On the effects of return migration on self-employment

Clotilde Mahé* Wim Naudé† Tilman Brück‡

January 7, 2019

Abstract

Using longitudinal data from a transition economy with significant international migration, we assess whether migrants are more likely to enter into and to persist in self-employment upon return to their origin countries than non-migrants. Results indicate that temporary migration affects self-employment upon return differently by types of self-employment. Migrants who were self-employed as own-account workers before migrating, are found less likely to be self-employed on return. In contrast, temporary migration is shown to increase the likelihood that, on return, self-employed who were employers persist in entrepreneurship. We conclude that return migration can improve the quality of entrepreneurship in origin countries.

JEL classifications: F22, J24, L26, P20
Keywords: Occupational choice, entrepreneurship, migration, transition economies

*Correspondeing author. Maastricht University, Netherlands; clotilde.mahe@maastrichtuniversity.nl.

†Maastricht School of Management and Maastricht University, Netherlands; IZA - Institute for the Study of Labor, Bonn, and RWTH Aachen University, Germany; and Saïd Business School, University of Oxford, United Kingdom; w.naude@maastrichtuniversity.nl.

‡ISDC - International Security and Development Center, Berlin; Leibniz Institute of Vegetable and Ornamental Crops, Großbeeren; and IZA - Institute for the Study of Labor, Bonn, Germany; brueck@isdc.org.

1 Introduction

One feature of migration is that it is often temporary, because it is undertaken in support of an initiative in origin countries, such as accumulating funds for household consumption or for investment in a business (Dustmann and Görlach, 2016).[1] The latter would result in return migrants choosing self-employment as an occupation. Indeed, a number of studies have found that return migrants are more likely to be self-employed than non-migrants (e.g. Wahba and Zenou, 2012; Wahba, 2015; Batista et al., 2017).

In the absence or inefficiency of markets, savings accumulated during migration (and remittances) have been found to act as substitutes for formal insurance, facilitating access to capital and promoting investments in new or existing ventures in origin communities (Wahba and Zenou, 2012). By inducing changes in occupation, emigrating and returning 'home' have also been shown to affect returnees' capabilities – their skills and know-how (Reinhold and Thom, 2013).

As a consequence, as Naudé et al. (2017) point out, there is a pervasive belief that return migration can facilitate entrepreneurship and economic development in sending countries. There are at least two reasons why such a belief may be mistaken. The first is that whether return migrants as self-employed provide growth-enhancing entrepreneurship depends on why they choose self-employment. Return migrants might not be able to find a wage-paying job immediately. Self-employment upon return might be a transitory phase for reintegrating into the domestic labour market in the spirit of Harris and Todaro's (1970) 'parking lot' hypothesis. Facing limited formal sector employment opportunities, returnees would often work in small, informal, self-employed activities until finding formal wage employment or migrating again (Piracha and Vadean, 2010). The latter is often an option in countries where migration is structural and a means to secure an economic basis at home, such as in the case of Kyrgyzstan (Thieme, 2014).

Furthermore, return migrants may opt for self-employment as a temporary occupation and not as a sustained career choice due to difficult business conditions or because self-employment can require strong social capital. The depreciation of migrants' social capital as a result of time spent abroad might not be offset by the resources they gained overseas (Marchetta, 2012; Wahba and Zenou, 2012).[2]

The second reason that explains why assuming return migrants' entrepreneurial prowess

[1] In this paper, migration is defined as the 'relocation of people within space that involves their permanent or temporary change of residence' (Mafukidze, 2006, p.106).

[2] There is nonetheless evidence suggesting that households with migrants are likely to insure their social networks against shocks and to redistribute wealth to poorer (non-migrant) households in sending communities in Kyrgyzstan (Chakraborty et al., 2015).

may be mistaken is that the occupation of migrants before they migrated may influence their choices of occupation upon return to their origin countries.Return migrants who were self-employed before they left might be less likely than non-migrants to choose self-employment as an occupation, if they discovered that they lacked sufficient entrepreneurial ability, as per Jovanovic's (1979) learning mechanism. This is consistent with a developing economy setting, where wage and self-employment coexist and individuals frequently transition between these two occupations (Koelle, 2016).[3]

If the reasons migrants choose self-employment on return are any of the above, the implications are that they will not sustain a career in self-employment. Any businesses they may create will not have a significant development impact. Return migrants may only be more likely to choose self-employment for entrepreneurial reasons if they discovered they had, or developed, entrepreneurial abilities. In addition, social capital deterioration would not be an insurmountable obstacle to thrive as entrepreneur. These conditions may be hard to hold, especially in developing economies, where market-supporting institutions are often absent. Hence, our hypotheses are that (i) if return migrants choose self-employment they may tend to do so as a temporary mechanism for eventually facilitating entry into wage employment or for future re-migration; and that (ii) they would be less likely to choose self-employment on return if they had been in self-employment before migration.

Very few studies have yet focused on these two hypotheses in integrated and consistent fashion. The occupational stability of migrants – whether return migrants sustain their entrepreneurial activities – has received little attention. To the best of our knowledge, only Marchetta (2012) has specifically studied the survival of returnees' entrepreneurial activities. *Entry* into and *exit* from self-employment have been either analysed separately or not specifically modeled as two separate decisions. There is thus a lacuna in the literature this study aims to fill, by expanding the evidence presented in particular by Marchetta (2012), Wahba and Zenou (2012), Wahba (2015) and Batista et al. (2017).

We intend to close these knowledge gaps by asking the following questions: (i) are return migrants more likely than non-migrants to enter into *and* to sustain a career in self-employment? And (ii) is this decision influenced by whether return migrants were self-employed before migrating?

We answer these questions by using a rich longitudinal database of a transition economy with significant international temporary migration, the 2010-2013 Life in Kyrgyzstan (LiK) Study. Although numbers vary by sources, emigration from Kyrgyzstan is, on all accounts,

[3] If they expect to receive better rewards as wage-employed, self-employed will shift to wage employment and shut down their firm. This very possibility to exit self-employment decreases incentives to invest in entrepreneurial activities, and so to persist in entrepreneurship over time (Koelle, 2016).

significant.[4] The structure of Kyrgyzstan's economy has led to high rates of emigration (Atamanov and van den Berg, 2012). Demographic pressure and (land) resource limitations coincided with economic opportunities in neighbouring countries to encourage migration. International migration became a natural response to economic challenges in Kyrgyzstan, to mainly Russia – hosting 92% of Kyrgyzstani migrants – and Kazakhstan – 8%. Migration is mainly economically driven and in essence temporary (NISI and IOM, 2016). Migrant workers to Russia do not plan to settle down there or to obtain permanent residence. They maintain close links with Kyrgyzstan by returning often and migrating again, and sending relatives remittances for their daily expenses or strategic needs such as housing, prestige goods or festivities.

The organisation of Soviet-era planned economies has simultaneously left Kyrgyzstan without market-supporting institutions, and difficult access to financial support. For instance, while, in 2014, individual entrepreneurs and small farmers contributed to respectively 18% and 9% of gross domestic product (GDP), they accounted for 90% of entrepreneurial activities;, and over 2001-2014, large enterprises were the main driver of GDP growth (Rudaz, 2017). Restrictions on private land ownership and state-led rent-seeking limit the growth of Kyrgyz family farms (Atamanov and van den Berg, 2012). Political instability, tax rates and corruption were recently listed as the first challenges that formal, non-agricultural enterprises faced in Kyrgyzstan (IBRD and World Bank, 2014). However, entrepreneurship has been found key to a successful transition from planned to market economy (McMillan and Woodruff, 2002).[5]

The extent of international migration in Kyrgyzstan, its dependence on destination country economic prospects and the role entrepreneurship could play in its unique transition setting point to the relevance of studying the potential implications of return migration for Kyrgyzstan's economic development through entrepreneurship. By following the same thousands of individuals across four consecutive years, the LiK can help shed light on these dynamics.

An important empirical challenge in studying the link between return migration and self-employment is that endogeneity might influence this relationship. Omitted variables could simultaneously explain migration and occupational choice upon return to origin communities. Emigrating itself, often viewed as a means to accumulate wealth, could help set up or expand

[4] With a population of about 5.7 million in 2013, the number of labour migrants approximated 200,000 to 1 million people depending on information sources. The subsequent growth in remittances has ranked Kyrgyzstan second worldwide after Tajikistan. Remittances represented about a third of its gross domestic product (GDP) in 2014 (Karymshakov et al., 2016).

[5] In China, Poland or Vietnam, new firms drove reforms by generating economic growth and jobs, offering goods and services otherwise inexistent, stimulating savings, and limiting the power of public firms (McMillan and Woodruff, 2002).

a business at home, when there are profitable investment opportunities in origin communities. Last, choosing and persisting in an occupation are likely to be related decisions. This suggests that a naïve estimation of the impact of return migration on self-employment might be biased.

To recover the causal effect of return migration on self-employment, we estimate a series of correlated random effects models. The robustness of these results is then assessed by resorting to matching techniques, akin to a matched difference-in-differences. Transition probabilities and baseline estimates indicate that return migrants in Kyrgyzstan are, as in other settings, more likely to be self-employed than non-migrants. However, their motive to do so is rather based on necessity than opportunity, and self-employment tends to be more of a temporary choice, as return migrants are more likely to exit from self-employment in subsequent periods than non-migrants. Correlated random effects model estimates suggest that, on average, return migration increases the likelihood of being self-employed by 18.5 percentage points for returnees who were self-employed in the previous time period; by 22.4, if they were not.

We also find that if return migrants were self-employed *before* migrating, they would be less likely to choose self-employment on their return. Matched difference-in-differences estimates reveal that, on average, return migration increases the likelihood of being self-employed by 1.19 percentage points for returnees who were self-employed before leaving Kyrgyzstan. This likelihood increases to 19.98 percentage points if they were not.

Baseline results are found to hold for individuals living in rural communities and own-account workers. However, return migration *increases* the likelihood to persist as employers, suggesting that migration can disrupt self-employment trajectories of limited economic prospect, but support job-generating entrepreneurial activities.

These novel findings contribute to the literature on the links between labour market outcomes and migration in a developing economy setting, and complement the growing empirical evidence for more advanced, destination countries.[6] By entailing a change in occupations, and, in particular, the exit from activities of potentially limited economic impact, migration could act thus as an 'experience good' (Nelson, 1970) that helps determine the quality of a worker's occupation-productivity match, by updating information on their tastes and abilities. This transition pattern could be viewed as part of a dynamic lifecycle sequence of sorting through gradual learning and experimentation, in which individuals self-select into their preferred options over time (Dillon and Stanton, 2017). Emigrating and returning to their source communities could lead return migrants who were initially self-employed, now better informed about their lower chances of success as entrepreneurs or with

[6] See for instance Ruiz and Vargas-Silva (2018) for evidence on the differences in labour market outcomes between natives, refugees and other migrants in the United Kingdom.

stronger interests in wage employment, to switch to wage employment upon return. In such a case, the 'disruption' caused by migration to self-employment can be viewed in a positive light. It enables career development by offering an escape from self-employment, and it reduces the number of entrepreneurs with low abilities in the entrepreneurial pool.

The rest of the paper is structured as follows. Section 2 sets out the estimation strategy and data. Section 3 presents the empirical results and robustness checks. Section 4 concludes.

2 Methodology

2.1 Estimation strategy

The empirical challenge in evaluating whether return migration influences entry into and persistence in self-employment is that endogeneity might drive this relationship. A naïve estimation of the impact of return migration on self-employment will be biased if, first, omitted variables simultaneously explain migration and occupational choice upon return to origin communities. For instance, both migration and self-employment require a higher propensity to take risk. Temporary migrants might self-select into temporary migration and self-employment, which might influence their chances of sustaining a career in self-employment. Second, there might be reverse causality if resources gained abroad help set up or expand a business at home, when profitable investment opportunities in origin communities were perceived before emigrating. Those self-employed might also decide to migrate if their entrepreneurial ventures were unsuccessful. Third, choosing and persisting in an occupation are likely to be related decisions.

To recover the causal effect of temporary migration on the decision to enter into and to exit self-employment, we proceed as in de Ree and Nillesen (2009) and Bleaney and Dimico (2011) by running a series of non-linear probability models, in which the outcome variable Y_{it} is a binary variable defined if a working-age (18-64-year-old) individual i has reported a primary occupation at time t. This variable takes value 1 if s/he is self-employed in year t, and zero otherwise – a (paid or unpaid) employee or member of a cooperative.[7]

Baseline model

The baseline specification follows what has usually been estimated in the literature, ignoring entry- and exit-specific dynamics:

[7] These three categories are lumped together because of their status as they all imply dependence on co-workers or supervisors, family members or not.

$$Y_{it} = \alpha_0 + \sum_{k=1}^{K} \beta_{0k} X_{kit} + u_{it} \tag{1}$$

where X_{kit} is a k-vector of explanatory variables, including 1 to model a constant, $returnee_{it}$, defined as a binary variable taking value 1 if an individual has lived abroad for at least one month since 2005, and $k-2$ exogenous variables. Exogenous variables control for individual- (gender, age, ethnicity, household headship, educational attainment), household- (household size, below 15 dependency ratio), location- and year-specific characteristics. α_j and β_{jk} are parameters to be estimated; u_{it}, an unobserved random disturbance. Standard errors are clustered at the household level to account for interdependence of labour market outcomes between household members.

Entry model

To disentangle the relationship between return migration and entrepreneurship in analysing entry into self-employment, individuals observed as self-employed in $t-1$ are excluded. Discarding observations for which $Y_{i,t-1} = 1$, the entry model can be written as the probability of being self-employed in t, conditional on not being self-employed in $t-1$ as:

$$Y_{it} = \alpha_1 + \sum_{k=1}^{K} \beta_{1k} X_{kit} + u_{it} \tag{2}$$

Estimating this model should provide an answer on whether return migrants are more likely to opt for self-employment compared to non-migrants.

'Survival' model

To analyse the likelihood for a return migrant to sustain a career in self-employment, individuals not observed as self-employed in $t-1$ are excluded. Conditioning the probability of self-employment at t on being self-employed in $t-1$ yields a 'survival' model:

$$Y_{it} = \alpha_2 + \sum_{k=1}^{K} \beta_{2k} X_{kit} + u_{it} \tag{3}$$

Combined model

While it is unlikely that $\alpha_1 = \alpha_2$, since an individual is more likely to be self-employed at time t if s/he was self-employed in $t-1$, the 'true' model might be such that $\beta_{1k} = \beta_{2k}$. The literature on business survival has tended to study enterprise survival on one hand, and on the other hand, the persistence of entrepreneurs' occupational choice, based on individual features. However, such a distinction is less clear when the sample of interest includes rather

micro and small units (Marchetta, 2012). We thus expect factors associated with continuing in self-employment to be similar to those associated with entry into self-employment, in particular in a context of small entrepreneurial units, as in our estimation sample. If this is the case, estimating equations (2) and (3) separately is inefficient, leading to relatively wide confidence intervals on estimated coefficients, specifically in smaller subsets of the data.

The model to be estimated on the whole data set should combine entry into and 'survival' in self-employment models, as in:

$$Y_{it} = \alpha_3 + \sum_{k=1}^{K} \beta_{3k} X_{kit} + \lambda_3 Y_{i,t-1} + u_{it} \tag{4}$$

Equation (4) reduces to equation (2) if $Y_{i,t-1} = 0$, and to equation (3) if $Y_{i,t-1} = 1$. However, equation (4) imposes β_{3k} to be the same in both cases and implies that the errors of the two different subsamples come from the same population. Relaxing this equality on the coefficients leads to the following model:

$$Y_{it} = \alpha_4 + \sum_{k=1}^{K} \beta_{4k} X_{kit} + \lambda_4 Y_{i,t-1} + \sum_{k=1}^{K} \theta_{4k} X_{kjt} Y_{i,t-1} + u_{it} \tag{5}$$

Equation (5) reduces to equation (4) if θ_{4k} is a vector of zeros. If it is not certain whether return migration is similarly related to entry into and persistence in self-employment, i.e. whether any element of θ_{4k} is zero, the appropriate model to estimate is (5) (combined model). We can then re-estimate this model, setting to zero any elements of θ_{4k} that are not significantly different from zero at the first stage (*parsimonious combined model*). Equation (5) allows testing parameter restrictions across entry and 'survival' model coefficient estimates.

Dealing with endogeneity

We account for potential endogeneity between return migration and self-employment by exploiting the longitudinal dimension of the data set to control for time-invariant unobserved heterogeneity. We do so by incorporating a Mundlak 'correction' (Mundlak, 1978). The Mundlak model assumes that the (individual) fixed effects are projected on the group means of time-varying variables, imposing a form on the relationship between the time-invariant random disturbance and the regressors. Regressions are estimated by adding individual-specific (group) means across time, and then estimated as random effects models. Rejecting the null hypothesis that the joint significance of the group means coefficient estimates is not different from zero is evidence of time-invariant unobserved heterogeneity at the individual level.

Since u_{it}, the unobserved random disturbance, has a time-invariant γ_i and a time-varying component ϵ_{it}, augmenting our baseline specification (equation (1)) with a Mundlak correction assumes that:

$$E[\gamma_i|X_{kit}] = g(X_{kit}) \tag{6}$$

where time-invariant effects are correlated with the other exogenous variables, X_{kit}. The expected value of γ_i, conditional on the exogenous regressors, is a linear combination of the average panel of the time-varying regressors. The Mundlak correction yields:

$$\gamma_i = \delta_0 + \sum_{k=1}^{K} \delta_{1k} \overline{X}_{ki} + \epsilon_i \tag{7}$$

$$Y_{it} = (\alpha_0 + \delta_0) + \sum_{k=1}^{K} \beta_{0k} X_{kit} + \sum_{k=1}^{K} \delta_{1k} \overline{X}_{ki} + u_{it} + \epsilon_i \tag{8}$$

where ϵ_i is assumed to be uncorrelated with X_i. Equation (5) is estimated as in Wooldridge (2005), who proposes an extension to the Mundlak estimator to estimate dynamic non-linear panel data models with unobserved heterogeneity. Wooldridge (2005) suggests modeling the distribution of the outcome variable $\{Y_{i1}, ..., Y_{iT}\}$, given Y_{i0} (its initial value) using conditional Maximum Likelihood (ML). The distribution of u_i, time-invariant, individual-specific heterogeneity is directly specified, given y_{i0} and other exogenous variables, such that:

$$g(\gamma_i|y_{i0}, X_{kit}) \sim N(\delta_0 + \delta_1 Y_{i0} + \sum_{k=1}^{K} \delta_{2k} \overline{X}_{ki}, \sigma_\epsilon^2) \tag{9}$$

This can be re-written as:

$$\gamma_i = \delta_0 + \delta_1 Y_{i0} + \sum_{k=1}^{K} \delta_{2k} \overline{X}_{ki} + \epsilon_i, \tag{10}$$

where $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ is independent of Y_{i0} and X_{ki}. As in Mundlak's approach, fixed effects are projected on the group means of time-varying variables. Regressions are run by regressing the outcome variable at time t on the set of explanatory variables, the individual-specific (group) means across time, the initial outcome value (at time $t = 1$, 2010), the lagged $t - 1$ outcome and the interaction of the lagged outcome with the set of explanatory variables, as in Bleaney and Dimico (2011). Estimated by a standard random effects probit, equation (5) can then be re-written as:

$$Y_{it} = (\alpha_4 + \delta_0) + \sum_{k=1}^{K} \beta_{4k} X_{kit} + \delta_1 Y_{i0} + \lambda_4 Y_{i,t-1} + \sum_{k=1}^{K} \theta_{4k} X_{kjt} Y_{i,t-1} + \sum_{k=1}^{K} \delta_{2k} \overline{X}_{ki} + u_{it} + \epsilon_i \quad (11)$$

We account for selection into reporting a primary occupation at time t, by adding to the below specifications Inverse Mills Ratios (IMR), based on the predictions of the probability to be observed working at time t. Our selection (dependent) variable is a binary variable taking unity if an individual has a primary occupation; zero, otherwise, i.e. either inactive or unemployed. The IMR for selection into working are calculated for each specification and then added to the corresponding models for self-employment probability. In line with the existing literature, we use being married as exclusion restriction.[8]

To conclude this subsection, we should indicate that controlling for individual fixed effects does not exclude measurement errors and time-varying endogeneity. Specifically, the explanatory variable of interest only records return migration up to 2005. There might be some return migrants in the data we are not able to identify. Finding a relevant and strong instrument will not prevent bias as we cannot identify migrants who migrated and returned to Kyrgyzstan before 2005. For this reason, we include time-varying variables and individual fixed effects in the benchmark specifications. We later estimate a matched difference-in-differences to check the robustness of these results.

2.2 Data source

We use data from the Life in Kyrgyzstan Study (LiK), a multi-topic longitudinal survey carried out annually from 2010 to 2013 in Kyrgyzstan, tracking the same households over time in all seven regions (oblasts) and the two major cities, Bishkek and Osh. Detailed information on the LiK can be found in Brück et al. (2014).

The unit of analysis is a working-age (18-64-year-old) individual born in Kyrgyzstan, interviewed in all four waves. This resulted in a total of 4,765 respondents across 2,195 original, non-splitting households. From the original sample of 3,000 households identified in 2010, 2,450 households (81.6%) participated in all four waves of the project.

The exclusion of 2,099 individuals from 557 households initially interviewed in 2010 could represent a threat to our analysis, if attrition across waves is structural, i.e. if the probability of attrition is different between returnees and non-migrants. For instance, return migrants who plan to re-migrate might take self-employment as a transitory occupation while waiting for future migration, and might not have any interest in lasting entrepreneurial activities.

[8] Estimates of sample selection regressions are available on request.

Alternatively, returnees whose entrepreneurial activities fail might be more likely to migrate again. In these cases, non-response might be selective. This would provide a non-random picture of the population and bias estimates.

In the absence of longitudinal and cross-sectional weights, and since this attrition problem can be viewed as a sample selection problem, we correct for potentially structural attrition in all specifications by adding IMR, based on the predictions of the probability of being interviewed in all four waves regressed on benchmark specification covariates as well as an exclusion restriction, (individual) perception of safety, on the 2010 sample of working-age individuals (cross-section). This exclusion restriction takes a value from 1 to 5, the greater its value, the less safe an individual feels walking alone in the neighbourhood at night.

As Table A1 shows, its statistical significance and negative sign suggests that the less safe an individual interviewed in the first wave of the survey feels, the less likely s/he will be successfully interviewed in subsequent waves. In other words, the more likely s/he might have left her place of residence during the first wave, or the more suspicious s/he might be, leading to refusal to being interviewed in subsequent waves.

The outcome of interest is a binary variable defined for individuals who reported a primary occupation in any industry, taking the value of 1 if an individual i is self-employed as primary occupation in year t, and zero otherwise. Independent variables include: return migration, defined as a binary variable taking value 1 if an individual has lived abroad for at least one month since 2005; gender; age and age squared; ethnicity (three major groups: Kyrgyz, Uzbek, Russian); household headship, marital status, household size and below 15 dependency ratio; secondary, vocational, university education; rural/urban location; and year-specific binary variables to capture wave-specific characteristics and control for year-specific date of fieldwork (month of interview was not recorded).

3 Empirical results

3.1 Descriptive results

Occupational status and profile of all respondents

Table 1 reports that 21.8% and 29.5% of the sample are respectively self- and wage-employed. This suggests that both sectors coexist in Kyrgyzstan, consistent with existing findings for developing economies (Koelle, 2016). Figure 1 plots the Kernel density estimates of self- and wage employment earning distributions. While Table A2 reports that self-employed earn on average relatively more than employees, this figure indicates that no sector provides earnings making it strictly superior.

In addition, earnings from self-employment show greater variability than earnings from wage employment. This could reflect some degree of unpredictability in self-employed activities in Kyrgyzstan and the coexistence of poverty- and opportunity-driven self-employment. This is supported by the fact that working-age individuals who reported being self-employed in some waves earn on average relatively less in a month (KGS8,311) than those who reported being self-employed in all waves of the survey (KGS6,082).

More generally, Table A2 indicates that those in the sample that report being 'Always' self-employed are older, more likely to be men, head of their household and married than individuals 'Sometimes' or 'Never' self-employed. 'Sometimes' self-employed tend to come from bigger households that have experienced relatively more shocks over the preceding 12 months. This may indicate that transitions between occupations occur out of necessity.

As for the labour market activity of self-employed respondents, Table A3 indicates that individuals who reported being 'Always' self-employed are slightly more likely to employ non-household members (3.2%) than 'Sometimes' self-employed (1.9%). They are also more likely to run registered businesses than those who report being 'Sometimes' self-employed (54.7% against 44.8%). They tend to work more in wholesale, retail and manufacturing sectors, while 'Sometimes' self-employed are more present in agriculture and construction. These statistics suggest that individuals who reported being self-employed in some waves of the survey, but not all, might be less successful in self-employment and establish less stable entrepreneurial activities than those who reported being self-employed in every wave.

Profile and occupational status of return migrants

Table 1 suggests that 6.30% of the estimation sample are return migrants. They have spent, on average, 13.22 months abroad. Relative to Kyrgyzstan's 5,362,816 inabitants in 2009, estimation sample figures suggest that there were 337,857 temporary labour migrants in Kyrgyzstan from 2010 to 2013. This is at the lower bound of the true number of migrants that vary from 200,000 to one million depending on the source, and, in particular, close to Chakraborty et al.'s (2015) 5% international migrants.

Table A4 indicates that return migrants are on average more likely to be self-employed compared to non-migrants, and Table A2, that return migrants are most present among those who report being 'Sometimes' (9.1%) and 'Always' (8.6%) self-employed. It is therefore no surprise that households with return migrants receive a greater share of their income from household entrepreneurial activities compared to individuals living in a non-migrant household, as shown in the upper panel of Table A5. Furthermore, apart from the heterogeneity present among self-employed described above, the almost equal representation of return migrants in these two self-employed types indicates some level of (self-employed)

diversity among return migrants.

While the LiK data set does not provide consistent information on migration motivations and experience abroad, external sources suggest that international migration is mainly driven by better job, earning opportunities abroad, and is temporary in nature (NISI and IOM, 2016). Kyrgyzstani migrants tend to work in manufacturing, retail and transport sectors; moreover, there has been a recent an increase (decrease) in unskilled (skilled) workers, who, in majority have only completed general secondary education.

Considering household expenditure patterns in the lower panel of Table A5, migration seems to be undertaken to meet consumption rather than 'productive' needs. For instance, households with return migrants are more likely to spend their income on 'celebrations' and 'clothing' than non-migrant households. This is consistent with existing qualitative research in Kyrgyzstan suggesting that a frequent reason for migration is to obtain funds to cover daily expenditures or events such as festivities (Thieme, 2014).

Occupational transitions over time

We consider respondents' labour market transitions across the four years of the survey, summarised in Tables 2 (full estimation sample), 3 (return migrants) and 4 (non-migrants). In these tables, a cell should be read as the probability in percentage to transition from a row employment status in year t to a column employment status in year $t + 1$. The main diagonal measures inertia – the proportion of individuals who stayed in the same occupation in two subsequent years.

While Table 2 shows general flexibility between occupations, the persistence of the working-age population in self-employment (64%) is lower than in wage employment (77%). Transitions thus appear more frequent from self- to wage employment (11%), and from unemployment to self- (11%) and wage employment (35.6%), than from wage to self-employment (7.9%). Last, we should note the relatively high propensity of inactivity, controlled for by including IMR, as explained in the previous section.

Return migrants (Table 3) have less stable occupations than non-migrants (Table 4), either over the four years ('ever' returnees, upper panel) or upon return (lower panel). While the persistence of both groups in self-employment is similar (around 63%), the persistence of returnees in wage employment is lower than that of non-migrants (about 59% against 78%). Moreover, if self-employed return migrants have a similar probability to transition to wage employment than non-migrants (about 10-11%), 18.7% of wage-employed return migrants are likely to be self-employed in the next year, compared to only 7.2% of wage-employed non-migrants. 18 and 28.2% of unemployed 'ever' returnees (upper panel) are likely to be respectively self- and wage-employed in the next year, compared to 10.3 and 36.5% of

unemployed never migrants. In contrast, of those unemployed return migrants upon return (lower panel), 9.1 and 40.9% are likely to be respectively self- and wage-employed in the next year, which is somewhat similar to never migrants.

Descriptive statistics indicate that self-employment in Kyrgyzstan is a rather transitory choice of occupation, consistent with Harris and Todaro's (1970) 'parking lot' hypothesis. It is also one seemingly more often selected by return migrants, who display lower occupational stability, both before and after migrating, than non-migrants. In the following subsections, we investigate the determinants of these occupational choices.

3.2 Benchmark regression results

In section 2, we presented our estimating strategy, deriving regression equations for a baseline model, a model for entry into self-employment, for 'survival' in self-employment and a combined model accounting for dynamics in and out of self-employment. Regression results for each of these models are presented in Table 5 and, in more details, in Table A6. Table A7 displays associated average marginal effects.

What can these estimates say about our research questions?

First, baseline model results (columns (1)-(2) of Table 5) show that the probability of being self-employed is significantly and positively correlated with return migration. Column (2) of Table A7 suggests that return migration increases the probability of being self-employed by 14.2 percentage points. This result is consistent with results elsewhere in the literature, as reported in section 1. Estimates are confirmed by the entry model (columns (3)-(4) of Table 5), indicating that opting for self-employment is positively correlated with return migration in both random and correlated random effects models. On the contrary, return migration is not significantly related to persistence in self-employment (columns (5)-(6)), which might be explained by the reduction in sample size – 2,620 observations out of 11,361 – inducing wide confidence intervals on individual coefficients.

Second, column (7) of Table 5 presents coefficient estimates of a combined model that takes into account dependence between the choice of and the duration in self-employment, as in Wooldridge (2005). Interactions between the lagged outcome variable, $Self\text{-}employed_{(t-1)}$, and all covariates are included.[9] Although coefficient estimates on interaction terms are jointly significantly different from their corresponding non-interacted variables, column (7) of Table 5 suggests that out of 16 interaction terms, 11 are statistically significant. Column (8) presents coefficient estimates of a parsimonious combined model including these 11

[9] This model assesses whether covariates are similarly related to entry into and persistence in self-employment. If these interaction terms are not significantly different from zero, we cannot reject the null hypothesis that they are equally associated with entry and persistence.

interaction terms, setting to zero any other interaction terms.

Combined and parsimonious combined models indicate that return migration is significantly and positively associated with entry into self-employment. Its interacted term, negatively associated with self-employment, is significant across both specifications and of a smaller magnitude than the non-interacted return migration term. This suggests that return migrants who were self-employed in $t-1$ are less likely to be self-employed in year t, i.e. to persist in self-employment over time.

Column (7) of Table A7 suggests that, on average, return migration increases the likelihood of entering into self-employment by 18.6 percentage points. While migrants who were not self-employed in $t-1$ are 23 percentage points more likely to be self-employed upon return at time t (column (8)), migrants who were self-employed in $t-1$ are 18.3 percentage points more likely to be self-employed at time t (column (9)). Marginal effects of coefficient estimates of a parsimonious combined model, columns (10)-(12), present similar estimates, as Figure 2 depicts.

With regard to our first research question, benchmark estimates indicate that return migrants are more likely to enter into self-employment than non-migrants in Kyrgyzstan. However, had they been self-employed in previous waves of the survey, they were found to be less likely to sustain a career in self-employment. The positive relationship between return migration and self-employment often found in the literature might thus be driven by entry into rather than persistence in self-employment. One likely reason for this we inferred from descriptive statistics is that self-employment can act as a 'parking lot', since labour market transitions are more frequent from self- to wage employment than from wage to self-employment. This indicates that self-employment in Kyrgyzstan is a rather transitory choice of occupation and one seemingly more often selected by return migrants than non-migrants.

3.3 Robustness checks

Matched difference-in-differences analysis

To assess the robustness of these estimates, we run benchmark specifications on a matched sample of control (non-returnees) and treated (returnees), following Egger et al. (2008) and Falvey and Foster-McGregor (2015). As the data set consists of four years (2010 to 2013), and there is no 'new' return migrant in 2013, we define controls as individuals who are not return migrants, and treated, as individuals who are reported as return migrants in year t, but who were not in the previous year, $t-1$, for each year t. Only those 'newly' treated in year t are used in the matching procedure; those existing treated are dropped.

Using either propensity score (PSM) or covariate matching (CVM), we match new

returnees to non-returnees at time t (2011 and 2012) on the basis of $t-1$ (respectively, 2010 and 2011) explanatory variables – working-age individuals, members of non-splitting households, who were born in Kyrgyzstan, and who reported a primary activity at time t. We obtain two-year pooled cross-sections of matched individuals, on which we run benchmark specifications applying matching weights as frequency weights. These 'matched' regressions should control for observed heterogeneity between returnees and non-migrants as well as self-selection into temporary migration. Although we cannot rule out the existence of reverse causality, estimates should bring potentially causal evidence on the effects of migration on self-employment propensity upon return in the short term.

PSM is successively applied with one nearest neighbour, five nearest neighbours, radius and kernel density matching, using a logit equation for the probability of being a return migrant;[10] CVM with one nearest neighbour and five nearest neighbours, with Abadie and Imbens's (2006) bias correction. The best matches are given using five nearest neighbours with PSM and CVM.[11] Matching quality is reported in Figures A1 and A2.

Table 6 presents coefficient estimates of specifications run with these two matched samples. Signs are similar to, but magnitudes and statistical significance slightly differ from, benchmark estimates. In particular, columns (4)-(5) present coefficient estimates of a 'combined' and parsimoniously combined model. Coefficient estimates on return migration and its interaction with lagged self-employment are statistically significant and differ in sign. This confirms benchmark estimates in showing that, controlling for observed heterogeneity and self-selection into temporary migration, migration increases the propensity to enter into self-employment, but decreases (increases) the probability to persist in (exit from) self-employment in the short term, in line with the 'parking lot' hypothesis.

Column (1) of Table A8 suggests that return migration significantly increases self-employment propensity by 10.94 to 12.40 percentage points on average. This is driven by the greater probability of entering into than of persisting in self-employment. When migrants had not been self-employed before leaving Kyrgyzstan, they were 19.98 to 20.09 percentage points more likely to be self-employed (column (2)). When they had been, they were however 4.32 less to 1.19 more likely to be self-employed (non-significant). Figure 3 presents marginal effects of return migration on the self-employment probability of the last specification, for both matched samples. Although marginal effects only give information at distribution average, these estimates confirm that temporary migration increases the likelihood of opting for self-employment upon return, but increases the likelihood of exiting from self-employment

[10] Since the matching procedure is performed to use weights that are generated to then run specifications, but not to compute average treatment effects of return migration on self-employment propensity, standard errors are not bootstrapped.

[11] Matching estimates for alternative techniques are available on request.

upon return, if returnees had been self-employed *before* leaving Kyrgyzstan.

Matched difference-in-differences estimates provide an answer to the second research question we asked. Return migrants were found to be more likely to enter into self-employment than non-migrants if they had not been self-employed *before* leaving Kyrgyzstan. If they had been, they were found to be less likely to be self-employed upon return. This result suggests that temporary migration could be more disruptive of self-employment trajectories than is often thought.

Household-level analysis

We then exploit the fact that the decision to migrate tends to be taken at the household level. Remittances sent to non-migrating household members or resources repatriated upon return such as savings could induce newly reunited families to switch to entrepreneurial activities or invest in existing ventures. Not only have migrants returning to their origin country been found to be more likely to start their own enterprises, they have also been found to stimulate the entrepreneurship of non-migrant family members, through spill-over effects (Mansuri, 2007; Giulietti et al., 2013).

In Kyrgyzstan, where international migration is typically temporary and involves repeated episodes, we could expect a household with returnees to be more likely to set up a family enterprise that might survive future migration episodes of its members. Indeed, descriptive statistics suggest that non-migrants living in a household with returnees are more likely to contribute to family work than those who do not (Table A4), and that they gain relatively more income from household enterprises (Table A5).

We run the above specifications on a panel of households. Potential endogeneity between migration and family enterprise is addressed by assuming that endogeneity mainly comes from unobserved time-invariant household characteristics, following Antman (2015).[12] We control for household panel attrition by including IMR as in previous regressions.

Tables 7 and 8 present coefficient estimates respectively on a full and matched estimation sample. Estimates of household specifications are similar to results of individual specifications, except for combined models. If a household has members who returned from abroad, having at least one self-employed member in $t-1$ does not affect the probability of at least one member being self-employed at time t any differently from if they had not had any self-employed member in $t-1$. However, estimates with matched samples show a statistically different effect, in line with individual specifications. This, in addition to descriptive statistics, suggests that, although households with returnees are more likely to

[12] Although we should remain cautious about any causal interpretations (see e.g. Steinmayr's (2015) case for invisible sample selection).

have members contributing to family work, they are less likely to persist in self-employment in the short term.

3.4 Heterogeneous effect analysis

Last, with reference to benchmark specifications, the sample is split up by location and type of self-employment, to understand which subgroups drive the effect of migration on self-employment we found.

Rural versus urban

Job opportunities in rural areas tend to differ from those in urban areas because farming might not be an option in the latter, while in the former, it might be the main activity. In addition, self-employment in cities might not be the most desirable option if wage works are offered and are more socially rewarded, as it is the case in Kyrgyzstan.

Table A2 indicates that of those 'always' self-employed, 32.53% live in urban communities, 22.62% of 'sometimes' self-employed and 44.25% of never self-employed; Table A4, that 31.72% of return migrants live in urban communities, compared to 36.75% of non-migrants. Table 9 suggests that baseline estimates hold for rural areas. However, temporary migration appears to increase both entry into and persistence in self-employment in cities. Although coefficient estimates are not statistically significant, this could reflect the heterogenity among self-employed, including both necessity- and poverty-driven self-employed in urban communities.

Own account workers versus employers

The self-employed are likely to be a heterogeneous category. It is composed of both own-account workers and employers, who often have different features. Moreover, not all entrepreneurial activities have lasting impacts on economic development. Being self-employed might not be a good indicator of entrepreneurship, since most self-employed neither innovate much nor generate jobs; many fail. The creation of jobs as self-employed might however be a precondition for a lasting, positive effect of migrants' activities upon return, in particular in a developing country context, where the turnover of MSEs is high (Marchetta, 2012).

Table A3 shows that among all self-employed, 97.72% are own-account workers; 2.28% are employers. Table 9 suggests that benchmark estimates hold for own-account workers, but temporary migration tends to *increase* persistence as employers upon return to Kyrgysztan. This implies that temporary migration could disrupt self-employment trajectories of limited

economic impact, and support job-generating entrepreneurial activities, more likely to stimulate origin countries' economic development. This is in line with Piracha and Vadean (2010), who find that, in Albania, return migrants are significantly more likely to create jobs as self-employed, and less likely to be own-account workers who have characteristics closer to non-participants in the labor market.

4 Concluding remarks

This paper questions and qualifies the assumption that return migration stimulates entrepreneurship in origin countries, by simultaneously dealing with the interdependence between entry into and exit from self-employment, and temporary migration and occupational choice, in a transition economy with prevalent international temporary migration, Kyrgyzstan.

Baseline estimates indicate that return migrants in Kyrgyzstan are found more likely to be self-employed than non-migrants, as in other contexts. One likely reason for this is that self-employment acts as a 'parking lot'. Labour market transitions appear more frequent from self- to wage employment than from wage to self-employment, and return migrants less likely to persist in self-employment over time compared to non-migrants. This suggests that self-employment in Kyrgyzstan is a rather transitory choice of occupation, and one seemingly more often selected by temporary migrants than non-migrants, either before migrating or upon return to their origin countries.

Robustness checks reveal that the decision of a return migrant to be self-employed is negatively affected if s/he had been self-employed *before* migrating. Analysing the heterogeneity of these effects, benchmark results are found to hold for individuals living in rural communities and own-account workers. However, return migration is found to *increase* the likelihood to persist as employer.

These novel findings suggest that migration can disrupt self-employment trajectories of limited economic prospect, but support those entrepreneurial activities that might stimulate origin countries' economic development through job creation. The finding that, if migrants were own-account workers before leaving, they are less likely to be so upon return is a 'disruption' that can be viewed in a positive light as it enables career development by offering an escape from poverty-driven self-employment. Temporary migration might help reveal preferences for wage employment or a better allocation of workers' capabilities as wage-employed. This 'disruption' thus reduces the number of self-employed with low abilities in the entrepreneurial pool. In this case, the appropriate policy response is to support smooth occupational transitions in labour markets and fast reintegration of return migrants to help

them make the best use of their resources and reduce their need for self-employment out of necessity.

However, the finding that self-employment might be a temporary occupational choice suggests that the 'disruption' caused by migration may hinder an economy from benefiting entrepreneurially from the experience and resources migrants might have accumulated abroad. In this case, the appropriate policy response is to improve the conditions for doing business in source countries. In a context of transition from planned to market economy, entrepreneurial success requires a mix of microeconomic reform, macroeconomic and institutional stability and efficiency. The evidence from Kyrgyzstan we presented suggests that temporary migration might substitute for an imperfect legal framework and weak financial markets. As such, support for formal market-supporting institutions is advised if firms are to grow, and countries are to harness the entrepreneurial acumen migrants might have accumulated abroad.

Figure 1: Kernel density estimates of earning distributions

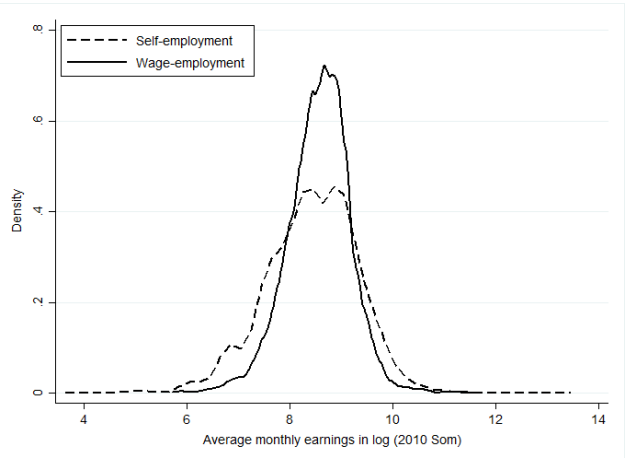

Figure 2: Average marginal effects of return migration with 95% confidence intervals of parsimonious combined model

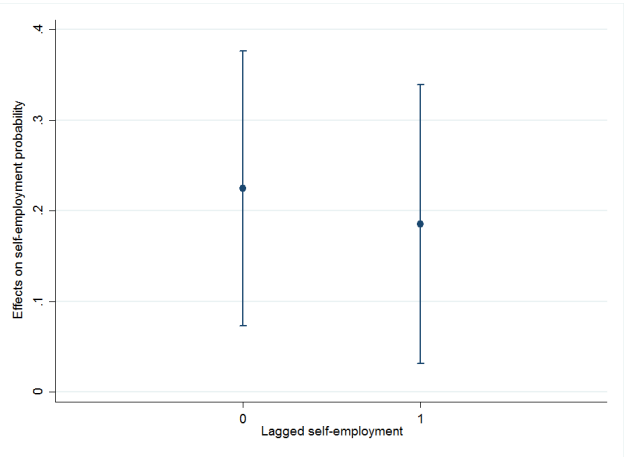

Figure 3: Average marginal effects of return migration with 95% confidence intervals of parsimonious combined model on PSM (l) and CVM (r) matched sample

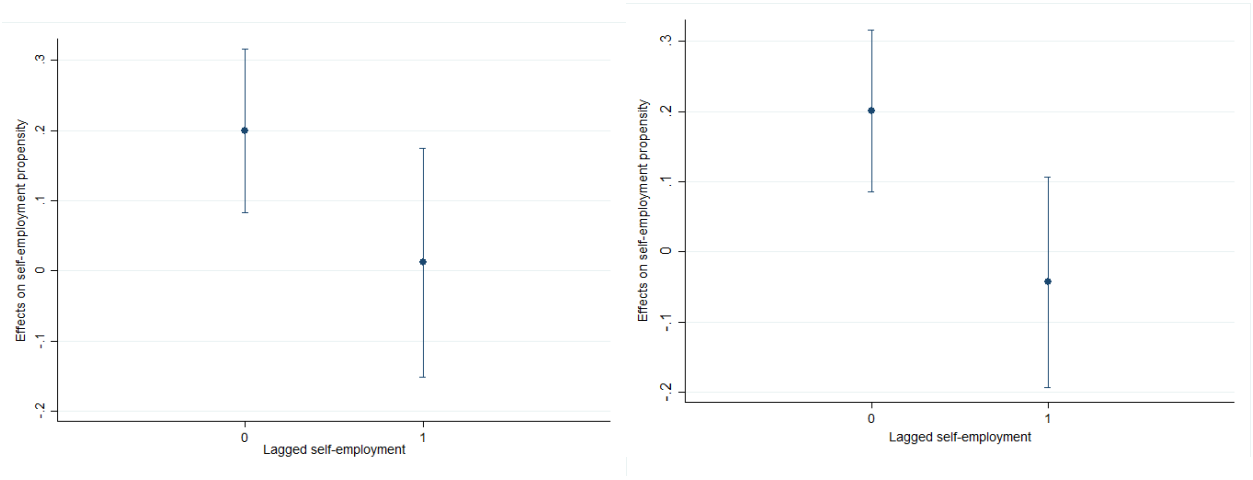

Table 1: Descriptive statistics of estimation sample

	Mean	S.D.
Individual		
Male	0.4717	0.4992
Age	39.5579	12.3972
Head	0.3612	0.4804
Married	0.7542	0.4305
Kyrgyz	0.6944	0.4607
Uzbek	0.1393	0.3463
Russian	0.0675	0.2508
Vocational	0.1572	0.3640
Secondary	0.5734	0.4946
University	0.1739	0.3790
Has worked over last week?	0.6002	0.4899
Self-employed	0.2182	0.4130
Wage-employed	0.2947	0.4559
Family worker	0.0846	0.2783
Has ever lived abroad	0.0630	0.2430
Number of months abroad	13.2198[a]	19.8593
Household		
Household size	5.6549	2.4114
0-15 dependency ratio	0.2689	0.2023
Living with returnee(s)	0.1637	0.3701
Member(s) currently working abroad	0.1287	0.3349
Owns land	0.7715	0.4199
Total land area (ha.)	0.8080	1.6696
Land area distributed (ha.)	0.7006	1.5892
Number of shocks	1.9453	2.2251
Location		
Urban	0.3644	0.4813
Total	19,060	

Notes: Means and standard deviations (S.D.) of variables of interest of balanced panel estimation sample of 4,765 individuals observed each four years of the survey.
[a]Based on 1,201 return migrants.

Table 2: Transition probabilities between occupations of all respondents

Employment status t	Employment status $t+1$					
	Inactive	Unemployed	Self-employed	Wage-employed	Other	Total
Inactive	71.29	2.68	8.74	10.59	6.69	100.00
Unemployed	36.67	13.85	11.03	35.64	2.82	100.00
Self-employed	19.67	0.94	63.97	10.95	4.47	100.00
Wage-employed	11.73	1.69	7.93	77.02	1.64	100.00
Other	33.74	0.85	12.29	15.55	37.57	100.00
Total $t+1$	36.92	2.11	21.71	31.66	7.60	100.00

Notes: Transition frequencies and probabilities between occupations of estimation sample. Rows reflect the initial t values; columns reflect the final $t+1$ values. *Inactive* includes individuals who have neither worked in the last seven days, nor have looked for a job. *Unemployed* includes individuals who have not worked in the last week, but have been looking for a job. *Self-employed* includes own-account workers and employers. *Wage-employed* includes wage-employed and members of a producer's cooperative. *Other* includes unpaid labour and unspecified.

Table 3: Transition probabilities between occupations of returnees

Employment status t	Employment status $t+1$					
	Inactive	Unemployed	Self-employed	Wage-employed	Other	Total
			Ever returnees			
Inactive	61.49	2.80	14.91	13.98	6.83	100.00
Unemployed	35.90	15.38	17.95	28.21	2.56	100.00
Self-employed	19.72	0.83	63.89	11.39	4.17	100.00
Wage-employed	15.81	1.72	18.56	59.11	4.81	100.00
Other	41.84	0.00	15.31	16.33	26.53	100.00
Total $t+1$	33.33	2.07	31.89	25.68	7.03	100.00
			Returnees upon return			
Inactive	65.04	2.65	13.72	10.62	7.96	100.00
Unemployed	36.36	9.09	9.09	40.91	4.55	100.00
Self-employed	22.18	0.70	63.03	10.21	3.87	100.00
Wage-employed	18.43	1.38	16.59	58.99	4.61	100.00
Other	42.68	0.00	12.20	17.07	28.05	100.00
Total $t+1$	35.26	1.56	31.05	24.55	7.58	100.00

Notes: See notes Table 3. The upper panel present transitions of ever returnees, i.e. individuals observed as return migrants at some point in time in the data; the lower panel, of individuals migrants specifically upon return.

Table 4: Transition probabilities between occupations of non-returnees

Employment status t	Employment status $t+1$					
	Inactive	Unemployed	Self-employed	Wage-employed	Other	Total
			Never returnees			
Inactive	71.98	2.68	8.31	10.35	6.68	100.00
Unemployed	36.75	13.68	10.26	36.47	2.85	100.00
Self-employed	19.67	0.95	63.98	10.90	4.50	100.00
Wage-employed	11.43	1.69	7.15	78.33	1.41	100.00
Other	33.13	0.92	12.06	15.50	38.40	100.00
Total $t+1$	37.22	2.11	20.86	32.17	7.65	100.00
			Never returnees and returnees before migration			
Inactive	71.85	2.67	8.39	10.46	6.62	100.00
Unemployed	36.94	13.89	10.56	35.83	2.78	100.00
Self-employed	19.57	0.97	63.93	11.04	4.48	100.00
Wage-employed	11.43	1.70	7.16	78.31	1.40	100.00
Other	33.08	0.91	12.09	15.51	38.40	100.00
Total $t+1$	37.16	2.13	20.93	32.17	7.61	100.00

Notes: See notes Table 3. The upper panel present transitions of never returnees, i.e. individuals observed as non return migrants in all years; the lower panel, of individuals observed as non-migrants at some point in time, e.g. never returnees and returnees who might have left after the first, second or third wave of the survey, and observed as returnees in subsequent waves.

Table 5: Coefficient estimates of benchmark specifications

	Baseline		Entry		'Survival'		Combined model	Parsimonious combined model
	RE	CRE	RE	CRE	RE	CRE	CRE	CRE
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Returnee	0.5659***	0.6244***	0.5571***	0.8097*	0.1200	0.5563	0.9162***	0.8977***
	(0.1436)	(0.2424)	(0.1461)	(0.4141)	(0.1351)	(0.3905)	(0.2873)	(0.2858)
Self-employed$_{(t=0)}$							1.0602***	1.0610***
							(0.1246)	(0.1237)
Self-employed$_{(t-1)}$							-0.4154	1.0144***
							(0.9229)	(0.2378)
... X Returnee							-0.2866*	-0.2678*
							(0.1597)	(0.1599)
Control variables	No	Yes	No	Yes	No	Yes	Yes	Yes
Time fixed effects	No	Yes	No	Yes	No	Yes	Yes	Yes
$IMR_{retention}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$IMR_{working}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group means	No	Yes	No	Yes	No	Yes	Yes	Yes
Interaction terms	No	No	No	No	No	No	All	Returnee, Male, Ethnicity, Vocational, University, Household size, Urban, Year
lnsig2u	1.2246***	1.2462***	0.3376*	0.3683*	-0.8696*	-0.7797*	-0.2255	-0.2156
	(0.0718)	(0.0728)	(0.2036)	(0.2045)	(0.4559)	(0.4463)	(0.2137)	(0.2119)
$\overline{X}_{kit} = 0$		54.99		21.96		144.91	20.19	21.18
		0.0000		0.0560		0.0000	0.0907	0.0694
$\beta_k = \theta_k$							46.92	34.47
							0.0000	0.0003
Observations	11,361	11,361	6,031	6,031	2,620	2,620	8,651	8,651
Number of groups	3,849	3,849	3,044	3,044	1,371	1,371	3,736	3,736

Notes: The dependent variable is a binary variable taking unity if a working-age individual is self-employed; 0, if employed, wage-employed or unpaid, or member of a cooperative. Observations are for working-age individuals, members of non-splitting households, who were born in Kyrgyzstan. Columns (1)-(2) present coefficient estimates of probit model of equation (1); columns (3)-(4) of equation (2); columns (5)-(6) of equation (3); and columns (7)-(8) coefficient estimates of a dynamic non-linear probability model with unobserved heterogeneity, as in equation (5). Tests of joint significance of group means and interaction terms when applicable report χ^2 with associated statistical significance. Standard errors clustered at the household level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Coefficient estimates of benchmark specifications on matched sample

Variables	Baseline (1)	Entry (2)	'Survival' (3)	Combined model (4)	Parsimonious combined model (5)
			Propensity score matching, 5 NN		
Returnee	0.2343	0.6435***	-0.0302	0.6462***	0.6325***
	(0.1456)	(0.1888)	(0.2737)	(0.1882)	(0.1841)
Self-employed$_{(t=0)}$				0.7408***	0.7357***
				(0.2135)	(0.2103)
Self-employed$_{(t-1)}$				-3.9184	1.2250***
				(2.4404)	(0.2094)
... X Returnee				-0.6415*	-0.5931*
				(0.3366)	(0.3344)
Control variables	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Interaction terms	No	No	No	All	Returnee Vocational Year
$\beta_k = \theta_k$				23.65	16.55
				0.0345	0.0009
Observations	1,190	743	434	1,177	1,190
			Covariate matching, 5 NN		
Returnee	0.1601	0.6274***	-0.3648	0.6195***	0.5895***
	(0.1414)	(0.1785)	(0.2905)	(0.1772)	(0.1792)
Self-employed$_{(t=0)}$				0.5734***	0.5927***
				(0.2220)	(0.2259)
Self-employed$_{(t-1)}$				-3.6235	1.4785***
				(2.3462)	(0.2361)
... X Returnee				-0.8953**	-0.7480**
				(0.3498)	(0.3316)
Control variables	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Interaction terms	No	No	No	All	Returnee Vocational Urban Year
$\beta_k = \theta_k$				30.10	20.88
				0.0074	0.0003
Observations	1,190	762	428	1,190	1,190

Notes: The dependent variable is a binary variable taking unity if a working-age individual is self-employed; 0, if employed, wage-employed or unpaid, or member of a cooperative. Observations are for working-age individuals, members of non-splitting households, who were born in Kyrgyzstan. In the upper panel of the Table, observations form a matched sample of control (non-returnees) and treated (returnees) applying propensity score matching techniques (5 nearest neighbours). In the lower panel, observations form a matched sample of control (non-returnees) and treated (returnees) applying covariate matching techniques (Mahalanobis metric, 5 nearest neighbours). Column (1) presents coefficient estimates of probit model of our baseline specification; column (2) of entry into self-employment; column (3) of persistence in self-employment; and columns (4)-(5) coefficient estimates of a dynamic, non-linear probability model. Tests of joint significance of interaction terms when applicable report χ^2 with associated statistical significance. Standard errors clustered at the household level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Coefficient estimates of household benchmark specifications

	Baseline		Entry		'Survival'		Combined model	Parsimonious combined model
Variables	RE (1)	CRE (2)	RE (3)	CRE (4)	RE (5)	CRE (6)	CRE (7)	CRE (8)
Returnee	0.2666***	0.2598***	0.2447***	0.2361***	0.1155	0.1148	0.2347***	0.2379***
	(0.0688)	(0.0700)	(0.0846)	(0.0860)	(0.0829)	(0.0851)	(0.0836)	(0.0833)
Family enterprise$_{(t=0)}$							0.5367***	0.5408***
							(0.0897)	(0.0888)
Family enterprise$_{(t-1)}$							0.7622***	0.7496***
							(0.2837)	(0.1763)
... X Returnee							-0.0728	-0.0781
							(0.1074)	(0.1070)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interaction terms	No	No	No	No	No	No	All	Returnee Household size Urban Years
lnsig2u	0.1909***	0.2022***	-1.2086***	-1.1838***	-0.7704**	-0.7684**	-1.1958***	-1.2016***
	(0.0673)	(0.0672)	(0.3582)	(0.3548)	(0.3019)	(0.3039)	(0.3269)	(0.3226)
$\overline{X}_{kit} = 0$		26.13		10.70		5.93	6.11	6.48
		0.0062		0.4688		0.8782	0.8657	0.8392
$\beta_k = \theta_k$							105.64	113.59
							0.0000	0.0000
Observations	9,112	9,112	3,587	3,587	3,244	3,244	6,831	6,831
Number of households	2,282	2,282	1,613	1,613	1,489	1,489	2,280	2,280

Notes: The dependent variable is a binary variable taking unity if at least one working-age individual is self-employed in a household; 0, if not. Observations are for non-splitting households, whose heads were born in Kyrgyzstan. Columns (1)-(2) present coefficient estimates of probit model of equation (1); columns (3)-(4) of equation (2); columns (5)-(6) of equation (3); and columns (7)-(8) coefficient estimates of a dynamic, non-linear probability model with unobserved heterogeneity, as in equation (5). Tests of joint significance of group means and interaction terms when applicable report χ^2 with associated statistical significance. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Coefficient estimates of household specifications on matched sample

Variables	Baseline (1)	Entry (2)	'Survival' (3)	Combined model (4)	Parsimonious combined model (5)
			Propensity score matching, 5 NN		
Returnee	0.1042*	0.4431***	0.1289	0.4362***	0.4390***
	(0.0546)	(0.0823)	(0.0865)	(0.0823)	(0.0821)
Self-employed$_{(t=0)}$				0.4008***	0.3994***
				(0.0845)	(0.0831)
Self-employed$_{(t-1)}$				2.5299***	2.2685***
				(0.3407)	(0.2650)
... X Returnee				-0.3002**	-0.2904**
				(0.1201)	(0.1198)
Control variables	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Interaction terms	No	No	No	All	Returnee Age Household size Location Year
$\beta_k = \theta_k$				75.70	71.86
				0.0000	0.0000
Observations	2,274	1,081	1,193	2,274	2,274
			Covariate matching, 5 NN		
Returnee	0.0038	0.3333***	-0.0241	0.3431***	0.3399***
	(0.0550)	(0.0817)	(0.0887)	(0.0816)	(0.0815)
Self-employed$_{(t=0)}$				0.3725***	0.3764***
				(0.0838)	(0.0836)
Self-employed$_{(t-1)}$				2.5849***	2.3635***
				(0.3363)	(0.2849)
... X Returnee				-0.3658***	-0.3486***
				(0.1211)	(0.1210)
Control variables	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Interaction terms	No	No	No	All	Returnee Vocational Urban Year
$\beta_k = \theta_k$				84.44	72.56
				0.0000	0.0000
Observations	2,270	1,093	1,177	2,270	2,270

Notes: The dependent variable is a binary variable taking unity if a working-age individual is self-employed in a household; 0, otherwise. Observations are for non-splitting households, whose heads were born in Kyrgyzstan. In the upper panel of the Table, observations form a matched sample of control (households with non-returnees) and treated (with returnees) applying propensity score matching techniques (5 nearest neighbours). In the lower panel, observations form a matched sample of control and treated applying covariate matching techniques (Mahalanobis metric, 5 nearest neighbours). Column (1) presents coefficient estimates of probit model of our baseline specification; column (2) of entry into self-employment; Column (3) of persistence in self-employment; and columns (4)-(5) coefficient estimates of a dynamic, non-linear probability model. Tests of joint significance of interaction terms when applicable report χ^2 with associated statistical significance. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Effect heterogeneity

	Location		Self-employed	
	Rural	Urban	Own account	Employers
Variables	(1)	(2)	(3)	(4)
Returnee	0.8799***	0.7461	0.8747***	-0.5915
	(0.3329)	(0.6894)	(0.2752)	(0.7770)
Self-employed$_{(t=0)}$	.8712***	2.1023***	1.0406***	1.2749
	(0.1268)	(0.5415)	(0.1218)	(0.9626)
Self-employed$_{(t-1)}$	-0.2416	2.6261	0.0800	-2.8966
	(1.002)	(2.4975)	(0.9069)	(13.9945)
... X Returnee	-0.3672*	0.4886	-0.2701*	3.0760**
	(0.1934)	(0.3864)	(0.1592)	(1.4351)
Control variables	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
$IMR_{retention}$	Yes	Yes	Yes	Yes
$IMR_{working}$	Yes	Yes	Yes	Yes
Group means	Yes	Yes	Yes	Yes
Interaction terms	All	All	All	All
lnsig2u	-0.2198	0.4083	-0.2525	-0.8639
	(0.2255)	(0.4917)	(0.2149)	(2.7894)
$\overline{X}_{kit} = 0$	25.79	9.64	19.34	5.49
	0.0115	0.6471	0.1128	0.9628
$\beta_k = \theta_k$	42.80	9.00	47.30	18.45
	0.0074	0.8308	0.0001	0.1412
Observations	5,382	3,012	8,651	6,205
Number of groups	2,415	1,259	3,736	3,561

Notes: The dependent variable is a binary variable taking unity if a working-age individual is self-employed; 0, if employed, wage-employed or unpaid, or member of a cooperative. Observations are for working-age individuals, members of non-splitting households, who were born in Kyrgyzstan. Columns (1)-(6) present coefficient estimates of a dynamic non-linear probability model with unobserved heterogeneity, as in equation (5). Tests of joint significance of group means and interaction terms when applicable report χ^2 with associated statistical significance. Standard errors clustered at the household level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Funding

This work was supported by the DFID/IZA GLM-LIC research grant 'Gender and Employment in Central Asia: Evidence from Panel Data' [grant number GA-C1-RA5-064]; and the European Union under the Marie Curie Initial Training Network (ITN), Transnational Migration, Citizenship and the Circulation of Rights and Responsibilities (TRANSMIC) (FP7-PEOPLE-2013-ITN) [grant number 608417].

Acknowledgements

The authors are grateful to various colleagues, in particular Neil Foster-McGregor, Eleonora Nillesen, Sergio Parra-Cely, and Pui-Hang Wong, as well as participants of the 2017 TRANSMIC conference, in Florence, Italy; of the 2017 Life in Kyrgyzstan conference, in Bishkek, Kyrgyzstan; of the 2017 GLM/LIC Research Network Conference, in Washington (DC), USA; of the 2017 AASLE Conference, in Canberra, Australia; and of the 2018 IZA/World Bank/NJD Conference on Jobs and Development, in Bogotá, Colombia. The usual disclaimer applies.

References

Abadie, A. and Imbens, G. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267.

Antman, F. (2015). Gender discrimination in the allocation of migrant household resources. *Journal of Population Economics*, 28(3):565–592.

Atamanov, A. and van den Berg, M. (2012). Heterogeneous effects of international migration and remittances on crop income: Evidence from the Kyrgyz Republic. *World Development*, 40(3):620–630.

Batista, C., McIndoe-Calder, T., and Vincente, P. (2017). Return migration, self-selection and entrepreneurship. *Oxford Bulletin of Economics and Statistics*, 79(5):797–821. Lisbon: Nova Africa Center for Business and Economic Development.

Bleaney, M. and Dimico, A. (2011). How different are the correlates of onset and continuation of civil wars? *Journal of Peace Research*, 48(2):145–155.

Brück, T., Esenaliev, D., Kroeger, A., Kudebayeva, A., Mirkasimov, B., and Steiner, S. (2014). Household survey data for research on well-being and behavior in Central Asia. *Journal of Comparative Economics*, 42:819–835.

Chakraborty, T., Mirkasimov, B., and Steiner, S. (2015). Transfer behaviour in migrant sending communities. *Journal of Comparative Economics*, 43(3):690–705.

de Ree, J. and Nillesen, E. (2009). Aiding violence or peace? The Impact of foreign aid on the risk of civil conflict in sub-Saharan Africa. *Journal of Development Economics*, 88:301–313.

Dillon, E. and Stanton, C. (2017). Self-employment dynamics and the returns to entrepreneurship. *NBER Working Paper No. 23168*.

Dustmann, C. and Görlach, J.-S. (2016). The Economics of temporary migrations. *Journal of Economic Literature*, 54(1):98–136.

Egger, H., Egger, P., and Greenaway, D. (2008). The Trade structure effects of endogenous reregion trade agreements. *Journal of International Economics*, 74:278–298.

Falvey, R. and Foster-McGregor, N. (2015). North-South FDI and bilateral investment treaties. *UNU-MERIT Working Papers No. 2015-010*. Maastricht: UNU-MERIT.

Giulietti, C., Wahba, J., and Zimmermann, K. (2013). Entrepreneurship of the left behind. *IZA Discussion Paper No. 7270*. Bonn: Institute for the Study of Labor.

Harris, J. and Todaro, M. (1970). Migration, unemployment and development: A Two-sector analysis. *American Economic Review*, 60(1):126–142.

IBRD and World Bank (2014). Kyrgyzstan country profile 2013: Enterprise survey. Technical report. Washington, D.C.: The World Bank.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, 87(5):972–990.

Karymshakov, K., Abdieva, R., Sulaimanova, B., and Sultakeev, K. (2016). Remittances impact on youth labour supply: Evidence from Kyrgyzstan. *PEP Working Paper 2016-05*. Nairobi: Partnership for Economic Policy.

Koelle, M. (2016). Firm dynamics and occupational choice. *Beitrïge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Labor Economics: Theory, No. E20-V2*.

Mafukidze, J. (2006). A Discussion of migration and migration patterns and flows in Africa. *In C. Cross, D. Gelderblom, N. Roux, and J. Mafukidze (eds). Views on Migration in Sub-Saharan Africa: Proceedings of An African Migration Alliance Workshop*, pages 103–29.

Mansuri, G. (2007). Does work migration spur investment in origin communities? Entrepreneurship, schooling and child health in rural Pakistan. In Özden, C. and Schiff, M., editors, *International migration, economic development and policy*, pages 99–140. World Bank and Palgrave MacMillan.

Marchetta, F. (2012). Return migration and the survival of entrepreneurial activities in Egypt. *World Development*, 40(10):1999–2013.

McMillan, J. and Woodruff, C. (2002). The Central role of entrepreneurs in transition economies. *Journal of Economic Perspectives*, 16(3):153–170.

Naudé, W., Siegel, S., and Marchand, K. (2017). Migration, entrepreneurship and development: Critical questions. *IZA Journal of Migration*, 6(5).

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2):311–329.

NISI and IOM (2016). Kyrgyzstan extended migration profile 2010-2015. Technical report. Bishkek: International Organisation for Migration.

Piracha, M. and Vadean, F. (2010). Return migration and occupational choice: Evidence from Albania. *World Development*, 38(8):1141–1155.

Reinhold, S. and Thom, K. (2013). Migration experience and earnings in the Mexican labor markets. *Journal of Human Resources*, 48(3):768–820.

Rudaz, P. (2017). The State of MSME development in Kyrgyzstan. *UNCTAD Working Paper Series on Informal Markets and Trade, No. 3*.

Ruiz, I. and Vargas-Silva, C. (2018). Differences in labour market outcomes between natives, refugees and other migrants in the UK. *Journal of Economic Geography,*, 18(4):855–885.

Steinmayr, A. (2015). When a random sample is not random. Bounds on the effect of migration on household members left behind. *Mimeo*.

Thieme, S. (2014). Coming home? Patterns and characteristics of return migration in Kyrgyzstan. *International Migration*, 52(5):127–143.

Wahba, J. (2015). Selection, selection, selection: The Impact of return migration. *Journal of Population Economics*, 28:535–563.

Wahba, J. and Zenou, Y. (2012). Out of sight, out of mind: Migration, entrepreneurship and social capital. *Regional Science and Urban Economics*, 42:890–903.

Wooldridge, J. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20(1):39–54.

For online publication

Appendices

A Appendix

Table A1: Investigating panel attrition

Variables	(1)
Returnee	-0.2597***
	(0.0794)
Male	-0.1123***
	(0.0315)
Head	0.0444
	(0.0413)
Married	0.2129***
	(0.0471)
Age	0.0282**
	(0.0116)
Age squared	-0.0001
	(0.0002)
Kyrgyz	-0.0890
	(0.0821)
Uzbek	0.1311
	(0.0985)
Russian	-0.2231**
	(0.1076)
Secondary	0.0063
	(0.0650)
Vocational	-0.0323
	(0.0770)
University	0.0384
	(0.0809)
Household size	-0.0141
	(0.0123)
<16 dependency ratio	0.5303***
	(0.1315)
Rural	0.0293
	(0.0515)
Safety	-0.0703***
	(0.0144)
Constant	-0.2713
	(0.2275)
Observations	6,910

Notes: Coefficient estimates of the probability of being interviewed in all four waves of the survey. Observations are for working-age individuals interviewed in 2010, members of non-splitting households, who were born in Kyrgyzstan. Robust standard errors clustered at the household level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Descriptive statistics of estimation sample by employment status

	Always self-employed	Sometimes self-employed	Never self-employed
Individual			
Male	0.8486	0.6666	0.3246
Age	42.1363	40.3391	38.8457
Head	0.7041	0.4753	0.2610
Married	0.8697	0.8196	0.7062
Kyrgyz	0.7764	0.6942	0.6852
Uzbek	0.0904	0.1655	0.1311
Russian	0.0309	0.0300	0.0916
Vocational	0.1747	0.1394	0.1648
Secondary	0.6107	0.6679	0.5188
University	0.1544	0.1017	0.2146
Has worked last week	1.0000	0.7025	0.4998
Self-employed	1	0.4591	0.0000
Wage-employed	0	0.1624	0.3990
Family worker	0	0.0767	0.0985
Average monthly profits (real, oblasts)	8,310.878*	6,082.414*	6,243.931*
Has ever lived abroad	0.0858	0.0909	0.0456
Household			
Household size	5.3102	5.9408	5.5421
0-15 dependency ratio	0.3016	0.2943	0.2516
Living with returnee(s)	0.1363	0.1937	0.1509
Member(s) currently working abroads	0.0828	0.1497	0.1228
Owns land	0.8170	0.8504	0.7242
Total land area (ha.)	1.2243	0.9623	0.6779
Land area distributed (ha.)	1.0761	0.8360	0.5853
Number of shocks	2.0813	2.1188	1.8373
Location			
Urban	0.3253	0.2262	0.4425
Total	1,328	6,164	11,568

Notes: Means of variables of interest of balanced panel estimation sample of 4,765 individuals observed each four years of the survey, by employment status of primary occupation. *Statistics reported for individuals always, sometimes or never self-employed who worked and received a salary (respectively 1,274, 3,676 and 4,646 individuals).

Table A3: Descriptive statistics of self-employed

	All self-employed	Always self-employed	Sometimes self-employed
Employer	0.0228	0.0316	0.0187
Own-account	0.9772	0.9684	0.9813
Months in activity	112.5171	124.7651	106.7696
Business is registered	0.4798	0.5467	0.4484
Employes non household members	0.1000	0.1408	0.0809
Number of employees (cond.)	3.5361	3.5989	3.4847
Average monthly profits (real, oblasts)	6,752.205*	8,310.878*	6,001.733*
Agriculture	0.5464	0.5256	0.5562
Wholesale and retail	0.2071	0.2470	0.1883
Transportation	0.1142	0.1175	0.1127
Manufacturing	0.0214	0.0377	0.0138
Construction	0.0418	0.0203	0.0519
Hotels and restaurants	0.0072	0.0060	0.0078
Other services	0.0618	0.0459	0.0693
Owns land	0.8523	0.8170	0.8689
Number of shock	2.1174	2.0813	2.1343
Total	4,158	1,328	2,830

Notes: Means of variables of interest of balanced panel estimation sample of self-employed individuals. The first column reports descriptive statistics of individuals observed as self-employed in the survey. The second column reports descriptive statistics of individuals always observed as self-employed throughout the survey. The last column reports descriptive statistics of individuals sometimes self-employed, when observed as self-employed. Real profits are computed at 2010 contact price for the whole country. Agriculture includes agriculture, hunting, forestry, fishing and extractive industry. Other services include financial intermediation, education, health, social work and private households. *Statistics reported for all, always and sometimes self-employed who reported a salary different from 0 (respectively 3,920, 1,274 and 2,646 individuals).

Table A4: Descriptive statistics of estimation sample by migration status

	Returnees	Non-migrants		
		All	With returnee(s)	Without returnee
Individual				
Male	0.6819	0.4576	0.3906	0.4657
Age	38.8493	39.6056	39.8021	39.5819
Head	0.4280	0.3567	0.2396	0.3708
Married	0.7802	0.7525	0.7802	0.7492
Kyrgyz	0.6495	0.6975	0.5875	0.7107
Uzbek	0.2256	0.1335	0.2526	0.1192
Russian	0.0300	0.0700	0.0422	0.0733
Vocational	0.1174	0.1599	0.1313	0.1634
Secondary	0.5970	0.5718	0.6469	0.5628
University	0.1807	0.1734	0.1104	0.1810
Has worked last week	0.6386	0.5976	0.5552	0.6027
Self-employed	0.3189	0.2114	0.1901	0.2139
Wage-employed	0.2406	0.2983	0.2417	0.3052
Family worker	0.0741	0.0853	0.1198	0.0812
Has ever lived abroad	1	0	0	0
Number of months abroad	13.2198[a]	0	0	0
Household				
Household size	6.1424	5.6221	6.9771	5.4589
0-15 dependency ratio	0.2802	0.2682	0.2469	0.2707
Living with returnee(s)	1	0.1075	1	0
Member(s) currently working abroad	0.1757	0.1255	0.1870	0.1181
Owns land	0.7968	0.7698	0.8391	0.7615
Total land area (ha.)	0.7034	0.8150	0.6919	0.8298
Land area distributed (ha.)	0.5827	0.7085	0.5952	0.7221
Number of shocks	2.0275	1.9398	1.9625	1.9371
Location				
Urban	0.3172	0.3675	0.2630	0.3801
Total	1,201	17,859	1,920	15,939

Notes: Means of variables of interest of balanced panel estimation sample of 4,765 individuals observed each four years of the survey, by individual and household migration status.
[a]Based on 1,201 return migrants.

Table A5: Income sources and expenditures by household migration status

	All households		With returnee(s)	Without returnee	t-test
	Mean	S.D.	Mean	Mean	
			Incomes		
Share from:					
Household enterprise	.3384	.3715	.37433	.33213	3.84***
Property	.0097	.0651	.00931	.00978	-0.24
Social transfers	.1853	.2746	.15697	.19043	-4.11***
Material aid	.0878	.2069	.12274	.08184	6.68***
Wage employment	.3544	.3937	.30877	.36205	-4.57***
Other incomes	.0244	.1152	.02789	.02377	1.21
Total	9,031		1,336	7,695	
			Expenditures		
Share from:					
Health	.1381	.1065	.11773	.14164	-7.64***
Housing and utilities	.3063	.1622	.30543	.30649	-0.22
Education	.0205	.0687	.0225	.02019	1.14
Transportation	.1030	.1006	.10661	.1024	1.42
Leisure	.1124	.0898	.11049	.11269	-0.83
Celebrations	.0987	.1323	.10757	.09714	2.67***
Clothing	.2210	.1485	.22968	.21945	2.34**
Total	9,106		1,349	7,757	

Notes: In upper panel, share of monthly average household income by sources of balanced panel households for non-missing values are reported. Household enterprises include income from agricultural and non-agricultural enterprises. Property includes income from building, land, interests and dividends. Social transfers include pensions, monthly benefits, compensatory and social payments. Material aid includes humanitarian aid, global and regional remittances. Share of monthly average household expenditures on non-food items of balanced panel households for non-missing values are reported in lower panel. Health expenditures include soap, detergents, personal care, medicine and medical care. Housing and utilities expenditures include energy, taxes, construction, maintenance, household goods and vehicles. Education expenditures exclude school expenditures. Transportation expenditures include transportation services and fuel used for transportation. Leisure expenditures include communication, entertainment, TV, radio, Internet and jewellery. Celebration expenditures include celebrations, funerals and rituals. Clothing expenditures include clothing, shoes and fabrics.

Table A6: Coefficient estimates of benchmark specifications

	Baseline		Entry		'Survival'		Combined model	Parsimonious combined model
	RE	CRE	RE	CRE	RE	CRE	CRE	CRE
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Returnee	0.5659***	0.6244***	0.5571***	0.8097*	0.1200	0.5563	0.9162***	0.8977***
	(0.1436)	(0.2424)	(0.1461)	(0.4141)	(0.1351)	(0.3905)	(0.2873)	(0.2858)
Male	1.7852***	0.5934	0.8801***	-1.9887	0.3700*	2.1122	-0.4583	-0.5841
	(0.2373)	(0.4510)	(0.2603)	(3.2860)	(0.2022)	(3.0860)	(2.2325)	(2.2436)
Age	0.2533***	0.2862***	0.0872	0.3043**	0.0800	0.1440	0.1804**	0.2119***
	(0.0636)	(0.0717)	(0.0760)	(0.1200)	(0.0498)	(0.0954)	(0.0757)	(0.0702)
Age squared	-0.0032***	-0.0032***	-0.0010	-0.0012	-0.0009	-0.0010	-0.0009	-0.0012**
	(0.0008)	(0.0008)	(0.0009)	(0.0010)	(0.0006)	(0.0006)	(0.0007)	(0.0005)
Kyrgyz	-0.2734*	0.1292	-0.0549	1.0359	-0.3123**	-4.6563	0.7556	0.7258
	(0.1495)	(0.7721)	(0.1338)	(1.5082)	(0.1341)	(4.5458)	(0.9856)	(1.0010)
Uzbek	-0.2179	-1.6129***	0.0896	-2.1362	-0.3531**	-4.6676	-1.2087	-1.2035
	(0.1881)	(0.6042)	(0.1725)	(1.6543)	(0.1758)	(3.9201)	(1.1490)	(1.1519)
Russian	-0.9143***	-0.4359	-0.4718**	0.9941	-0.6587**	-5.6670	0.1022	0.1407
	(0.2550)	(0.7060)	(0.2345)	(0.9988)	(0.2617)	(4.5837)	(0.8684)	(0.8859)
Head	0.7333***	0.7299***	0.3067***	0.6892*	0.2035**	0.6510	0.5866**	0.6072**
	(0.1124)	(0.2737)	(0.1138)	(0.3879)	(0.1020)	(0.4158)	(0.2974)	(0.2934)
Secondary	0.3672**	-0.0250	0.0544	0.6101	0.3459**	-0.1866	0.0071	0.2124
	(0.1824)	(0.3619)	(0.2036)	(0.7346)	(0.1636)	(0.5528)	(0.4497)	(0.4244)
Vocational	0.1374	0.2900	-0.5758*	0.2855	0.3349*	0.0933	-0.1778	0.0498
	(0.2510)	(0.3862)	(0.3194)	(0.7466)	(0.1964)	(0.5385)	(0.4621)	(0.4306)
University	-0.4364	0.0900	-0.8918**	0.4356	0.1695	-0.1875	-0.1964	0.0647
	(0.3200)	(0.5038)	(0.4044)	(0.8787)	(0.2155)	(0.7010)	(0.5646)	(0.5323)
Household size	0.0074	0.0093	0.0200	0.0207	-0.0218	0.0788	0.0454	0.0442
	(0.0221)	(0.0404)	(0.0223)	(0.0494)	(0.0229)	(0.0627)	(0.0378)	(0.0379)
<16 dependency ratio	-0.4111*	-0.6841**	0.0546	-0.0615	-0.0607	-0.6775	-0.2864	-0.3528
	(0.2380)	(0.3373)	(0.2765)	(0.4943)	(0.2633)	(0.6084)	(0.3756)	(0.3567)
Urban	-0.6899***	-2.1461***	-0.5442***	-1.4259**	0.3679***	13.3892***	-0.6429*	-0.6449*
	(0.1303)	(0.7270)	(0.1444)	(0.6276)	(0.1098)	(1.2764)	(0.3820)	(0.3757)
2011	-0.0558	-0.0972	-0.3530***	0.0797	0.3782	0.5599*	0.0391	0.0598
	(0.0704)	(0.0789)	(0.1264)	(0.2346)	(0.2564)	(0.3203)	(0.1515)	(0.1493)
2012	-0.0367	-0.0824	-0.1599	0.0616	0.1171	0.1995	0.0016	0.0358
	(0.0995)	(0.1258)	(0.1457)	(0.1705)	(0.2638)	(0.2805)	(0.1199)	(0.1144)
2013	-0.2465***	-0.3161**						
	(0.0794)	(0.1359)						
Group means	No	Yes	No	Yes	No	Yes	Yes	Yes
lnsig2u	1.2246***	1.2462***	0.3376*	0.3683*	-0.8696*	-0.7797*	-0.2255	-0.2156
	(0.0718)	(0.0728)	(0.2036)	(0.2045)	(0.4559)	(0.4463)	(0.2137)	(0.2119)
$\overline{X}_{kit} = 0$		54.99		21.96		144.91	20.19	21.18
		0.0000		0.0560		0.0000	0.0907	0.0694
$\beta_k = \theta_k$							46.92	34.47
							0.0000	0.0003
Observations	11,361	11,361	6,031	6,031	2,620	2,620	8,651	8,651
Number of groups	3,849	3,849	3,044	3,044	1,371	1,371	3,736	3,736

Notes: Please, refer to Table 5.

Table A6: Coefficient estimates of benchmark specifications (continued)

	Baseline		Entry		'Survival'		Combined model	Parsimonious combined model
	RE	CRE	RE	CRE	RE	CRE	CRE	CRE
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self-employed$_{(t=0)}$							1.0602***	1.0610***
							(0.1246)	(0.1237)
Self-employed$_{(t-1)}$							-0.4154	1.0144***
							(0.9229)	(0.2378)
... X Returnee							-0.2866*	-0.2678*
							(0.1597)	(0.1599)
... X Male							-0.2242*	-0.2843**
							(0.1340)	(0.1144)
... X Age							0.0553	
							(0.0423)	
... X Age squared							-0.0006	
							(0.0005)	
... X Kyrgyz							-0.3271**	-0.2532
							(0.1656)	(0.1590)
... X Uzbek							-0.4441**	-0.4404**
							(0.2128)	(0.2131)
... X Russian							-0.5558*	-0.5037
							(0.3352)	(0.3367)
... X Head							0.0028	
							(0.1350)	
... X Secondary							0.3209	
							(0.2144)	
... X Vocational							0.8134***	0.5215***
							(0.2727)	(0.1556)
... X University							0.8143***	0.4896***
							(0.3069)	(0.1877)
... X Household size							-0.0430*	-0.0475**
							(0.0248)	(0.0221)
... X <16 dependency ratio							-0.0293	
							(0.2979)	
... X Urban							0.7692***	0.7812***
							(0.1563)	(0.1551)
... X 2011							0.6980***	0.7282***
							(0.1847)	(0.1823)
... X 2012							0.3441**	0.3758***
							(0.1402)	(0.1364)
IMR$_{\text{retention}}$	-1.8509***	-1.4439**	-0.6906	-0.5274	-1.0095*	-0.7027	-0.6307	-0.5622
	(0.5365)	(0.6269)	(0.5263)	(0.6009)	(0.5849)	(0.6672)	(0.4341)	(0.4296)
IMR$_{\text{working}}$	1.2281**	1.1892**	0.1753	0.2269	-0.2035	-0.1905	0.1710	0.3467
	(0.5381)	(0.5399)	(0.6373)	(0.6468)	(0.6356)	(0.6453)	(0.4325)	(0.3946)
Constant	-6.0365***	-6.4416***	-2.9000	-3.6806	-0.6423	-1.1039	-2.7632*	-3.6925***
	(1.8174)	(1.8742)	(2.2104)	(2.2738)	(1.5853)	(1.6569)	(1.5408)	(1.3214)
Group means	No	Yes	No	Yes	No	Yes	Yes	Yes
lnsig2u	1.2246***	1.2462***	0.3376*	0.3683*	-0.8696*	-0.7797*	-0.2255	-0.2156
	(0.0718)	(0.0728)	(0.2036)	(0.2045)	(0.4559)	(0.4463)	(0.2137)	(0.2119)
$\overline{X}_{kit} = 0$		54.99		21.96		144.91	20.19	21.18
		0.0000		0.0560		0.0000	0.0907	0.0694
$\beta_k = \theta_k$							46.92	34.47
							0.0000	0.0003
Observations	11,361	11,361	6,031	6,031	2,620	2,620	8,651	8,651
Number of groups	3,849	3,849	3,044	3,044	1,371	1,371	3,736	3,736

Notes: Please, refer to Table 5.

Table A7: Marginal effects of benchmark coefficient estimates

	Baseline		Entry		'Survival'		Combined model			Parsimonious combined model		
	RE	CRE	RE	CRE	RE	CRE	Average	$SE_{(t-1)}=0$	$SE_{(t-1)}=1$	Average	$SE_{(t-1)}=0$	$SE_{(t-1)}=1$
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Returnee	0.132***	0.142**	0.109***	0.168	0.0273	0.106*	0.186***	0.230***	0.183**	0.182***	0.224***	0.185**
	(0.0346)	(0.0572)	(0.0331)	(0.104)	(0.0293)	(0.0574)	(0.0687)	(0.0780)	(0.0790)	(0.0678)	(0.0774)	(0.0785)
Male	0.409***	0.134	0.140***	-0.307	0.0962*	0.585	-0.0968	-0.0971	-0.181	-0.121	-0.122	-0.223
	(0.0450)	(0.106)	(0.0416)	(0.386)	(0.0561)	(0.598)	(0.384)	(0.455)	(0.487)	(0.372)	(0.443)	(0.428)
Age	-0.00165	0.00613	0.000848	0.0334**	0.000964	0.0152	0.0217**	0.0242*	0.0356**	0.0215*	0.0240*	0.0352**
	(0.00145)	(0.00790)	(0.00106)	(0.0152)	(0.00189)	(0.0191)	(0.0110)	(0.0125)	(0.0173)	(0.0110)	(0.0124)	(0.0177)
Kyrgyz	-0.0608*	0.0273	-0.00888	0.136	-0.0694**	-0.354***	0.113	0.150	0.129	0.114	0.144	0.145
	(0.0336)	(0.162)	(0.0219)	(0.164)	(0.0274)	(0.0101)	(0.169)	(0.174)	(0.293)	(0.174)	(0.178)	(0.301)
Uzbek	-0.0467	-0.260***	0.0148	-0.169***	-0.0926*	-0.734***	-0.237	-0.201	-0.406**	-0.237	-0.199	-0.416**
	(0.0393)	(0.0589)	(0.0292)	(0.0472)	(0.0497)	(0.0116)	(0.144)	(0.128)	(0.175)	(0.146)	(0.128)	(0.180)
Russian	-0.176***	-0.0884	-0.0620**	0.215	-0.194**	-0.811***	-0.0228	0.0227	-0.139	-0.0103	0.0313	-0.114
	(0.0407)	(0.135)	(0.0243)	(0.263)	(0.0883)	(0.0118)	(0.192)	(0.197)	(0.278)	(0.197)	(0.202)	(0.294)
Head	0.180***	0.174**	0.0514***	0.121	0.0495**	0.163	0.122*	0.136*	0.188*	0.127*	0.140**	0.196**
	(0.0291)	(0.0696)	(0.0195)	(0.0742)	(0.0252)	(0.109)	(0.0653)	(0.0717)	(0.0979)	(0.0656)	(0.0706)	(0.0941)
Secondary	0.0811**	-0.00533	0.00866	0.0927	0.0862**	-0.0424	0.0245	0.00156	0.100	0.0411	0.0463	0.0663
	(0.0403)	(0.0771)	(0.0323)	(0.109)	(0.0419)	(0.122)	(0.0837)	(0.0982)	(0.127)	(0.0824)	(0.0926)	(0.133)
Vocational	0.0304	0.0629	-0.0778**	0.0490	0.0715*	0.0212	0.0174	-0.0379	0.185*	0.0414	0.0109	0.170
	(0.0561)	(0.0847)	(0.0359)	(0.138)	(0.0368)	(0.119)	(0.0693)	(0.0962)	(0.109)	(0.0750)	(0.0946)	(0.115)
University	-0.0935	0.0193	-0.111***	0.0779	0.0379	-0.0466	0.0135	-0.0419	0.174	0.0416	0.0141	0.160
	(0.0664)	(0.108)	(0.0396)	(0.175)	(0.0451)	(0.184)	(0.0844)	(0.117)	(0.128)	(0.0922)	(0.117)	(0.133)
Household size	0.00163	0.00200	0.00320	0.00328	-0.00517	0.0184	0.00574	0.00990	0.000730	0.00518	0.00960	-0.00101
	(0.00485)	(0.00863)	(0.00356)	(0.00781)	(0.00540)	(0.0147)	(0.00720)	(0.00825)	(0.0127)	(0.00720)	(0.00824)	(0.0125)
<16 dependency ratio	-0.0903*	-0.146**	0.00874	-0.00975	-0.0144	-0.159	-0.0572	-0.0625	-0.0960	-0.0681	-0.0766	-0.109
	(0.0523)	(0.0722)	(0.0442)	(0.0783)	(0.0624)	(0.142)	(0.0695)	(0.0821)	(0.125)	(0.0689)	(0.0776)	(0.111)
Urban	-0.150***	-0.388***	-0.0801***	-0.192**	0.0804***	0.358***	-0.0645	-0.136*	0.0383	-0.0638	-0.136*	0.0422
	(0.0271)	(0.0868)	(0.0195)	(0.0786)	(0.0215)	(0.00994)	(0.0658)	(0.0767)	(0.120)	(0.0647)	(0.0750)	(0.120)
2011	-0.0122	-0.0206										
	(0.0154)	(0.0166)										
2012	-0.00804	-0.0175										
	(0.0217)	(0.0266)										
2013	-0.0530***	-0.0657**										
	(0.0168)	(0.0275)										
Self-employed$_{(t-1)}$							0.3061***	0.290***	0.290***	0.2897***	0.306***	0.306***
							(0.0448)	(0.0472)	(0.0472)	(0.0472)	(0.0448)	(0.0448)
Self-employed$_{(t=0)}$							0.2047***	0.231***	0.322***	0.2042***	0.230***	0.329***
							(0.0177)	(0.0207)	(0.0296)	(0.0178)	(0.0206)	(0.0305)
Observations	11,361	11,361	6,031	6,031	2,620	2,620	8,651	8,651	8,651	8,651	8,651	8,651

Notes: Columns (1)-(7) and (10) present average marginal effects corresponding of coefficient estimates presented in Table 10. Columns (8) and (11) present corresponding marginal effects when lagged self-employment is equal to 0; columns (9) and (12), when lagged self-employment is equal to 1. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure A1: Matching quality of 2011 (l) and 2012 (r) propensity score matching with 5 nearest neighbours

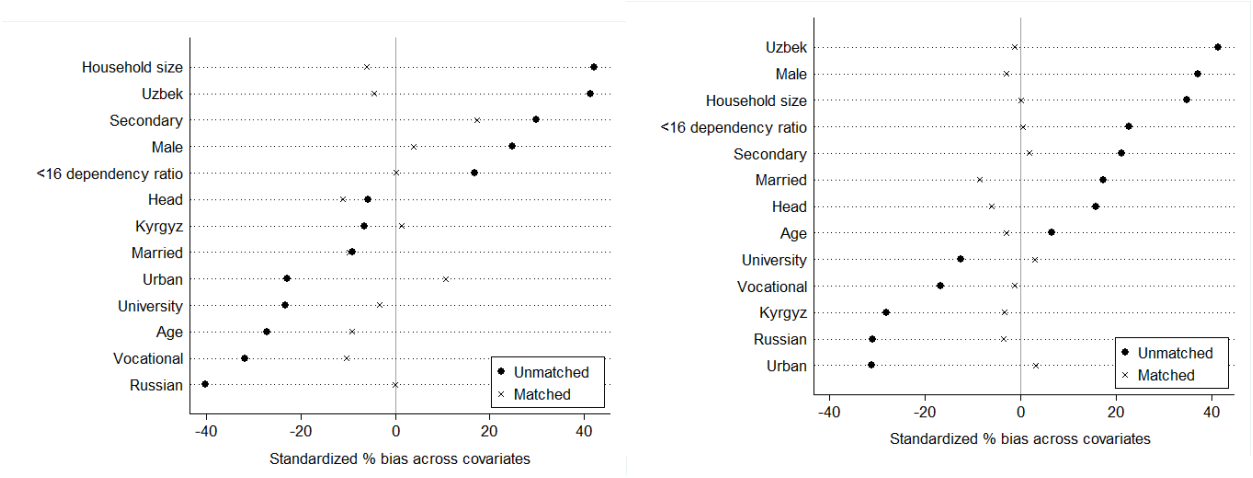

Figure A2: Matching quality of 2011 (l) and 2012 (r) covariate matching with 5 nearest neighbours

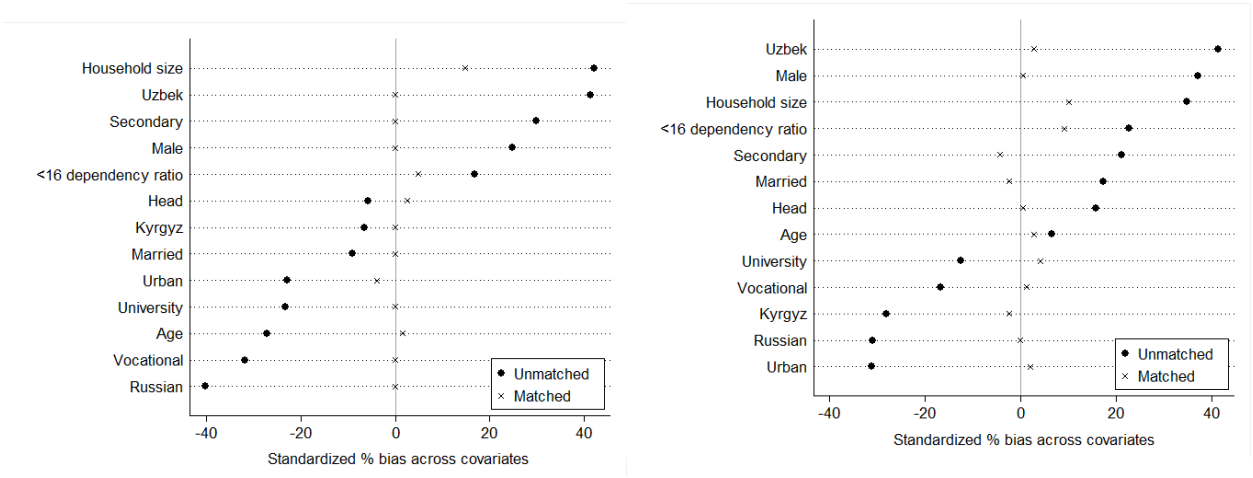

Table A8: Marginal effects of parsimonious combined model coefficient estimates on matched sample

Variables	Average (1)	$SE_{(t-1)}=0$ (2)	$SE_{(t-1)}=1$ (3)
	Propensity score matching, 5 NN		
Returnee	0.1240***	0.1998***	0.0119
	(0.0443)	(0.0593)	(0.0832)
Self-employed$_{(t=0)}$	0.2048***	0.2304***	0.2219***
	(0.0550)	(0.0598)	(0.0611)
Self-employed$_{(t-1)}$	0.2509***	0.2509***	0.2509***
	(0.0677)	(0.0677)	(0.0677)
Observations	1,190	1,190	1,190
	Covariate matching, 5 NN		
Returnee	0.1094**	0.2009***	-0.0432
	(0.0436)	(0.0587)	(0.0766)
Self-employed$_{(t=0)}$	0.1588***	0.1864***	0.1554**
	(0.0597)	(0.0699)	(0.0633)
Self-employed$_{(t-1)}$	0.3613***	0.3613***	0.3613***
	(0.0677)	(0.0677)	(0.0677)
Observations	1,190	1,190	1,190

Notes: Column (1) presents average marginal effects corresponding of coefficient estimates presented in Table 10, column (5). Column (2) presents corresponding marginal effects when lagged self-employment is equal to 0; column (3) when lagged self-employment is equal to 1. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure A3: Matching quality of 2011 (l) and 2012 (r) propensity score matching with 5 nearest neighbours of household sample

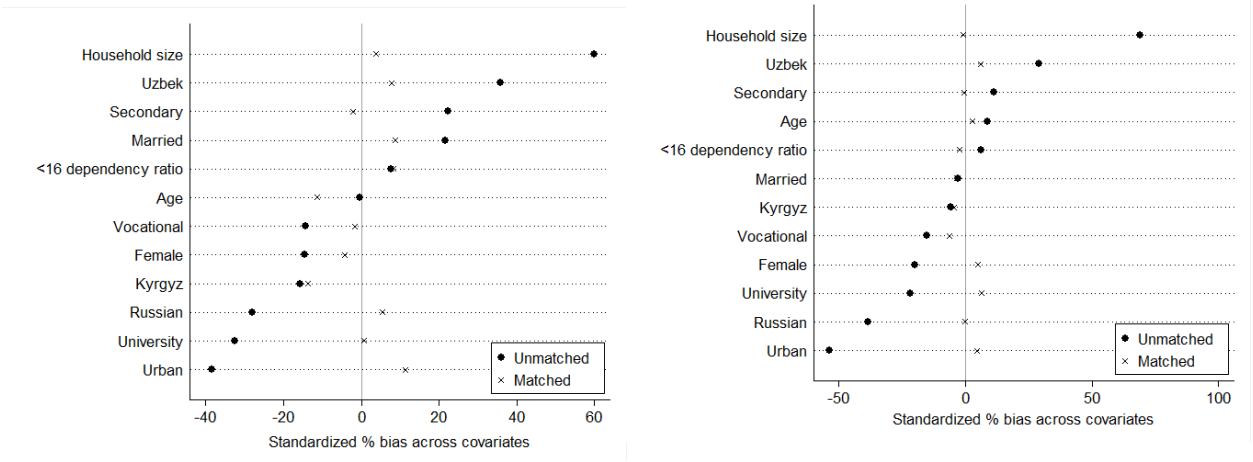

Figure A4: Matching quality of 2011 (l) and 2012 (r) covariate matching with 5 nearest neighbours of household sample

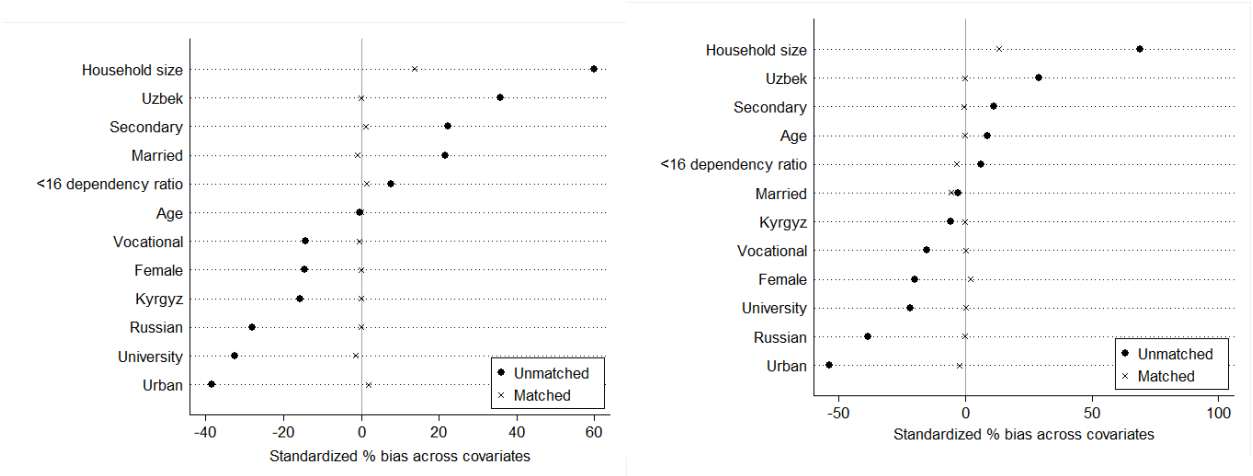

Table A9: Marginal effects of parsimonious combined model household specification coefficient estimates

Variables	Average (1)	$FE_{(t-1)}=0$ (2)	$FE_{(t-1)}=1$ (3)
	Full estimation sample		
Returnee	.0525***	.0675***	.0477**
	(.0159)	(.0240)	(.0226)
Family enterprise$_{(t=0)}$	.1540***	.1564***	.1654***
	(.0234)	(.0236)	(.0243)
Family enterprise$_{(t-1)}$	.2608***	.2608***	.2608***
	(.0362)	(.0362)	(.0362)
Observations	6,831	6,831	6,831
	Matched sample, PSM		
Returnee	.0944***	.1588***	.0433*
	(.0179)	(.0292)	(.0252)
Family enterprise$_{(t=0)}$	.1294***	.1455***	.1172***
	(.0278)	(.0303)	(.0261)
Family enterprise$_{(t-1)}$	.3083***	.3083***	.3083***
	(.0278)	(.0278)	(.0278)
Observations	2,274	2,274	2,274
	Matched sample, CVM		
Returnee	.0577***	.1236***	-.0024
	(.0180)	(.0294)	(.0252)
Family enterprise$_{(t=0)}$	.1190***	.1374***	.1071***
	(.0272)	(.0304)	(.0255)
Family enterprise$_{(t-1)}$	.3162***	.3162***	.3162***
	(.0275)	(.0275)	(.0275)
Observations	2,270	2,270	2,270

Notes: Column (1) presents average marginal effects corresponding of coefficient estimates presented in Tables 14 (upper panel) and 15 (middle and lower panels), column (5). Column (2) presents corresponding marginal effects when lagged self-employment is equal to 0; column (3) when lagged self-employment is equal to 1. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure A5: Average marginal effects of return migration with 95% confidence intervals of parsimonious combined model on household sample

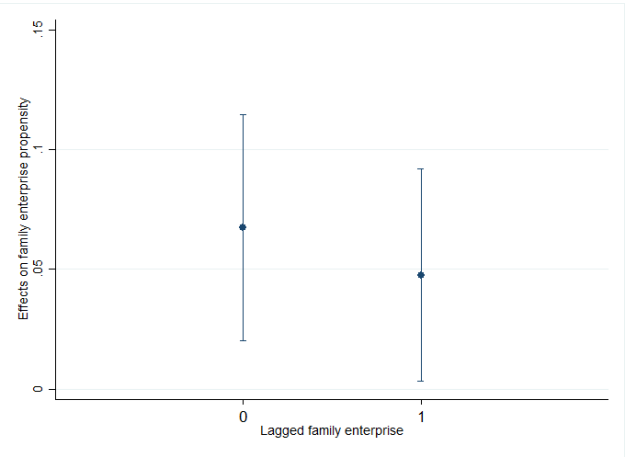

Figure A6: Average marginal effects of return migration with 95% confidence intervals of parsimonious combined models on PSM (l) and CVM (r) household matched sample