UNIVERSITE DE LAUSANNE
ECOLE DES HAUTES ETUDES COMMERCIALES

Applied Econometrics II

MODELLING EMIGRATION FROM PORTUGAL TO SWITZERLAND AND A WAGE EQUATION FOR EMIGRANTS

Cátia Felisberto

May 2003

Professor: Marius Brülhart
Assistant: Aleksandar Georgiev

Abstract

Many theories have been developed around the reasons that lead an individual to take the decision to emigrate. The purpose of this paper is, in a first stage, to analyse what are the determinants of the decision to emigrate, namely from Portugal to Switzerland. After, we progress to the study of one of the most important factors in that decision, the wage of an individual as emigrant.

We have found that young, married and low educated Portuguese males have greater probability to emigrate to Switzerland. Having an emigrant brother strongly influences people to take the decision to emigrate. Wage in Portugal was not found to be a determinant factor. Concerning the wage equation for the Portuguese emigrants in Switzerland we have arrived to the conclusion that Portuguese males are likely to be better paid than females. Furthermore, emigrants with a higher education level have, in general, higher wages. An interesting point is that professional experience obtained in Portugal has little or no significance. However, the professional experience acquired in Switzerland can increase the wage level.

I. Introduction

Portugal has a long history of emigration. In the last centuries the overseas territories were the main destinations. During the sixties, Europe started to be more attractive, namely France, Switzerland and Germany. In 1999, the total number of Portuguese emigrants summed up 4,7 millions[1], from which 153 thousands were living in Switzerland (approximately 1,5% of the Portuguese population).

The focus of this work, in a first stage, is precisely to analyse what are the main determinants of the decision to emigrate, namely from Portugal to Switzerland. Firstly, we follow the paper by Hoddinott (1994) entitled "A Model of Migration and Remittances Applied to Western Kenya", which uses a *probit model* to estimate the decision to emigrate. In Appendix I, is described and explained the theoretical model in which the estimations, performed in the cited paper, are based in. After, some explanatory variables will be included in the base model in order to achieve better results. In the final model we just keep the significant variables. Secondly, motivated by the results obtained in the estimations of the emigration model, we progress to the estimation of one of the most important factors in the decision to emigrate, the wage of an individual as an emigrant. The procedure used to perform this last estimation (*Heckman procedure*) allows us to combine the emigrant wage estimation with the estimation of the decision to emigrate (here we use the results from the previous section).

Before proceeding, it is useful to do a brief review of the literature concerning the migration models. Typically, migration is modelled as the result from the individual utility maximisation: individuals emigrate when this decision increases their expected utility. One of the most important and primary contributes to understand the motivations of emigration was Ravenstein's (1885, 1889) sixth law which says that individuals emigrate in order to increase their well being. The formalisation of this idea appears in 1969 by the hand of Todaro. After, some authors (Connel et al. (1976)) stressed that most of the times the emigration decision is not taken exclusively by the emigrant[2] himself but by the emigrant and

[1] Source: DGACCP/MNE

[2] Through all the paper the word emigrant is used in his general sense, i.e., it refers to both genders, feminine and masculine.

his family together. This idea was developed and applied by Low (1986), Rempel (1981) and Levhari (1982). Later, Hoddinott (1994) argues that still this is an "oversimplification" because it neglects the role of the potential emigrant and does not allow us to clearly see "how the costs and benefits of emigration are shared". He presents a model, in which the emigration decision results from the utility maximisation of the potential emigrant and the household together, treating the previous models as particular cases of his more general model.

This paper is organised into four main parts. In the next section, the emigration determinants are analysed, namely of the Portuguese emigration to Switzerland. Section III is devoted to the estimation of a wage equation for the Portuguese emigrants in Switzerland. Within each of these three sections the theoretical framework is described briefly and the data, estimation procedures and results are presented. Finally, in section IV the main conclusions are reported.

II. Estimating a Model of Emigration

A Model of Emigration

As already mentioned, the theoretical foundation of the emigration model that will be estimated is Hoddinott's model. In Appendix I, firstly the appropriateness of Hoddinott's model to our framework is discussed and secondly, the model itself is described. In this section, just the main results are presented.

Hoddinott's model concludes that the emigration decision depends on the wage of the individual as an emigrant (w_e), the wage of the individual as a non-emigrant (w_n), the wage of his parents (w^p) and the net value of money transfers made by other descendants to the parents (OT). Formally, we have:

$$E = f_e\,(w_e,\ w_n,\ w^p,\ OT)$$

where E = 1 if the individual emigrates and E = 0 if the individual does not emigrate.

Data Specification

The data that will be used was collected by the authoress, between February and April of 2003. For the emigrants, interviews were done in the Portuguese Consulate at Geneva and questionnaires were sent to Portuguese Associations in Switzerland for the associates to fill in. The data about the Portuguese non-emigrants was collected in small towns around Lisbon, by requesting people to fill in the questionnaires themselves. These questionnaires were delivered with an envelope, which should be sent, with the filled questionnaire inside, to the University of Lausanne.

Because of constraints and difficulties in collecting the data, it was not possible to have both sub-samples constituted by people original from the same village or city. Hence, it is not possible to control for certain factors specific to each village or town. We should also keep in mind that it will be used an urban sample and that emigrants are not a random sub-sample of the population.

The total sample has 190 observations, 85 non-emigrants (45%) and 105 emigrants (55%). The respondents are individuals with more than 18 years and non-students. The sub-sample of non-emigrants includes individuals who have married and constituted their own household, individuals who have emigrated but know are living in Portugal and individuals who still live with their parents.

As shown before, the decision to emigrate (or not) depends on w_e , w_n , w^p and *OT*. The variable w_e was only observed for emigrants, and for *OT* there is no information available. However, w_e is assumed to be a function of the age, the gender, the level of education and the number of hours of work. Concerning OT, we can observe, directly from the dataset, that only 3% of the respondents declare to send remittances for their parents. Therefore, assuming that the other family members have a similar behaviour, it does not seem to be appropriate to included proxies for this variable in the regression. The equation for the decision to emigrate assumes the form:

$$E = f_e\,(age,\ gender,\ level\ of\ education,\ number\ of\ hours\ of\ work,\ w_n,\ w^p)$$

Afterwards, in order to try to improve the model, more explanatory variables which, intuitively, are believed to affect the decision to emigrate, are included in the model. These variables are the followings: marital status (*single*), the age of the household head (*agehhhead*), the average education level of the parents (*ayearsschoolp*), the number of emigrant brothers the individual has (*brothersemigrants*), the per capita income in the household (*pcincome*) and the number of children the prospective emigrant has (*ndependents*). This last model will be called expanded emigration model.

All the variables used in this section are summarized in Table 1 (Appendix II).

Before proceeding it is useful to explain the explanatory variables within the context of the outcomes which are hypothesised by the theory.

Age (*age*) is used here as a proxy for work experience, which influences positively the wage of the individual as an emigrant. Generally, diminishing returns to experience are believed to exist, therefore we will also use a quadratic form of age (*age2*). This variable can also give us information about other features. For instance, age is generally expected to be inversely correlated with emigration, since younger people expect higher lifetime returns from emigration.

The variable gender (*gender*) assumes the value 1 if male and 0 if female. The sign of this variable is difficult to be hypothesised. The effect gender can have on the earnings is uncertain but since we believe that man are more likely to emigrate than woman, a positive sign is expected.

The relationship between the level of education and emigration is also ambiguous. Some authors argue that a positive relationship should exist between emigration and education, since people with more years of school are more likely to obtain better employment and earning opportunities in the destination areas. On the other hand, it can be the case that unskilled workers are better paid in the foreign country, and in this case the less educated they are, the more likely they are to emigrate. The influence of this variable will be kept an open question. We will enter in the model two variables: the number of years the individual has studied (*yearsschool*) and a dummy variable (*graduation*) to test how the decision to

emigrate is related with the fact that the individual has a graduation. This variable takes the value 1 if the individual has a graduation and 0 otherwise.

Obviously, the wage depends positively on the number of hours worked. Thus, a positive relationship between hours of work (*hourswork*) and emigration is expected.

The wage in Portugal (*wagePTperhour*) for the non emigrants is the real wage the individual earns in Portugal and for the emigrant is the wage the individual supposes he would be earning if he was working in Portugal. This variable is supposed to have a negative relationship with emigration, because the higher the wage a person has in the home country the less incentive she has to search for a higher wage, namely in a foreign country.

The parent's wage (*pwage*) can have an ambiguous effect on the decision to emigrate. On the one hand, families with higher income can proportionate better conditions for the sons to emigrate. One the other hand, descendents of families with fewer income are more likely to fell the necessity to go abroad to search for a better life.

The dummy variable for the marital status (*single*) assumes the value 1 if the individual is single and 0 otherwise. We expect a positive relationship between this variable and the decision to emigrate, since the individual who did not yet constituted his own family, more easily takes the decision to emigrate.

For the sake of simplicity, the household head is assumed to be the man when both elements of the couple are alive; otherwise it is the widow(er). In what concerns the effect of the age of the household head (*agehhhead*) some authors defend that migration is in part a life-cycle phenomenon. Some studies (Lipton, 1980; Connel et al., 1976) suggest that more mature families are more likely to produce emigrants. Thus, we expect the coefficient of this variable to be positive.

The parent's average level of education (*ayearsschoolp*), like the parent's wage, can have an ambiguous effect on the decision to emigrate. Households where the parents are less educated are more likely to have less disposable income and hence, more likely to have children looking for a better live abroad. On the other hand households with higher income can provide more conditions for the children to emigrate than households with less income.

Another measure of wealth will also be introduced, the per capita income in the household (*pcincome*). Here the analysis is the same as before.

Regarding the number of dependents the prospective emigrant has (*ndependents*), the effect of this variable is not, *a priori*, easy to guess.

Finally, having brothers who are living in a foreign country (*brothersemigrants*) is thought to increase the probability to take the decision to emigrate, since the individual can suffer the influence of the emigrant brother(s).

Estimation Procedures and Results

As already mentioned three models will be estimated, the first one based in the theoretical model presented before, the second one which includes some more explanatory variables and, finally, a restricted model. The aim of this second estimation is to test if some of the variables which seem to be important in the decision to emigrate, and that are not included in the theoretical model, are indeed relevant.

In these models we do not have a continuous dependent variable, we have a dichotomous dependent variable, and hence a *probit* or a *logit* model should be used. The main difference between both models is the distribution that is assumed. The *probit* model assumes a normal distribution and the *logit* model assumes a logistic distribution. These distributions are different except in the tails; the logistic distribution has heavier tails. To analyse the determinants of emigration (for the three models) a probit *model is used*. Some estimations were also performed with a *logit* model and the results do not differ substantially from the ones obtained with a *probit*.

One of the limitations of the *probit* model (and the same applies for the *logit* model) is that the estimated coefficients do not represent the marginal change of the dependent variable given a unit change in the explanatory variables. The *probit* coefficients can only be used to analyse the signs and the statistical significance of the parameters. Another limitation is that *probit* (and *logit*) models are sensitive to misspecifications. Thus, the estimators will be inconsistent if an explanatory variable is omitted or if there is heteroskedasticity.

Now we turn to the analysis of the estimations results.

In Table 2 (Appendix I) we have the results for the first estimation of the emigration model. In a general way, the results are not very good. The chi-squared statistic, testing the null hypothesis that all regressors are jointly zero, is strongly rejected. However, only the variables gender and education are statistically significant, at a 5% and 1% level, respectively. The dummy variable graduation and the parents' wage are only marginally significant. Concerning the signs of these variables, gender has a positive sign, as it was expected. The level of education has a negative sign which confirms that the more educated an individual is the higher the probability that he stays in the home country. In the same way, the less educated the individual is, more likely he is to decide to emigrate. This reflects the big discrepancy that exists between the wages and work conditions for the low skilled workers in Portugal and Switzerland. This discrepancy tends to diminish as the level of education increases. The positive sign of the variable graduation does not confirm the previous result; however this variable does not seem to be very significant. The parents' wage has a negative sign, which confirms the scenario that individuals whose parents have higher wages are less likely to emigrate. The variables age, age squared, hours of work and wage in Portugal are not significant, which does not give support to the theory.

Since the results with this model are not very good, another model incorporating more explanatory variables was estimated. Concerning the estimation of this expanded model the results do not improve significantly (Table 3, Appendix II). Once more, the hypothesis that all regressors are jointly zero is strongly rejected. Regarding the variables which are common to the first estimate presented, gender now turns to be not significant as well as graduation. Education is again significant (now at a 5% level) and as before has a negative sign. Wage in Portugal (per hour) and the parents' wage are now significant at a 10% level, but we can not reject the hypothesis of this coefficients being equal to zero. Wage in Portugal, has a positive sign, which was not expected. This result can have two meanings. First, that the higher the wage in Portugal the higher the probability that the individual decides to emigrate. Second, since the wage in Portugal for the emigrants corresponds to the expected wage in Portugal, it may just say that the emigrants'

expectations about the level of wages in Portugal are higher than the reality. As before the parents' wage has a negative sign. In relation to the coefficients that were added to the previous model, the only one that is significant is the number of emigrant brothers that the individual has. This variable presents a positive sign and is very significant (at a 1% level). This result is in accordance with our expectations.

Combining both models and discarding the non significant variables we obtain the restricted model shown in Table 4 (Appendix II). Once more the hypothesis that all regressors are jointly zero is strongly rejected. All the variables are significant and regarding the signs just "single" has a non expected sign. Married people are more likely to emigrate than single people, which was not expected. It is worth to analyse this model more deeply.

Before performing some tests, it is interesting to observe the "odds ratio" for each of the variables. To obtain these ratios we have to resort to a logistic model (Table 5, Appendix II). The "odds ratio" is interpreted as a relative probability, i.e., the probability of deciding to emigrate relatively to the probability of deciding not to emigrate. If the ratio is bigger than 1 then the decision to emigrate is more likely then the decision to stay in Portugal. From Table 5 we can observe that the variable which most affects the decision to emigrate is the number of emigrant brothers the individual has. This variable is followed by gender. On the other hand, single and education are the variables with the major effect on the decision to stay in Portugal.

Complementarily, we can analyse the marginal effects of each explanatory variable on the probability of the decision to emigrate. Table 6 (Appendix II) shows us that the higher marginal effect on the predicted probability to emigrate is 0,407 and corresponds to the number of emigrant brothers the individual has. "Single" has the second most important effect on the predicted probability to emigrate (but this time is a negative effect).

To test the capacity our model has to reproduce the reality, Pearson "goodness-of-fit" test is performed. This test compares the empiric distribution function of the dependent variable with the theoretical distribution function obtained through the probit model. As we can observe in Table 7 (Appendix II) we

can not reject the null hypothesis of equality between the empiric and theoretical values. The model has a high predictive power.

In Table 8 (Appendix II) is shown the summary statistics, including the classification table. This table gives us information about the predictive power of the model. We can observe that the model is able to correctly classify 78% of the respondents, what suggest the model to be quite significant[3].

In the Graph 1 (Appendix II) we can observe the ROC curve which confirms the results obtained in the previous test. Since the ROC curve is distant from the 45° line we can say that the model has a good predictive power. Graph 2 shows the evolution of the sensitivity[4] and of the specificity[5] across the probability cut off (which varies from 0 to 1).

From the analyses performed above we can say that the restricted model has some predictive power.

III. Estimating a Wage Equation for the Portuguese Emigrants in Switzerland

A Wage equation

In this part, we describe the wage equation that will be used to estimate the emigrant's wage.

Generally, the individual earnings are assumed to be determined by the level of education and the professional experience the individual has. Unfortunately, sometimes differences in the labour market status of women compared to men still persist. Thus, we also include gender as an explanatory variable.

The individual earnings are defined as gross hourly wage (*wagechperhour*).

---

[3] This test codes (by default) the decision to emigrate to be equal to one if the predicted probability is higher or equal to 0,5.

[4] "Sensitivity" represents the share of the observations, corresponding to the decision to emigrate, correctly predicted by the model.

[5] "Specificity" represents the share of the observations, corresponding to the decision not to emigrate, correctly predicted by the model.

Formally we have:

WageCHperhour = f (level of education, professional experience, gender)

Data Specification

The dataset used in this section is the same presented in section III.

As mentioned before the hourly wage depends on the level of education, professional experience and gender. The variable professional experience was not observed. Hence, as a proxy, we use the individuals' potential years in working life (*profexperience*), defined as age minus years of schooling minus age of school start (generally 6). The number of years living in Switzerland (*yearslivingch*) also gives us a measure of professional experience, or more precisely, a measure of the experience within the Swiss labour market. The dummy variable gender (*gender*) assumes the value 1 if male and 0 if female. Concerning the level of education, besides the number of years the individual has frequented school (*yearsschool*) we will also include in the regression the square of this variable (*yearsschool2*) and a dummy variable (*graduation*) which assumes the value 1 if the individual has a graduation and 0 otherwise.

In terms of the expected signs, we should find a positive relationship between hourly wage and all the explanatory variables.

As we will see, to estimate the emigrant wage equation we will resort to the determinants of the decision to emigrate, which were described and explained in the previous section.

The data used in this section is summarised in Table 9 (Appendix III).

Estimation Procedures and Results

The estimation of wage equations comprises special issues which do not allow us to use the usual techniques, namely Ordinary Least Squares (OLS). In this context, OLS estimates would be biased. Generally, in the estimation of a wage equation we have a problem of censored data, since we will not observe the real wage of the people whose reservation wage is above the wage offered in the market (they will not work, thus the observed wage will be zero). In face of this situation

we should use a *Tobit* model (which combines OLS with *Probit*). Concerning the emigrants wage an additional issue arises: a sample selection problem. The emigrant wage is only observed if the individual takes the decision to emigrate. This means that, to estimate the emigrant wage, two equations have to be estimated: one modelling the decision to emigrate and another one modelling the wage. The bias will arise from the fact that the errors of both equations will be correlated. Suppose emigrants have an unobserved characteristic which is dynamism. This characteristic will appear in the error term of both the emigration equation and the emigrant wage equation, leading to biased parameter estimates in the later. To control for this selectivity bias *Heckman's procedure* should be used. We can either work with Heckman's two-step[6] consistent estimator or with full maximum-likelihood. We will choose the maximum likelihood procedure.

Now we are going to analyse the estimations results. The decision to emigrate is used has the selection variable and the emigration model used here is the last one estimated in the previous section (the restricted model). The hypothesis that all regressors are jointly zero is strongly rejected. As we sow before all the variables in the restricted emigration model are statistically significant. Concerning the wage equation, only the gender and the number of years living in Switzerland are statistically significant, at a 5% and 10% levels, respectively. The fact that the variable professional experience is not significant and the number of years living in Switzerland is, shows that for the emigrant wage it is more important the professional experience obtained in Switzerland than the total number of years of professional experience. Restricting the model we obtain the estimation in Table 11. Now, education turns to be significant. As expected the signs of the variables education ("yearsschool"), gender and the number of years living in Switzerland are positive. This implies that, given the fact that the individual is an emigrant, the

[6] In the first stage the decision equation (in our case the equation which models the decision to emigrate) is estimated using a *Probit* model. The fitted values of the estimated parameters ($\hat{\beta}$) are then used to estimate $X_i \hat{\beta}$ (being X_i the vector of regressors) and the *inverse Mill's ratio*. In the second stage, the *inverse Mill's ratio* is plugged in the wage equation, which is then estimated using OLS.

more educated the individual is and the more professional experience he has in Switzerland, the higher will be his wage as an emigrant. We can also observe that being a male also contributes for a higher wage, for the Portuguese emigrants in Switzerland. Rho, which gives the correlation between the residuals in the selection equation and the residuals in the wage equation, is not significant which means that we can not reject the hypothesis that the decision to emigrate is independent of the wage in Switzerland. In the case where rho = 0 the likelihood function can be split into two parts: a *probit* for the probability of being selected and an OLS regression for the expected value of wage in the selected sub-sample.

In order to check if there are significant changes in the results we have also estimated the model using Heckman two-steps procedure (Table 12, Appendix III) and using the option robust (to estimate robust variance estimates) (Table 13, Appendix III). Indeed, the results do not change significantly.

IV. Conclusions

Many theories have been developed around the reasons that lead an individual to take the decision to emigrate. In this paper we try to apply, to a dataset composed by Portuguese emigrants in Switzerland and residents in Portugal, one of the most recent models of migration: Hoddinott's model. This model is based on the assumption that the decision to emigrate is the result of the joint utility maximization of the household (prospective emigrant and his parents). Since results of this estimation are a little bit disappointing we try the estimation of other models. The model that seems to perform better is the one where the decision to emigrate depends on the following characteristics of the individual: age, gender, single, education he has and the number of emigrant brothers the individual has. We have found that young, married and low educated Portuguese males have greater probability to emigrate to Switzerland. Having an emigrant brother strongly influences people to take the decision to emigrate. Wage in Portugal was not found to be a determinant factor.

Secondly, we try to estimate a wage equation for the Portuguese emigrants in Switzerland. We arrive to the conclusion that Portuguese males are likely to be better paid than females. Furthermore, emigrants with a higher education level have, in general, higher wages. An interesting point is that professional experience obtained in Portugal has little or no significance. However, the professional experience acquired in Switzerland can increase the wage level.

It would be interesting to extend this analysis to a sample of individuals coming from the same region or city in order to control for certain factors specific to which village or town. Another interesting improvement would be to include in the questionnaires questions concerning the risk aversion of the respondents in order to analyse the importance of this factor in the decision to emigrate.

REFERENCES

Adams, Jr., 1993, "The Economic and Demographic Determinants of International Migration in Rural Egypt", Journal of Development Studies, Vol. 30, N°1

Agrawal, Reena, 1999, "Are International Remittances Altruism or Insurance? Evidence from Guyana Using Multiple-Migrant Households", Journal of Economic Literature

Bauer, T., Pereira P., Vogler, M. and Zimmermann, K., 1998, "Portuguese Migrants in the German Labor Market: Performance and Self-selection", IZA

Green, William H., 2003, *Econometric Analysis*, 5th edition, Prentice Hall

Heckman, James, 1979, "Sample Selection Bias as a Specification Error", Econometrica, Vol. 47, Issue 1, 153-162

Hoddinott, John, 1994, "A Model of Migration and Remittances Applied to Western Kenya", Oxford Economic Papers, Vol. 46, Issue 3, 459-476

Kennedy, Peter, 1992, *A Guide to Econometrics*, Basil Blackwell Ltd

Maddala, G.S., 2001, *Introduction to Econometrics*, 3rd edition, Wiley

APPENDIX I

Firstly, we discuss the appropriateness of the Hoddinott's model to our framework and secondly, the model itself is presented.

Hoddinott presents three justifications for the use of the joint utility maximisation. Two of them are only applicable for a rural sample, which is not necessarily the present case. They are related with the fact that the agreement allows the farming households to have another kind of revenues (namely the remittances the emigrant may send) that do not depend on their main activity, agriculture, which is very uncertain and with the possibility of suppressing "imperfections in rural capital markets" (Hoddinot, 1994). The other justification is more general and may contain financial and personal motivations (like altruistic reasons): "agreement yields substantial benefits to both parties" (Hoddinot, 1994). Even if the emigrant does not have a financial benefit, the psychological support at the beginning of live abroad and the undoubted support in case of failure, are already a benefit. Of course for households members that are not so united this statement looses importance. Other reasons are presented and explained by Hoddinott but since they are specific to a rural environment it is not worthy to describe them here[7].

Undoubtedly Hoddinott's model is the more accurate one. However, when we analyse the data collected for Portugal we observe that only 3% of the respondents send remittances for the parents and nobody receives remittances from the parents, which can lead us to conclude that the joint utility maximization is not present. Nevertheless, even if money transfers between parents and sons/daughters are not observed, the parent's income is expected to play a role in the emigration decision of the descendants. For instance, individuals whose parents are relatively wealthier are less likely to emigrate. For this reason and for the one mentioned before, Hoddinott's model will still be used. In order to make it more representative of the Portuguese reality nowadays, the reward function will not be considered. The next lines present the Hoddinott's model.

[7] For details see Hoddinott, 1994.

The model considers two agents, the potential emigrant and his parents. It is assumed that the agents obtain utility from consumption of goods and leisure and that together they maximise a single utility function. Whether explicitly or implicitly, the decision of emigrating or not is taken jointly by the potential emigrant and his parents[8]. Formally, the global utility function depends on the utility both potential emigrant (*i*) and parents (*p*) take from the consumption of a composite commodity (*c*) and from leisure (*l*):

$$U_Z^X = U_Z^X\ (c_Z^X, l_Z^X) \tag{1}$$

$$= \left[U_e^i - U_n^i\right]^{\alpha x} . \left[U_e^p - U_n^p\right]^{\alpha p} \tag{2}$$

where $X = i, p$ and $Z = e, n$

The utility functions of each agent are assumed to be strictly quasi-concave and they consider the case where the individual emigrates (*e*) and the case where the individual decides not to emigrate (*n*). The weights attached to the individual's utility and to the parents' utility are, respectively, denoted by αi and αp and they sum up to one. For instance, if $\alpha i = 0$ and $\alpha p = 1$, then the parental utility function is the only one that is maximised. On the other hand, if $\alpha i = 1$ and $\alpha p = 0$, then only the individual utility function is maximised.

The budget constraint of he individual, as an emigrant, is assumed to be:

$$w_e . T_e^i + R_e^{pi} = w_e . l_e^i + P_S . c_e^i + R_e^{ip} \tag{3}$$

where w_e is the wage as emigrant, T_e^i is the total time available for the emigrant, P_S is a price index of goods consumed in Switzerland, R_e^{pi} denotes the value of transfers the individual, as an emigrant, receives from his parents and R_e^{ip}represents the value of transfers the individual, as an emigrant, makes for his parents.

If the individual does not emigrate his budget constraint assumes the form:

$$w_n . T_n^i + R_n^{pi} = w_n . l_n^i + P_P . c_n^i + R_n^{ip} \tag{4}$$

[8] This is the so-called "migration contract" (Lucas and Stark, 1985).

where w_n is the supposed wage as a non-emigrant and P_P is the price index in Portugal. The remaining notation as the same meaning as before but now for the case where the individual does not emigrate.

The parental budget constraint when the individual decides to emigrate is the following:

$$w^p.T^p + R_e^{ip} + OT = w^p.l_e^p + P_P.c_e^p + R_e^{pi} \quad (5)$$

where OT is the net value of money transfers made by other descendants to the parents, w^p is the wage of the parents and T^p denotes the total time available for the parents.

When the individual decides to stay in his country the parental budget constraint is:

$$w^p.T^p + R_n^{ip} + OT = w^p.l_n^p + P_P.c_n^p + R_n^{pi} \quad (6)$$

It is known that the level of prices in Switzerland is much higher then the level of prices in Portugal. But since the results do not change if we assume that prices are the same in both locations (Hoddinott, 1994), for the sake of simplicity we will assume $P_P = P_S = P$.

The full budget constraint of the individual is:

$$w_e.T_e^i + w_n.T_n^i + R_e^{pi} + R_n^{pi} = w_e.l_e^i + w_n.l_n^i + P.c_e^i + P.c_n^i + R_e^{ip} + R_n^{ip} \quad (7)$$

assuming $T_e^i = h_e^i + l_e^i$, $T_n^i = h_n^i + l_n^i$, being h_e^i the supply of labour if the individual emigrates and h_n^i the supply of labour if the individual does not emigrate, and denoting $g^i = c_e^i + c_n^i$, equation (7) becomes,

$$\begin{aligned} w_e.(h_e^i + l_e^i) + w_n.(h_n^i + l_n^i) + R_e^{pi} + R_n^{pi} = \\ = w_e.l_e^i + w_n.l_n^i + P.g^i + R_e^{ip} + R_n^{ip} \end{aligned} \quad (8)$$

The full parental budget constraint is:

$$w^p.T^p + R_e^{ip} + R_n^{ip} + OT = w^p.l_e^p + w^p.l_n^p + P.c_e^p + P.c_n^p + R_e^{pi} + R_n^{pi} \quad (9)$$

Similarly, denoting $g^p = c_e^p + c_n^p$ and $l^p = l_e^p + l_n^p$ the last equation becomes,

$$w^p.T^p + R_e^{ip} + R_n^{ip} + OT = w^p.l^p + P.g^p + R_e^{pi} + R_n^{pi} \quad (10)$$

Finally, to obtain the full household budget constraint it is necessary to combine equation (8) and (10):

$$w_e.(h_e^i + l_e^i) + w_n.(h_n^i + l_n^i) + w^p.T^p + OT = w_e.l_e^i + w_n.l_n^i + P.g^i + w^p.l^p + P.g^p \quad (11)$$

Maximising (2) subject to (11) gives us the demand function for goods and leisure for the individual and for the parents. Recalling that $T^i = h_e^i + h_n^i + l_e^i + l_n^i$ $\Leftrightarrow$ $h_e^i + h_n^i = T^i - l_e^i - l_n^i$ is possible to obtain the individual supply of labor as an emigrant and as a non emigrant:

$$h_e = h\ (P,\ w_e,\ w_n,\ w^p,\ OT) \quad (12)$$

$$h_n = h\ (P,\ w_e,\ w_n,\ w^p,\ OT) \quad (13)$$

If the individual decides to emigrate $h_e = 1$ and $h_n = 0$, and if he decides not to emigrate $h_e = 0$ and $h_n = 1$. Hence, from equation (12) and (13) together:

$$E = f_e\ (w_e,\ w_n,\ w^p,\ OT) \quad (14)$$

where E = 1 if $h_e = 1$ (the individual emigrates) and E = 0 if $h_n = 1$ (the individual does not emigrate).

APPENDIX II

Table 1: Summary of the variables used in the estimation of the decision to emigrate

Variable	Obs	Mean	Std. Dev.	Min	Max
emigrant	191	0.586387	0.493775	0	1
age	191	34.11518	8.797567	18	58
age2	191	1240.838	624.8176	324	3364
gender	191	0.528796	0.500482	0	1
yearsschoc	191	9.931937	4.637442	4	22
graduation	191	0.230367	0.422174	0	1
hourswork	189	41.89153	6.514928	10	70
wageptperh	189	21.01058	28.34457	0	356
pwage	186	1550.704	2044.513	0	12000
single	191	0.350785	0.47847	0	1
agehhheac	176	222.9489	855.9114	38	7000
ayearsschc	187	157.5882	841.194	0	7000
pcincome	191	435.6894	703.9075	0	4000
ndependen	191	0.874346	1.013032	0	4
nbrotherse	191	0.680628	1.352421	0	8

Table 2: First estimation of the emigration model

```
Iteration 0:   log likelihood = -123.58679
Iteration 1:   log likelihood = -96.295266
Iteration 2:   log likelihood =   -95.4636
Iteration 3:   log likelihood = -95.456519
Iteration 4:   log likelihood = -95.456514

Probit estimates                  Number of obs   =        184
LR chi2(8)      =      56.26
Prob > chi2     =     0.0000
Log likelihood = -95.456514       Pseudo R2       =     0.2276

------------------------------------------------------------------------------
    emigrant |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0119189   .097953     0.12   0.903    -.180064    .203902
        age2 |  -.0005218   .001411    -0.37   0.711    -.003287    .002243
      gender |   .5127331   .231387     2.22   0.027     .059222    .966244
  yearsschoc |  -.1806509   .047317    -3.82   0.000    -.27339    -.087911
  graduation |   .7481404   .451415     1.66   0.097    -.136616    1.632897
  hourswork  |  -.0009949   .017091    -0.06   0.954    -.034493    .032503
  wageptperh |  -.0037054   .005141    -0.72   0.471    -.013782    .006371
       pwage |  -.0001316   7.63E-05   -1.72   0.085    -.000281    .000018
       _cons |   2.226916   1.835017     1.21   0.225    -1.369652   5.823484
------------------------------------------------------------------------------
```

Table 3: Estimation of the expanded emigration model

```
Iteration 0:   log likelihood = -116.74769
Iteration 1:   log likelihood = -75.722074
Iteration 2:   log likelihood = -65.729072
Iteration 3:   log likelihood = -59.761535
Iteration 4:   log likelihood = -57.915002
Iteration 5:   log likelihood = -56.903483
Iteration 6:   log likelihood = -56.759517
Iteration 7:   log likelihood = -56.741479
Iteration 8:   log likelihood = -56.739818
Iteration 9:   log likelihood = -56.739782

Logit estimates                 Number of obs   =        173
LR chi2(14)     =     120.96
Prob > chi2     =     0.0000
Log likelihood = -56.269403     Pseudo R2       =     0.5180

------------------------------------------------------------------------------
    emigrant |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0634221   .141914     0.45   0.655    -.214725    .341569
        age2 |  -.0014257   .001906    -0.75   0.454    -.005161    .002309
      gender |   .4739472   .323978     1.46   0.143    -.161038    1.108933
  yearsschoo |  -.1247322   .062342       -2   0.045    -.246921   -.002543
  graduation |   -.47054    .622377    -0.76    0.45    -1.690377   .749297
  hourswork  |   .0077958   .020571     0.38   0.705    -.032522    .048114
  wageptperh |     .0089    .004677      1.9   0.057    -.000266    .018067
       pwage |  -.0006743   .000358    -1.88    0.06    -.001376    2.75E-05
      single |  -.3777614   .397998    -0.95   0.343    -1.157823      .4023
  agehhheac  |   .0235689   .024961     0.94   0.345    -.025353    .072491
  ayearsscho |  -.0181765   .023969    -0.76   0.448    -.065156    .028803
    pcincome |   .0013847   .000936     1.48   0.139    -.000449    .003218
  ndependen  |   .003925    .196458     0.02   0.984    -.381126    .388976
  nbrotherse |    1.94048   .406709     4.77       0     1.143346    2.737614
       _cons |  -1.000824   2.514448    -0.4   0.691    -5.929051    3.927403
------------------------------------------------------------------------------

note: 0 failures and 19 successes completely determined.
```

Table 4: Estimation of the restricted emigration model

```
Iteration 0:   log likelihood = -129.52597
Iteration 1:   log likelihood = -93.306737
Iteration 2:   log likelihood = -85.577001
Iteration 3:   log likelihood = -81.097463
Iteration 4:   log likelihood = -80.268823
Iteration 5:   log likelihood = -80.244867
Iteration 6:   log likelihood = -80.244842

Probit estimates                  Number of obs   =        191
LR chi2(5)      =      98.56
Prob > chi2     =     0.0000
Log likelihood = -80.244842       Pseudo R2       =     0.3805

-------------  -  ------------  -------------  --------  --------  --------------  ---------
emigrant       |        Coef.   Std. Err.       z       P>|z|     [95% Conf. Interval]
-------------  +  ------------  -------------  --------  --------  --------------  ---------
age            |     -0.0400191   0.016443    -2.43     0.015    -0.072247   -0.007791
gender         |      0.4599355   0.235151     1.96      0.05    -0.000951    0.920822
yearsschoc     |     -0.1138634    0.02808    -4.05         0    -0.168899   -0.058828
single         |      -0.570851   0.290329    -1.97     0.049    -1.139886   -0.001816
nbrotherse     |      1.618568   0.338278     4.78         0     0.955555    2.281581
_cons          |       2.31258   0.763709     3.03     0.002     0.815738    3.809423
-------------  -  ------------  -------------  --------  --------  --------------  ---------

note: 0 failures and 12 successes completely determined.
```

Table 5: Estimation of the restricted model with logistic

Logit estimates Number of obs = 191
LR chi2(5) = 98.76
Prob > chi2 = 0.0000
Log likelihood = -80.147221 Pseudo R2 = 0.3812

emigrant	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	0.9375665	0.027236	-2.22	0.026	0.885676	0.992498
gender	2.219726	0.885363	2	0.046	1.015752	4.850771
yearsschoc	0.8298909	0.042569	-3.64	0	0.750514	0.917663
single	0.4009265	0.201362	-1.82	0.069	0.149815	1.072936
nbrotherse	19.69838	14.86211	3.95	0	4.489651	86.42681

note: 0 failures and 2 successes completely determined.

Table 6: Marginal effects after probit estimation

Marginal effects after probit
y = Pr(emigrant) (predict)
= .83173501

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
age	-0.0100605	0.00441	-2.28	0.023	-0.018708	-0.001413	34.1152
gender*	0.1169858	0.06653	1.76	0.079	-0.013411	0.247382	0.528796
yearss~l	-0.0286244	0.00878	-3.26	0.001	-0.045829	-0.01142	9.93194
single*	-0.1546862	0.08291	-1.87	0.062	-0.317193	0.007821	0.350785
nbrot~ts	0.4068951	0.04582	8.88	0	0.317088	0.496702	0.680628

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 7: Goodness-of-fit test

Probit model for emigrant, goodness-of-fit test

number of observations = 191
number of covariate patterns = 173
Pearson chi2(167) = 152.39
Prob > chi2 = 0.7845

Table 8: Summary statistics including classification table

Probit model for emigrant

	-------- True --------		
Classified	D	~D	Total
+	89	19	108
-	23	60	83
Total	112	79	191

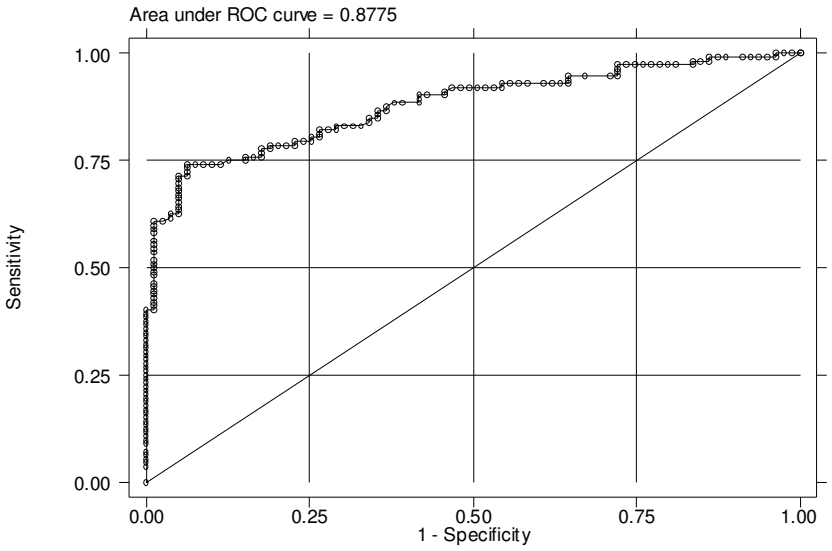
Classified + if predicted Pr(D) >= .5
True D defined as emigrant ~= 0

Sensitivity	Pr(+| D)	79.46%
Specificity	Pr(-|~D)	75.95%
Positive predictive value	Pr(D| +)	82.41%
Negative predictive value	Pr(~D| -)	72.29%
False + rate for true ~D	Pr(+|~D)	24.05%
False - rate for true D	Pr(-| D)	20.54%
False + rate for classified +	Pr(~D| +)	17.59%
False - rate for classified -	Pr(D| -)	27.71%
Correctly classified		78.01%

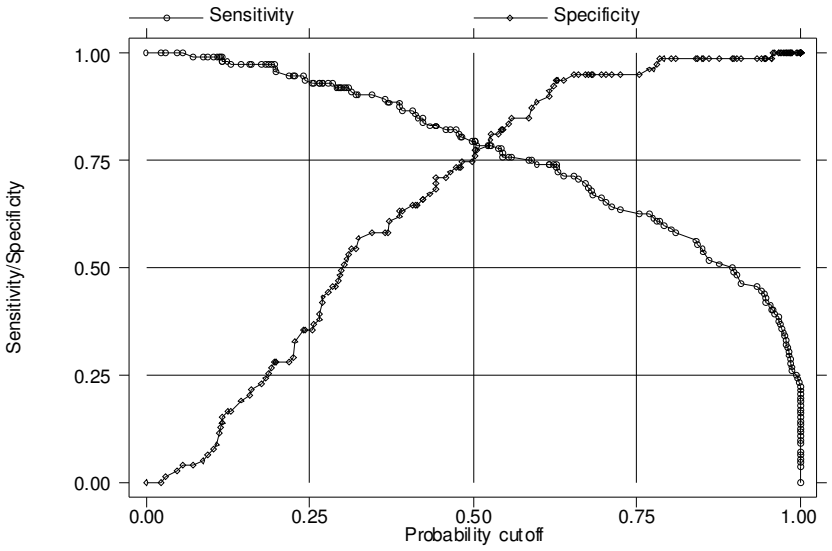
Graph 1: ROC curve

Probit model for emigrant

number of observations = 191
area under ROC curve = 0.8775



Graph 2: Sensitivity and specificity vs. probability cut off



APPENDIX III

Table 9: Summary of the variables used in the estimation of the emigrant wage

Variable	Obs	Mean	Std. Dev.	Min	Max
wagechperh~r	112	68.69643	25.54764	0	172
yearsschool	191	9.931937	4.637442	4	22
yearsschool2	191	120.0366	102.0565	16	484
graduation	191	0.230367	0.422174	0	1
gender	191	0.528796	0.500482	0	1
yearslivin~h	191	7.424084	8.449431	0	30
profexperi~e	191	18.18325	11.23542	-2	48
emigrant	191	0.586387	0.493775	0	1
age	191	34.11518	8.797567	18	58
gender	191	0.528796	0.500482	0	1
yearsschool	191	9.931937	4.637442	4	22
single	191	0.350785	0.47847	0	1
nbrotherse~s	191	0.680628	1.352421	0	8

Table 10: First estimation of the emigrants' wage equation

```
Iteration 0:   log likelihood = -592.32368
Iteration 1:   log likelihood =   -592.321
Iteration 2:   log likelihood =   -592.321

Heckman selection model                         Number of obs     =        191
(regression model with sample selection)        Censored obs      =         79
                                                Uncensored obs    =        112
                                                Wald chi2(6)      =      19.72
Log likelihood =  -592.321                      Prob > chi2       =     0.0031

-------------+----------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
wagechper    |
yearsschoc   |   5.026741   4.680293     1.07   0.283    -4.146464    14.19995
yearsschoc   |  -0.189484    0.27775    -0.68   0.495    -0.733864    0.354897
graduation   |   11.77092   24.14301     0.49   0.626    -35.54851    59.09035
gender       |   10.15266   4.775013     2.13   0.033     0.793805    19.51151
yearslivin~  |   0.831418   0.434218     1.91   0.056    -0.019634    1.682469
profexperi~  |   0.376383   0.432844     0.87   0.385    -0.471975    1.224741
_cons        |   16.96087   22.31603     0.76   0.447    -26.77776    60.69949
-------------+----------------------------------------------------------------
emigrant     |
age          |  -0.039238   0.016854    -2.33   0.02    -0.072271   -0.006204
gender       |   0.456513   0.235518     1.94   0.053    -0.005094    0.918119
yearsschoc   |  -0.113754   0.028055    -4.05      0     -0.16874   -0.058768
single       |  -0.551059   0.305926     -1.8   0.072    -1.150663    0.048545
nbrotherse   |   1.615902    0.33833     4.78      0     0.952787    2.279016
_cons        |   2.278671   0.779485     2.92   0.003     0.750908    3.806433
-------------+----------------------------------------------------------------
/athrho      |   0.056604   0.276718      0.2   0.838    -0.485754    0.598962
/lnsigma     |   3.153707   0.066945    47.11      0     3.022496    3.284917
-------------+----------------------------------------------------------------
rho          |   0.056544   0.275834                     -0.45084     0.53631
sigma        |   23.42273   1.568042                     20.54251    26.70677
lambda       |   1.324404   6.467031                    -11.35074    13.99955
-------------+----------------------------------------------------------------
LR test of indep. eqns. (rho = 0):   chi2(1) =     0.04   Prob > chi2 = 0.8352
------------------------------------------------------------------------------
```

Table 11: Estimation of the restricted emigrants' wage equation

```
Iteration 0:   log likelihood = -592.84733
Iteration 1:   log likelihood = -592.83566
Iteration 2:   log likelihood = -592.83566

Heckman selection model                         Number of obs     =        191
(regression model with sample selection)        Censored obs      =         79
                                                Uncensored obs    =        112

                                                Wald chi2(3)      =      18.49
Log likelihood = -592.8357                      Prob > chi2       =     0.0003

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
wagechper    |
yearsschoc   |    1.20051   0.708357     1.69   0.090    -0.187844    2.588863
gender       |   10.45863   4.730255     2.21   0.027     1.187503    19.72976
yearslivin~  |    1.09782   0.320107     3.43   0.001     0.470421    1.725218
_cons        |   37.08976   8.133014     4.56   0.000     21.14935    53.03018
-------------+----------------------------------------------------------------
emigrant     |
age          |  -0.038618   0.016649    -2.32   0.020    -0.071249   -0.005987
gender       |   0.454102   0.234837     1.93   0.053     -0.00617    0.914374
yearsschoc   |  -0.113673   0.027985    -4.06   0.000    -0.168524   -0.058823
single       |  -0.526396   0.303374    -1.74   0.083    -1.120998    0.068207
nbrotherse   |   1.610832   0.338332     4.76   0.000     0.947714     2.27395
_cons        |   2.248356   0.771133     2.92   0.004     0.736963    3.759749
-------------+----------------------------------------------------------------
/athrho      |   0.129548   0.266481     0.49   0.627    -0.392745     0.65184
/lnsigma     |   3.160482   0.067343    46.93   0.000     3.028492    3.292472
-------------+----------------------------------------------------------------
rho          |   0.128828   0.262058                     -0.373724    0.572907
sigma        |   23.58196   1.588082                      20.66605     26.9093
lambda       |   3.038012   6.209532                     -9.132447    15.20847
-------------+----------------------------------------------------------------
LR test of indep. eqns. (rho = 0):   chi2(1) =     0.25   Prob > chi2 = 0.6162
------------------------------------------------------------------------------
```

Table 12: Estimation of the restricted emigrants' wage equation, using Heckman two steps procedure

Heckman selection model -- two-step estimat
(regression model with sample selection)

Number of obs = 191
Censored obs = 79
Uncensored obs = 112

Wald chi2(5) = 42.35
Prob > chi2 = 0.0000

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
wagechper						
yearsschoc	1.204969	0.69831	1.73	0.084	-0.163694	2.573632
gender	10.45386	4.726877	2.21	0.027	1.189356	19.71837
yearslivin~	1.097202	0.320141	3.43	0.001	0.469738	1.724667
_cons	37.09245	8.133233	4.56	0	21.15161	53.0333
emigrant						
age	-0.040019	0.016443	-2.43	0.015	-0.072247	-0.007791
gender	0.459936	0.235151	1.96	0.05	-0.000951	0.920822
yearsschoc	-0.113863	0.02808	-4.05	0	-0.168899	-0.058828
single	-0.570851	0.290329	-1.97	0.049	-1.139886	-0.001816
nbrotherse	1.618568	0.338278	4.78	0	0.955555	2.281581
_cons	2.31258	0.763709	3.03	0.002	0.815738	3.809423
mills						
lambda	2.967881	5.920658	0.5	0.616	-8.636395	14.57216
rho	0.12586					
sigma	23.58038					
lambda	2.967881	5.920658				

Table 13: Estimation of the restricted emigrants' wage equation, using robust option

```
Iteration 0:   log likelihood = -592.84733
Iteration 1:   log likelihood = -592.83566
Iteration 2:   log likelihood = -592.83566

Heckman selection model                         Number of obs     =        191
(regression model with sample selection)        Censored obs      =         79
                                                Uncensored obs    =        112

                                                Wald chi2(3)      =      24.08
Log likelihood = -592.8357                      Prob > chi2       =     0.0000

------------------------------------------------------------------------------
             |                 Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
wagechper    |
yearsschoc   |     1.20051   0.6563063     1.83   0.067    -0.085827    2.486846
gender       |     10.45863    4.379486     2.39   0.017     1.874998    19.04227
yearslivin~  |     1.09782   0.3004299     3.65   0.000     0.508988    1.686652
_cons        |     37.08976   7.071012     5.25   0.000     23.23084    50.94869
-------------+----------------------------------------------------------------
emigrant     |
age          |    -0.038618   0.0179981    -2.15   0.032    -0.073893   -0.003342
gender       |     0.454102   0.2285244     1.99   0.047     0.006202    0.902001
yearsschoc   |    -0.113673   0.0316287    -3.59   0.000    -0.175664   -0.051682
single       |    -0.526396   0.3071454    -1.71   0.087     -1.12839    0.075598
nbrotherse   |     1.610832   0.4337647     3.71   0.000     0.760669    2.460995
_cons        |     2.248356   0.8540494     2.63   0.008      0.57445    3.922262
-------------+----------------------------------------------------------------
/athrho      |     0.129548   0.2256066     0.57   0.566    -0.312633    0.571729
/lnsigma     |     3.160482    0.111769    28.28   0.000     2.941419    3.379545
-------------+----------------------------------------------------------------
rho          |     0.128828   0.2218623                     -0.302831    0.516628
sigma        |     23.58196   2.635732                      18.94271    29.35742
lambda       |     3.038012   5.347057                     -7.442028    13.51805
-------------+----------------------------------------------------------------
Wald test of indep. eqns. (rho = 0): chi2(1) =     0.33   Prob > chi2 = 0.5658
------------------------------------------------------------------------------
```