Working Paper
T11

Who is Most Likely to Migrate from Albania? Evidence from the Albania Living Standards Measurement Survey

ADRIANA CASTALDO, JULIE LITCHFIELD AND BARRY REILLY
(*Sussex Centre for Migration Research)*

June 2005

Issued by the Development Research Centre on
Migration, Globalisation and Poverty

The UK Department of for International Development (DFID) supports policies, programmes and projects to promote international developmentpoverty reduction globally. DFID provided funds for this study as part of that objective goal but the views and opinions expressed are those of the author(s) alone.

Development Research Centre on
Migration, Globalisation and Poverty
Arts C-226, University of Sussex
Brighton BN1 9SJ

Website: http://www.migrationdrc.org
Email: migration@sussex.ac.uk
Tel: +44 (0)1273 873394
Fax: +44 (0)1273 873158

C O N T E N T S

ABSTRACT

This paper uses the Albania Living Standards Measurement Survey from 2002 to examine the factors that render an individual most prone to the risk of international migration. The analysis uses novel data on whether individuals ever considered migrating abroad. The econometric models used provide good descriptions of the data and are well specified on the basis of a battery of diagnostic tests conducted. The estimated results are generally consistent with findings from the empirical literature on the willingness to migrate. The usual suspects emerge as determining factors with age, gender, employment status and education exerting predictable influences on migration risk. There is also a strong role for local labour market conditions and community level variables capturing, among other things, the prevalence of crime.

1. INTRODUCTION

It is over ten years since the first centrally planned economy implemented a comprehensive programme of economic reforms designed to expedite transformation to a market-based economy. The final decade of the twentieth century witnessed the largest economic experiment of recent times as other former communist countries embarked on similar endeavours to transform their economies. The transformation process influenced the direction of economic policies and shaped social policies, business practices and institutions. The collapse of the central planning system in Europe and the former Soviet Union also provided the erstwhile citizens of communist regimes with greater opportunities to migrate abroad.

The mass exodus anticipated in some of the early writings on the transition[1] did not materialize and over the decade migration flows from east to west were generally modest in comparison to original expectations. However, Albania proved something of an exception to this general rule and experienced a steady increase in migration over the early period of its transition. By the end of the decade over one-fifth of the Albanian population were estimated to be living abroad representing the largest outflow relative to population of any transitional economy.[2]

[1] For example, see Layard, Blanchard, Dornbusch and Krugman (1992).
[2] The estimates reported in table 5.2.4 of the UNECE *Economic Survey of Europe* (2003) places the Albanian experience in the broader context of other transitional economies.

In more recent years, however, there has been modest progress as the government of Albania, under the framework of the Growth and Poverty Reduction Strategy (GPRS), embarked on reforms designed to stimulate economic growth and improve the living standards of the most vulnerable. In the early years of the current decade, Albania has registered a stronger economic performance with steady economic growth, reductions in the unemployment rate and more stable inflation (see Table 1). Structural programmes have been introduced to tackle financial regulation, land reform and privatization. In addition, there has been a strengthening of governance systems and an anti-corruption plan is in the process of implementation. In spite of some positive economic developments, poverty remains high in Albania and per capita income is one of the lowest of all the transitional countries. The World Bank's recent poverty assessment of Albania estimated that over one-quarter of the population were below an absolute poverty line.[3]

Table 1. Main Economic Indicators for Albania 1996-2002

	1996	1997	1998	1999	2000	2001	2002
Annual GDP Growth Rate (%)	9.1	-10.2	12.7	10.1	7.3	7.2	3.4
Annual CPI Inflation Rate (%)	17.4	42.1	8.7	-1.0	4.2	3.5	1.7
Unemployment Rate (%)	12.3	14.9	17.7	18.4	16.8	16.4	15.8

Source: INSTAT website

Migration, whether it is internal or external in character, provides an important strategy for coping with poverty in Albania. The scale of internal migration has induced a radical demographic transformation within the country with Tirana as the primary recipient of internal migratory movements. However, for a sizeable portion of internal migrants, the process represents nothing more than a prelude to an external move. The economic conditions, though improving, are unlikely to alter the external migration pattern that has marked Albania over the last fifteen years. In addition, migrant remittances represent an important source of foreign exchange for Albania and,

[3] The absolute poverty line used in the World Bank assessment was 4891 Leks per capita per month. The report noted that a large number of individuals are clustered around the poverty line and thus changes to the poverty induce large changes to the numbers defined as poor (see World Bank 2003).

as noted by Kule *et al.* (2002), migrants are generally well positioned to find a job or establish a business on their return to Albania.

A number of studies have documented and profiled the nature of the Albanian migration experience over the last decade and even earlier.[4] For instance, the study by Kule *et al*. (2002) used data drawn from a specially commissioned survey conducted in 1998 to elicit responses from both individuals and enterprises. The study provides useful and informative insights into the profile of Albanian migrants.[5] They tend to be young, disproprtionately male, better educated, and they tend to migrate to either Greece or Italy.

The recent Albania Living Standards Measurement Survey (ALSMS) undertaken in 2002 represents an opportunity to complement our understanding of migration at a more recent point in Albania's transitional experience. The survey's primary objective is to delineate poverty profiles within Albania, but it also contains detailed information on migration and important labour market information at the individual level. In addition, this survey also contains a community questionnaire that informs on community level factors that potentially impact on migration behaviour.

The purpose of the current paper is to explore the factors that influence whether someone who has not migrated in the last five years is currently at risk of doing so. Our approach uses the 2002 ALSMS and exploits a combination of individual level, household level, district level and community level information. In particular, we are interested in exploring the key factors that determine whether an individual considered migrating abroad from Albania with a particular emphasis on the role of age, gender, education, and household welfare. In addition, we are also keen to examine the role of district level unemployment and wage rates and selected community level variables in determining whether the migration option was considered.

[4] For example, see Mancellari, Papapanagos and Sanfey (1996), Papapanagos and Sanfey (2002), Kule et al. (2002) for a contemporary account of external migration. In addition, King and Vullnetari (2003) provide a detailed account of the more recent migration experience and earlier historical episodes.

[5] Although the survey was conducted in a careful and systematic way, the authors acknowledge a number of limitations asscoiated with the survey work.

The structure of the paper is now briefly outlined. Section One discusses a number of conceptual issues that attach to employing the migration measure used in this study and is followed by a section describing the data. Section Three outlines the econometric methodology and Section Four details the empirical results. A final section offers a summary and some conclusions.

2. CONCEPTUAL ISSUES

A systematic analysis of migration behaviour is invariably constrained by the partial nature of information available to the investigator. Surveys of actual migrants in labour markets can be effective at profiling their demographic and labour force characteristics but are generally poorer at generating accurate information on the structure, welfare status and composition of the households from which they originated. The reverse is generally the case for household level surveys where limited information tends to be available on migrants living abroad. These data limitations have prompted economists to explore migrant behaviour more indirectly through the use of intentions data, and cast the empirical analysis in terms of migration willingness.[6]

The use of such intentions data, though popular among social psychologists, is not without its critics within the economics profession (for example, see Manski 1990). The discipline emphasizes how behaviour can generally be inferred from actual revealed preference (and not intended) outcomes. Future behaviour is then predicted from past behaviour through judicious use of appropriate econometric models. Manski (1990) concludes that although there is some informational content in intentions-based survey questions, researchers should not expect too much from such data. In contrast, Louviere, Hensher and Swait (2000) take a different view and argue that stated (rather than revealed) preferences are capable of generating data consistent with economic theory and facilitate the estimation of econometric models that are almost indistinguishable from those using revealed preference data. In addition, they argue that the empirical track record of stated preference data models is as impressive as their revealed preference counterparts.

[6] For example, see Hughes and McCormick (1985), Papapanagos and Sanfey (2002), Ahn, De la Rica, Ugidos (1999), Ahn, Jimeno, Garcia (2002), Drinkwater (2003a), Drinkwater (2003b).

It is worth noting that in the context of migration where follow-up surveys were conducted, there is some evidence that intentions are strongly correlated with actual outcomes (see De Jong 2000). There are also many reasons why the migration intentions of individuals are not actually realised. The deficit between intended and actual migration outcomes is likely to be larger in regard to external rather than internal migration (see Gardner, De Jong, Arnold and Carino 1986). Intentions may also evolve over time if preferences or personal circumstances change. In addition, formulating expectations about the costs and benefits associated with migration is extremely difficult when information is unobservable on the destination country's labour market conditions and when individuals are asked to operate in an uncertain and unpredictable environment in the source country.[7] The theoretical work of O'Connell (1997) demonstrates why such uncertainties may induce a 'wait and see' approach to migration and thus partly explains the observed wedge between intentions and outcomes.

The current study does not use responses to a question on intentions *per se* to explore the propensity of individuals to re-locate abroad. The responses used are based on the question, '*Has individual X ever considered moving abroad, even temporarily?*' The question is thus more retrospective and, arguably, may be more immune to criticism than a question based on intentions. In our view it captures whether someone could be considered to be at risk of migrating abroad.

There are, however, a number of potential problems that attach to the use of responses based on this particular question and to whom the question was asked in the Albanian Living Standards Measurement Survey (ALSMS) in 2002. Firstly, the timeframe to which the question relates is open-ended and it is unclear whether a positive response to this question relates to a more recent period of time or is based on an individual's longer historical perspective. This poses a potential alignment problem in that the covariates used to explain an individual's disposition to migrate abroad comprise contemporary realizations of the measures available and not historical ones. We assume, however, that respondents are more likely to attach a heavier weight to the more recent past in furnishing a response to this question. Secondly, individuals who have actually migrated abroad in the past five years were not asked this particular question for obvious reasons. Thus, the

[7] The existence of this type of uncertainty could be taken to characterise many post-socialist economies in the early phase of their transition to a market based economy.

sample that we use is conditioned on those that have not recently migrated abroad. This potentially poses a sample selection problem but one that is not tractable for us given the absence of information on those migrants who are currently based abroad. However, we may be in a position to offer some insights on whether selection bias poses a serious problem for our econometric estimates when we subject our estimated models to a battery of diagnostic tests. Finally, given the tendency of those in rural areas to engage in internal rather than external migration in the first instance, they may not have a tendency to offer a positive response to this question. We do not believe that this poses too much of a problem for us in this case as internal migration is likely to represent a stepping stone to an external move abroad.

It is acknowledged that there are legitimate concerns in making links between whether an individual considered or intended migrating at some point in time and actual migration behaviour. The general scepticism among economists about the reliability of subjective statements is warranted to some degree as, it is argued, respondents have little or no incentive to provide honest responses.[8] However, as noted by Manksi (1995: 151 fn 12), there is something of an asymmetry in the profession's concern on this matter given an almost unquestioning willingness to use, *inter alia*, respondents' self-reported data on earnings, socio-economic status and other demographic characteristics in micro-econometric empirical work. These concerns and limitations aside, given the absence of data reflecting actual outcomes, we believe an examination of whether an individual considered migrating is a useful exercise in its own right. In particular, the approach provides the basis for a systematic analysis of the profile of individuals considered most at risk of future migration from Albania. This type of analysis is clearly fruitful and potentially incorporates a strong policy content.

[8] The scepticism within the economics discipline is softening somewhat with an expanding literature that uses subjective measures to inform on household poverty (see Revallion and Lokshin 2000 and Pradham and Revallion 2000).

3. DATA

There is a dearth of household-level and/or individual-level data available on migrants for Albania and this clearly constrains understanding of the phenomenon. As noted earlier, Kule et *al.* (2002) conducted surveys in the late 1990s at both an individual and enterprise level to inform on both the profile of migrants and the role of migration in enterprise development in Albania. The more nationally representative and first Albania Living Standards Measurement Survey (ALSMS), conducted in spring 2002, contains a module on migration and provides researchers with the opportunity to investigate in more detail some limited aspects of Albanian migration.

The ALSMS was recently exploited by the World Bank to undertake that country's poverty assessment (see World Bank 2003).[9] The chosen sample of 3,599 households was selected using a two-stage cluster design with 450 primary sampling units selected from a 2001 pre-census list of census enumeration areas. The sampling frame was stratified across four regions (Tirana, Coastal, Central and Mountain) and 1,640 of the households were situated in rural settlements.[10]

The selected sample of households elicited responses on 16,521 individuals. Given the objectives of our research, this sample has been refined to include only those individuals aged between 15 and 60 who are defined to be in the labour force. These are by the conventional definition those who are either employed or unemployed but actively looking for work. The sample is further conditioned on those who have not migrated from Albania in the last five years but expressed an opinion on whether they had ever considered migrating or not. Once we exclude individuals on whom there are missing values for either this key variable or other variables of interest, we are left with an overall sample of 5,423 individuals on whom usable data are available.

Given the structure of the LSMS, a rich array of individual level, household level, and community level information is available to us. The set of individual level variables that can be constructed from the survey include, *inter alia*, age, gender, marital status, health status, educational level achieved, and current employment status. The set of household level variables that can be

[9] The ALSMS (2002) is the product of a collaborative project involving the Albanian National Institute of Statistics (INSTAT), the World Bank and the UK's Department for International Development (DfID).
[10] There are additional details on the survey methodology and fieldwork available in the Albanian 2002 LSMS Basic Information Document on www.worldbank.org/lsms.

constructed from the data comprise the demographic structure and welfare status of the household (i.e., household size and age structure of dependent children within the household etc., and total household consumption expenditure[11]). Measures relating to the quality of the dwelling in which the individuals reside (i.e., the dwelling size, the age of the dwelling and the presence of an internal water closet) are also available. A set of community level variables can be constructed from responses to the community questionnaire and these capture, *inter alia*, the quality of the community environment in regard to crime and physical conditions and whether land disputes are a characteristic feature of the community. Finally, in order to construct measures that capture the influence of local labour market conditions, an unemployment rate and an hourly wage rate measure is constructed using district level data from the LSMS.[12]

Table A1 of the appendix provides a full description of the variables used in our analysis and Table A2 provides summary statistics for the variables, which are also reported by gender. The summary statistics reveal that almost one-third of the sample considered migration abroad and that men are almost twice as likely as women to consider this option.

4. ECONOMETRIC METHODOLOGY

The dependent variable in this application assumes either a value of one or zero depending on whether an individual considered migrating abroad or not. A probit model is used in estimation given the limited binary nature of the dependent variable.[13] The binary probit is generally motivated by reference to a latent (or unobservable) dependent variable (y_i^*) and is usually expressed as a linear function of a set of explanatory variables as follows:

[11] Household expenditure is preferred to household income as a measure of household welfare given that it is less prone to shocks and respondents are less diffident at providing this information than income.
[12] An attempt to construct local labour market measures using the community as the basis of dis-aggregation was abandoned due to small cell size problems. This prompted the use of district level data. There are 36 districts used in this analysis.
[13] Alternatively, a logistic model could be used in estimation here. Although there are practical interpretational benefits to using the logit, no process, experiment or model naturally engenders the underlying logistic distribution. This is not the case for the underlying normal distribution inherent in the probit model. This explains our preference for the probit model in this application.

$$y_i^* = x_i'\gamma + u_i \qquad [1]$$

where $u_i \sim N(0,\sigma^2)$, i=1,……n, x_i is a column vector of realizations on k explanatory variables for individual i and γ is a column vector of k unknown parameters. The values of the latent variable are measured on the real line and in this case reflect the underlying propensity of an individual to consider migrating. The error term is assumed normally distributed with a mean of zero and a constant variance σ^2. A threshold (assumed zero in this case) is used to delineate whether the individual considers migrating or not. The probability can be linked to the latent variable as follows:

$$\text{Prob}[y_i^* > 0] = \text{Prob}[y_i = 1] = \Phi(z_i) \qquad [2]$$

where y_i is the dichotomous realization of the latent dependent variable, $\Phi(\cdot)$ denotes the cumulative distribution function for the standard normal, and $z_i = \frac{x_i'\gamma}{\sigma}$. For identification purposes it is conventional to normalize σ =1. Given this restriction, the estimated probit coefficients can be interpreted by reference to their effect on the standardized probit index.

The i^{th} generalized residual for the probit model is computed as[14]:

$$\hat{u}_i = \frac{y_i - \Phi(\hat{\theta}_i)}{\Phi(\hat{\theta}_i)\times[1-\Phi(\hat{\theta}_i)]}\phi(\hat{\theta}_i) \qquad [3]$$

where $\phi(\cdot)$ denotes the probability distribution function (or density function) for the standard normal and the 'hat' denotes estimates based on the maximum likelihood technique. The set of generalized (or pseudo) residuals are important in evaluating the inherent assumptions of the probit model (e.g., normality and homoscedasticity) since they are used to compute the score contributions of the estimated parameters for the model's diagnostic tests.[15]

[14] The generalized residual is obtained as the first order derivative of the log-likelihood function with respect to the probit model's constant term. It is interpretable as the score contribution for the constant term.

[15] They also can be used to construct a robust variance-covariance matrix for the probit model as per Huber (1967).

The maximum likelihood estimation of probit-type models assumes an exact knowledge of the probability distribution function up to the set of unknown parameters that are the subject of the estimation procedure. The estimates are sensitive to departures from the specification of the likelihood function. Thus, the failure of the normality assumption may have implications for model specification. However, even if the maximum likelihood (ML) estimator is based on a mis-specified likelihood function, the estimates could be viewed as consistent through the interpretation of the estimator within the less stringent Generalized Methods of Moments (GMM) framework. In this case, consistency relies less restrictively on the validity of first order conditions relating to the maximum likelihood problem.[16]

The presence of heteroscedasticity in the probit model, on the other hand, yields inconsistent parameter estimates since, as noted above, the assumption of a constant variance is a necessary though not sufficient condition to identify the parameters of the mean function.[17] However, if the primary concern of the investigator, as in our case, is an examination of the effect of a covariate on the probability of the event occurring, it matters little whether such an effect is mediated through the mean or the variance function. It is thus sometimes argued that in most cases the problem of heteroscedasticity in a probit model may be more '…apparent than real' (see Johnston and DiNardo 1999: 426–427). Nevertheless, determining whether a regression model is characterized by heteroscedasticity is useful in its own right.

Given that the normality and/or the homoscedasticty assumption may be vitiated in many cases, the probit maximum likelihood estimator is now increasingly interpreted as a quasi-maximum or pseudo-maximum likelihood estimator (see Gouriéoux, Monfort and Trognon 1984) and requires computation and use of a robust variance-covariance matrix (see White 1980 or Huber 1967).[18]

[16] It is for this reason that the ML estimator in the normal linear regression model retains the property of consistency even if the assumption of normality is actually violated.
[17] This inconsistency reflects the fact that heteroscedasticity in the probit model is analogous to functional form mis-specification in a linear regression model.
[18] It is acknowledged, as noted by Greene (2000: 823-4), that such a correction to the variance-covariance matrix for an otherwise inconsistent estimator may not be sufficient to redeem it.

It is thus of some importance to evaluate the estimated models in regard to their key econometric assumptions. The adequacy of the estimated models is assessed in this paper using efficient score tests of the type originally suggested in Chesher and Irish (1987) and these are expressed in their matrix form as $\mathbf{i}'R(R'R)^{-1}R'\mathbf{i}$. In this case, $\mathbf{i}$ is an $n\times1$ vector of ones, and R is an $n\times q$ matrix of score contributions computed for each of the k parameters from the original specification and the k+1,….q parameters that capture the form of the alternative hypothesis, assumed zero under the null. The resultant test statistics are all distributed as chi-squared with $p = q - k$ degrees of freedom. The test represents the outer-product gradient (OPG) form of the score (or Lagrange Multiplier) test.

The efficient score tests undertaken are designed to assess the reported specifications in terms of functional form and omitted variables (RESET), homoscedastic errors, and a symmetric and meso-kurtic distribution of the generalized residuals. Given the interpretation of the RESET as a test for omitted variables it may prove useful in this context as a test for the selection bias if the selection bias is interpretable as an omitted variables problem.[19]

Orme (1990) has questioned the use of OPG-based tests and, using Monte Carlo simulations, demonstrated their poor finite sample properties. Orme's findings suggest that efficient score tests constructed using the OPG covariance matrix tend to reject the correct null hypothesis far too frequently. The poor performance of score tests in relatively small samples is acknowledged but our analysis is based on a reasonably large sample, thus allowing for some degree of confidence in the efficient score tests used in this application.[20]

[19] This was how Heckman (1979) originally interpreted the selection bias problem in the linear regression framework. It is conceded that the detection of selection bias in this context sets the RESET a very challenging task and there is no guarantee that it can meet the challenge in all circumstances.

[20] Even if we take the size bias from the simulation studies at face value, it is clear we are setting our empirical models fairly stern tests in requiring them to satisfy the set of econometric assumptions. In particular, the direction of the bias would not alter an efficient score test inference that indicated non-rejection of a particular null hypothesis.

5. EMPIRICAL RESULTS

Table 2 reports the probit estimates for a number of different regression models. The marginal and impact effects are also reported in Table A3 of the appendix to this paper, and should be viewed in conjunction with the set of probit estimates. As outlined in Section 3, a number of diagnostic tests are also computed for each model to determine if the underlying econometric assumptions are satisfied.

5.1 The Pooled Model's Estimates for Considering Migration

Our emphasis in the first instance focuses on the estimates reported in the first column of Table 2 using a model that pools data points across gender. The RESET relating to potential omitted variables and functional form is passed at a respectable level of statistical confidence[21], as is the normality assumption. However, the null hypothesis of homoscedasticity in the probit model is rejected prompting use of the robust variance-covariance matrix as *per* Huber (1967). The McFadden Pseudo-R^2, the estimated squared correlation coefficient between the predicted probabilities and the actual outcomes (denoted r^2), and the Cramer (1999) measure (denoted λ) are in concordance with each other and suggest satisfactory fits.[22]

[21] Peters (2000) provides Monte-Carlo simulation evidence on the power of the RESET in cross-sectional applications.
[22] A desirable feature of the Cramer (1998) statistic is that it measures the predictive failure of the model and heavily penalizes incorrect predictions.

Table 2: Probit Estimates for the Determinants of Migration Risk in Albania

	Pooled Model	Male Model	Female Model	Pooled Model with Gender Interactions
Constant	-0.2206 (1.4688)	3.2657* (1.9756)	-5.7635** (2.4976)	-8.1558*** (2.2835)
Age Spline: 15 – 25 years	0.0270** (0.0118)	0.0434*** (0.0164)	0.0176 (0.0192)	0.0311** (0.0123)
Age Spline: 16 – 35 years	-0.0284*** (0.0092)	-0.0537*** (0.0138)	-0.0178 (0.0141)	-0.0322*** (0.0093)
Age Spline: 36 – 45 years	-0.0121 (0.0086)	-0.0195* (0.0117)	-0.0165 (0.0135)	-0.0173** (0.0086)
Age Spline: 46 – 55 years	-0.0488*** (0.0107)	-0.0402*** (0.0133)	-0.0351* (0.0192)	-0.0411*** (0.0108)
Age Spline: 56 – 60 years	-0.1341*** (0.0439)	-0.1716*** (0.0481)	0.0158 (0.0907)	0.0127 (0.0836)
Male	0.6541*** (0.0521)	§	§	13.060*** (2.5247)
Head of Household	0.0023 (0.0614)	0.2473*** (0.0888)	0.1721 (0.1547)	0.1792*** (0.0700)
Married	0.1036 (0.0767)	-0.0378 (0.1110)	0.2553** (0.1158)	0.3294*** (0.0956)
Divorced/Separated	0.2633* (0.1526)	0.6455** (0.3106)	0.0863 (0.2121)	0.2536* (0.1543)
Single	*f*	*f*	*f*	*f*
Primary: ≤ 4 grades	*f*	*f*	*f*	*f*
Primary: 5 to 8 grades	0.1975* (0.1105)	0.2272 (0.1454)	0.3029* (0.1763)	0.2324** (0.1107)
Secondary	0.2536** (0.1190)	0.1445 (0.1567)	0.5241*** (0.1903)	0.2772** (0.1195)
Vocational	0.3627*** (0.1191)	0.3157** (0.1555)	0.5879*** (0.1914)	0.4034*** (0.1193)
University	0.1464 (0.1299)	0.1189 (0.1695)	0.3685* (0.2106)	0.1909 (0.1301)
Born in the Municipality	0.0620 (0.0427)	-0.0158 (0.0624)	0.1015 (0.0624)	0.0640 (0.0433)
Health Disability	0.0557 (0.0605)	-0.0303 (0.0821)	0.1066 (0.0919)	0.0346 (0.0609)
Unemployed	*f*	*f*	*f*	*f*
Employee	-0.1753*** (0.0650)	-0.1930** (0.0863)	-0.1433 (0.1005)	-0.1739*** (0.0654)
Farmer	-0.2295*** (0.0782)	-0.1011 (0.1034)	-0.2893** (0.1314)	-0.2151*** (0.0791)
Self-Employed	-0.2356*** (0.0857)	-0.0969 (0.1080)	-0.3197** (0.1507)	-0.1928** (0.0858)
Temporary Layoff	-0.2708* (0.1545)	-0.0229 (0.2483)	-0.4661** (0.2179)	-0.2794* (0.1581)
Household Size	-0.0152 (0.0139)	-0.0153 (0.0196)	0.0049 (0.0214)	-0.0077 (0.0142)

Children in Household: Aged $\leq$ 4 years	0.0671 (0.0521)	0.0386 (0.0701)	0.1302 (0.0829)	0.0705 (0.0527)
Children in Household: Aged 5 $\leq$ years $\leq$ 8	0.0183 (0.0460)	0.0907 (0.0628)	-0.0721 (0.0727)	-0.1048 (0.0676)
Children in Household: Aged 9 $\leq$ years $\leq$ 14	-0.0083 (0.0461)	-0.0062 (0.0620)	-0.0020 (0.0742)	-0.0061 (0.0465)
Log Consumption: First Quintile	-0.0132 (0.1663)	-0.2266 (0.2240)	0.4769* (0.2840)	0.6924*** (0.2661)
Log Consumption: Second Quintile	0.4647* (0.2693)	0.2803 (0.3607)	0.7307* (0.4215)	0.4018 (0.2722)
Log Consumption: Third Quintile	0.2976 (0.3458)	0.5384 (0.4555)	0.0169 (0.5380)	0.3350 (0.3451)
Log Consumption: Fourth Quintile	0.1657 (0.2594)	-0.2069 (0.3452)	0.5596 (0.3956)	0.1305 (0.2590)
Log Consumption: Fifth Quintile	0.0755 (0.1432)	0.0508 (0.1866)	-0.0078 (0.2148)	0.0425 (0.1408)
Residence Dwelling Area: $\leq$ 69 Sq.Metres	*f*	*f*	*f*	*f*
Residence Dwelling Area: 70 $\leq$ Sq.Metres $\leq$ 130	-0.0902** (0.0400)	-0.0601 (0.0534)	-0.1312** (0.0624)	-0.0893** (0.0404)
Residence Dwelling Area: Sq.Metres > 130	-0.2855*** (0.1061)	-0.3118** (0.1414)	-0.2292 (0.1591)	-0.2740*** (0.1063)
Residence constructed after 1990	0.1545*** (0.0466)	0.1799*** (0.0624)	0.1295* (0.0732)	0.1515*** (0.0470)
Residence contains Internal Watercloset	-0.0908* (0.0494)	-0.0699 (0.0666)	-0.1404* (0.0774)	-0.0953* (0.0500)
Central Region	0.1916*** (0.0529)	0.2191*** (0.0718)	0.1511* (0.0804)	0.1806*** (0.0531)
Coastal Region	*f*	*f*	*f*	*f*
Mountain Region	-0.0281 (0.0608)	-0.0152 (0.0846)	-0.0839 (0.0917)	-0.0459 (0.0617)
Tirana	-0.0966 (0.0734)	-0.0973 (0.0949)	-0.0402 (0.1157)	-0.0798 (0.0732)
Urban Settlement Type	0.1037 (0.0635)	0.0666 (0.0800)	0.1898* (0.1132)	0.1329** (0.0641)
District Level Unemployment Rate (%)	0.0114*** (0.0026)	0.0160*** (0.0035)	0.0068 (0.0042)	0.0118*** (0.0027)
District Level Hourly Wage (log)	-0.3126*** (0.1217)	-0.6115*** (0.1701)	-0.0270 (0.1813)	0.0894 (0.1559)
Land Disputes in Commune	0.0824** (0.0391)	0.1780*** (0.0529)	-0.0326 (0.0606)	-0.0531 (0.0590)
Thefts in Commune	0.0799* (0.0430)	-0.0399 (0.0575)	0.2407*** (0.0654)	0.3263*** (0.0622)
Stagnant Water in Commune	0.1341*** (0.0465)	0.0932 (0.0623)	0.1911*** (0.0706)	0.1371*** (0.0465)
No Migration in Commune	-0.3295* (0.1802)	-0.5386** (0.2347)	-0.0959 (0.2562)	-0.3370* (0.1771)
Gender Interactions:				
Male×[Age Spline: 56 – 60 years]	§	§	§	-0.1762* (0.0906)

Male×[Children in Household:Aged $5 \leq$ years ≤ 8]	§	§	§	0.1976** (0.0860)
Male×[Married]	§	§	§	-0.4810*** (0.1065)
Male×[Log Consumption: Spline: First Quintile]	§	§	§	-1.0443*** (0.2977)
Male× [District Level Hourly Wage (log)]	§	§	§	-0.7970*** (0.1863)
Male×[Land Disputes in Commune]	§	§	§	0.2401*** (0.0772)
Male×[Thefts in Commune]	§	§	§	-0.4349*** (0.0787)
Number of Observations	5423	2841	2582	5423
Log Pseudo Likelihood Value	-3119.307	-1779.567	-1252.228	-3060.482
Mc Fadden's Pseudo-R^2	0.092	0.081	0.091	0.109
r^2	0.111	0.104	0.094	0.130
Cramer's λ	0.111	0.103	0.095	0.130
Efficient Score Tests:				
1.RESET ~ χ^2_3	4.418 (0.220)	2.035 (0.565)	4.622 (0.202)	2.177 (0.537)
2. Normality ~ χ^2_2	3.623 (0.163)	1.844 (0.398)	4.599 (0.101)	0.322 (0.851)
3. Homoscedasticity ~ χ^2_k	125.57 (0.000)	57.004 (0.049)	47.044 (0.239)	80.932 (0.002)

Notes to Table 2

(a) ***, ** and * denote statistical significance at the 0.01, 0.05 and 0.10 level respectively using two-tailed tests.
(b) The variance-covariance matrix for the estimated probit model is corrected for the presence of heteroscedasticity of unknown form (see Huber(1967)).
(c) The diagnostics are computed using the efficient score test approach outlined in Chesher and Irish (1987) (see text). The RESET uses the predicted standardized indices raised to the second, third and fourth powers. The normality diagnostic tests for departures from zero skewness and a kurtosis value of 3 in the generalized residuals. The homoscedasticity test uses the set of explanatory variables from the original model and is interpreted as a variant of the White test in a linear regression model. The degrees of freedom for the test are equal to the number of explanatory variables in the estimated model. The significance levels for all the tests are reported in parentheses.
(d) The Pseudo-R^2 is defined as $1 - L^{restricted}/L^{unrestricted}$ where L denotes the maximised value of the log-likelihood function. This is more commonly known as the McFadden R^2.

(e) The r^2 is the squared correlation coefficient between the predicted probabilities $\Phi(x_i \hat{\gamma})$ and the actual outcomes y_i.

(f) Cramer's $\lambda = \overline{\Phi(x_i \hat{\gamma})} \mid y_i = 1 - \overline{\Phi(x_i \hat{\gamma})} \mid y_i = 0$ where y_i is the dependent variable in this case (see Cramer 1999).
(g) § denotes not applicable in estimation and *f* denotes category omitted in estimation.

We now provide comment on the probit estimates themselves. In order to introduce piece-wise linear effects for an individual's age, the variable is splined using the nodes reported in Table 2. The splined age effects are well determined and signed as anticipated, with the youngest labour force members (those aged less than 25 years) more favourably disposed to migration, on average and *ceteris paribus*. This is consistent with estimates reported by Papapanagos and Sanfey (2002) who used Albanian data from 1992 drawn from the Central and Eastern Europe Eurobarometer. In addition, although the more educated considered migrating abroad, this is generally confined to those with secondary and vocational level rather than those with the highest university education. In particular, an individual with a vocational qualification is, on average and *ceteris paribus*, over 13 percentage points more likely to consider migrating relative to individuals with a primary grade four education or less. Not surprisingly, the unemployed (the base group in this case) are more likely to consider migrating than individuals in any of the other employment status groups. It is also interesting to note that self-employed individuals, perhaps the least risk-averse and most entrepreneurial, are as reticent as most others in employment to consider migration abroad.

The strongest estimated effect is reserved for gender. The probit estimate suggests that men are more willing to consider migration than women.[23] The standardized probit index is 0.65 of a standard deviation higher for a male relative to a female. This suggests that men, on average and *ceteris paribus*, are over 22 percentage points more likely to consider migrating than women. The interesting issue of gender is revisited in more detail below.

In general, the household demographic variables perform less well in explaining the phenomenon of interest here than the individual level variables. The estimated coefficient for the head of household variable is poorly determined and although the corresponding coefficient for the household size variable is negatively signed it is imprecisely determined. None of the estimated effects for the age structure variables for the presence of dependent children in the household is statistically significant at a conventional level.

In order to explore the relationship between household welfare and migration risk, the log of total household expenditure is splined with the quintiles of log consumption determining the nodes. The

[23] This finding also resonates with that of Papapanagos and Sanfey (2002).

probit estimates reveal that individuals in households situated within the bottom quintile of the consumption distribution are unlikely to consider migrating. It is worth noting that Papapanagos and Sanfey (2002) detected no independent effect for their income measure on Albanian migration intentions. However, the estimated effect for the second spline suggests that changes in consumption expenditure for individuals in households located within the second quintile of the consumption distribution induces an increased risk of migration. On the basis of the marginal effect reported in Table A3, a 5 percent increase of household expenditure raises the probability of migration risk by 0.8 of a percentage point. Thus, the estimated effect is only significant at the 10 percent level, but the finding provides some tentative evidence on the relationship between household impoverishment and the propensity to migrate.

In contrast to the results obtained for the various household demographic variables, the measures for residence quality provide more precisely estimated coefficients. These measures may be viewed as proxying for household wealth, and the sign of the estimates are consistent with this notion. For instance, those in dwellings with a water closet located within the residence are, on average and *ceteris paribus*, over three percentage points less likely to consider migrating than those without such a facility. The size of the dwelling area also appears important, with those individuals living in the smallest sized dwelling areas more likely to consider migrating, on average and *ceteris paribus*. However, there is an inverse relationship between the inclination to migrate and the age of the property, with those residing in recently constructed dwellings more likely to have considered the possibility of migration than those in older dwellings. This may be attributable to the fact that more recently constructed dwellings could be inferior in quality to older ones and occupied by individuals with less deeper roots to the communities within which they reside.

There is limited variation in migration risk across settlement types and regions. Individuals residing in urban settlements are more inclined to express an interest in migrating abroad, though the estimated effect is on the borderline of statistical significance. Those residing in the central region are more disposed towards external migration than in any of the other broadly defined regions used here.

The estimated effects for the district-level labour market variables are consistent with our priors. The estimated effect for the unemployment rate is positive, with a negative coefficient reported for the wage effect. The computed marginal effect for the unemployment rate suggests that a one percentage point rise in the district level unemployment rate increases the probability that the average individual considers migrating by 0.4 of a percentage point. The implied elasticity, computed at the means of the data, is 0.155.[24] The estimated marginal effect for the wage suggests that a 5 percent rise in the district level hourly wage reduces the probability that an average individual considers migrating by 0.55 of a percentage point. The implied elasticity, again computed at the means of the data for this measure, is –0.356. On balance, the inclination to migrate appears more sensitive to changes in the district-level wage than to changes in the district-level unemployment rate.

The inclusion of variables constructed from community level information in the regression model yielded estimated effects that are generally plausible. For instance, individuals in communities where acquisitive crime is considered a problem are more likely to consider migrating, as are those in communities where land disputes are prevalent. It is also revealing that communities with poor environmental quality, as measured by the presence of stagnant water sources, are also likely to induce individuals to consider migrating.

5.2 Gender Differentials in Considering Migration

The separation of the data by gender is an empirically testable proposition. Given the use of the robust variance-covariance matrix, the test for separation is conducted using a Wald on the set of 41 gender interaction terms in a pooled model. The resultant test statistic is computed as χ^2_{41} = 182.2 and the null hypothesis of constant coefficients across gender is decisively rejected by the data.[25] Thus, the data are separable by gender status, and the second and third columns of Table 2 duly provide probit estimates for these separate gender models. The set of diagnostic tests again

[24] A positive local unemployment rate effect was also detected by Drinkwater (2003a) using data for six Central and Eastern European countries and by Ahn, de la Rica and Ugidos (1999) for willingness to migrate within Spain.
[25] The corresponding likelihood ratio test (LRT) yields a value of 175.6 and the 0.05 critical value with 41 degrees of freedom is 56.94. Given the presence of heteroscedasticity in the pooled probit models, however, the LRT is deemed inappropriate in this application (see text).

reveal no problems with the reported specifications and the null hypothesis of homoscedasticity is actually upheld in the case of the female model.

One way we could examine the differential in outcomes between the male and female sub-samples is by decomposing the standardized probit index into two component parts – one attributable to differentials in endowments and the other to differentials in treatment effects between the two groups. The gender differential at the average could be expressed either as:

$$\bar{Z}_m - \bar{Z}_f = (\bar{X}_m - \bar{X}_f)'\hat{\gamma}_m + \bar{X}_f'(\hat{\gamma}_m - \hat{\gamma}_f) \quad \text{or} \qquad [4]$$

$$\bar{Z}_m - \bar{Z}_f = (\bar{X}_m - \bar{X}_f)'\hat{\gamma}_f + \bar{X}_m'(\hat{\gamma}_m - \hat{\gamma}_f) \qquad [4']$$

where $\bar{X}_m$ and $\bar{X}_f$ denote vectors containing the sample mean values for the males and females respectively, and $\hat{\gamma}_m$ and $\hat{\gamma}_f$ denote vectors containing the estimated probit coefficients for males and females respectively.

This decomposition is based on the familiar index number approach as applied to a linear regression model (see Oaxaca 1973). The sampling variances for the two component parts are easily computable within this linear framework. However, the approach does not provide an exact decomposition for the average differential in probabilities between the two sub-samples.[26]

Using the coefficients from columns two and three of Table 2, the following estimates corresponding to expressions [4] and [4′] (with asymptotic standard errors in parentheses) are obtained:

[26] Gomulka and Stern (1990), Even and McPherson (1993) and Dorion and Riddell (1994) provide more elaborate decompositions based on differentials in the average probabilities. Nothing of substance is gained through use of these procedures in our application.

	Expression [4]		Expression [4′]	
Total Differential	Endowment Effect	Treatment Effect	Endownment Effect	Treatment Effect
$\bar{z}_m - \bar{z}_f$	$(\bar{x}_m - \bar{x}_f)'\hat{\gamma}_m$	$\bar{x}_f'(\hat{\gamma}_m - \hat{\gamma}_f)$	$(\bar{x}_m - \bar{x}_f)'\hat{\gamma}_f$	$\bar{x}_m'(\hat{\gamma}_m - \hat{\gamma}_f)$
0.615	–0.025 (0.052)	0.640*** (0.065)	0.136 (0.098)	0.479*** (0.015)

*** denotes statistical significance at the 0.01 level.

The average gender differential in the standardised probit index is 0.615, which is relatively close to the gender effect reported in column one of Table 1 for the pooled model. Regardless of which coefficient vector is used to weight the gender differences in average characteristics, the endowment effect is not found to be statistically significant at a conventional level in either case. In contrast, the treatment effects are well determined and suggest that the gender difference in outcomes is largely attributable to gender differences in coefficients and not to gender differences in characteristics.

In order to investigate the gender issue further we explore in more detail the statistically significant interaction terms in the pooled model. The significant result encountered in regard to separation by gender may be attributable to statistically significant differentials in a sub-set of coefficients rather than across the entire set. The model reported in the fourth column of Table 2, containing a limited set of gender interactions, was found to best fit the data in our application. The diagnostic score tests again suggest an adequately specified model with a satisfactory performance recorded in terms of the goodness of fit measures.

The estimated coefficients for the seven additional interaction terms provide some useful insights into gender differences in the estimated effects for certain variables. The age interaction suggests that older men are much less inclined than older women to consider migration. The presence of dependent children within the household in the five to eight years age-category reduces a woman's

migration risk by about two percentage points[27] but raises a man's by over three percentage points. Married women are, on average and *ceteris paribus*, more disposed to consider migration relative to single women but the reverse is the case for married men. Women in households where consumption expenditure is situated within the bottom quintile of the consumption distribution are more inclined to consider migration but the reverse is again the case for men. Although the district-level unemployment rate retains a positive effect for both males and females, local labour market wages are found, on average and *ceteris paribus*, to exert no independent effect on a woman's disposition to migrate. In contrast, a 5 percent increase in the local wage reduces the migration risk of the average man by about 1.2 percentage points.[28] The implied elasticity in this case, again computed at the means of the data, is about –0.8 – more than twice that reported for the pooled sample in column one of Table 2. There is also some degree of gender asymmetry in terms of the community level variables. For instance, whereas the presence of land disputes within a community exerts no independent effect on a woman's migration risk, it raises the probability that a man considers migrating by over six percentage points, on average and *ceteris paribus*. In addition, the presence of acquisitive criminal activity within a community reduces a man's propensity to migrate but increases the probability that a woman is disposed to migration by over six percentage points.

6. SUMMARY AND CONCLUSIONS

This paper uses data drawn from the Albania Living Standards Measurement Survey from spring 2002 to contribute to what is currently a limited empirical literature on Albanian migration. The primary purpose of our analysis is to examine the factors that render an individual prone to external migration. The econometric models used provide fairly good descriptions of the data and are well

[27] It is acknowledged that the estimated female effect is poorly determined given that it only commands a t-ratio in absolute terms of 1.55.

[28] The use of gender interaction terms in conjunction with a logged explanatory variable requires some caution in interpretation. Using the marginal effects reported in Table A3, the estimated effect for women is provided by the value reported for the log hourly district wage itself (0.0313), which is not statistically different from zero. The male marginal effect is the sum of the female marginal effect and the gender differential effect associated with the interaction term for this variable (0.0313 – 0.2796 = –0.2483). If we induce a 5 percent change in the wage, this is equivalent to a 0.05 change in the log wage and the marginal effect for this change is thus obtained as $-0.2483 \times 0.05 = -0.0124$. The percentage point equivalent of this probability point outcome is what is reported in the text.

specified on the basis of a battery of diagnostic tests conducted. The estimated results are generally in comport with findings from the empirical literature on the willingness to migrate. The usual suspects emerge as determining factors, with age, employment status and education exerting predictable influences on migration risk. In respect of education, however, the most qualified (i.e., university educated) appear less at risk of migration than those with either secondary or vocational education.[29] In addition, the quality of the household's residential property appears more important than the demographic structure of the household itself in determining whether an individual considered external migration. There is a strong gender dimension to migration risk in Albania and this is not adequately captured by use of a standard intercept shift in the pooled regression model. Although men are considerably more likely to consider migration, women in households where consumption expenditure is situated within the bottom quintile of the overall consumption distribution are more inclined to consider migration abroad than men. This type of gender effect is concealed from view within the more restrictive pooled model. It should be noted that this particular finding provides the only piece of tentative evidence on the relationship between household poverty and an individual's migration risk.

An important contribution of this paper is the explicit examination of the role local labour market and community conditions exert on migration risk in Albania. In regard to the former, both the unemployment rate and the wage rate effects had anticipated signs and registered plausible elasticities.[30] There is some evidence that local labour market wages are more important than unemployment in determining the overall propensity to migrate. There is evidence that the migration risk of men responded more elastically to district level wages than women.[31] The prevalence of land disputes within a community appeared a more potent determinant of migration risk for men, but female migration risk appeared more informed by criminal activity within the community than land disputes.

[29] However, there is some evidence that university educated females have a relatively high risk of migration (see column 3, Table 1).

[30] The elasticity with respect to the local unemployment rate, however, is somewhat lower than that reported for Spain by Ahn *et al* (1999) using regional data on the willingness to migrate.

[31] This finding is not entirely inconsistent with the findings of Agesa (2003) in the study of actual migration patterns in Kenya.

We believe that the approach adopted here and the analysis undertaken is fruitful and informative. On the basis of our empirical work the short-term prognosis is somewhat mixed. Even after some modest recent economic progress and reform in Albania, young males and the vocationally educated still present a higher tendency to consider migration than most other categories. It is clear that the incidence of Albanian migration is not likely to fall sharply anytime soon. However, whether an individual considers migration is strongly linked to economic conditions, and improvements in employment rates and wage levels are thus likely to reduce the risk of migration in the future. These improvements may emerge if there is a continued commitment to pursuing the type of economic and other reform policies embraced in the more recent years of the Albanian transitional experience and that have borne some fruit. Nevertheless, it is acknowledged that economic reforms cannot occur in a vacuum, and are unlikely to be successful in the absence of strong government institutions. The probit estimates reported here for the community level variables provide a reminder of the role these institutions play, emphasizing the importance of the rule of law and the protection of property rights in reducing migration risk.

REFERENCES

Agesa, R. (2003) 'Gender Differences in the Urban to Rural Wage Gap and the Prevalence of the Male Migrant', *Journal of Developing Areas* 37: 13-34.

Ahn, M, de la Rica, S. and A.Ugidos (1999) 'Willingness to Move for Work and Unemployment Duration', *Economica* 66: 335 – 358.

Ahn, N., J.Jimeno and E.Garcia (2002) *Migration Willingness in Spain: Analysis of Temporal and Regional Differences.* Spain: FEDEA.

Chesher, A. and Irish, M. (1987) 'Residual Analysis in the Grouped and Censored Normal Linear Model', *Journal of Econometrics* 34: 33-61.

Cramer, J. (1999) 'Predictive Performance of the Binary Logit Model in Unbalanced Samples', *Journal of the Royal Statistical Society*, Series D, 48: 85-94.

DeJong, G. (2000) 'Expectations, Gender and Norms in Migration Decision-Making', *Population Studies* 54: 307-319.

Dorion, D.J. and Riddell, C. (1994) 'The Impact of Unionization on Male-Female Earnings Differences in Canada', *The Journal of Human Resources*, Vol. XXIX: .504-534.

Drinkwater, S. (2003 a) *Go West? Assessing the Willingness to Move from Central and Eastern European Countries.* Mimeo. University of Surrey: Department of Economics.

Drinkwater, S. (2003 b) *Estimating the Willingness to Move within Great Britain: Importance and Implications.* Mimeo. University of Surrey: Department of Economics.

Even, W.E. and McPherson, D.A. (1993) 'The Decline of Private-Sector Unionism and the Gender Wage Gap', *The Journal of Human Resources*, Vol. XXVIII: 279-296.

Gardner, R., G.DeJong, F.Arnold, B.Carino (1986) 'The Best-Laid Schemes: an Analysis of the Discrepancies Between Migration Intentions and Behaviour', *Population and Environment*. 8: 63-77.

Gomulka, J. and Stern, N. (1990) 'The Employment of Married Women in the United Kingdom 1970-83', *Economica* 57: 171-200.

Gouriéoux, C., A. Montfort, and A. Trognon (1984) 'Pseudo-Maximum Likelihood Methods: Theory', *Econometrica* 42: 681 – 700.

Greene, W.H. (2000) *Econometric Analysis*, 4th Edition, Prentice-Hall.

Huber, P.J. (1967) 'The Behaviour of Maximum Likelihood Estimates Under Non-Standard

Conditions' in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, CA: University of California Press* 1: 221 – 223.

Hughes, G. and B.McCormick (1985) 'Migration Intentions in the UK: Which Households Want to Migrate and Which Succeed', *Economic Journal*, Vol.95 (Supplement: Conference Papers): 113-123.

Heckman, J. (1979) 'Sample Selection Bias as a Specification Error', *Econometrica* 47: 153-161.

Johnston, J and J. DiNardo (1999) *Econometric Methods*, 4th Edition, McGraw-Hill International Editions.

Kule, D., A.Mancellari, H.Papapanagos, S.Qirici and P.Sanfey (2002) 'The Causes and Consequences of Albanian Emigration During Transition: Evidence from Micro Data', *International Migration Review* 36:.229-239.

King, R. and J.Vullnetari (2003) 'Migration and Development in Albania', Working Paper C5, Development Research Centre on Migration, Globalisation and Poverty, University of Sussex.

Layard, R., O.Blanchard, R.Durnbush and P.Krugman (1992) *East-West Migration: The Alternatives.* MIT: Cambridge.

Mancellari, A., H. Papapanagos and P.Sanfey (1996) 'Job Creation and Temporary Emigration: The Albanian Experience', *Economics of Transition* 4: 471-490.

Manski, C. (1990) 'The Use of Intentions to Data to Predict Behaviour: A Best Case Analysis', *Journal of the American Statistical Association* 85: 934-940.

_____. (1995) *Identification Problems in the Social Sciences*, Harvard University Press.

Louviere, J.J., D.Henscher, and J.D.Swait (2000) *Stated Choice Methods: Analysis and Applications.* Cambridge: Cambridge University Press.

Oaxaca, R. (1973) 'Male-Female Wage Differentials in Urban Labor Markets', *International Economic Review*, 693-705.

O'Connell, P. (1997) 'Migration under Uncertainty: Try Your Luck or Wait and See', *Journal of Regional Science* 37: 331-347.

Orme, C. (1990) The Small-Sample Performance of the Information Matrix Test', *Journal of Econometrics* 46: 309–331.

Peters, S. (2000) 'On the Use of the RESET Test in Microeconometric Models', *Applied Economics*

Letters 7: 361–365

Papapanagos, H. and P.Sanfey (2002) 'Intentions to Emigrate in Transition Countries: The Case of Albania, *Journal of Population Economics* 14: 491-504.

Pradham, M. and M. Revallion (2000) 'Measuring Poverty using Qualitative Perceptions of Consumption Adequacy', *Review of Economics and Statistics* 82:. 462-471.

Revallion, M. and M. Lokshin (2001) 'Identifying Welfare Effects from Subjective Questions', *Economica* 68: 335 – 358.

UNECE (2003) *Economic Survey of Europe* No.1. Geneva: United Nations.

World Bank (2003) *Albania Poverty Assessment* – Report Number 26213-AL. Washington DC: Human Development Sector Unit Europe and the Central Asian Republics.

White, H. (1980) 'A Heteroscedastic-consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity', *Econometrica* 48: 817-838.

APPENDIX

Table A1: Description of Variables Used in the Analysis

Variable	Variable Description
Considered Migrating	=1 if the individual considered migrating; = 0 otherwise.
Age (years)	The age of the individual in years
Male	=1 if the individual is male; = 0 otherwise.
Head of Household	=1 if the individual is household head; = 0 otherwise.
Married	=1 if the individual is married; = 0 otherwise.
Divorced/Separated	=1 if the individual is divorced or separated; = 0 otherwise.
Single	=1 if the individual is single; = 0 otherwise.
Primary: $\leq$ 4 grades	=1 if the individual has no education or achieved four or less primary grades; = 0 otherwise.
Primary: 5 to 8 grades	=1 if the individual achieved between five and eight primary grades; = 0 otherwise.
Secondary	=1 if the individual achieved secondary level; = 0 otherwise.
Vocational	=1 if the individual achieved vocational level; = 0 otherwise.
University	=1 if the individual achieved university level; = 0 otherwise.
Born in the Municipality	=1 if the individual was born in the municipality; = 0 otherwise
Health Disability	=1 if the individual has a health disability; = 0 otherwise.
Unemployed	=1 if the individual is unemployed; = 0 otherwise.
Employee	=1 if the individual is an employee; = 0 otherwise.
Farmer	=1 if the individual is a farmer; = 0 otherwise.
Self-Employed	=1 if the individual is self-employed; = 0 otherwise.
Temporary Layoff	=1 if the individual is a temporary layoff; = 0 otherwise.
Household Size	The total number of individuals in the household.
Children in Household: Aged $\leq$ 4 years	=1 if the individual's household has any children aged between four or less; = 0 otherwise.
Children in Household: Aged $5 \leq$ years ≤ 8	=1 if the individual's household has any children aged between five and eight; = 0 otherwise.
Children in Household: Aged $9 \leq$ years ≤ 14	=1 if the individual's household has any children of age between nine and fourteen; = 0 otherwise.
Log of Total Household Consumption:	The logarithm of the total (monthly) expenditure of the household.
Residence Dwelling Area: $\leq$ 69 Sq.Metres	=1 if the area of the dwelling is less than 69 square metres; = 0 otherwise.
Residence Dwelling Area:70 $\leq$ Sq.Metres $\leq$ 130	=1 if the area of the dwelling is between 70 and 130 square metres; = 0 otherwise.
Residence Dwelling Area:Sq.Metres > 130	=1 if the area of the dwelling is over 130 square metres; = 0 otherwise.
Residence Constructed after 1990	=1 if the dwelling was built after 1990; = 0 otherwise.

Residence Contains Internal Watercloset	=1 if the dwelling contains an internal watercloset; = 0 otherwise.
Central Region	=1 if the individual resides in the Central region; = 0 otherwise.
Coastal Region	=1 if the individual resides in the Coastal region; = 0 otherwise.
Mountain Region	=1 if the individual resides in the Mountain region; = 0 otherwise.
Tirana	=1 if the individual resides in Tirana; = 0 otherwise.
Urban Settlement Type	=1 if the individual resides in an urban settlement; = 0 otherwise.
District Level Unemployment Rate (%)	The unemployment rate (%) in the district where the individual resides.
District Level Hourly Wage (log)	The logarithm of the average hourly wage in the district where the individual resides.
Land Disputes in Commune	=1 if there are any problems related to land disputes in the community; = 0 otherwise.
Thefts in Commune	=1 if thefts occur in the community where the individual resides; = 0 otherwise.
Stagnant Water in Commune	=1 if there are stagnant waters in the community where the individual resides; = 0 otherwise.
No Migration in Commune	=1 if people do not emigrate from the community where the individual resides; = 0 otherwise.

Table A2: Summary Statistics for the Variables Used in the Analysis

	Pooled	Male	Female
Considered Migrating	0.3292	0.4234	0.2254
Age (years)	36.486	38.061	34.754
Male	0.5239	1.000	0.000
Head of Household	0.3594	0.6375	0.0534
Married	0.7531	0.7677	0.7370
Divorced/Separated	0.0223	0.0063	0.0399
Single	0.2246	0.2260	0.2231
Primary: $\leq$ 4 grades	0.0415	0.0387	0.0445
Primary: 5 to 8 grades	0.4900	0.4586	0.5244
Secondary	0.1886	0.1925	0.1844
Vocational	0.1698	0.1873	0.1507
University	0.1060	0.1190	0.0918
Born in the Municipality	0.6655	0.7339	0.5902
Health Disability	0.1287	0.1225	0.1356
Unemployed	0.1210	0.1214	0.1204
Employee	0.3487	0.4252	0.2645
Farmer	0.4256	0.3182	0.5438
Self-Employed	0.0889	0.1246	0.0496
Temporary Layoff	0.0159	0.0106	0.0217
Household Size	5.2558	5.2144	5.3013
Children in Household: Aged $\leq$ 4 years	0.3113	0.3157	0.3064
Children in Household: Aged $5 \leq$ years ≤ 8	0.3070	0.2964	0.3187
Children in Household: Aged $9 \leq$ years ≤ 14	0.4639	0.4579	0.4706
Log of Total Household Consumption	8.8429	8.8531	8.8317
Residence Dwelling Area: $\leq$ 69 Sq.Metres	0.4959	0.4944	0.4977
Residence Dwelling Area: $70 \leq$ Sq.Metres ≤ 130	0.4628	0.4636	0.4620
Residence Dwelling Area: Sq.Metres > 130	0.0413	0.0422	0.0403
Residence constructed after 1990	0.2261	0.2306	0.2211
Residence contains Internal Watercloset	0.6305	0.6610	0.5968
Central Region	0.2749	0.2703	0.2800
Coastal Region	0.2775	0.2728	0.2827
Mountain Region	0.3044	0.2858	0.3249

Tirana	0.1431	0.1711	0.1123
Urban Settlement Type	0.4573	0.5076	0.4020
District Level Unemployment Rate (%)	11.880	12.118	11.619
District Level Hourly Wage (log)	4.2761	4.2902	4.2607
Land Disputes in Commune	0.5322	0.5294	0.5352
Thefts in Commune	0.4077	0.4256	0.3881
Stagnant Water in Commune	0.2976	0.3101	0.2839
No Migration in Commune	0.0135	0.0134	0.0136

Table A3: Marginal and Impact Effects for the Determinants of Migration Risk

	Pooled Model	Male Model	Female Model	Pooled Model with Gender Interactions
Age Spline: 15 – 25 years	0.0095**	0.0169***	0.0049	0.0109**
	(0.0042)	(0.0064)	(0.0054)	(0.0043)
Age Spline: 16 – 35 years	-0.0100***	-0.0209***	-0.0050	-0.0113***
	(0.0032)	(0.0054)	(0.0040)	(0.0032)
Age Spline: 36 – 45 years	-0.0043	-0.0076*	-0.0046	-0.0061**
	(0.0030)	(0.0046)	(0.0038)	(0.0030)
Age Spline: 46 – 55 years	-0.0172***	-0.0157***	-0.0099*	-0.0144***
	(0.0038)	(0.0052)	(0.0054)	(0.0038)
Age Spline: 56 – 60 years	-0.0473***	-0.0668***	0.0044	0.0045
	(0.0155)	(0.0187)	(0.0255)	(0.0293)
Male	0.2260***	§	§	†
	(0.0174)			
Head of Household	0.0008	0.0953***	0.0515	0.0636***
	(0.0217)	(0.0338)	(0.0489)	(0.0251)
Married	0.0361	-0.0148	0.0682**	0.1101***
	(0.0263)	(0.0434)	(0.0292)	(0.0302)
Divorced/Separated	0.0979*	0.2519**	0.0251	0.0937*
	(0.0590)	(0.1134)	(0.0636)	(0.0594)
Single	*f*	*f*	*f*	*f*
Primary: $\leq$ 4 grades	*f*	*f*	*f*	*f*
Primary: 5 to 8 grades	0.0697*	0.0885	0.0846*	0.0815**
	(0.0389)	(0.0565)	(0.0487)	(0.0387)
Secondary	0.0926**	0.0567	0.1661***	0.1009**
	(0.0446)	(0.0619)	(0.0659)	(0.0448)
Vocational	0.1342***	0.1245**	0.1907***	0.1492***
	(0.0456)	(0.0615)	(0.0686)	(0.0457)
University	0.0530	0.0467	0.1160*	0.0693
	(0.0482)	(0.0670)	(0.0725)	(0.0486)
Born in the Municipality	0.0218	-0.0062	0.0284	0.0223
	(0.0149)	(0.0243)	(0.0173)	(0.0150)
Health Disability	0.0199	-0.0118	0.0310	0.0122
	(0.0218)	(0.0318)	(0.0275)	(0.0216)
Unemployed	*f*	*f*	*f*	*f*
Employee	-0.0610***	-0.0748**	-0.0392	-0.0601***
	(0.0223)	(0.0332)	(0.0267)	(0.0223)
Farmer	-0.0802***	-0.0392	-0.0822**	-0.0747***
	(0.0270)	(0.0399)	(0.0376)	(0.0272)
Self-Employed	-0.0788***	-0.0374	-0.0790**	-0.0647**
	(0.0269)	(0.0413)	(0.0321)	(0.0274)

Temporary Layoff	-0.0886* (0.0463)	-0.0089 (0.0962)	-0.1064** (0.0383)	-0.0905* (0.0466)
Household Size	-0.0054 (0.0049)	-0.0060 (0.0076)	0.0014 (0.0060)	-0.0027 (0.0050)
Children in Household: Aged $\leq$ 4 years	0.0238 (0.0186)	0.0151 (0.0274)	0.0374 (0.0243)	0.0249 (0.0187)
Children in Household: Aged $5 \leq$ years ≤ 8	0.0065 (0.0163)	0.0355 (0.0246)	-0.0201 (0.0200)	-0.0364 (0.0232)
Children in Household: Aged $9 \leq$ years ≤ 14	-0.0029 (0.0163)	-0.0024 (0.0241)	-0.0006 (0.0209)	-0.0021 (0.0163)
Log Consumption: First Quintile	-0.0047 (0.0587)	-0.0883 (0.0872)	0.1343* (0.0800)	0.2429*** (0.0933)
Log Consumption: Second Quintile	0.1640* (0.0950)	0.1092 (0.1405)	0.2057* (0.1187)	0.1409 (0.0955)
Log Consumption: Third Quintile	0.1050 (0.1220)	0.2097 (0.1774)	0.0048 (0.1515)	0.1175 (0.1210)
Log Consumption: Fourth Quintile	0.0585 (0.0915)	-0.0806 (0.1345)	0.1575 (0.1114)	0.0458 (0.0909)
Log Consumption: Fifth Quintile	0.0266 (0.0505)	0.0198 (0.0727)	-0.0022 (0.0605)	0.0149 (0.0494)
Residence Dwelling Area: $\leq$ 69 Sq.Metres	*f*	*f*	*f*	*f*
Residence Dwelling Area: $70 \leq$ Sq.Metres ≤ 130	-0.0318** (0.0141)	-0.0234 (0.0208)	-0.0368** (0.0174)	-0.0312** (0.0141)
Residence Dwelling Area: Sq.Metres > 130	-0.0934*** (0.0317)	-0.1161** (0.0496)	-0.0588 (0.0367)	-0.0893*** (0.0317)
Residence Constructed after 1990	0.0556*** (0.0171)	0.0707*** (0.0247)	0.0375* (0.0218)	0.0542*** (0.0171)
Residence Contains Internal Watercloset	-0.0322* (0.0176)	-0.0273 (0.0261)	-0.0400* (0.0223)	-0.0336* (0.0177)
Central Region	0.0689*** (0.0194)	0.0860*** (0.0283)	0.0437* (0.0239)	0.0646*** (0.0193)
Coastal Region	*f*	*f*	*f*	*f*
Mountain Region	-0.0099 (0.0214)	-0.0059 (0.0329)	-0.0233 (0.0252)	-0.0160 (0.0214)
Tirana	-0.0335 (0.0250)	-0.0376 (0.0364)	-0.0112 (0.0317)	-0.0276 (0.0249)
Urban Settlement Type	0.0367 (0.0225)	0.0259 (0.0311)	0.0542* (0.0328)	0.0467** (0.0226)
District Level Unemployment Rate (%)	0.0040*** (0.0009)	0.0062*** (0.0014)	0.0019 (0.0012)	0.0041*** (0.0009)
District Level Hourly Wage (log)	-0.1103*** (0.0429)	-0.2382*** (0.0662)	-0.0076 (0.0510)	0.0313 (0.0547)
Land Disputes in Commune	0.0290** (0.0138)	0.0692*** (0.0205)	-0.0092 (0.0171)	-0.0186 (0.0207)

Thefts in Commune	0.0283* (0.0153)	-0.0155 (0.0224)	0.0692*** (0.0192)	0.1158*** (0.0222)
Stagnant Water in Commune	0.0479*** (0.0168)	0.0364 (0.0244)	0.0556*** (0.0212)	0.0487*** (0.0167)
No Migration in Commune	-0.1058* (0.0515)	-0.1896** (0.0710)	-0.0259 (0.0664)	-0.1070* (0.0498)
Male×[Age Spline: 56 – 60 years]	§	§	§	-0.0618* (0.0318)
Male×[Children in Household: Aged $5 \leq$ years ≤ 8]	§	§	§	0.0715** (0.0320)
Male×[Married]	§	§	§	-0.1635*** (0.0348)
Male×[Log Consumption: First Quintile]	§	§	§	-0.3663*** (0.1043)
Male× [District Level Hourly Wage (log)]	§	§	§	-0.2796*** (0.0654)
Male×[Land Disputes in Commune]	§	§	§	0.0862*** (0.0283)
Male×[Thefts in Commune]	§	§	§	-0.1418*** (0.0235)

Notes to table A3

(a) ***, ** and * denote statistical significance at the 0.01, 0.05 and 0.10 level respectively using two-tailed tests. The reported standard errors are corrected for the presence of heteroscedasticity of unknown form (see Huber(1967)).
(b) § denotes not applicable in estimation, *f* denotes category omitted in estimation and † denotes not readily interpretable given the use of interaction terms.