bank. 2008b. A Study on International Migrants' Remittances in Central Asia and South Caucasus, Country Report on Remittances of International Migrants in Tajikistan.

Bourguignon, F., M. Fournier and M. Gurgand. (2007). "Selection bias corrections based on multinomial logit model: Monte Carlo comparisons," *Journal of Economic Surveys* 21(1):174-205.

De Haas, H. 2007. Remittances, Migration and Social development. A conceptual review of the literature. Social policy and development programme paper №34. United Nations Development Institute for Social Development.

Dimova, R. and I. Gang. 2007. "Self-selection and wages during volatile transition," *Journal of Comparative Economics* 35(3):612-639.

Dubin, J. and D. McFadden. 1984. "An econometric analysis of residential electric appliance holdings and consumption," *Econometrica* 52(2): 345-362.

Hausman, J. and D. McFadden. 1984. "Specification Tests for the Multinomial Logit Model," *Econometric Society* 52(5): 1219-1240.

Heckman, J. 1979. "Sample bias as a specification error," *Econometrica*, 47(1): 153-161.

International Labour Organization. 2010. Migration and Development in Tajikistan –Emigration, Return and Diaspora. Moscow.

Lanzona, L. 1998. "Migration, self-selection and earnings in Philippine rural communities," *Journal of Development Economics* 56(1):27-50.

Lee, L. 1983. Generalized econometrics models with selectivity. *Econometrica*, 51(2): 507-512.

Martin, P. and E. Taylor. 2001. *Human Capital: Migration and Rural Population Change*. Chapter 9 in *Handbook of Agricultural Economics*, 2001, 1(1): 457-511.

McKenzie, D and H. Rapoport. 2007. "Network effects and the dynamics of migration and inequality: theory and evidence from Mexico," *Journal of Development Economics* 84(1): 1–24.

Mincer, J. 1974. *Schooling, Experience and Earnings*. Columbian University Press.

Mohapatra, S., D. Ratha and A. Silwal. 2010. Outlook for remittances flows 2011-2012. Recovery after the crisis, but risks lie ahead. Migration and development brief, November 8. The World Bank.

Mora, J. and E. Taylor. 2006. Determinants of migration, destination, and sector choice: Disentangling individual, household, and community effects. In Ozden, C. & Schiff, M. (eds.), *International Migration, Remittances, and the Brain Drain*. New York: Palgrave Macmillan.

Mughal, A. 2006. Migration, remittances, and living standards in Tajikistan. A Report Based on Khatlon Remittances and Living Standards Measurement Survey (KLSS 2005), Stanford University.

Reardon, T., J. Berdegue and G. Escobar. 2001. "Rural nonfarm employment and incomes in Latin America: Overview and policy implications," *World Development* 29(3): 395-409.

Reardon, T., J. Berdegue, C. Barrett, and K. Stamoilis, 2006. Household Income Diversification into Rural Nonfarm Activities Chapter 8 in Haggblade, S., Hazell, P. and Reardon, T. (eds.). *Transforming the Rural Nonfarm Economy*. Baltimore: Johns Hopkins University Press.

Rosenzweig, M. and K. Wolpin. 1988. "Migration selectivity and the effect of public programs," *Journal of Public Economics* 37(3): 265-289.

Shi, H., N. Heerink and F. Qu. 2007. "Choices between different off-farm employment sub-categories: An empirical analysis for Jiangxi Province, China," *China Economic Review* 18(4): 438-455.

Small, K. and C. Hsiao. 1985. "Multinomial logit specification tests," *International Economic Review* 26(3), 619-27

Stark, O. 1991. *The Migration of Labor*. Massachusetts: Blackwell Publishers.

Taylor, E. 1999. "The New Economics of Labour Migration and the role of remittances in the migration process," *International Migration* 37(1): 63-88.

USAID. 2004. Land reform and farm reorganization in Tajikistan. Policy issues paper.

World Bank. 2009. Republic of Tajikistan. Poverty Assessment. Report №51341-TJ.

Wu, Z. 2010. "Self-selection and earnings of migrants: Evidence from Rural China," *Asian Economic Journal* 24(1): 23-44.

## 6. Appendix

Table A.1. Results of Small-Hsiao tests IIA assumptions (N=7184)

| Omitted | lnL(full) | lnL(omit) | chi2 | df | P>chi2 | Evidence |
|---|---|---|---|---|---|---|
| 1 | -1234.4 | -1219.9 | 29.0 | 23.0 | 0.18 | for Ho |
| 3 | -838.9 | -833.8 | 10.2 | 23.0 | 0.99 | for Ho |
| 2 | -455.6 | -447.4 | 16.5 | 23.0 | 0.83 | for Ho |
| 3 | -833.9 | -820.5 | 26.7 | 23.0 | 0.27 | for Ho |
| 1 | -1263.0 | -1250.3 | 25.3 | 23.0 | 0.33 | for Ho |
| 2 | -451.8 | -437.7 | 28.1 | 23.0 | 0.21 | for Ho |

Table A.2. Results of Hausman tests of IIA assumptions (N=7184)

| Omitted | chi2 | df | P>chi2 | Evidence |
|---|---|---|---|---|
| 1 | 53.6 | 21 | 0 | against Ho |
| 3 | -34.5 | 22 | --- | --- |
| 2 | 327.7 | 21 | 0 | against Ho |
| 3 | -34.5 | 22 | --- | --- |
| 1 | 53.6 | 21 | 0 | against Ho |
| 2 | 327.7 | 21 | 0 | against Ho |

Table A.3. Descriptive statistics of variables

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Logarithm of nonfarm labour income | 4364 | 0.09 | 0.42 | -1.40 | 1.98 |
| Logarithm of farm labour income | 1946 | -0.44 | 0.46 | -1.70 | 1.83 |
| Logarithm of migrants' income | 874 | 2.97 | 0.23 | 2.14 | 3.71 |
| Dummy for secondary education | 7184 | 0.48 | 0.50 | 0 | 1 |
| Dummy for vocational education | 7184 | 0.18 | 0.39 | 0 | 1 |
| Dummy for higher education | 7184 | 0.19 | 0.39 | 0 | 1 |
| Dummy, worked at this job longer than 6 years | 6310 | 0.38 | 0.49 | 0 | 1 |
| Dummy, worked at this job during 3-5 years | 6310 | 0.23 | 0.42 | 0 | 1 |
| Dummy, worked at this job during 1-2 years | 6310 | 0.18 | 0.39 | 0 | 1 |
| Dummy, worked at this job during 7-12 month | 6310 | 0.09 | 0.28 | 0 | 1 |
| Dummy, migrated in 2006 | 874 | 0.45 | 0.50 | 0 | 1 |
| Dummy, migrated in 2005 | 874 | 0.18 | 0.39 | 0 | 1 |
| Dummy, migrated earlier than 2005 | 874 | 0.25 | 0.43 | 0 | 1 |
| Dummy for gender, male is the base | 7184 | 0.68 | 0.47 | 0 | 1 |
| Dummy for Tajik | 7184 | 0.77 | 0.42 | 0 | 1 |
| Dummy, urban | 7184 | 0.31 | 0.46 | 0 | 1 |
| Share of migrants in adult population at the district level in 2003, % | 7184 | 6.11 | 4.46 | 0 | 21.35 |
| Age, years | 7184 | 35.6 | 11.9 | 15 | 65 |
| Total land per capita at the community level , in hundredth parts of a hectare | 7184 | 2.92 | 3.63 | 0 | 30.58 |
| Father of the head of household had higher education | 7184 | 0.07 | 0.26 | 0 | 1 |
| Father of the head of household had vocational education | 7184 | 0.10 | 0.30 | 0 | 1 |
| Father of the head of household had secondary education | 7184 | 0.14 | 0.35 | 0 | 1 |
| Dummy, altitude less than 400 meters | 7184 | 0.15 | 0.36 | 0 | 1 |
| Dummy, altitude between 400 and 1000 metres | 7184 | 0.62 | 0.49 | 0 | 1 |
| Dummy, altitude between 1000 and 1500 metres | 7184 | 0.10 | 0.30 | 0 | 1 |
| Distance to regional centre in kilometres | 7184 | 197 | 183 | 0 | 1095 |
| Dummy for central water supply at the community level | 7184 | 0.49 | 0.50 | 0 | 1 |

**The UNU-MERIT WORKING Paper Series**

2011-01 *Mitigating 'anticommons' harms to research in science and technology* by Paul A. David
2011-02 *Telemedicine and primary health: the virtual doctor project Zambia* by Evans Mupela, Paul Mustard and Huw Jones
2011-03 *Russia's emerging multinational companies amidst the global economic crisis* by Sergey Filippov
2011-04 *Assessment of Gender Gap in Sudan* by Samia Satti Osman Mohamed Nour
2011-05 *Assessment of Effectiveness of China Aid in Financing Development in Sudan* by Samia Satti Osman Mohamed Nour
2011-06 *Assessment of the Impacts of Oil: Opportunities and Challenges for Economic Development in Sudan* by Samia Satti Osman Mohamed Nour
2011-07 *Labour Market and Unemployment in Sudan* by Samia Satti Osman Mohamed Nour
2011-08 *Social impacts of the development of science, technology and innovation indicators* by Fred Gault
2011-09 *User innovation and the market* by Fred Gault
2011-10 *Absorptive capacity in technological learning in clean development mechanism projects* by Asel Doranova, Ionara Costa and Geert Duysters
2011-11 *Microeconomic analysis of rural nonfarm activities in the Kyrgyz Republic: What determines participation and returns?* By Aziz Atamanov and Marrit van den Berg
2011-12 *Immigration and growth in an ageing economy* by Joan Muysken and Thomas Ziesemer
2011-13 *State-led technological development: A case of China's nanotechnology development* by Can Huang and Yilin Wu
2011-14 *A historical perspective on immigration and social protection in the Netherlands* by Melissa Siegel and Chris de Neubourg
2011-15 *Promoting return and circular migration of the highly skilled* by Metka Hercog and Melissa Siegel
2011-16 *Voluntary agreements and community development as CSR in innovation strategies* by Vivekananda Mukherjee and Shyama V. Ramani
2011-17 *Strengthening the roles of political parties in Public Accountability - A case study of a new approach in political party assistance* by Renée Speijcken
2011-18 *The elusive quest for the golden standard: Concepts, policies and practices of accountability in development cooperation* by Renée Speijcken
2011-19 *Are health care payments in Albania catastrophic? Evidence form ALSMS 2002, 2005 and 2008* by Sonila Tomini and Truman G. Packard
2011-20 *On India's plunge into Nanotechnology: What are good ways to catch-up?* By Shyama V. Ramani, Nupur Chowdhury, Roger Coronini and Susan Reid
2011-21 *Emerging country MNEs and the role of home countries: separating fact from irrational expectations* by Rajneesh Narula and Quyen T.K. Nguyen
2011-22 *Beyond knowledge brokerage: An exploratory study of innovation intermediaries in an evolving smallholder agricultural system in Kenya* by Catherine W. Kilelu, Laurens Klerkx, Cees Leeuwis and Andy Hall
2011-23 *Dynamics of biosciences regulation and opportunities for biosciences innovation in Africa: Exploring regulatory policy brokering* by Ann Kingiri and Andy Hall

2011-24 *The when and where of research in agricultural innovation trajectories: Evidence and implications from RIU's South Asia projects* by Vamsidhar Reddy, T.S., Andy Hall and Rasheed Sulaiman V.

2011-25 *Innovation and Diffusion of Clean/Green Technology: Can Patent Commons Help?* By Bronwyn H. Hall and Christian Helmers

2011-26 *Technology alliances in emerging economies: Persistence and interrelation in European firms' alliance formation* By Rene Belderbos, Victor Gilsing, Jojo Jacob

2011-27 *Innovation pathways and policy challenges at the regional level: smart specialization* By René Wintjes and Hugo Hollanders

2011-28 *Innovation and productivity* by Bronwyn H. Hall

2011-29 *Mapping the interdisciplinary nature and co-evolutionary patterns in five nano-industrial sectors* by Lili Wang and Ad Notten

2011-30 *Assessment of industrial performance and the relationship between skill, technology and input-output indicators in Sudan* by Samia Satti Osman Mohamed Nour

2011-31 *Assessment of skill and technology indicators at the macro-micro levels in Sudan* by Samia Satti Osman Mohamed Nour

2011-32 *Education, training and skill development policies in Sudan: Macro-micro overview* by Samia Satti Osman Mohamed Nour

2011-33 *Estimating the rate of return to education in Sudan* by Samia Satti Osman Mohamed Nour

2011-34 *The importance (impact) of knowledge at the macro-micro levels in Sudan* by Samia Satti Osman Mohamed Nour

2011-35 *Angus Maddison and Development Economics* by Adam Szirmai

2011-36 *Managerial ownership and urban water utilities efficiency in Uganda* by Dorcas Mbuvi and Achraf Tarsim

2011-37 *Immigration and growth in an ageing economy* by Joan Muyskens and Thomas Ziesemer

2011-38 *The Development of Diaspora Engagement Policies in Burundi and Rwanda* by Sonja Fransen & Melissa Siegel

2011-39 *Understanding the changing role of the Turkish diaspora* by Özge Bilgili & Melissa Siegel

2011-40 *Understanding Ethiopian diaspora engagement policy* by Katie Kuschminder and Melissa Siegel

2011-41 *Engaging the diaspora in India* by Metka Hercog and Melissa Siegel

2011-42 *Protecting Vulnerable Families in Central Asia: Poverty, vulnerability and the impact of the economic crisis* by Franziska Gassmann

2011-43 *Innovation performance and embeddedness in networks: evidence from the Ethiopian footwear cluster* by Mulu Gebreeyesus and Pierre Mohnen

2011-44 *The power of the strong state: A comparative analysis of the diaspora engagement strategies of India and Ethiopia* by Katie Kuschminder and Metka Hercog

2011-45 *New insights on the role of location advantages in international innovation* by Rajneesh Narula and Grazia D. Santangelo

2011-46 *Preferences for conditioning and being conditioned - experimental & survey evidence from Zambia* by Esther Schüring

2011-47 *International migration and local employment: analysis of self-selection and earnings in Tajikistan* by Aziz Atamanov and Marrit van den Berg