Migration and Agricultural Efficiency –

Empirical Evidence for Kosovo

Johannes Sauer
Department of Agricultural Economics, University of Kiel, Kiel, Germany

Matthew Gorton
Newcastle University Business School, Newcastle University,
Newcastle upon Tyne, UK

Sophia Davidova
School of Economics, University of Kent, Canterbury, UK

jsauer@ae.uni-kiel.de

GEWISOLA

2013

Vortrag anlässlich der 53. Jahrestagung der GEWISOLA

„Wie viel Markt und wie viel Regulierung braucht eine nachhaltige Agrarentwicklung?“

Berlin, 25.-27. September 2013

Migration und Effizienz landwirtschaftlicher Produktion - Empirische Evidenz für Kosovo

Abstrakt

Wie die meisten Länder in Mittel und Osteuropa hat auch das Kosovo eine substantielle Abwanderung in den letzten Jahren zu verzeichnen. Dies führte zu Diskussionen im Hinblick auf die Effekte solcher Migration auf die Effizienz landwirtschaftlicher Produktion. Der vorliegende Beitrag adressiert dieses Problem indem ein umfangreiches und repräsentatives Sample ländlicher Haushalte analysiert wird (n=2217). Ein zweistufiges Schätzverfahren wird hierzu verwandt: eine Frontiertechnik um den Effekt von Migration auf die Effizienz landwirtschaftlicher Betriebe zu untersuchen, gefolgt von einem matching Schätzverfahren um den Effizienzeffekt verschiedener Migrationsintensitäten zu ermitteln. Wir finden, dass Migration die landwirtschaftliche Produktionseffizienz verringert. Dieser Effekt ist umso stärker, je älter und besser ausgebildet die jeweiligen Migranten sind.

Schlüsselworte: Migration, Technische Effizienz, Ländliche Haushalte, Kosovo

Migration and Agricultural Efficiency - Empirical Evidence for Kosovo

Abstract

Kosovo, like most of rural Central and Eastern Europe, has witnessed substantial out-migration in recent years, prompting debates on the effect of migration on agricultural efficiency. This paper addresses this issue, drawing on a large (n=2217) and representative sample of agricultural households. A two-stage estimation procedure is followed: a frontier technique to estimate the effect of migration on farm efficiency, followed by a matching estimation approach to robustly estimate the sample average effect on efficiency for different levels of migration intensity. Migration has an efficiency decreasing effect which is amplified for better educated and older workers.

Keywords: Migration, Technical Efficiency, Agricultural Households, Kosovo

JEL: J43, J61, Q12

1. Introduction

Rural areas in many developing and transitional economies have witnessed significant outmigration in recent years. Outmigration has tended to be relatively greatest from the most impoverished regions, which also tend to be most reliant on agriculture as a source of income and employment (Bolganschi, 2011). The impact on rural areas can be considerable, for instance studies for Bulgaria (Dittrich and Jeleva, 2009), Romania (Surd, 2010) and Ukraine (Peacock, 2012) describe villages either almost entirely depopulated or consisting of elderly residents and their grandchildren after those of working age migrated in search of better paid employment. This leads to the important question of what has been the impact of migration on agricultural efficiency?

This paper analyses the impact of migration on farm technical efficiency in Kosovo, drawing on an extensive and representative survey of agricultural households. Kosovo was selected as an exemplary case where outmigration has been particularly high (Gashi and Haxhikadrija, 2012) and the majority of rural households engage in farming. The impact of migration on farm efficiency is assessed using a two-stage estimation procedure: a frontier technique to estimate the effect of migration on farm technical efficiency, followed by a matching estimation approach to robustly estimate the sample average effect on efficiency for different levels of migration intensity. This two stage approach accounts for empirical identification problems and lagged decisions, and the paper provides a more robust and nuanced analysis of the impact of migration on agricultural efficiency than present in most previous studies by considering the percentage of total available work time per household per year accounted for by migration (migration intensity). Distinctions are also drawn between male and female, and skilled and unskilled migrants, as well as assessing if the impact of migration varies by age, wealth of household and region.

The study contributes to the literature, particularly within the context of the New Economics of Labour Migration (NELM), and the questions of whether migration affects technical efficiency and if there is a relationship whether it is positive or negative. While Kosovo can be considered an extreme case, most of rural Central and Eastern Europe has witnessed significant out-migration in recent years, particularly of adults of working age, the better educated and from the poorest regions (OECD, 2012). Assessing the impact of migration on farm efficiency is thus of wider importance within the region.

2. Case Context

Over several decades rural Kosovo has witnessed substantial outmigration. Not surprisingly, internal outmigration has been relatively greatest from the poorest regions, whilst there was an inflow of migrants to the more developed regions, particularly the capital city of Pristina (Vathi and Black, 2007). Although in general there is no agreement in the literature on whether migration has changed the educational composition of the labour force in Kosovo, since on average migrants only have completed secondary education (Gashi and Haxhikadrija, 2012), analysis indicates outmigration of more educated people from rural to urban areas, constituting a "brain drain" (e.g. Haskuka *et al.*, 2004).

The statistical information on international migration from Kosovo is unreliable. A country report prepared for the European Commission (Gashi and Haxhikadrija, 2012) quotes two estimates which vary from 415,000 to 800,000 migrants from a resident population in 2011 of 1.74 million (ASK, 2012). Although it is often claimed that migration out of Kosovo was forced due to the military conflict in 1999, a UNDP (2010) survey of the reasons for migration identified that in only 18.2% of cases was the motive related to this, another 23.8% involved other political reasons, but the most important impetus was economic (42.9%). The latter is reflected in the pattern of emigration from Kosovo from the 1960s to 2011. The largest share of emigration, 53.6%, took place post-conflict (UNDP, 2012). Moreover, intentions to migrate remain widespread: the UNDP survey reports that in 2011 15% of household heads intended to migrate, 70% of which for economic reasons.

Due to its scale, migration (internal and international) has potentially significant ramifications for rural Kosovo, bearing in mind that 62% of the resident population lives in rural areas and that the share of the labour force engaged in agriculture is 49% (ARCOTRASS, 2006).

3. Theoretical Framework

Various macro- and micro-economic theories and models of migration have been proposed and tested over several decades (Massey *et al.*, 1993). In neo-classical theory, individuals decide to migrate or not based on a comparison of expected costs and benefits. More recently, the New Economics of Labour Migration (NELM) relates migration to production and incomes in the households (communities) from where migrants originate. It challenges the neo-classical assumption of individual decision making, arguing instead for a collective household perspective on the spatial allocation of labour (Stark and Bloom, 1985). The NELM also recognises that migration typically occurs under conditions of market failure.

In the case of missing or imperfect credit and insurance markets, the migrant acts as a financial intermediary who, through remittances, enables rural households, particularly those poor in liquid assets, to overcome credit and risk constraints (Taylor and Wyatt, 1996; Rozelle *et al.*, 1999). The impact of migration on technical efficiency will be positive where remittances relax credit constraints and enable efficiency improving investments. However, the impact of remittances on farm efficiency may be negative where they provide rural household members with an income that lessens their incentives to engage in agricultural production. Therefore, remittances may change preferences between work and leisure. On the other hand, labour market imperfections may weaken technical efficiency in the absence of perfect substitutes for lost household labour (Arslan and Taylor, 2012). There is a debate in the academic literature as to whether migration increases or depletes human capital. Theoretically, Stark *et al.* (1998) argue that the opportunity to migrate increases human capital in migrants who invest to increase their opportunities upon migration, as well as in workers who stay in home country ("a brain gain") in comparison to a situation of a closed economy. However, households are heterogeneous and it cannot be claimed that some of them do not lose (at least temporarily) human capital embedded in the migrant family labour.

The theory relating to the relationship between migration and agricultural efficiency is, thus, ambivalent and empirical evidence conflicting. Under perfect markets, migration should not affect farm efficiency since there are no transaction costs, perfect substitutes for family labour are instantaneously available, and credit and liquidity constraints are absent. However, it is assumed here that, as in most emerging economies, markets in Kosovo (labour, credit, insurance) are underdeveloped, with high transaction costs. Therefore, a statistically significant relationship between farm household technical efficiency and migration intensity is expected. Whether this relationship is positive or negative is a case of empirical estimation.

To date the few empirical studies of the impact of migration on technical efficiency fail to provide consistent results concerning the direction of the relationship. For instance, Mochebelele and Winter-Nelson (2000) found that technical inefficiency was greater amongst non-migrant households in Lesotho, suggesting that migrant households benefited from cash resources that allowed them to buy inputs when required and improve overall farm management. Similarly, Nonthakot and Villano (2008) in their study of efficiency of maize farms in Northern Thailand estimated that technical efficiency on migrant farms was 10% higher than on non-migrant ones. However, Rozelle *et al.* (1999) found that the net impact of migration on incomes from maize production in China was negative although remittances partially offset this loss. Jokisch (2002), while not formally testing the impact of migration on technical efficiency, argues that outmigration in Ecuador had little impact on farm production and land use.

One reason for the inconsistency in findings may stem from the treatment of migration. In some studies it has been treated as a homogenous act, failing to differentiate between types (male / female, skilled / unskilled, young / old etc.). Moreover, much analysis has depended on a binary variable (non-migrant versus migrant households) that fails to capture what can be termed migration intensity (percentage of household members absent and for how long).

3. Data and Definition of Variables

The data employed in the study were obtained from the annual Agricultural Household Surveys (AHS) conducted by the Statistical Office of Kosovo (SOK) between 2005 and 2008. To construct the samples, SOK (2010) applied a two-stage sampling process, first stratifying by region and then by farm size (cultivated area). Within each category, agricultural households were randomly selected for face to face interview.[1]

The survey provides, for each household member, information on age, gender, educational attainment and the number of months, if at all, the family member lived away from the household in the previous 12 months. This was used to calculate migration intensity (the % of total available household work time accounted for by migration). Detailed information, on a plot by plot basis, relating to crops grown, yields, plot sizes and inputs used were collected.

Outputs included in the multi-output multi-input directional distant function for the estimation of technical efficiency were wheat, hay, potatoes, tomatoes, peppers and onions. These were chosen since they are the most common products in Kosovo for which a sufficiently large sample (2,217 households) could be built with all farm households producing some output. The survey collected data relating to the following inputs: land, labour, seeds, fertilizers, plant protection chemicals, fuel and machinery. Machinery value was estimated as the expected resale value expressed in Euros. These inputs have been included in the distant function. Land was quantified in hectares. The remaining inputs were measured as expenditure in Euros. All input values were deflated.

Kosovo is divided into seven regions (*Ferizaj*, *Gjakove*, *Gjilan*, *Mitrovice*, *Peje*, *Prishtine* and *Prizren*). Regions were included as dummies to control for agro-environmental conditions and variations in infrastructure. To capture land fragmentation for each farm household, we calculated a Simpson Index (Blarel *et al.*, 1992), which can be expressed as:

$$1 - \sum_1^i A_i^2 \,/ A^2 \qquad (1)$$

where A_i, is the area of the i^{th} plot and A is the total farm area. The *SI* is defined between the values of 0 and 1, where a value of zero indicates no fragmentation of farm land into spatially separated plots. The larger the index score, the greater the level of land fragmentation. Table 1 presents key descriptive statistics for the sample (see appendix). The average sampled farm utilised 2.61 ha, with production very fragmented (mean of 8.38 plots per farm). The majority of land is given over to wheat and hay production. By Western European standards (European Commission, 2011), farms are poorly capitalized with the total (resale) value of machinery per agricultural household equating to €3551 in 2005 values.

4. Empirical Modelling

Directional distance function to evaluate the impact of migration on technical efficiency

A directional distance function (Chambers *et al.*, 1998) is employed to model technological processes and used to derive measures of technical (in)efficiency. We assume a farm household uses a vector of input levels $x = (x_1, \dots, x_N) \in R_+^N$ to produce a vector of output quantities $y = (y_1, \dots, y_M) \in R_+^M$. The relationship between inputs and outputs is represented by the set:

$$T = \{(x, y): x \text{ can produce } y\} \qquad (2)$$

where T is the set of technically feasible input and output combinations, assuming that T satisfies free disposability of inputs and outputs (Färe and Primont, 1995). A functional representation of *T* is the directional output distance function, defined as:

$$\overrightarrow{D_O}(x, y, g) = sup\{\beta: (x, y + \beta g) \in T\} \qquad (3)$$

[1] SOK (2008, p.14), for the purpose of the survey, defined a household 'as a union of persons that live together, and pool their income'. Agricultural households were defined as those that cultivate more than 0.10 ha utilised arable land or less than 0.10 ha of utilised arable land but had at least: 1 cow or 5 sheep/goats or 3 pigs or 50 poultry or 20 beehives.

where $g = (g_1, \dots, g_M) \in R_+^M$ This distance function maps the input-output vector (x,y) into a scalar of value. If free disposability holds, the distance function

$$\overrightarrow{D_O}(x, y, g) \geq 0 \quad if, and\ only\ if\ (x, y) \in T \tag{4}$$

gives a complete characterization of the technology to be approximated (Chambers *et al.*, 1996). The translation property of the directional distance function allows its use for empirical work:

$$\overrightarrow{D_O}(x, y + \mu g; g) = \overrightarrow{D_O}(x, y, g) - \mu; \mu \in R \tag{5}$$

This property states that if outputs are translated by μg, then the value of the distance function is reduced by the scalar μ. To empirically estimate the directional output distance function a quadratic functional form can be chosen which makes $\overrightarrow{D_O}(\cdot)$ a second-order approximation of the underlying technology T. Imposing symmetry in parameters, the distance function is given by:

$$\overrightarrow{D_O}(x, y; g) = \alpha_0 + \sum_{i=1}^{M}(\alpha_i y_i + 0.5\alpha_{ii} y_i^2) + \sum_{i=1}^{M}\sum_{j=i+1}^{M}\alpha_{ij} y_i y_j + \sum_{i=1}^{N}(\beta_i x_i + 0.5\beta_{ii} x_i^2) + 0.5\sum_{i=1}^{N}\sum_{j=i+1}^{N}\beta_{ij} x_i x_j + \sum_{i=1}^{M}\sum_{j=1}^{N}\gamma_{ij} y_i x_j \tag{6}$$

Translation then requires:

$$\overrightarrow{D_O}(x, y + \mu g; g) = \alpha_0 + \sum_{i=1}^{M}(\alpha_i y_i + \mu g_i) + \sum_{i=1}^{M} 0.5\alpha_{ii}(y_i + \mu g_i)^2 + \sum_{i=1}^{M}\sum_{j=i+1}^{M}\alpha_{ij}(y_i + \mu g_i)\left(y_j + \mu g_j\right) + \sum_{i=1}^{N}(\beta_i x_i + 0.5\beta_{ii} x_i^2) + 0.5\sum_{i=1}^{N}\sum_{j=i+1}^{N}\beta_{ij} x_i x_j + \sum_{i=1}^{M}\sum_{j=1}^{N}\gamma_{ij}(y_i + \mu g_i) x_j - \mu \tag{7}$$

To measure the efficiency of individual farms a parametric stochastic frontier approach is used. In this paper the Battese and Coelli (1995) estimator on the distance function described in (7) is applied using an unbalanced panel data specification. The corresponding likelihood function and efficiency derivations are given in Kumbhakar and Lovell (2000). The stochastic specification of the directional output distance frontier takes the form:

$$0 = \overrightarrow{D_O}(x, y + \mu g; g) + \varepsilon \tag{8}$$

where $\varepsilon = v - u$; $v \sim N(0, \sigma_v^2)$ and $u \sim N^+(u, \sigma_u^2)$. To estimate (8), the translation property of the directional output distance function is exploited. Following common practice (Färe *et al.*, 2005) we set $g = 1$, resulting in:

$$\overrightarrow{D_O}(x, y + \mu; 1) + \mu = \overrightarrow{D_O}(x, y; 1) \tag{9}$$

By substituting $\overrightarrow{D_O}(x, y + \mu; 1) + \mu$ in (13) and rearranging, the following equation is obtained:

$$-\mu = \overrightarrow{D_O}(x, y + \mu; 1) + \varepsilon \tag{10}$$

Choosing $\mu = y_1$, which is farm household specific, a sufficient variation on the left-hand side is obtained to estimate the specification given in (10). The output vector used is y = (wheat, hay, pepper, tomatoes, onions, and potatoes) whereas the input vector is x = (land, full-time labour, part-time labour, machinery, fuel, rented services, fertilizer, chemicals and seed). The final specification estimated is:

$$-y_w = \alpha_0 + \sum_{i=1}^{M}(\alpha_i y_i') + \sum_{i=1}^{M} 0.5\alpha_{ii}(y_i')^2 + \sum_{i=1}^{M}\sum_{j=i+1}^{M}\alpha_{ij}(y_i')\,(y_i') + \sum_{i=1}^{N}(\beta_i x_i + 0.5\beta_{ii} x_i^2) + 0.5\sum_{i=1}^{N}\sum_{j=i+1}^{N}\beta_{ij} x_i x_j + \sum_{i=1}^{M}\sum_{j=1}^{N}\gamma_{ij}(y_i') x_j + v - u \tag{11}$$

where $y_i' = y_i + y_w$ with y_w as the quantity of wheat produced and abstracting from farm household and time related variation.

The vector of technical inefficiency effects u in the stochastic frontier model outlined in (11) is specified as:

$$u = z\delta + w \tag{12}$$

with, according to the conceptual framework, the following components of the vector z: migration intensity, average education of household members, average age of household members, educational level of the head of the household, age of the head of the household, female to male ratio, Simpson index (SI), total income, number of plots, region and year. The random variable w is defined by the truncation of the normal distribution with mean of zero and variance, σ_w^2, such that the point of truncation is $-z\delta$, i.e. $w \geq -z\delta$ (see Battese and Coelli, 1995). Abstracting from farm households and time related variations, technical efficiency is defined by:

$$TE = \exp(-u) = \exp(-z\delta - w) \quad (13)$$

Coelli *et al.* (2005) detail the corresponding likelihood function and its partial derivatives with respect to the individual parameters.

Matching estimation approach

The second stage of the empirical analysis consists of a matching approach to robustly estimate the sample average effect of migration on efficiency as well as the effect of different levels of migration intensity. As farm households are defined by a multitude of different characteristics over space and time, a sophisticated matching approach is needed to accurately determine the effect of migration on them in a statistically robust way (Guo and Fraser, 2010). As we use survey based non-experimental data collected through the observation of agricultural household farming systems as they operate in practice (Rubin, 1997) this type of method allows for reducing multi-dimensional covariates to a one-dimensional score. Appendix 1 outlines the approach in greater depth and Table 2 summarizes the two matching models estimated (see appendix).

5. Results

Before presenting the efficiency estimations, Table 3 details the scale of migration within the sample (see appendix). Overall, migration is widespread: 45.8% of sampled households have witnessed some degree of migration. While migration has occurred however it is most likely to be limited to one household member. Few households have witnessed high levels of migration intensity, for example it is 50% or higher in only 3.8% of cases. The most common level of migration intensity is between 5 and 10% of total household work time available.

The overall model quality of the estimated distance frontier and the estimated matching models are largely satisfactory indicating the robustness of our empirical results.[2] Table 4 presents the estimations for the determinants of inefficiency (see appendix). Migration intensity (based on % of total available work time per household per year) has an efficiency decreasing effect. This effect is strongly significant even when region, year, socio-economic characteristics of the household (age, education, gender, income) and farm characteristics (number of plots, cattle etc.) are accounted for. The interaction effects indicate that the efficiency decreasing effect of migration is amplified in better educated and older households and where the female to male ratio is higher. This suggests that older, better educated and male farm workers who have migrated are more difficult to replace (absence of perfect substitutes) so that the impact of migration of such workers is relatively greater. Total household income is not a significant determinant of technical efficiency.

Fragmentation of production, captured by both the Simpson Index and the number of plots, has a significant, negative effect on efficiency. This is consistent with recent findings on small-scale agriculture in Bangladesh (Rahman and Rahman, 2009), Bulgaria (Di Falco *et al.*, 2010) and Vietnam (Hung *et al.*, 2007). One important insight from the technical efficiency estimations is that human capital (approximated by education) and physical capital (farm equipment), decreases technical inefficiency. From this point of view it is disappointing that only 4.6% of remittances are used for investment in education and 3.9% for business investment, including 0.8% for purchase of land (UNDP, 2012).

[2] The overall model quality of the estimated distance frontier are evaluated using the value of the log-likelihood functions, the Lagrange Multiplier test statistics, the Akaike Information Criterion and the R-Squared test values. The statistical quality of the estimated matching models is judged by the values of the standard errors for the estimated sample average treatment effect estimates.

Table 5 reports the sample average treatment effects for changes in technical efficiency at household level for different levels of migration intensity (matching estimation, see appendix). As may be expected, the impact on technical efficiency is greatest for those households with the greatest level of migration intensity (migration accounts for more than 80% of total available work time of the household in a particular year). Interestingly, migration has a significant, efficiency lower effect even at low levels of intensity (migration accounting for 5 or less per cent of total available work time per household in a particular year).

Where migration accounts for between 30 and 60% of total available work time of the household, however, the effect of migration on technical efficiency is either fairly minor or not significant. Considering the farm and household characteristics for each category of migration intensity reveals some explanations for this. At low levels of migration intensity, households rarely adjust their farming activities, which given the labour intensive nature of farming in Kosovo means that even relatively small adjustments in labour input affect technical efficiency. However, those households with medium levels of migration intensity have significantly lower numbers of cattle (daily, labour intensive farming activity), adjusting their farming operations to account for migration. This pattern is consistent with the findings of De Brauw (2010), who found that seasonal migration in Vietnam prompted a shift from labour intensive to land-intensive crops, rather than changes in total factor productivity. However, at the highest levels of migration intensity in Kosovo, such adjustments are insufficient to compensate, and the deleterious effect of migration on technical efficiency is greatest.

6. Conclusions

Rural outmigration in Kosovo, as in much of Central and Eastern Europe, has been widespread and this paper tackles the important question of the impact of such outmigration on agricultural efficiency. The paper extends previous analysis by calculating migration intensity (rather than relying on crude, dichotomies measures of whether migration occurred or not) and applying a two-stage estimation procedure (frontier technique followed by a matching estimation approach).

The analysis identifies that there is a significant and negative 'lost labour effect' on farm efficiency. The negative effect of migration on technical efficiency is amplified for households with better educated and older workers. This suggests the presence of labour market imperfections with a lack of suitable alternative workers to replace such migrants. While remittances may partially compensate for the lost labour effect in some cases (Taylor *et al.*, 2003), for Kosovo total household income is not a significant determinant of technical efficiency and the proportion of remittances spent on upgrading human and physical capital appears small (UNDP, 2012). Migration has a significant negative effect on technical efficiency even at low levels of intensity although at moderate levels of intensity switching to less labour intensive types of farming may mitigate the effect. Overall, however, the findings for Kosovo support more pessimistic assessments (Wouterse, 2010) of the impact of outmigration on farm household efficiency.

7. References

Abadie, A., Drukker, D., Herr, J.L. and Imbens, G.W. (2004) 'Implementing matching estimators for average treatment effects in Stata', *Stata Journal*, 4(3), pp. 290-311.

Abadie, A. and Imbens, G. (2002) *Simple and bias-corrected matching estimators for average treatment effects*. Cambridge, MA: National Bureau of Economic Research.

Abadie, A. and Imbens, G.W. (2011) 'Bias-Corrected Matching Estimators for Average Treatment Effects', *Journal of Business & Economic Statistics*, 29(1), pp. 1-11.

ARCOTRASS (2006) *Study on the State of Agriculture in Five Applicant Countries, Kosovo Report*. Available at: http://ec.europa.eu/agriculture/analysis/external/applicant/kosovo_en.pdf.

Arslan, A. and Taylor, J.E. (2012) 'Transforming Rural Economies: Migration, Income Generation and Inequality in Rural Mexico', *Journal of Development Studies*, 48(8), pp. 1156-1176.

ASK (2012) *Kosovo Population and Housing Census 2011*. Prishtinë.

Battese, G.E. and Coelli, T.J. (1995) 'A model for technical inefficiency effects in a stochastic frontier production function for panel data', *Empirical Economics*, 20(2), pp. 325-332.

Blarel, B., Hazell, P., Place, F. and Quiggin, J. (1992) 'The economics of farm fragmentation: Evidence from Ghana and Rwanda', *The World Bank Economic Review*, 6(2), pp. 233-254.

Bolganschi, D. (2011) *Rural Out Migration and Land Use in Moldova*. Budapest, Hungary.

Chambers, R.G., Chung, Y. and Färe, R. (1996) 'Benefit and distance functions', *Journal of Economic Theory*, 70(2), pp. 407-419.

Chambers, R.G., Chung, Y. and Färe, R. (1998) 'Profit, directional distance functions, and Nerlovian efficiency', *Journal of Optimization Theory and Applications*, 98(2), pp. 351-364.

Coelli, T.J., Rao, D.S.P., O'Donnell, C.J. and Battese, G.E. (2005) *An Introduction to Efficiency and Productivity Analysis*. New York, NY: Springer.

De Brauw, A. (2010) 'Seasonal Migration and Agricultural Production in Vietnam', *Journal of Development Studies*, 46(1), pp. 114-139.

Di Falco, S., Penov, I., Aleksiev, A. and van Rensburg, T.M. (2010) 'Agrobiodiversity, farm profits and land fragmentation: Evidence from Bulgaria', *Land Use Policy*, 27(3), pp. 763-771.

Dittrich, E. and Jeleva, R. (2009) 'The future of rural communities in Bulgaria', *Research in Rural Sociology and Development*, 14, pp. 107-126.

European Commission (2011) *EU Farm Economics Overview*. Brussels. [Online]. Available at: http://ec.europa.eu/agriculture/rica/pdf/report_2008.pdf.

Färe, R., Grosskopf, S., Noh, D.W. and Weber, W. (2005) 'Characteristics of a polluting technology: theory and practice', *Journal of Econometrics*, 126(2), pp. 469-492.

Färe, R. and Primont, D. (1995) *Multi-Output Production and Duality: theory and applications*. Norwell, MA: Kluwer.

Gashi, A. and Haxhikadrija, A. (2012) *Social Impact of Emigration and Rural-Urban Migration in Central and Eastern Europe, Final Country Report: Kosovo*. [Online]. Available at: ec.europa.eu/social/BlobServlet?docId=8859&langId=en.

Guo, S. and Fraser, M.W. (2010) *Propensity Score Analysis*. London: Sage.

Haskuka, M., Malazogu, L., Luci, N. and Dugolli, I. (2004) *Human Development Report Kosovo 2004*. Pristina. [Online]. Available at: http://hdr.undp.org/en/reports/nationalreports/europethecis/kosovo/kosovo_2004_en.pdf

Hung, P.V., MacAulay, T.G. and Marsh, S.P. (2007) 'The economics of land fragmentation in the north of Vietnam', *Australian Journal of Agricultural and Resource Economics*, 51(2), pp. 195-211.

Jokisch, B.D. (2002) 'Migration and agricultural change: The case of smallholder agriculture in highland Ecuador', *Human Ecology*, 30(4), pp. 523-550.

Kumbhakar, S.C. and Lovell, C.A.K. (2000) *Stochastic frontier analysis*. Cambridge, UK: Cambridge University Press.

Massey, D.S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A. and Taylor, J.E. (1993) 'Theories of international migration: a review and appraisal', *Population and Development Review*, 19(3), pp. 431-466.

Mochebelele, M.T. and Winter-Nelson, A. (2000) 'Migrant labor and farm technical efficiency in Lesotho', *World Development*, 28(1), pp. 143-153.

Nonthakot, P. and Villano, R.A. (2008) *2008 Conference (52nd), February 5-8, 2008, Canberra, Australia*. Australian Agricultural and Resource Economics Society.

OECD (2012) *International Migration Outlook*. Paris. [Online]. Available at: http://www.oecdbookshop.org/oecd/display.asp?K=5K9CSF15F9XW&CID=&LANG=EN.

Peacock, E.A. (2012) 'The Authentic Village and the Modern City: The Space-Time of Class Identities in Urban Western Ukraine', *Anthropology of East Europe Review*, 30(1), pp. 213-236.

Rahman, S. and Rahman, M. (2009) 'Impact of land fragmentation and resource ownership on productivity and efficiency: The case of rice producers in Bangladesh', *Land Use Policy*, 26(1), pp. 95-103.

Rozelle, S., Taylor, J.E. and deBrauw, A. (1999) 'Migration, remittances, and agricultural productivity in China', *American Economic Review*, 89(2), pp. 287-291.

Rubin, D.B. (1997) 'Estimating causal effects from large data sets using propensity scores', *Annals of Internal Medicine*, 127(8(S)), pp. 757-763.

SOK (2010) *Agricultural Household Survey 2008*. Pristina. [Online]. Available at: http://esk.rks-gov.net/ENG/agriculture/publications.

Stark, O. and Bloom, D.E. (1985) 'The New Economics of Labor Migration', *American Economic Review*, 75(2), pp. 173-178.

Stark, O., Helmenstein, C. and Prskawetz, A. (1998) 'Human capital depletion, human capital formation, and migration: a blessing or a “curse”?', *Economics Letters*, 60(3), pp. 363-367.

Surd, V. (2010) 'Restraints and opportunities of the Romanian rural areas', *Revija za geografijo-Journal for Geography*, 5(1), pp. 55-66.

Taylor, J.E., Rozelle, S. and de Brauw, A. (2003) 'Migration and incomes in source communities: A new economics of migration perspective from China', *Economic Development and Cultural Change*, 52(1), pp. 75-101.

Taylor, J.E. and Wyatt, T.J. (1996) 'The shadow value of migrant remittances, income and inequality in a household-farm economy', *The Journal of Development Studies*, 32(6), pp. 899-912.

UNDP (2010) *Kosovo Remittance Study 2010*. Pristina. [Online]. Available at: www.kosovo.undp.org/repository/docs/Final-english.pdf.
UNDP (2012) *Kosovo Remittance Study*. Available at: http://www.kosovo.undp.org/repository/docs/KRS2012_English_858929.pdf.
Vathi, Z. and Black, R. (2007) *Migration and poverty reduction in Kosovo*. Brighton, UK. [Online]. Available at: http://94.126.106.9/R4D/PDF/Outputs/MigrationGlobPov/WP-C12.pdf.
Wouterse, F. (2010) 'Migration and technical efficiency in cereal production: evidence from Burkina Faso', *Agricultural Economics*, 41(5), pp. 385-395.

8. Appendix

Table 1: Descriptive Statistics for the Sample

Variable	Mean	Min	Max
n= 2217			
Average land area used for wheat production (ha)	1.25	0.0300	150.0
Average land area used for hay production (ha)	1.24	0.0050	30.7
Average land area used for pepper production (ha)	0.03	0.0003	3.0
Average land area used for tomatoes production (ha)	0.01	0.0003	0.9
Average land area used for onions production (ha)	0.02	0.0004	5.2
Average land area used for potatoes production (ha)	0.05	0.0004	10.2
Age of household head (years)	55.61	19	98
Gender of household head (1-male, 2-female)	1.02	1	2
Education of household head (level)	3.98	1	9
Average age of household members (years)	29.41	13	76.5
Average education of household members (category 1-9)	3.36	1.5	7.4
Full-time labour per year (no of household members)	1.13	0	21
Part-time labour per year (no of household members)	1.50	0	14
Utilised land area (ha)	2.61	0.20	151.66
Machinery value (in 2005 values in Euro)	3550.64	0	101826.5
Simpson Index	0.75	0.020	0.941
Number of plots	8.38	2	28
Product diversity index	14.30	6	43

Table 2: Overview of Matching Models

W_i	Y_i	X_i	*N*	*M*	*wm*	*bc*	*rm*
Model 1 'migration intensity' W_{MIG} - level of migration intensity (0 - falls not in specific migration category, 1 - falls in specific migration category, migration categories: >0%<=5% of total work time per hh and year used by migrants >5%<=10% >10%<=15% >15%<=20% >20%<=30% >30%<=40% >40%<=50% >50%<=60% >60%<=70% >70%<=80% >80%<=90%)	Technical efficiency per farm household and year	Value of machinery, age of household head, educational level of household head, average age of household members, average educational level of household members, female to male ratio, year dummies for 2006, 2007 and 2008 (year 2005 as reference), regional dummies for Gjakove, Gjilan, Mitrovice, Peje, Prishtine, Prizren (region Ferizaj as reference), Simpson index, product diversity index, number of plots, market integration measure, ownership of car, cattle production	2152	4	inverse variance	4	10

Model 2 'migration' W_{MIG} – indicator for migration (categories: 0 – no migration for hh and year 1 – migration for hh and year)	Technical efficiency per farm household and year	Value of machinery, age of household head, educational level of household head, average age of household members, average educational level of household members, female to male ratio, year dummies for 2006, 2007 and 2008 (year 2005 as reference) regional dummies for Gjakove, Gjilan, Mitrovice, Peje, Prishtine, Prizren (region Ferizaj as reference), Simpson index, product diversity index, number of plots, market integration measure, ownership of car, cattle production	2152	4	inverse variance	4	10

W_i: treatment condition, Y_i: indicator variable, N: number of observations, X_i: covariates; M: number of matches, wm: weighting matrix, rm: number of robust matches.

Table 3: Extent of Migration from Farm Households

	Number	% of sample
Households from which migration occurred	1016	45.8
Households without migration	1201	54.2
Households with one migrant	663	29.9
Households with more than migrant	353	15.9
Households with up to 5% migration intensity[a]	31	1.4
Households with >= 5 < 10% migration intensity[a]	401	18.1
Households with >= 10 < 15% migration intensity[a]	84	3.8
Households with >= 15 < 20% migration intensity[a]	86	3.9
Households with >= 20 < 30% migration intensity[a]	160	7.2
Households with >= 30 < 40% migration intensity[a]	112	5.1
Households with >= 40 < 50% migration intensity[a]	56	2.5
Households with >= 50 < 60% migration intensity[a]	58	2.6
Households with >= 60 < 70% migration intensity[a]	20	0.9
Households with >= 70 < 80% migration intensity[a]	5	0.2
Households with >= 80 < 90% migration intensity[a]	3	0.1

[a] Migration intensity expressed as % of total available work time per household per year

Table 4: Determinants of Inefficiency

Determinant	coefficient	t-statistic
Migration Intensity *(% of total available work time per household, per year)*	1.319***	5.75
Migration Intensity * Educational Level of Household Members	0.061***	6.93
Migration Intensity * Average Age of Household Members	0.011***	3.61
Migration Intensity * Female-to-Male-Ratio	0.178**	2.12
Migration Intensity * Total Income	0.000	3.09
Migration Intensity * Cattle	-0.093	-1.07
Migration Intensity * Farm Equipment	0.000	0.60
Average Educational Level of Household Members	-0.219***	-3.93
Average Age of Household Members	0.051***	7.71
Educational Level of Household Head	-0.254***	-11.23
Age of Household Head	0.018***	6.40
Female-to-Male Ratio	0.083*	1.94
Farm Equipment / Machinery	-0.000***	-18.73
Total Income	0.000	-0.47

Cattle	0.025	1.11
Children-to-Adult Ratio	0.372***	3.10
Simpson Index (SI)	11.739***	22.68
Number of Plots	0.457***	23.44
Measure of Products used for Household Consumption	-0.011	-0.19
Product Diversity Index	0.031***	3.39
Region Ferizai	-0.283***	-2.19
Region Prizren	-0.578***	-5.49
Region Gjakove	-0.684***	-5.21
Region Peje	-0.350***	-3.05
Region Mitrovice	0.492***	4.15
Region Prishtine	-0.789***	-6.51
Year 2006	-0.301***	-3.02
Year 2007	-0.979***	-9.56
Year 2008	-0.314***	-3.31
Constant	9.998***	19.99

*10%, **5%, ***1% significance; benchmark year: 2005; benchmark region: Gjilan

Table 5: Sample Average Treatment Effects (SATE) in Efficiency for Different Levels of Migration Intensity

Migration Share/Intensity *(% of total available worktime per hh and year used by migrants)*	Change in Technical Efficiency due to Migration at Household Level (SATE)		
	mean	min	max
0% >= 5%	-0.164***	-0.275	-0.053
5% >= 10%	-0.034**	-0.062	-0.006
10% >= 15%	-0.041*	-0.104	0.023
15% >= 20%	0.009	-0.045	0.064
20% >= 30%	-0.044**	-0.091	0.002
30% >= 40%	3.489e-04	-0.049	0.049
40% >= 50%	-0.081**	-0.155	-0.006
50% >= 60%	0.009	-0.076	0.096
60% >= 70%	-0.189***	-0.203	-0.175
70% >= 80%	-0.105**	-0.285	0.075
80% >= 90%	-0.364***	-0.531	0.197
Migration (Yes/No)	-0.052***	-0.073	-0.031

*, **, *** - significant at 10, 5, 1%-level.

Technical Appendix: Matching Estimation Approach

The underlying framework of analysis refers to the Neyman-Rubin's model of matching methods for causal inference counterfactual framework (Guo and Fraser, 2010) where farm households selected into treatment and non-treatment groups have potential outcomes (Y_0, Y_1) in both states (W=0,1): the one in which the outcomes are observed ($E[Y_1|W=1]$, $E[Y_0|W=0]$) and the one in which the outcomes are not observed ($E[Y_1|W=0]$, $E[Y_0|W=1]$). Unobserved potential outcomes under either condition are missing data. A matching estimator directly imputes the missing data at the unit level by using a vector norm. Specifically, it estimates the values of $Y_i(0)|W_i = 1$, i.e. the potential outcome under the condition of control for the treatment participant, and $Y_i(1)|W_i = 0$ as the potential outcome under the condition of treatment for the control participant.

The central challenge is the dimensionality of covariates or matching variables, because as their number increases, the difficulty of finding matches for treated farm households also rises. Matching estimators use the vector norm to calculate distances on observed covariates between treated case and each of its potential control cases (i.e. counterfactuals). However, two assumptions are critical: the assignment to a specific treatment group is independent of outcomes and that there is sufficient overlap in the distribution of observed covariates (Abadie and Imbens, 2011).

Let the unit-level treatment (i.e. migration) effect for farm observation i be

$$\tau_i = Y_i(1) - Y_i(0) \tag{A1}$$

As one of the outcomes is always missing, the matching estimator (ME) imputes this missing value based on the average outcome for farm households with "similar" values on observed covariates. A simple ME is:

$$\hat{Y}_i(0) = \begin{cases} Y_i & if\ W_i = 0 \\ \frac{1}{\#J_M(i)} \sum_{l \in J_M(i)} Y_l & if\ W_i = 1 \end{cases} \quad \hat{Y}_i(1) = \begin{cases} \frac{1}{\#J_M(i)} \sum_{l \in J_M(i)} Y_l & if\ W_i = 0 \\ Y_i & if\ W_i = 1 \end{cases} \tag{A2}$$

where $J_M(i)$ as the set of indices for the matches for farm household observation i and $\#J_M(i)$ as the number of elements of $J_M(i)$. In the case of more than one observed covariate, the ME uses the vector norm to calculate distances between treated case and each of its multiple possible control cases. Consequently, M matches are chosen using the vector norm based on the condition of nearest distances applying

$$J_M(i) = \{l = 1, \ldots, N | W_l = 1 - W_i, \|X_l - X_i\| v \le d_M(i)\} \tag{A3}$$

with $d_M(i)$ as the distance from the covariates for unit i, X_i, to the Mth nearest match with the opposite treatment. Then point estimates for various treatment effects (i.e. migration levels) are obtained e.g. by the sample average treatment effect (SATE):

$$\hat{\tau}^{average} = \frac{1}{N} \sum_{i=1}^{N} \{\hat{Y}_i(1) - \hat{Y}_i(0)\} = \frac{1}{N} \sum_{i=1}^{N} (2W_i - 1)\{1 + K_M(i)\} Y_i \tag{A4}$$

where $K_M(i)$ are the number of times farm household observation i is used as a match, with M matches per unit i, and W_i as the treatment condition for unit i. Abadie *et al.* (2004) recommend using four matches for each unit since the drawback of using only one match is that the process uses too little information in matching. As we use continuous covariates, a bias-corrected matching estimator (Abadie and Imbens, 2002) is required which uses a least square regression to adjust for potential bias. Further, the assumption of a constant treatment and homoscedasticity may not be valid for certain types of covariates. To account for such potential heteroscedasticity we use a second matching procedure, matching treated to treated and control to control cases (Abadie *et al.*, 2004).