Patience goes a long way: Evidence from Spain.

González Chapela, Jorge

Centro Universitario de la Defensa de Zaragoza

17 February 2020

PATIENCE GOES A LONG WAY: EVIDENCE FROM SPAIN[*]

Jorge González Chapela[†]

Centro Universitario de la Defensa de Zaragoza

Abstract: This study uses newly available data from the Survey of Financial Competences to investigate whether cross-region migrants in Spain are more patient than individuals who choose to remain in their birth region. The empirical model incorporates predicted probabilities of underreporting and overreporting of the migrant status. Less patient individuals appear to be less likely to be migrants. This result is robust to controlling for a variety of demographic and economic factors, as well as for cognitive ability.

Keywords: Internal migration, time preference, measurement error, probit, Spain.

JEL codes: C35, D91, J60.

[*] I am grateful to Ernesto Villanueva and the ECF team for diligent assistance with ECF data.

[†] *Address*: Academia General Militar, Ctra. de Huesca s/n, 50090 Zaragoza, Spain. *Email*: jorgegc@unizar.es. *Ph*: (+34) 976739834.

1. INTRODUCTION

Standard models of migration recognize that a change of residential location is costly (Molloy et al. 2011). Some costs of migrating, such as out-of-pocket expenses and the psychic costs of changing environment, occur in the short term, whereas the benefits of migrating are collected in the future. Hence, and as argued by Gibson and McKenzie (2011), we might expect more patient individuals to be more likely to migrate. Nowotny (2014), however, shows that if potentially mobile persons expect benefits in the home region to exceed benefits in the destination region in the future, then the more patient will be less likely to migrate. Therefore, the extent to which migrants have above- or below-average levels of patience is to be determined empirically.

The time preference composition of migration flows may have important consequences for both sending and receiving regions. Individuals' level of patience correlates with behaviors involving intertemporal tradeoffs like savings rates, educational attainment, and medical adherence, with personality traits such as cognitive ability and agreeableness, and with economic outcomes including income level and personal unemployment (Cohen et al. forthcoming). The time preference composition of migration flows may also be relevant from a purely scientific point of view. For example, given the positive role of patience on human capital formation (e.g., Golsteyn et al. 2014, Cadena and Keys 2015), a positive effect of patience on migration would contribute to explaining the college migration premium (Malamud and Wozniak 2012).

The lack of questions on time preference in the main data sources for constructing migration rates has prevented the development of research in this area. Hence, the very few studies on the empirical relationship between time preference and migration have developed their own specialized surveys (Gibson and McKenzie 2011, Arcand and Mbaye 2013, Nowotny 2014, Goldbach and Schlüter 2018). These studies have found that more patient individuals are more likely to migrate internationally and internally, and less likely to migrate illegally. These patterns have generally been observed in small samples, which questions their generalizability and economic relevance, and without controlling for individuals' cognitive ability.[1] Previous research indicates that patience and cognitive ability are positively correlated even after removing variation due to common factors (e.g., Frederick 2005, Burks et al. 2009, Dohmen et al. 2010, Benjamin et al. 2013). In addition, if individuals dislike what they do not perceive precisely and cognitive skills reduce the noise in perceiving the utility of complex options (Burks et al. 2009), individuals with higher cognitive skills may be more likely to perceive the benefits accruing to the migrant and hence more likely to migrate. As a result, previous estimates of the relationship between time preference and migration could be biased in the positive direction.

Newly available data from the Survey of Financial Competences (ECF by its Spanish abbreviation) make it possible to investigate further the patience migration premium. The ECF collects nationally and regionally representative information of financial knowledge and practices in Spain following a questionnaire proposed by the International Financial Education Network. The questionnaire of the international organization, however, has been complemented in several respects, including the addition of a question on residence at birth plus some items designed to measure respondents' time preferences and cognitive skills. A Money Earlier or Later (MEL) task (Cohen et al. forthcoming) is used to measure time preference experimentally, while cognitive skills are assessed using questions validated in international studies

[1] The samples analyzed by Nowotny (2014) are large but the migration information is limited to migration willingness. The entire samples analyzed by Gibson and McKenzie (2011) consist of very high ability individuals.

answered by the respondent in private, so that no other household member could help her/him. Thus, the ECF allows investigating the extent to which the empirical relationship between time preference and migration owes to individuals' cognitive ability. Furthermore, it allows extending the existing literature by analyzing for the first time the relationship between time preference and internal migration in a developed country.

The residential information gathered by the ECF is limited to the birth and the time the survey was conducted. Discrepancies between these two residences provide a "reduced-form" measure of lifetime migration (Carlson 2007), which contains both errors of omission (false nonmigrants) and errors of commission (false migrants). As argued by Molloy et al. (2011), some true migrants will have returned to their birth region after having spent time elsewhere, while individuals who moved when they were still a member of their parents' household are indistinguishable in the data from individuals who moved during their adult lives. These errors imply that probit models of the migrant status will be biased because they do not take classification errors of this status into account (e.g., Hausman et al. 1998).

This research employs the predicted probabilities estimator of Bollinger and David (1997) to estimate a probit model accounting for misclassification in the survey measure of the migrant status. Models of classification errors are first estimated on a sample of the same population extracted from an alternative, administrative data set: the Continuous Sample of Work Histories (MCVL by its Spanish abbreviation). In addition to information on residence at birth and at data compilation, the MCVL gathers geographical information for each period of affiliation with Spain's Social Security. The estimated individual probabilities of misclassification are then incorporated into a modified probit likelihood function which is maximized on the ECF sample, as only this sample contains the information on migration determinants that is necessary to this research.

The rest of the paper is organized as follows. Section 2 describes the data and the construction of the samples. Section 3 defines the main measures and presents descriptive evidence on the relationship between time preference and migration. Section 4 reviews the econometric specification. Section 5 presents the regression results. Finally, Section 6 gathers the main conclusions, considerations important to their interpretation, and avenues for future research.

2. DATA AND SAMPLE SELECTION

The data for this study derive from two publicly available, individual-level data sets: the ECF and the MCVL. The ECF provides the primary sample of analysis, whereas the MCVL provides a validation sample for the migrant status. The two samples may contain some common individuals, but since these cannot be identified, they are treated as if they contained no common units.

2.1. ECF[2]

The ECF (Banco de España and National Securities Market Commission 2018) is a survey aimed at assessing knowledge and understanding of financial concepts in Spain. Sampling for the ECF is intended to be representative of the population of individuals aged 18–79 living in private households. It is also meant to be representative of each of the 17 regions of Spain. Since 1995, Spain is organized in 17 regions (Autonomous Communities) plus two Autonomous Towns (Ceuta and Melilla, on the north coast of Africa). This organization corresponds to the NUTS 2 level of territorial aggregation in the EU. Of the original sample of 21,250 individuals, 16,025 were contacted personally during the fieldwork period. Of these, 6,708 refused to answer and 763 were unable to

[2] A complete description of the ECF and its methods is provided in Bover et al. (2019).

give any type of information. The number of valid interviews completed is, therefore, 8,554.

Computer-assisted interviews were conducted in sample members' own homes between end-September 2016 and end-May 2017, although by the end of December 78% of the total number of valid interviews were already completed. The ECF provides the region or country of birth plus the region of residence at the time of the interview. No other former places of residence are asked for. With this information, it is not possible to know whether immigrants migrated between regions within Spain, so immigrants (986 persons) are excluded from the analysis. So are natives who were born outside the 17 regions surveyed by the ECF (457 persons), as they can only be migrants. Another 346 persons are dropped because they have missing information in some variable used in the analysis. Military personnel (10 persons) are also removed as their migration decisions are probably non-autonomous. For these reasons, the ECF sample comprises 6,755 individuals.

2.2. MCVL

The MCVL is an administrative data set compiled annually by Spain's Social Security. It comprises a 4% random sample of the population who, in a given year, are affiliated with the Social Security, namely individuals who are working, receiving unemployment benefits, or receiving a pension. The MCVL gathers these individuals' complete work career since the year 1967,[3] plus information for all periods of retirement, disability, or orphan pension receipt since 1996. For each period of work (including self-employment) the MCVL provides the province of the workplace establishment. For each period of unemployment benefits or pension receipt the MCVL gives the province

[3] The work career includes periods of unemployment benefits receipt. The work career prior to 1967 may be incomplete.

of the administrative office managing the payment. The MCVL also provides the province or country of birth plus the province of residence in approximately April of the year following that of the MCVL edition, both gathered from the Municipality Population Registry.

Spain's 17 regions are divided into a total of 50 provinces, which correspond to the NUTS 3 level of territorial aggregation in the EU. I aggregate the province data at the region level. While unemployment benefits and pensions are managed by the administrative office of the province associated with the recipient's home address, a worker may reside in a region other than that of her/his workplace establishment. Using data from the Spanish Labor Force Survey for the year 2016,[4] I calculated that the proportion of workers commuting to a different region is just 1.8%. Thus, even though it may contain errors, I treat the geographic information provided by the MCVL as truth, and use residence at each affiliation period to reveal interim moves between birth and data compilation.

For correspondence with the ECF sampling period, I use the MCVL for the year 2015, in which information on residence at data compilation refers approximately to April 2016. The MCVL 2015 encompasses over 1 million persons. After discarding immigrants, natives born or residing outside the 17 regions surveyed by the ECF, and individuals with missing data in some variable used in the analysis, the number of persons aged 18–79 is 917,879.[5]

3. MEASURES AND DESCRIPTIVE EVIDENCE

3.1. Lifetime Migration

[4] The source of data for the Spanish Labor Force Survey is the National Statistics Institute (www.ine.es).

[5] In the MCVL, it is not possible to identify military personnel.

Individuals who at survey or data compilation reside in a different region where they were born are classified as migrants. Nonmigrants are individuals who reside in the same region where they were born. This reduced-form measure of lifetime migration can be calculated in both the ECF and the MCVL, but as argued in the Introduction, may contain both errors of omission (false nonmigrants) and errors of commission (false migrants). I use the residential history provided by the MCVL to reveal and model both types of classification errors.

The residential history of the MCVL allows identifying false nonmigrants in the whole population. However, the identification of false migrants (i.e. individuals who left birth region as a child and never changed region as an adult) is restricted to the population of orphans. In Spain, single and double orphans are generally entitled to an orphan pension until the age of 21. The receipt of an orphan pension is recorded in the MCVL, with the province of the administrative office managing the pension reflecting that of the legal guardian's home address. (The orphan pension ceases in case of adoption.) For orphans younger than 18 years of age, I consider discrepancies between the region of birth and the region of orphan pension receipt to be non-autonomous migrations. Orphans tend to be more mobile than non-orphans (e.g., see Thomas 2012), but the MCVL residential history does not allow identifying false migrants in the whole population. For example, discrepancies between the regions of birth and of labor market entry may reflect both autonomous and non-autonomous migrations.

A full-form (i.e. "true") migrant is an individual who resided in more than one region between (and including) birth and data compilation, with the exception of orphans who left their birth region as a child and never changed region as an adult. These orphans plus individuals who always resided in the region of birth are considered to be full-form nonmigrants. This full-form measure of lifetime migration offers a more precise identification of individuals who have ever migrated internally, but it can only be calculated in the MCVL.

Columns (1) and (2) of Table 1 compare the population proportion of reduced-form migrants calculated in the ECF and MCVL samples, for the total population as well as stratified by sex, age, and education. The proportion of lifetime migrants is 14.5% in the ECF and 18.7% in the MCVL. The differences in lifetime migration rates across strata observed in the MCVL are also detectable in the ECF, but rates are smaller in the latter. The smaller rates in the ECF concur with the tendency of surveys to underestimate migration rates in comparison with administrative data sources (e.g., see Martí and Ródenas 2007). While this is a potential problem, estimates of the regression relationship between time preference and the migrant status will be unbiased if the form of the relationship is correctly specified.

Column (3) of Table 1 lists the proportion of full-form lifetime migrants in the MCVL sample. Therefore, columns (2) and (3) display the discrepancy between the reduced-form and full-form measures of lifetime migration. The discrepancy is large. In the total population, the full-form lifetime migration rate (i.e. the proportion of individuals who have ever migrated internally) is 36.2%, a value which is 94% larger than the 18.7% reduced-form estimate. Across strata, the degree of discrepancy ranges from 38% (individuals aged 65+) to 185% (individuals aged 18–24).

For true migrants, column (4) presents the percentage of migrants as calculated with the reduced-form measure. Therefore, one hundred minus the value in column (4) gives the percentage of omission errors. False nonmigrants are 48.3%, but this value is larger among men (52%), the youngest (66%), and those with exactly high school education (53%). Column (5) shows the extent of errors of commission among orphans. (Note that since the information about pensions in the MCVL is available starting in

1996, all orphans are younger than 42 years of age at data compilation.) Specifically, for orphan true nonmigrants, column (5) displays the percentage of migrants given by the reduced-form measure. False migrants are 4.2%, a figure which varies little across strata.

3.2. Time Preference

The ECF includes a MEL task to measure respondents' time preference. These are presented with two interdependent hypothetical binary choices between immediate and delayed financial rewards. (The wording of the questions, translated from Spanish, is reported in Appendix A.) In the first choice, they have to decide between receiving €2,000 today or €2,200 in a year's time. If they opt for the payment today, in the second choice the payment in a year's time is increased to €3,000, whereas if they first choose the payment in a year's time, this is decreased to €2,100 in the second choice. Such a "staircase" structure is also used in, for example, the MEL tasks of Goldbach and Schlüter (2018) and of the Global Preference Survey (Falk et al. 2018), although the series of binary choices is longer in these studies.

A 1.2% of the respondents answer "don't know" in the first binary choice, while, of those who do choose a payment, on average 0.6% answer "don't know" in the second choice. A "don't know" response may indicate either that the respondent is unable to choose (as confounding factors may complicate the choice: Frederick et al. 2002), or that he/she is indifferent between the two payments. The presence of cases in which the "don't know" response occurs in the second choice led me to stick to the latter interpretation. However, I will assess the sensitivity of the results to the exclusion of respondents who answer "don't know" in any of the two choices.

The responses to the sequence of two binary choices allow classifying respondents into four categories, which are described in Table 2 in terms of required rates of return (*RRR*s):[6] below 4.9%; between 4.9% and 9.8%; between 9.8% and 44.9%; and above 44.9%. Table 2 also presents sample percentages for each category. We see, for example, that 10.6% of the sample has *RRR*s between 4.9% and 9.8%, as either they prefer €2,200 in a year's time to €2,000 today, but €2,000 today to €2,100 in a year's time, or answer "don't know" in the first binary choice.

I fitted a normal distribution to the *RRR* data. The interval regression estimates of the mean and variance are 0.31 and 0.15. The appropriateness of the normal assumption is tested using a chi-square goodness-of-fit test. The predicted number of individuals in each category is listed in Table 3 under the heading of Normal model. The test statistic is 764.25. The critical value at 10% level with 1 df is 2.71. Clearly, the null hypothesis of normality is rejected. Then, I approximated the distribution of *RRR*s with a lognormal curve. In this case, the interval regression estimates of the mean and variance are -1.46 and 4.15. The chi-square statistic is 2.21. Therefore, the distribution of *RRR*s appears to be lognormal.[7] Under the hypothesis that *RRR*s are lognormally distributed with mean and variance given above, the (unconditional) mean *RRR* is 185%. The mean *RRR* for each of the four intervals of *RRR*s, calculated using the formula developed in Wang et al. (2012), is given in column (3) of Table 3.

---

[6] When a respondent is indifferent between (€d_1, today) and (€d_2, 1 year), the *RRR* necessary to induce her/him to forgo d_1 euros immediately is $2\left(\left(d_2/d_1\right)^{1/2}-1\right)$. The definition of *RRR* given in Cohen et al. (forthcoming) assumes continuous compounding of the annual interest rate. I choose semiannual compounding for comparability with Dohmen et al. (2010).

[7] Lognormality is rejected if respondents who answer "don't know" in any of the two binary choices are excluded from sample (p-value 0.00).

Figure 1 compares the population distribution of *RRR*s in Spain to that obtained for Germany by Dohmen et al. (2010). The German data are also developed with a time horizon of 1 year, but the size of the financial rewards is smaller than in the ECF (the early reward, for example, is always €100), so *RRR*s might not be comparable due to the "magnitude effect" (e.g., Cohen et al. forthcoming). Consistent with the magnitude effect, a higher fraction of Spaniards is found to be in the lowest category of impatience. However, the fraction found in the highest category is also higher, which seems to suggest that the distribution of *RRR*s in Spain is more spread.

Some caveats are needed related to two issues: the use of MEL tasks to assess time preference and the role of incentives in MEL tasks. First, the use of MEL tasks is probably a good choice if alternative income streams determine individuals' migration decisions, but less so if it is the alternative streams of utility that determine their decision to migrate. In the latter case, for the (annual) discount rate to equal the *RRR*, financial rewards are to be consumed at the date of receipt and the utility function is to be locally linear (Cohen et al. forthcoming). But if individuals smooth consumption over the life cycle, financial rewards at date t will be only weakly related to utility at date t. The ECF contains a question asking how much the respondent would spend of an unexpected windfall gain which may help to control for the type of consumer (on-receipt or optimizer),[8] plus an index of willingness to take risks in financial matters which may help to control for the degree of concavity of the utility function. In addition, the analysis will be conducted with dummy variables for the four categories of *RRR*s identified in the ECF as well as with a binary indicator for whether the *RRR* is greater

[8] The wording of the question is in Appendix A. Jappelli and Pistaferri (2014) analyze the responses to a similar question included in the 2010 edition of the Italian Survey of Household Income and Wealth.

than 9.8%. The use of the binary indicator may mitigate errors in the classification of respondents whose *RRR* is not reflecting their actual time preference. These errors may stem from consumption reallocation, concave utility, and the other confounding factors discussed by Frederick et al. (2002).

Another issue is whether the use of a hypothetical MEL task might produce biased preferences because respondents have no incentives to elicit their true preferences. In their review of the MEL literature, Cohen et al. (forthcoming) conclude that there exists little evidence of systematic differences between *RRR*s obtained in incentivized and unincentivized experiments, although they recommend conducting more research on this issue.

3.3. Risk Attitude

The ECF contains the following agree-disagree statement assessing attitudes toward risk in financial matters: “I’m prepared to risk a little money on saving or investing, if I can then obtain a better return in the future.” Responses are coded on a scale from 1 to 5, with 1 indicating complete disagreement and 5 indicating complete agreement. I use the response to this statement as an index of willingness to take risks.

This financial-specific measure of risk may be useful to predict migration, as this activity is typically viewed as an investment (Sjaastad 1962, Borjas 1987). But even if respondents do not view migration as an investment, Dohmen et al. (2011) have found that context-specific measures of risk taking predict risky behaviors in multiple contexts, a finding which they view as suggestive of the existence of a single underlying risk trait.

3.4. Cognitive Skills

The ECF includes three items measuring cognitive skills. The first item is a sample question from the Survey of Adult Skills assessing numeracy (OECD 2009).

Respondents are presented a card with a line plot showing the number of births in the U.S. every ten years from 1957 to 2007, and asked during which period or periods births fell. The second item is adapted from a task booklet of the International Adult Literacy Survey (OECD and Statistics Canada 2000). It consists of a 193-word news article followed by three questions assessing reading comprehension. Two of the questions test for content explicitly mentioned in the news, whereas the other tests for a concept implied by the news. The third item is a question taken from Frederick's (2005) Cognitive Reflection Test (CRT): "Imagine that to produce five pieces of equipment you need five machines working for five minutes. How long would 100 machines take to produce 100 pieces of equipment?" As argued by Frederick (2005), the suppression of the incorrect intuitive answer (100 minutes) requires cognitive reflection, namely the ability or disposition to resist reporting the response that first comes to mind.

I measure performance in each of these three items with a variable counting the number of correct responses: 0 or 1 in the first and third items; 0, 1, 2, or 3 in the second item. As shown in Table 4, the three scores correlate positively with one another, as performance in the three items likely reflects common cognitive factors. However, the strength of the correlations is not large, which suggests that the items are measuring conceptually different traits. Hence, in the regression analysis I will control for the three scores separately.

3.5. Descriptive Evidence

Figure 2 compares the distribution of *RRR*s for migrants and nonmigrants in the ECF sample. The distribution for migrants has more weight on the second more patient group, but also on the least patient class. Differences between migrant status are small and a two-sample chi-square test does not reject that both samples come from a common distribution (p-value 0.23).

Table 5, which is constructed analogously to Jaeger et al. (2010, Table 1), presents mean *RRR*s for nonmigrants and migrants calculated using the conditional means listed in Table 3 as ordered scores, as well as the percentage of the sample with *RRR* > 9.8%, stratified by demographics, cognitive skills, and risk attitude. The sample mean *RRR* for migrants is 189.4%, while that for nonmigrants is 184.4%, showing a more patient behavior among the latter. However, the percentage with *RRR* > 9.8% is smaller among migrants (65.2 vs. 67.1), indicating a more patient behavior among migrants. This result holds for nearly all of the strata when *RRR*s are measured with the binary indicator. However, when the full set of four categories is used, the pattern observed across strata is mixed. This outcome suggests that the estimated relationship between time preference and the migrant status may be influenced by the aggregation of categories of *RRR*s.

Table 5 also shows that women are less patient than men on average, and that patience decreases almost monotonically with age and increases monotonically with education level and cognitive skills. (The exception to the pattern observed in cognitive skills involves 33 observations.) Patience also tends to increase with the willingness to take risks in financial matters, although the increase is non-monotonic. We also see that the probability of lifetime migration increases monotonically with age (which captures the exposure to the risk of moving), is larger for women, the most educated, and those scoring 0 in the numeracy or CRT items, exhibits an inverted-U-shaped relationship in the reading comprehension score, and declines in the risk score (except for the most risk taking class).

The existence of common factors influencing both time preference and the migrant status demands the use of regression analysis to characterize better the effect of the former on the latter. Beyond the factors shown in Table 5, the ECF allows estimating this relationship controlling for other individual characteristics found to be correlated with the propensity to migrate internally, such as labor force status and housing tenure (e.g., Greenwood 1997). Table 6 presents summary statistics for the variables used in the estimations conducted on the ECF sample. Furthermore, and given the evidence presented in Table 1, it is necessary to incorporate information about classification errors of the migrant status in order to obtain unbiased estimates of the migrant determinants.

4. SPECIFICATION

Let y_i^*, an unobserved propensity of individual *i* to migrate internally, be given by

$$y_i^* = x_i'\beta + \varepsilon_i \tag{1}$$

where x_i is a vector of observed regressors, β is an unknown parameter vector, and ε_i is a standard normally distributed error term. Without misclassification of the migrant status, the true migrant indicator

$$\tilde{y}_i = 1\left(y_i^* > 0\right) \tag{2}$$

where the function $1(\cdot)$ is the usual indicator function, would be observed. When the true migrant status may be misclassified, the observed migrant indicator, y_i, is to be distinguished from $\tilde{y}_i$.

Let the conditional probabilities of misclassification be defined as:

$$P\left(y_i = 1 \middle| \tilde{y}_i = 0\right) = \alpha_{0i} \tag{3}$$

$$P\left(y_i = 0 \middle| \tilde{y}_i = 1\right) = \alpha_{1i} \tag{4}$$

The total probability theorem is used to derive the observed migrant probability:

$$P\left(y_i = 1\right) = \alpha_{0i} + \left(1 - \alpha_{0i} - \alpha_{1i}\right)\Phi\left(x_i'\beta\right) \tag{5}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function. Equation (5) implies the following log likelihood function:

$$\ln L(\alpha_0, \alpha_1, \beta) = \sum_{i=1}^{N} y_i \ln\left(\alpha_{0i} + (1-\alpha_{0i}-\alpha_{1i})\Phi(x_i'\beta)\right) + (1-y_i)\ln\left(\alpha_{1i} + (1-\alpha_{0i}-\alpha_{1i})(1-\Phi(x_i'\beta))\right) \tag{6}$$

where α_0 and α_1 are stacked vectors of α_{0i} and α_{1i}. Expressions (5) and (6) collapse to the probit ones when there is no misclassification ($\alpha_{0i} = \alpha_{1i} = 0$ for all i). Hence, if a probit model is estimated when misclassification exists, the resulting estimate of β will be generally inconsistent, with a tendency for the asymptotic bias to attenuate the estimated parameters (e.g., Hausman et al. 1998, Meyer and Mittag 2017). On the other hand, if one allows for misclassification when it is absent, a loss of efficiency occurs. These statistical properties suggest comparing estimators of β accounting and non accounting for misclassification of the migrant status using a Hausman (1978) specification test.

The unknown parameters $(\alpha_0, \alpha_1, \beta)$ are unidentified as there are $2N + \dim(\beta)$ parameters. Following Bollinger and David (1997), a two-step procedure is applied to estimate β. First, probit models for errors of commission and errors of omission

$$\alpha_{0i} = \Phi(x_{0i}'\beta_0) \tag{7}$$

$$\alpha_{1i} = \Phi(x_{1i}'\beta_1) \tag{8}$$

are estimated, respectively, on the orphan full-form nonmigrant sample and the full-form migrant sample of the MCVL. Besides an intercept, x_0 and x_1 include variables measured in both the MCVL and the ECF, so after estimating β_0 and β_1 on the MCVL samples, $\hat{\alpha}_{0i} = \Phi\left(x_{0i}'\hat{\beta}_0\right)$ and $\hat{\alpha}_{1i} = \Phi\left(x_{1i}'\hat{\beta}_1\right)$ can be obtained for the ECF sample.

Second, after replacing α_{0i} and α_{1i} with $\hat{\alpha}_{0i}$ and $\hat{\alpha}_{1i}$ in equation (6), the resulting expression is maximized with respect to β on the ECF sample. Under the assumption that $\hat{\alpha}_{0i}$ and $\hat{\alpha}_{1i}$ are consistent estimates of α_{0i} and α_{1i}, this "predicted probabilities estimator" of β (Meyer and Mittag 2017) is consistent and asymptotically efficient.

Standard errors of $\hat{\beta}$ are estimated using the bootstrap to account for the fact that $\hat{\alpha}_{0i}$ and $\hat{\alpha}_{1i}$ are, respectively, functions of $\hat{\beta}_0$ and $\hat{\beta}_1$, and therefore that the asymptotic variance of $\hat{\beta}$ is a function of the asymptotic variances of $\hat{\beta}_0$ and $\hat{\beta}_1$. I specify 200 bootstrap replications, a number which, for asymptotically normal estimators, ensures that the deviation from the ideal bootstrap standard errors is less than 10% with probability at least 95% (Andrews and Buchinsky 2000).

5. REGRESSION RESULTS

5.1. Errors of Commission

The orphan full-form nonmigrant sample (Table 1, column 9) is used to estimate the probit model for the probability of being a false migrant (7). I started by including in x_0 indicators for sex, education, and region of birth, plus a function of age. Since all orphans included in this sample are younger than 42 years of age, the estimated function of age has to be used to predict cohort effects on α_{0i} for individuals aged 42+ included in the ECF sample. I abandoned this approach because the resulting $\hat{\alpha}_{0i}$ seemed to me too high. For example, using the best data-fitting function of age (the linear function),[9] the mean, standard deviation, and maximum value of $\hat{\alpha}_{0i}$ calculated for the ECF sample

[9] I experimented with semi-log, polynomial (up to degree four), and inverse functions of age.

were 0.077, 0.052, and 0.366, respectively. Hence, I will assume constant α_{0i}'s across birth cohorts.

Column (1) of Table 7 presents average marginal effects (AMEs) yielded by a probit regression of the indicator for being a false migrant on an intercept plus indicators for sex, education, and region of birth. The effects of the male and college graduate dummies are very small. The largest effects are observed across regions. The effect for Cantabria cannot be calculated because there are no false migrants born in this region. The estimated effect for La Rioja may be imprecise as it is calculated with between 20 and 49 observations.[10] Another complication is created by the low prevalence of false migrants, which may give rise to bootstrap samples in which some region(s) contain no false migrants. For these reasons, I calculate $\hat{\alpha}_{0i}$ simply as the proportion of false migrants in each region. The proportion for La Rioja is taken from Navarre, a neighboring region with a relatively high rate of false migrants (10.0%). (Replacing the observation for La Rioja with the country mean hardly changes the results.) The resulting $\hat{\alpha}_{0i}$ will be treated as known population parameters, so they will not contribute to the variance of $\hat{\beta}$. Descriptive statistics for $\hat{\alpha}_{0i}$ calculated in the ECF sample are presented in Table 6.

5.2. Errors of Omission

Column (2) of Table 7 presents the AMEs of a probit regression of the indicator for being a false nonmigrant on an intercept plus indicators for sex, age, education, and region of birth, estimated on the full-form migrant sample (Table 1, column 8). The effect of age is modeled flexibly with dummy variables for each year of age (excluded

[10] Due to confidentiality restrictions in the contract for use of MCVL data, I cannot disclose the exact number of observations.

category: 50 years old). Male migrants have higher probability of being misclassified as nonmigrants than female migrants. The effects of age are shown in Figure 3. The incidence of errors of omission is highest for the 24–30-year-old group and declines steadily thereafter. This pattern, which mirrors migration propensities by age (Greenwood 1997), suggests that, as in other countries, return migration in Spain is not a phenomenon of retirement but follows closely upon the original migration (Lee 1974, DaVanzo 1983, Newbold and Bell 2001). Having a college degree leads to lower probability of being misclassified as nonmigrant. With respect to migrants born in Andalusia, migrants born in for example the Canary Islands or Valencia Region are more likely to be misclassified as nonmigrants. Descriptive statistics for $\hat{\alpha}_{1i}$ calculated in the ECF sample are presented in Table 6.

5.3. Time Preference and Migrant Status: Baseline Results

Tables B1 and B2 in Appendix B contain, respectively, the estimated parameters yielded by predicted probabilities and probit regressions of the indicator for being a migrant on six different specifications of x. As a rule, estimated probit parameters appear to be attenuated. For the parameters associated to the time preference dummies, the degree of attenuation averages 80%. Probit standard errors are also much smaller, as they are biased downwards when misclassification exists (Hausman et al. 1998). The null hypothesis that probit and predicted probabilities estimates of the slope parameters of β have the same probability limit is strongly rejected in any of the six specifications (p-values 0.00). Hence, I discuss the results of the predicted probabilities estimator.

Table 8 presents estimated AMEs yielded by the predicted probabilities estimator.[11] In columns (1)–(3), time preference is measured with dummy variables for the categories of *RRR*s identified in the ECF, while results in columns (4)–(6) are developed with the indicator for *RRR* > 9.8%. The covariates in columns (1) and (4) (sex, age, and region of birth) are conceivably exogenous to the decision to migrate. Columns (2) and (5) add covariates that may be jointly determined with migration decisions (college graduate, whether the respondent has ever worked, current employment status, whether living with spouse/partner, and housing tenure). Columns (3) and (6) control, additionally, for cognitive skills, the marginal propensity to consume (MPC) from windfall income, and the willingness to take risks in financial matters.

In column (1), the incidence of migration is smaller in the two least patient groups. Results suggest that individuals with *RRR*s between 9.8% and 44.9% are 4.4 (*S.E.* 2.1) percentage points less likely to have ever migrated than individuals with *RRR*s below 4.9% (the base category). This effect, which is statistically different from zero at 5%, represents 12% of the average probability of ever migrating (36.2%). The effect for those with *RRR*s above 44.9% is a little larger (-5.3 percentage points, *S.E.* 2.1). For individuals with *RRR*s between 4.9% and 9.8% the effect is positive but very small and measured imprecisely (0.6 percentage points, *S.E.* 3.6). Overall, the three time preference dummies are jointly significant at 5% for explaining lifetime migration: the Wald test for joint significance gives *p*-value 0.02.

[11] The ME of a regressor on the true migrant probability, $P\left(\tilde{y}_i = 1\right)$, is calculated using the finite-difference method when the regressor is binary and using calculus when it is continuous. AMEs are obtained by averaging MEs across observations, with standard errors calculated using the delta method.

In column (2), the addition of controls for education, employment and marital status, and housing tenure diminishes the size of the effect for the two least patient groups, which become -3.2 and -3.5 percentage points (*S.E.* 2.1 and 2.0), respectively. The effect for those with *RRR*s between 4.9% and 9.8% suffers little change. As a consequence, only the effect for the least patience group attains significance at 10%. When considered jointly, the three time preference dummies reach significance at 10% (p-value 0.09). The addition of further controls in column (3) leaves the estimated effects almost unchanged, and gives a p-value of 0.08 for the test of joint significance. Overall, therefore, when time preference is measured with the full set of dummies for *RRR*s, the evidence in favor of an effect of time preference on the likelihood of ever migrating is not conclusive.

Classifying respondents into just two categories of *RRR*s may reduce biases stemming from classification errors in terms of true time preference. Furthermore, even if this type of error was absent in the data, the use of four categories to represent the regression relationship between time preference and migrant status when this relationship can be accounted for by just two categories of *RRR*s represents a loss of information. In this respect, the difference in the incidence of migration between the two most patient groups does not attain significance in any of the regressions shown in columns (1)–(3) (p-values > 0.57). The same conclusion applies to the two least patient groups (p-values > 0.64). Measuring time preference with a binary indicator for *RRR* > 9.8% also facilitates comparison with Gibson and McKenzie (2011).

In column (4), the incidence of migration is 5.1 percentage points (*S.E.* 1.7) smaller among the least patient respondents. This effect, which represents 14% of the average probability of ever migrating, attains statistical significance at 1%. In column (5), the addition of controls for education, employment and marital status, and housing tenure reduces its size to -4.0 percentage points (*S.E.* 1.6). The addition of further controls in column (6) leaves an effect of -4.1 percentage points (*S.E.* 1.6), representing 11% of the average probability of ever migrating and being statistically different from zero at 2%. Therefore, when time preference is measured with the binary indicator for *RRR* > 9.8%, impatience decreases the probability of lifetime migration by 11 to 14 percent of this variable's unconditional probability. Gibson and McKenzie (2011) estimate a somewhat larger effect. In their samples of top students, individuals with *RRR*s > 9.8% are found to be 12 to 13 percentage points less likely to have ever migrated, which amounts approximately to 20 percent of the average probability of lifetime migration in their samples (about 64 percent).

Table 8 also lists the estimated AMEs of the covariates. The college migration premium is robust to the inclusion of a measure of impatience in the specification. In comparison with individuals not having a college degree, college graduates are approximately 6.5 (*S.E.* 2.3) percentage points more likely to have ever migrated, which represents an 18 percent increase of the average probability of ever migrating. Excluding the indicator(s) for *RRR*s from the set of regressors yields an AME of college education of 7.3 (*S.E.* 2.1) percentage points. Therefore, behavior in the MEL task accounts for about 11% of the lifetime migration impact of college education.

If we look at the cognitive skills measures, only the reading comprehension item attains significance at 10%. Its AME indicates that a one-standard-deviation increase in the reading comprehension score leads to a 1.4 (*S.E.* 0.8) percentage points increase in the probability of lifetime migration. When consider jointly, the three cognitive skills measures appear to be statistically insignificant for explaining migrant status (p-values > 0.15). Perhaps more importantly, excluding these measures from the specification leaves the estimated relationship between time preference and migrant status almost unchanged. For example, the estimated AME of the binary indicator for *RRR* > 9.8% becomes -4.0 (*S.E.* 1.7). This conclusion is reassuring for the results of studies that identify the effect of patience on migration without information on individuals' cognitive ability.

As expected, the incidence of migration increases with age. The self-employed are 7.4 (*S.E.* 2.5) percentage points less likely to have ever migrated than those not in the labor force (and, generally, than any other type of worker). Living with spouse/partner increases the probability of having ever migrated by some 3.4 percentage points (*S.E.* 1.7). In comparison with renters, owners who have inherited their dwelling are 19.0 percentage points less likely to have ever migrated, while among owners who have purchased their dwelling the negative effect associated to home ownership is substantial but much smaller: about 7.7 (*S.E.* 2.7) percentage points. As to the MPC from windfall income, this variable appears to be insignificant at conventional levels. If the MPC depends on individuals' resources, this result agrees with previous evidence showing that greater wealth does not increase the likelihood of migration (McHenry 2015). The estimated effect of the risk index is negative but fails to gain significance at 10 percent. Although most previous studies estimate a positive correlation between attitudes toward risk in general and migration (Jaeger et al. 2010, Gibson and McKenzie 2011, Nowotny 2014, Akgüç et al. 2016, Dustmann et al. 2017, and Huber and Nowotny 2018), Jaeger et al. (2007) find that attitudes towards risk in financial matters are essentially unrelated to the probability of migration.

5.4. Time Preference and Migrant Status: Robustness to Sensitivity Analyses

I analyze the robustness of these results to alternative specification choices. First, I exclude respondents who answer "don't know" in any of the binary choices of the MEL task. Second, following Jaeger et al. (2010), I re-estimate the models from Table 8 controlling for age and education flexibly by using dummy variables for 31 two-year age groups (18–19, 20–21, etc.) and 8 educational attainment levels. (Disaggregating these groups further produces extremely large bootstrap standard errors for some parameters, which are then transferred to standard errors of AMEs of key regressors.) With this much less parsimonious specification, it is unlikely that the effect of time preference may be contaminated by neglected effects of age or education. And third, following Kirby et al. (2002), who find an inverse relationship between discount rates and recent income, I re-estimate the models from Table 8 (except those in columns 1 and 4) with household income included among the regressors.[12]

The main results are shown in the three panels of Table 9. In all panels, the sets of covariates by column are as in Table 8, except that age and education are controlled for flexibly in Panel 2 and that household income is included among the regressors in columns (2), (3), (5), and (6) of Panel 3. The use of the binary indicator for $RRR > 9.8\%$ unveils an inverse relationship between impatience and the probability of being a migrant in the three panels. The exclusion of "don't know" respondents to the MEL task yields slightly less precise AMEs. With age and education controlled for flexibly, the AME of the binary indicator is somewhat smaller, ranging from -4.5 to -3.3 percentage points, and measured less precisely, but it still attains significance at or around 5%. The dummies representing the interval to which household income belongs are jointly insignificant for explaining migrant status at conventional levels (p-values > 0.16).

6. CONCLUSIONS

[12] Annual household income in the ECF is recorded in six categories. About 10 percent of respondents do not provide this variable. For each missing value, the ECF provides five imputed values. Following Little and Rubin (2002), I conduct multiple imputation estimations in each of the bootstrap samples.

The Survey of Financial Competences conducted in Spain in 2016 offers a unique alternative for assessing the relationship between time preference and internal migration in a developed country. The residential history collected by this survey, however, is limited to a baseline comparison of residence at birth and at survey, which introduces significant misclassification of the migrant status. In an administrative data set representative of the same population, considering residence at the time of each period of Social Security affiliation increases the proportion of migrants by 94 percent over the baseline. Modeling the migrant status conditional on individual probabilities of misclassification produces parameter estimates that are significantly different from estimates that do not condition on probabilities of misclassification.

The results reveal that required rates of return (*RRR*s) for financial flows and lifetime migration tend to be inversely related, even after controlling for several measures of individuals' cognitive ability. The nature of the relationship manifests most clearly when *RRR*s are classified into just two categories, and less so when the original classification into four categories is used. When time preference is measured with a binary indicator for requiring a rate of return greater than 9.8%, being impatient decreases the probability of lifetime migration. Over a range of specifications, the size of this effect ranges between 9 and 15 percent of the unconditional probability of lifetime migration (0.362).

The results of this study have been developed assuming that time preference is an intrinsic characteristic of individuals which is not affected by their migration decisions. Goldbach and Schlüter (2018) review the literature investigating the influence of past decisions (migration, in particular) on individuals' attitudes, and conclude that risk and time preferences appear to be rather "deep parameters" and hence mostly exogenous to past decisions.

Cadena and Keys (2015) provide compelling evidence that impatient individuals are more likely to exhibit preference reversals in educational investment. To investigate the role of time-inconsistent preferences in the migration decision, additional variables such as the willingness to migrate or the level of regret for inappropriately deciding in the past would be needed.

APPENDIX A. ECF QUESTIONS ON TIME PREFERENCE AND PROPENSITY TO SPEND.

“We’re now going to put to you several hypothetical situations. The following questions do not have a right or wrong reply. We would just like to know what your choice faced with different options would be.

Suppose you are offered €2,000 today. However, if you wait a year, you would be offered €2,200. In both cases, you would be fully certain to receive the money. What would you choose: €2,000 today or €2,200 in a year’s time?

(If respondents choose €2,000 today, they are asked:)

Now suppose that, if you wait a year, you would be offered €3,000. In both cases, you would be fully certain to receive the money. What would you choose: €2,000 today or €3,000 in a year’s time?

(If, on the other hand, respondents choose €2,200 in a year’s time, they are asked:)

Now suppose that, if you wait a year, you would be offered €2,100. In both cases, you would be fully certain to receive the money. What would you choose: €2,000 today or €2,100 in a year’s time?

(This part of the ECF questionnaire ends with the following question:)

Imagine you were to win (e.g. in the Christmas lottery) an amount of money equivalent to your household’s monthly income. What percentage would you spend during the following 12 months, instead of saving it or using it to repay outstanding loans?”

APPENDIX B. ESTIMATED PARAMETERS OF MIGRATION MODELS.

Table B1. Predicted probabilities estimates of lifetime migration.

Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)
$1(4.9\% < RRR \leq 9.8\%)$	.028	.081	.090			
	(.160)	(.159)	(.162)			
$1(9.8\% < RRR \leq 44.9\%)$	-.211**	-.165	-.165			
	(.101)	(.105)	(.106)			
$1(44.9\% < RRR)$	-.256**	-.181*	-.184*			
	(.105)	(.103)	(.104)			
$1(9.8\% < RRR)$				-.246***	-.202**	-.207***
				(.077)	(.079)	(.080)
Male	.079	.077	.085	.080	.075	.084
	(.069)	(.075)	(.079)	(.069)	(.074)	(.079)
Age	.021***	.025***	.026***	.021***	.025***	.026***
	(.003)	(.003)	(.004)	(.002)	(.003)	(.004)
College graduate		.337***	.322***		.336***	.321***
		(.102)	(.107)		(.101)	(.106)
Ever worked		.170	.160		.172	.162
		(.156)	(.162)		(.155)	(.161)
Self employed		-.446***	-.451***		-.446***	-.450***
		(.167)	(.173)		(.167)	(.173)
Employee		.049	.041		.046	.037
		(.102)	(.103)		(.102)	(.103)
Unemployed		.062	.064		.056	.057
		(.139)	(.142)		(.138)	(.140)
Living with spouse/partner		.185**	.181**		.184**	.180**
		(.088)	(.091)		(.088)	(.090)
Owner (purchase)		-.342***	-.353***		-.343***	-.354***
		(.111)	(.112)		(.112)	(.113)
Owner (inheritance)		-1.167***	-1.162***		-1.167***	-1.162***
		(.226)	(.227)		(.226)	(.226)
Numeracy item			.105			.103
			(.093)			(.093)
Reading comprehension item			.083*			.084*
			(.046)			(.047)
CRT item			-.106			-.109
			(.099)			(.099)
MPC (÷ 10)			-.015			-.015
			(.013)			(.013)
Risk score			-.048			-.048
			(.031)			(.031)
Intercept	-1.367***	-1.594***	-1.614***	-1.352***	-1.561***	-1.581***
	(.155)	(.246)	(.313)	(.154)	(.245)	(.315)
Log-likelihood	-2,291.31	-2,249.49	-2,244.45	-2,291.46	-2,249.68	-2,244.68
Joint significance of time preference dummies	[.017]	[.087]	[.078]			

Notes: The number of observations is 6,755 in all columns. All estimations include a complete set of dummies for region of birth. $1(\cdot)$ is the usual indicator function. Bootstrap standard errors are in parentheses and probability values in brackets. *: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table B2. Probit estimates of lifetime migration.

Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)
$1(4.9\% < RRR \leq 9.8\%)$	.060	.097	.101			
	(.075)	(.076)	(.077)			
$1(9.8\% < RRR \leq 44.9\%)$	-.122**	-.090	-.082			
	(.059)	(.060)	(.060)			
$1(44.9\% < RRR)$	-.119**	-.063	-.056			
	(.057)	(.059)	(.060)			
$1(9.8\% < RRR)$				-.141***	-.110**	-.104**
				(.044)	(.046)	(.046)
Male	-.061	-.054	-.050	-.062	-.058	-.054
	(.042)	(.043)	(.044)	(.042)	(.043)	(.044)
Age	.021***	.021***	.022***	.021***	.021***	.022***
	(.001)	(.002)	(.002)	(.001)	(.002)	(.002)
College graduate		.214***	.206***		.208***	.201***
		(.051)	(.053)		(.051)	(.053)
Ever worked		.144	.140		.145	.142
		(.099)	(.099)		(.099)	(.099)
Self employed		-.249***	-.259***		-.248***	-.256***
		(.084)	(.085)		(.084)	(.084)
Employee		-.031	-.043		-.032	-.045
		(.061)	(.061)		(.061)	(.061)
Unemployed		-.008	-.010		-.006	-.009
		(.078)	(.078)		(.078)	(.078)
Living with spouse/partner		.111**	.108**		.109**	.106**
		(.049)	(.049)		(.049)	(.049)
Owner (purchase)		-.172***	-.177***		-.174***	-.179***
		(.061)	(.061)		(.061)	(.061)
Owner (inheritance)		-.673***	-.678***		-.673***	-.678***
		(.101)	(.102)		(.101)	(.101)
Numeracy item			.049			.046
			(.047)			(.047)
Reading comprehension item			.048*			.047*
			(.027)			(.027)
CRT item			-.030			-.031
			(.053)			(.053)
MPC (÷ 10)			-.015**			-.015**
			(.007)			(.007)
Risk score			-.023			-.024
			(.017)			(.017)
Intercept	-1.788***	-1.895***	-1.891***	-1.768***	-1.859***	-1.847***
	(.096)	(.144)	(.174)	(.092)	(.141)	(.171)
Log-likelihood	-2,278.48	-2,233.55	-2,227.69	-2,278.80	-2,234.47	-2,228.66

Notes: The number of observations is 6,755 in all columns. All estimations include a complete set of dummies for region of birth. $1(\cdot)$ is the usual indicator function. Standard errors are in parentheses. *: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

REFERENCES

Akgüç, Mehtap, Xingfei Liu, Massimiliano Tani, and Klaus F. Zimmermann. 2016. Risk attitudes and migration. *China Economic Review* 37: 166–176.

Andrews, Donald W.K., and Moshe Buchinsky. 2000. A three-step method for choosing the number of bootstrap repetitions. *Econometrica* 68(1): 23–51.

Arcand, Jean-Louis, and Linguère Mously Mbaye. 2013. Braving the waves: The role of time and risk preferences in illegal migration from Senegal. IZA Discussion Paper No. 7517.

Banco de España and National Securities Market Commission. 2018. Survey of Financial Competences (ECF) 2016. Available at https://app.bde.es/pmk/en/ecf/2016.

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro. 2013. Who is 'behavioral'? Cognitive ability and anomalous preferences. *Journal of the European Economic Association* 11(6): 1231–1255.

Bollinger, Christopher R., and Martin H. David. 1997. Modeling discrete choice with response error: Food Stamp participation. *Journal of the American Statistical Association* 92(439): 827–835.

Borjas, George J. 1987. Self-selection and the earnings of immigrants. *American Economic Review* 77(4): 531–553.

Bover, Olympia, Laura Hospido, and Ernesto Villanueva. 2019. The Survey of Financial Competences (ECF): Description and methods of the 2016 wave. Banco de España Occasional Paper No. 1909.

Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini. 2009. Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences of the United States of America* 106(19): 7745–7750.

Cadena, Brian C., and Benjamin J. Keys. 2015. Human capital and the lifetime costs of impatience. *American Economic Journal: Economic Policy* 7(3): 126–153.

Carlson, Elwood. 2007. Internal migration from Turkish birth histories. *Genus* 63(3/4): 99–125.

Cohen, Jonathan, Keith Marzilli Ericson, David Laibson, and John Myles White. Forthcoming. Measuring time preferences. *Journal of Economic Literature*. Available at https://scholar.harvard.edu/laibson/publications/measuring-time-preferences.

DaVanzo, Julie. 1983. Repeat migration in the United States: Who moves back and who moves on? *Review of Economics and Statistics* 65(4): 552–559.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2010. Are risk aversion and impatience related to cognitive ability? *American Economic Review* 100: 1238–1260.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* 9(3): 522–550.

Dustmann, Christian, Francesco Fasani, Xin Meng, and Luigi Minale. 2017. Risk attitudes and household migration decisions. Centro Studi Luca d'Agliano Development Studies Working Paper No. 423. Available at: https://ssrn.com/abstract=3044497.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. Global evidence on economic preferences. *Quarterly Journal of Economics* 133(4): 1645–1692.

Frederick, Shane. 2005. Cognitive reflection and decision making. *Journal of Economic Perspectives* 19(4): 25–42.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue. 2002. Time discounting and time preference: A critical review. *Journal of Economic Literature* 40: 351–401.

Gibson, John, and David McKenzie. 2011. The microeconomic determinants of emigration and return migration of the best and brightest: Evidence from the Pacific. *Journal of Development Economics* 95: 18–29.

Goldbach, Carina, and Achim Schlüter. 2018. Risk aversion, time preferences, and out-migration. Experimental evidence from Ghana and Indonesia. *Journal of Economic Behavior and Organization* 150: 132–148.

Golsteyn, Bart H.H., Hans Grönqvist, and Lena Lindahl. 2014. Adolescent time preferences predict lifetime outcomes. *Economic Journal* 124: F739–F761.

Greenwood, Michael J. 1997. Internal migration in developed countries. In *Handbook of Population and Family Economics*, edited by Mark R. Rosenzweig and Oded Stark, 647–720. Amsterdam: North Holland.

Hausman, Jerry A. 1978. Specification tests in econometrics. *Econometrica* 46(6): 1251–1271.

Hausman, Jerry A., Jason Abrevaya, and Fiona M. Scott-Morton. 1998. Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics* 87: 239–269.

Huber, Peter, and Klaus Nowotny. 2018. Risk aversion and the willingness to migrate in 30 countries. WIFO Working Paper No. 569.

Jaeger, David A., Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, and Holger Bonin. 2010. Direct evidence on risk attitudes and migration. *Review of Economics and Statistics* 92(3): 684–689.

Jaeger, David A., Armin Falk, Holger Bonin, David Huffman, Thomas Dohmen, and Uwe Sunde. 2007. Direct evidence on risk attitudes and migration. IZA Discussion Paper No. 2655.

Jappelli, Tullio, and Luigi Pistaferri. 2014. Fiscal policy and MPC heterogeneity. *American Economic Journal: Macroeconomics* 6(4): 107–136.

Kirby, Kris N., Ricardo Godoy, Victoria Reyes-García, Elizabeth Byron, Lilian Apaza, William Leonard, Eddy Pérez, Vincent Vadez, and David Wilkie. 2002. Correlates of delay-discount rates: Evidence from Tsimane' Amerindians of the Bolivian rain forest. *Journal of Economic Psychology* 23: 291–316.

Lee, Anne S. 1974. Return migration in the United States. *International Migration Review* 8(2): 283–300.

Little, Roderick J.A., and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data*. Second edition. Wiley-Interscience: Hoboken, New Jersey.

Malamud, Ofer, and Abigail Wozniak. 2012. The impact of college on migration. Evidence from the Vietnam generation. *Journal of Human Resources* 47(4): 913–950.

Martí, Mónica, and Carmen Ródenas. 2007. Migration estimation based on the Labour Force Survey: An EU-15 perspective. *International Migration Review* 41(1): 101–126.

McHenry, Peter. 2015. Does low wealth constrain long-distance migration? *Industrial and Labor Relations Review* 68(1): 79–119.

Meyer, Bruce D., and Nikolas Mittag. 2017. Misclassification in binary choice models. *Journal of Econometrics* 200: 295–311.

Molloy, Raven, Christopher L. Smith, and Abigail Wozniak. 2011. Internal migration in the United States. *Journal of Economic Perspectives* 25(3): 173–196.

Newbold, K. Bruce, and Martin Bell. 2001. Return and onwards migration in Canada and Australia: Evidence from fixed interval data. *International Migration Review* 35(4): 1157–1184.

Nowotny, Klaus. 2014. Cross-border commuting and migration intentions: the roles of risk aversion and time preference. *Contemporary Economics* 8(2): 137–156.

OECD. 2009. Sample questions and questionnaire. Available at https://www.oecd.org/skills/piaac/samplequestionsandquestionnaire.htm.

OECD and Statistics Canada. 2000. Literacy in the information age. Final report of the International Adult Literacy Survey. Available at http://www.oecd.org/education/innovation-education/39437980.pdf.

Sjaastad, Larry A. 1962. The costs and returns of human migration. *Journal of Political Economy* 70(5): 80–93.

Thomas, Kevin J.A. 2012. Migration, household configurations, and the well-being of adolescent orphans in Rwanda. *Population Research and Policy Review* 31(4): 587–607.

Wang, Yibing, Wei Dong, Liangqi Zhang, David Chin, Markos Papageorgiou, Geoffrey Rose, and William Young. 2012. Speed modeling and travel time estimation based on truncated normal and lognormal distributions. *Transportation Research Record: Journal of the Transportation Research Board* 2315: 66–72.

TABLES AND FIGURES

Table 1. Cross-region migrants in Spain (%), by sample and definition of migrant. MCVL 2015 and ECF.

	All			Full-form migrants	Orphan full-form nonmigrants	Observations			
	Reduced form		Full form	Reduced form	Reduced form	All		Full-form migrants	Orphan full-form nonmigrants
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	ECF	MCVL	MCVL	MCVL	MCVL	ECF	MCVL	MCVL	MCVL
Total population	14.5	18.7	36.2	51.7	4.2	6,755	917,879	332,245	7,006
Sex									
Female	15.7	19.0	33.2	57.1	4.2	3,369	420,324	139,746	3,492
Male	13.3	18.5	38.7	47.7	4.3	3,386	497,555	192,499	3,514
Age									
18–24	2.7	6.7	19.1	33.8	3.7	655	45,603	8,752	2,861
25–44	8.9	13.2	34.7	37.8	4.6	2,261	354,920	123,427	4,145
45–64	16.7	20.6	37.4	55.1		2,686	352,543	131,881	0
65+	27.8	30.0	41.4	72.5		1,153	164,813	68,185	0
Education									
Less than high sch.	16.0	19.1	34.9	54.5	4.2	2,922	545,720	190,559	4,431
High school	10.3	16.4	34.6	47.4	4.5	2,253	220,185	76,233	1,701
College graduate	17.6	20.8	43.1	48.3	4.1	1,580	151,974	65,453	874

Notes: Population estimates. Samples include individuals aged 18–79 born and residing in the 17 regions of Spain surveyed by the ECF. Lifetime migration in the MCVL is measured in 2016. Reduced-form migrants are individuals who at survey (ECF) or data compilation (MCVL) reside in a region other than where they were born. Full-form migrants are individuals who resided in more than one region between (and including) birth and data compilation, with the exception of orphans who left birth region as a child and never changed region as an adult.

Table 2. Required rates of return (%) in the MEL task.

	First binary choice: (€2,000, today)	(€2,200, 1 year)
Second binary choice:		
(€2,000, today)	*RRR* > 44.9 [37.2]	4.9 < *RRR* ≤ 9.8 [10.6]
Money in 1 year[a]	9.8 < *RRR* ≤ 44.9 [29.7]	*RRR* ≤ 4.9 [22.5]

Notes: [a]: €3,000 if the respondent first choose (€2,000, today); €2,100 if the respondent first choose (€2,200, 1 year). The number of observations is 6,755. The sample includes individuals aged 18–79 born and residing in the 17 regions surveyed by the ECF. Sample percentages are in brackets.

Table 3. Comparison of fitted and actual distributions.

(1) Normal model			(2) Lognormal model			(3)
*RRR*s range (%)	Actual	Model[a]	Logarithm of *RRR*s range	Actual	Model[b]	Mean *RRR* (%)
$RRR \leq 4.9$	1,519	1,647	$\ln RRR \leq \ln 4.9$	1,519	1,510	2.1
$4.9 < RRR \leq 9.8$	718	279	$\ln 4.9 < \ln RRR \leq \ln 9.8$	718	753	7.2
$9.8 < RRR \leq 44.9$	2,003	2,386	$\ln 9.8 < \ln RRR \leq \ln 44.9$	2,003	1,971	23.1
$44.9 < RRR$	2,515	2,444	$\ln 44.9 < \ln RRR$	2,515	2,521	475.2
Total	6,755	6,756[c]		6,755	6,755	
χ^2 statistic		764.25			2.21	
$\chi^{2\,10}_{1\,\mathrm{df}}$ critical value		2.71			2.71	

Notes: [a]: *RRR* ~ $N(0.31, 0.15)$. [b]: *RRR* ~ $LN(-1.46, 4.15)$. [c]: Totals may not sum to actuals because of rounding. The sample includes individuals aged 18–79 born and residing in the 17 regions surveyed by the ECF.

Table 4. Pearson correlation coefficients between cognitive measures.

	Numeracy	Reading comprehension
Reading comprehension	.23	
Cognitive reflection	.18	.17

Notes: The number of observations is 6,755. The sample includes individuals aged 18–79 born and residing in the 17 regions surveyed by the ECF.

Table 5. Average required rates of return (%) for individuals classified as migrants and nonmigrants in the ECF.

	Mean *RRR*		% with *RRR* > 9.8		Observations		
	Nonmigrants	Migrants	Nonmigrants	Migrants	Nonmigrants	Migrants	% migrants
All	184.4	189.4	67.1	65.2	5,890	865	12.8
Sex							
Female	196.8	206.4	68.6	67.2	2,918	451	13.4
Male	172.2	170.9	65.7	63.0	2,972	414	12.2
Age							
18–24	132.3	194.2	60.6	50.0	635	20	3.1
25–44	153.9	160.2	61.6	53.1	2,069	192	8.5
45–64	202.3	186.8	70.3	69.6	2,305	381	14.2
65+	246.5	213.3	76.5	68.8	881	272	23.6
Education							
Less than high school	241.9	238.3	76.8	75.1	2,517	405	13.9
High school	161.5	190.2	64.2	67.1	2,034	219	9.7
College graduate	110.8	106.5	53.5	46.9	1,339	241	15.3
Numeracy item[a]							
0	215.7	217.8	72.5	70.5	3,386	526	13.4
1	142.0	145.4	59.9	56.9	2,504	339	11.9
Reading comprehension item[a]							
0	261.9	166.3	76.0	60.6	262	33	11.2
1	233.3	230.8	74.3	70.7	681	123	15.3
2	201.5	204.0	69.8	68.4	1,879	307	14.0
3	156.4	167.5	63.2	61.4	3,068	402	11.6
CRT item[a]							
0	199.8	198.8	69.4	67.8	4,436	676	13.2
1	137.2	155.7	60.2	56.1	1,454	189	11.5
Risk index							
1	241.7	264.2	75.7	73.8	1,458	294	16.8
2	186.2	157.0	68.0	61.9	1,143	189	14.2
3	156.0	134.3	62.9	55.6	1,275	151	10.6
4	153.4	160.9	62.5	64.7	1,350	150	10.0
5	172.8	149.4	64.3	60.5	664	81	10.9

Notes: Sample estimates. The overall sample includes individuals aged 18–79 born and residing in the 17 regions surveyed by the ECF. Average *RRR*s are calculated using the mean *RRR*s listed in Table 3 as ordered scores. [a]: Number of correct answers. The risk index is coded on a scale from 1 to 5, with 1 indicating unwilling to take financial risks and 5 indicating very willing to take financial risks.

Table 6. Descriptive statistics. ECF sample.

	Mean	Std. dev.	Min	Max
Lifetime migrant	0.128		0	1
Male	0.501		0	1
Age	47.505	15.734	18	79
College graduate	0.234		0	1
Ever worked	0.927		0	1
Self employed	0.114		0	1
Employee	0.418		0	1
Unemployed	0.130		0	1
Living with spouse/partner	0.655		0	1
Owner (purchase)	0.739		0	1
Owner (inheritance)	0.113		0	1
Numeracy item	0.421		0	1
Reading comprehension item	2.307	0.844	0	3
CRT item	0.243		0	1
MPC	39.545	32.024	0	100
Risk score	2.727	1.350	1	5
$\hat{\alpha}_{0i}$	0.047	0.028	0	0.100
$\hat{\alpha}_{1i}$	0.504	0.167	0.095	0.891

Notes: The number of observations is 6,755. The distribution of individuals across categories of *RRR*s is given in Table 2.

Table 7. Probit models of commission errors and omission errors. Average marginal effects.

	(1) Commission errors		(2) Omission errors	
Explanatory variables	AME	S.E.	AME	S.E.
Male	.002	.005	.111***	.002
Age			See Figure 3	
College graduate	-.003	.007	-.036***	.002
Region				
Andalusia	Ref.		Ref.	
Aragon	.013	.015	.006	.005
Asturias	.033*	.018	.018***	.005
Balearic Islands	.075***	.024	.112***	.008
Canary Islands	-.001	.008	.189***	.007
Cantabria	Not estimable		.058***	.007
Castile-Leon	.059***	.016	-.084***	.003
Castile-La Mancha	.057***	.018	-.085***	.003
Catalonia	.021***	.007	-.005	.004
Valencia Region	-.001	.006	.134***	.004
Extremadura	.013	.014	-.125***	.004
Galicia	.005	.008	.085***	.004
Madrid Region	.058***	.010	-.038***	.003
Murcia Region	-.001	.011	.086***	.005
Navarre	.079**	.032	.105***	.007
Basque Country	.047***	.014	-.040***	.004
La Rioja	.170**	.077	.060***	.008

Notes: Column (1): Estimated on the orphan full-form nonmigrant sample of the MCVL (Table 1, column 9). Dependent variable = 1 if false migrant. Mean of dependent variable is 0.042. Age is not included in the regressors. Column (2): Estimated on the full-form migrant sample of the MCVL (Table 1, column 8). Dependent variable = 1 if false nonmigrant. Mean of dependent variable is 0.483. Age is included in the regressors as dummy variables for each year of age. *: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 8. Predicted probabilities estimates of lifetime migration. Average marginal effects.

Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)
$1(4.9\% < RRR \leq 9.8\%)$	.006	.017	.019			
	(.036)	(.034)	(.035)			
$1(9.8\% < RRR \leq 44.9\%)$	-.044**	-.032	-.032			
	(.021)	(.021)	(.021)			
$1(44.9\% < RRR)$	-.053**	-.035*	-.036*			
	(.021)	(.020)	(.020)			
$1(9.8\% < RRR)$				-.051***	-.040**	-.041**
				(.017)	(.016)	(.016)
Male	.016	.015	.016	.016	.014	.016
	(.014)	(.014)	(.015)	(.014)	(.014)	(.015)
Age	.004***	.005***	.005***	.004***	.005***	.005***
	(.000)	(.001)	(.001)	(.000)	(.001)	(.001)
College graduate		.069***	.065***		.068***	.065***
		(.022)	(.023)		(.022)	(.023)
Ever worked		.031	.029		.031	.029
		(.027)	(.028)		(.027)	(.028)
Self employed		-.073***	-.074***		-.073***	-.074***
		(.024)	(.025)		(.024)	(.025)
Employee		.010	.008		.009	.007
		(.020)	(.020)		(.020)	(.020)
Unemployed		.012	.013		.011	.011
		(.028)	(.029)		(.028)	(.028)
Living with spouse/partner		.035**	.034**		.034**	.034**
		(.016)	(.017)		(.016)	(.017)
Owner (purchase)		-.075***	-.077***		-.076***	-.078***
		(.026)	(.027)		(.027)	(.027)
Owner (inheritance)		-.190***	-.190***		-.190***	-.190***
		(.031)	(.031)		(.031)	(.032)
Numeracy item			.020			.020
			(.018)			(.018)
Reading comprehension item			.016*			.016*
			(.009)			(.009)
CRT item			-.020			-.020
			(.018)			(.018)
MPC (÷ 10)			-.003			-.003
			(.002)			(.002)
Risk score			-.009			-.009
			(.006)			(.006)
$1(9.8\% < RRR \leq 44.9\%)$ $= 1(44.9\% < RRR)$	[.643]	[.871]	[.850]			

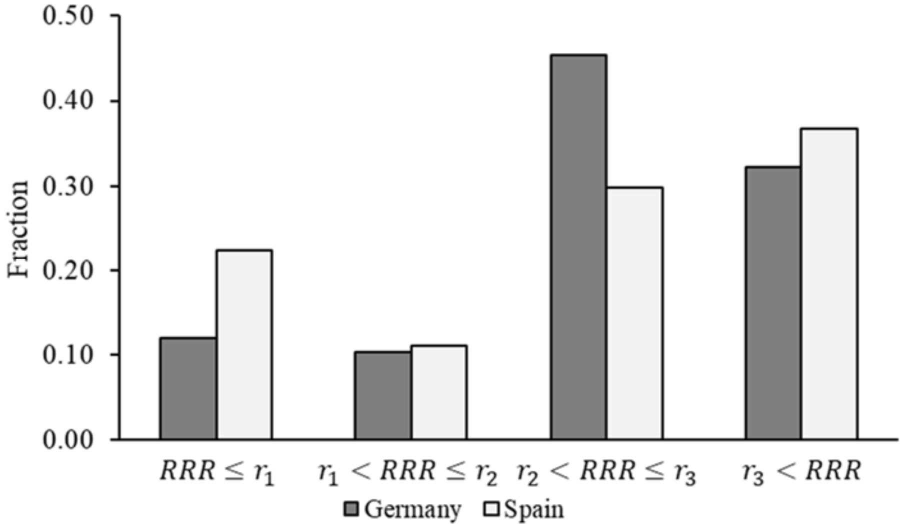
Notes: The number of observations is 6,755 in all columns. All estimations include a complete set of dummies for region of birth. $1(\cdot)$ is the usual indicator function. Bootstrap standard errors are in parentheses and probability values in brackets. *: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Table 9. Robustness of predicted probabilities estimates of lifetime migration to alternative specifications. Selected average marginal effects.

Panel 1: "Don't know" responses in MEL task excluded from sample (observations: 6,646).						
	(1)	(2)	(3)	(4)	(5)	(6)
$1(4.9\% < RRR \le 9.8\%)$	.026	.034	.035			
	(.038)	(.037)	(.038)			
$1(9.8\% < RRR \le 44.9\%)$	-.042*	-.031	-.031			
	(.022)	(.022)	(.022)			
$1(44.9\% < RRR)$	-.051**	-.035	-.035			
	(.022)	(.022)	(.022)			
$1(9.8\% < RRR)$				-.055***	-.043**	-.044**
				(.018)	(.017)	(.017)
Panel 2: Age and educational attainment controlled for flexibly.						
	(1)	(2)	(3)	(4)	(5)	(6)
$1(4.9\% < RRR \le 9.8\%)$	.015	.023	.025			
	(.035)	(.034)	(.035)			
$1(9.8\% < RRR \le 44.9\%)$	-.037	-.023	-.023			
	(.024)	(.024)	(.024)			
$1(44.9\% < RRR)$	-.045**	-.027	-.029			
	(.022)	(.022)	(.023)			
$1(9.8\% < RRR)$				-.046**	-.033*	-.035*
				(.018)	(.018)	(.019)
Panel 3: Household income included among the regressors.						
	(1)	(2)	(3)	(4)	(5)	(6)
$1(4.9\% < RRR \le 9.8\%)$	.006	.017	.019			
	(.036)	(.032)	(.033)			
$1(9.8\% < RRR \le 44.9\%)$	-.044**	-.031	-.031			
	(.021)	(.021)	(.021)			
$1(44.9\% < RRR)$	-.053**	-.034	-.035			
	(.021)	(.022)	(.022)			
$1(9.8\% < RRR)$				-.051***	-.039**	-.040**
				(.017)	(.018)	(.018)

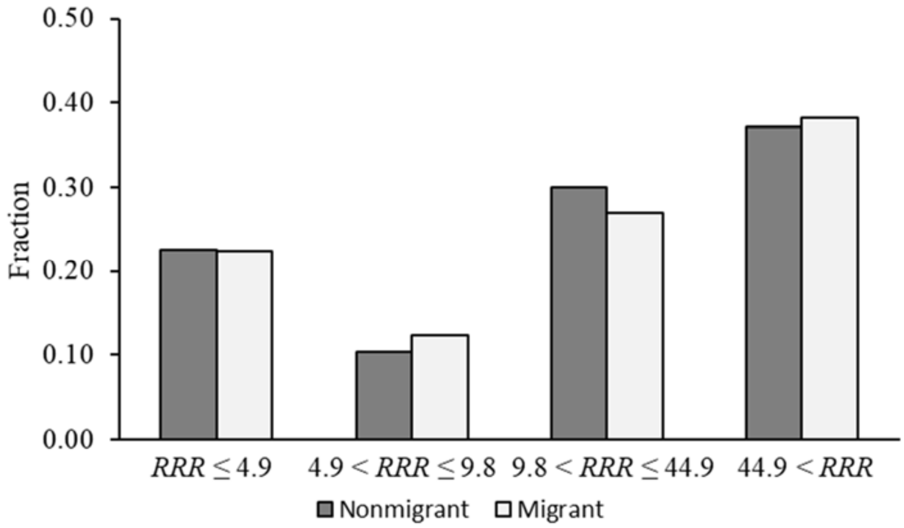
Notes: Except when noted, the number of observations is 6,755. In all panels, the set of controls in columns (1) and (4) comprises sex, age, and region of birth; columns (2) and (5) add controls for college graduate (education level in Panel 2), whether the respondent has ever worked, current employment status, whether living with spouse/partner, and housing tenure; columns (3) and (6) add controls for cognitive skills, MPC, and the risk index. Columns (2), (3), (5), and (6) of Panel 3 control additionally for household income. $1(\cdot)$ is the usual indicator function. Bootstrap standard errors are in parentheses. *: Significant at 10%. **: Significant at 5%. ***: Significant at 1%.

Figure 1. Population distribution of required rates of return (%).



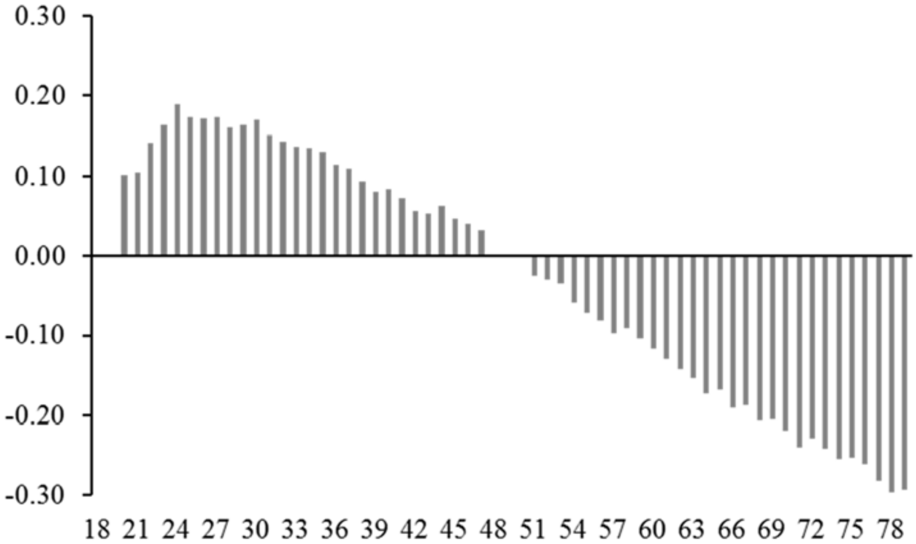
Notes: The data for Germany are from Dohmen et al. (2010) and pertain to the resident population of Germany aged 17 and older. The data for Spain are from the ECF and pertain to individuals aged 18–79 born and residing in the 17 regions surveyed by the ECF. Thresholds for the German (Spanish) data: $r_1 = 5.0$, $r_2 = 10.0$, and $r_3 = 45.0$ ($r_1 = 4.9$, $r_2 = 9.8$, and $r_3 = 44.9$).

Figure 2. Required rates of return (%) for individuals classified as migrants and nonmigrants in the ECF.



Notes: Sample estimates. The number of observations is 5,890 nonmigrants and 865 migrants. Both samples include individuals aged 18–79 born and residing in the 17 regions surveyed by the ECF.

Figure 3. Average marginal effects of each year of age on the probability of being misclassified as nonmigrant.



Notes: Estimated on the full-form migrant sample of the MCVL (Table 1, column 8). Dependent variable = 1 if false nonmigrant. Mean of dependent variable is 0.483. Excluded age: 50 years old. The effects shown are those achieving significance at 1%.