

# Intra-household Selection into Migration : Evidence from a Matched Sample of Migrants and Origin Households in Senegal

Isabelle Chort (Université Paris Dauphine, DIAL)  
Jean-Noël Senne (PSE, DIAL)

*MAFE Conference*  
12-14 december 2012

# Motivation

- ▶ Who migrates? Migrant selection is most often modelled as resulting from an individual decision
  - ▶ Based on wage differentials between origin and destination countries
- ▶ But migration has been studied as a household strategy, especially in developing countries
  - ▶ Importance of remittances in the resources of the sending household
- ▶ Is migrant selection based on a household decision different from individual self-selection?
  - ▶ Portfolio allocation : potential trade-off between migrants' expected gains at destination and the share of it that will come to the household through remittances

# Objectives of the paper

- ▶ Extend the Roy model of selection to a household based migration decision (similar to portfolio allocation)
  - ▶ Roy (1951), Borjas (1987), extended to multiple locations by Dahl (2002), all consistent with individual location decisions (Harris and Todaro, 1970; Sjaastad, 1962)
- ▶ Address the issue of intra-household selection : who migrates within the household ?
  - ▶ In line with the New Economics of Migration : mobility choices as household decisions (Stark and Bloom, 1985)
- ▶ Empirical application using survey data on a matched sample of migrants and their origin households (MIDDAS)
  - ▶ Counterfactual income of migrants, *had they not migrated*, based on characteristics of non-migrant members of migrant households : account for unobservable characteristics at the household level

## A household selection model for migration

- ▶ Intuition : the household is a portfolio of members, allocated depending on their return (expected income) and risk (of not remitting) at each location
- ▶ Household level additively separable utility function : with  $I$  potential migrant members within the household and  $J$  possible locations, the random utility of household  $h$  having a migrant  $i$  in country  $j$  writes :

$$U_{hij} = \alpha y_h + \beta(y_{ij} - r_{ij}) + \gamma r_{ij} + z'_{hi} \delta_j + \epsilon_{hij} \quad (1)$$

- ▶  $y_h$  : total log home earnings of remaining household members
  - ▶  $y_{ij}$  : log earnings of migrant  $i$  in country  $j$
  - ▶  $r_{ij}$  : log amount of remittances to household  $h$  of migrant  $i$  in country  $j$
  - ▶  $z_{hi}$  : vector of individual and/or household characteristics
- ▶ Which component mostly drives allocation decision ? Estimation of weights  $\alpha$ ,  $\beta$  and  $\gamma$

## Allocation choice

- ▶ Household chooses among  $I \times J$  alternatives the geographical allocation of its members that maximizes its collective utility
- ▶ Household  $h$  decides to locate member  $i$  in country  $j$  if :

$$U_{hij} = \max(U_{h11}, \dots, U_{h1J}, \dots, U_{hI1}, \dots, U_{hIJ})$$

- ▶ Household level selection equations :

$$M_{hij} = \begin{cases} 1 & \text{if } U_{hij} > U_{hkl} \\ 0 & \text{otherwise} \end{cases} \quad \forall (k, l) \neq (i, j)$$

- ▶ Interpretation :  $M_{hij}$  equals one if alternative  $\{ij\}$  is chosen and observed

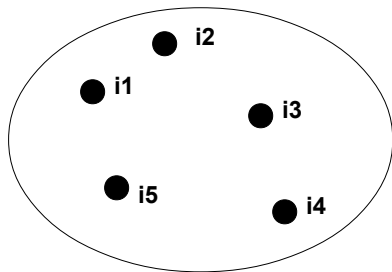
# Estimation

- ▶ Objective : estimate the the set of structural parameters  $(\alpha, \beta, \gamma, \delta_j)$  in equation 1 (different weights on incomes and remittances in the household utility function)
- ▶ Identification issues :
  - ▶ Earnings and remittances are only observed at one location for each household member
  - ▶ Households choosing a specific utility-maximizing geographical allocation are not a random subsample of the population
- ▶ Earnings and remittances in other locations must be imputed, accounting for observed and unobserved individual and household characteristics driving location choices

# Econometric challenges

- ▶ Households are not of equal size : the pool of potential migrant members is not the same in all households
  - ▶ Number of allocation choices  $I \times J$  varies across households
- ▶ Estimate a conditional logit model with a varying number of alternatives (clogitVNA) across observations
  - ▶ Marketing models : Berry, Levinsohn, and Pakes (2004) ; Allenby and Rossi (1998).
  - ▶ Elections : Yamamoto (2012)

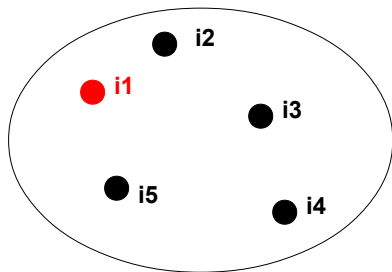
## Household selection : why a varying number of alternatives ?



- ▶ 5 possible allocation choices

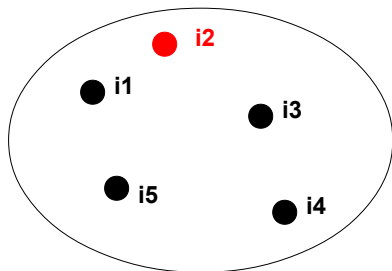


## Household selection : why a varying number of alternatives ?



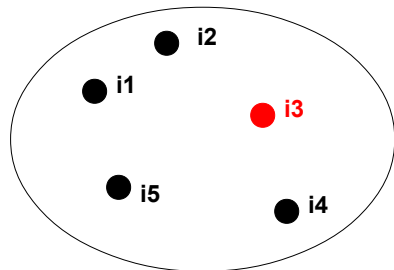
- ▶ **Alternative 1** : {i1 migrates, i2, i3, i4, i5 stay in Senegal}

## Household selection : why a varying number of alternatives ?



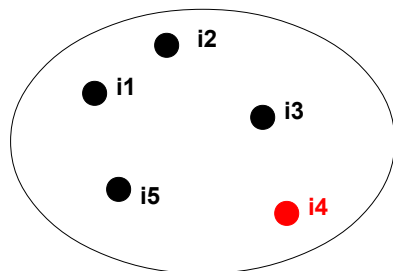
- ▶ Alternative 1 : {i1 migrates, i2, i3, i4, i5 stay in Senegal}
- ▶ **Alternative 2** : {i2 migrates, i1, i3, i4, i5 stay in Senegal}

## Household selection : why a varying number of alternatives ?



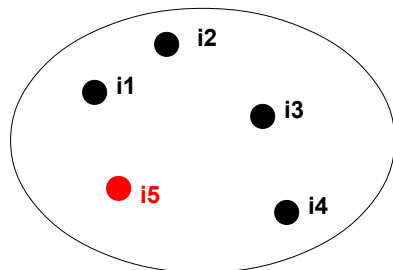
- ▶ Alternative 1 : {i1 migrates, i2, i3, i4, i5 stay in Senegal}
- ▶ Alternative 2 : {i2 migrates, i1, i3, i4, i5 stay in Senegal}
- ▶ **Alternative 3** : {i3 migrates, i1, i2, i4, i5 stay in Senegal}

## Household selection : why a varying number of alternatives ?



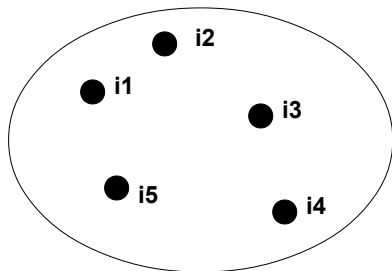
- ▶ Alternative 1 : {i1 migrates, i2, i3, i4, i5 stay in Senegal}
- ▶ Alternative 2 : {i2 migrates, i1, i3, i4, i5 stay in Senegal}
- ▶ Alternative 3 : {i3 migrates, i1, i2, i4, i5 stay in Senegal}
- ▶ **Alternative 4** : {i4 migrates, i1, i2, i3, i5 stay in Senegal}

## Household selection : why a varying number of alternatives ?



- ▶ Alternative 1 : {i1 migrates, i2, i3, i4, i5 stay in Senegal}
- ▶ Alternative 2 : {i2 migrates, i1, i3, i4, i5 stay in Senegal}
- ▶ Alternative 3 : {i3 migrates, i1, i2, i4, i5 stay in Senegal}
- ▶ Alternative 4 : {i4 migrates, i1, i2, i3, i5 stay in Senegal}
- ▶ **Alternative 5** : {i5 migrates, i1, i2, i3, i4 stay in Senegal}

## Household selection : why a varying number of alternatives ?



- ▶ 5 household members : 5 possible allocation choices
- ▶  $N$  household members :  $N$  possible allocation choices
- ▶ The number of alternatives varies across households

## 3 steps

1. Estimate a conditional logit model with a varying number of alternatives (clogitVNA) for selection into locations at the household level
  - ▶ Reduced form selection equation
2. Use results from step 1 to correct for endogenous selection in individual earnings and remittances equations and compute unconditional average earnings and remittance predictions, for each individual at each possible location.
  - ▶ Individual Mincer-type earnings and remittances equations
  - ▶ Instruments : precipitations interacted with individual characteristics (gender, age, education)
  - ▶ Household home earnings = sum of individual earnings of remaining working members
3. Using results from step 2, estimate a structural conditional logit model to recover the set of parameters
  - ▶ Identification relies on instruments affecting earnings/remittances but not location choices.

# Data

- ▶ MIDDAS : Surveys among Senegalese migrants in France, Italy, Mauritania, Côte d'Ivoire and their origin household in Senegal.
- ▶ Over 300 migrants surveyed in each destination country in 2009 and 2010
  - ▶ Use of census data to approach representativeness of migrant samples
- ▶ Origin households tracking in Senegal in 2009 and 2010 (except for Côte d'Ivoire)
  - ▶ 35,3% matching rate
  - ▶ 326 migrant-household pairs
- ▶ Rainfall data taken from the Climatic Research Unit of University of East Anglia
  - ▶ Gridded dataset : definition 0.5 degree
- ▶ Sample representativeness is assessed at three different stages



## Descriptive statistics : Individual characteristics by location

	Senegal (1)	France (2)	Italy (3)	Mauritania (4)
Age	33.9	36.9	35.9	35.8
Gender (%)				
... <i>Male</i>	42.6	75.7	77.2	63.6
... <i>Female</i>	57.4	24.3	22.8	36.4
Schooling level (%)				
... <i>No schooling</i>	40.4	18.0	10.9	40.6
... <i>Primary</i>	28.4	18.0	12.3	29.4
... <i>Middle School</i>	14.2	16.3	25.2	16.7
... <i>High School and more</i>	17.0	47.7	51.7	13.3
Marital status (%)				
... <i>Single</i>	39.6	37.3	22.5	27.2
... <i>Married</i>	52.7	47.7	70.2	60.6
... <i>Divorced</i>	4.1	13.7	6.3	8.4
... <i>Widowed</i>	3.6	1.3	1.0	3.8
Labour status (%)				
... <i>Unemployed/Non-working</i>	48.1	25.7	29.5	17.0
... <i>Working</i>	51.9	74.3	70.5	83.0
Monthly earnings (XOF/euros/MRO)	66 552.8	1 281.4	1 163.2	52 531.8
Monthly earnings (PPP)	216.7	1 408.1	1 368.5	367.4
Monthly remittances (XOF)	(/)	125 209.9	130 637.0	36 928
Monthly remittances (PPP)	(/)	407.7	425.3	120.2
<b>Observations</b>	<b>1,929</b>	<b>300</b>	<b>302</b>	<b>324</b>

# Estimation

- ▶ Step 1 : selection at the household level (who migrates within the household ?)
  - ▶ Conditional logit with a varying number of alternatives (clogitVNA) of location and member choice (*who migrates where ?*)
- ▶ Step 2 : counterfactual earnings and remittances (accounting for selection into migration)
  - ▶ Mincer-type equations for earnings in all 4 countries and remittances for the 3 destinations including correction terms from step 1
- ▶ Step 3 : recover structural parameters of the household utility function
  - ▶ Estimate coefficients on log-earnings/remittances in the structural household model of location choices.

# Step 1 : Household selection equation (reduced form)

	clogit (1)	clogit (2)
Male	1.146*** (0.147)	1.266*** (0.178)
Age	0.401*** (0.042)	0.456*** (0.048)
Age squared	-0.005*** (0.001)	-0.005*** (0.001)
Elementary school	0.320 (0.211)	0.319 (0.216)
Middle school	1.107*** (0.240)	1.098*** (0.250)
High school and more	1.660*** (0.250)	1.979*** (0.295)
Rain zscore x male		0.353* (0.208)
Rain zscore x age		0.0478*** (0.010)
Rain zscore x high school		0.687** (0.303)
Observations	2225	2159
Rain zscore variables joint significance ( $\chi^2$ )		24.51*** (0.000)

# Estimation

- ▶ Step 1 : selection at the household level (who migrates within the household?)
  - ▶ Conditional logit with a varying number of alternatives (clogitVNA) of location and member choice (*who migrates where?*)
- ▶ Step 2 : counterfactual earnings and remittances (accounting for selection into migration)
  - ▶ Mincer-type equations for earnings in all 4 countries and remittances for the 3 destinations including correction terms from step 1
- ▶ Step 3 : recover structural parameters of the household utility function
  - ▶ Estimate coefficients on log-earnings/remittances in the structural household model of location choices.

# Step 2 : Earnings/Remittances equations

	Home earnings		Destination earnings		Remittances	
	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected
Male	0.616*** (0.070)	0.620*** (0.072)	0.426*** (0.089)	0.479*** (0.104)	0.305* (0.163)	0.420** (0.165)
Age	0.0872*** (0.019)	0.0887*** (0.017)	0.0571** (0.029)	0.0768** (0.030)	0.0313 (0.042)	0.0606 (0.051)
Age squared	-0.001*** (0.000)	-0.001*** (0.000)	-0.001 (0.000)	-0.001** (0.000)	-0.004 (0.001)	-0.003 (0.001)
Elementary school	0.115 (0.086)	0.111 (0.107)	0.266** (0.104)	0.277*** (0.105)	0.144 (0.206)	0.154 (0.210)
Middle school	0.521*** (0.120)	0.518*** (0.107)	0.263** (0.112)	0.321*** (0.123)	-0.009 (0.205)	-0.021 (0.221)
High school and more	1.095*** (0.131)	1.104*** (0.119)	0.421*** (0.114)	0.444*** (0.115)	0.177 (0.193)	0.180 (0.201)
Capital/big cities	0.325*** (0.091)	0.336*** (0.089)	-0.069 (0.079)	-0.066 (0.080)	-0.001 (0.149)	-0.002 (0.157)
Koranic schooling					0.554*** (0.213)	0.507** (0.253)
Resident spouse/child					0.418*** (0.155)	0.395*** (0.129)
Son/daughter of head					0.107 (0.148)	0.0908 (0.135)
Italy (d)			-0.0441 (0.090)	-0.0673 (0.108)	0.131 (0.173)	0.129 (0.192)
Mauritania (d)			-1.171*** (0.092)	-1.175*** (0.091)	-1.447*** (0.163)	-1.448*** (0.177)
Constant	2.247*** (0.362)	2.154*** (0.329)	5.368*** (0.571)	4.856*** (0.611)	9.905*** (0.841)	9.683*** (0.911)
Selection probability		0.239** (0.109)		-0.385** (0.294)		-2.206** (0.974)
Selection probability <sup>2</sup>		-0.052** (0.023)		0.220** (0.098)		2.343** (0.940)
Observations	813	813	255	255	249	249

# Estimation

- ▶ Step 1 : selection at the household level (who migrates within the household ?)
  - ▶ Conditional logit with a varying number of alternatives (clogitVNA) of location and member choice (*who migrates where ?*)
- ▶ Step 2 : counterfactual earnings and remittances (accounting for selection into migration)
  - ▶ Mincer-type equations for earnings in all 4 countries and remittances for the 3 destinations including correction terms from step 1
- ▶ Step 3 : recover structural parameters of the household utility function
  - ▶ Estimate coefficients on log-earnings/remittances in the structural household model of location choices.

## Step 3 : Household selection equation (structural form)

	clogit
Home earnings	2.585* (1.112)
Destination earnings	-24.59*** (2.945)
Remittances	16.36*** (2.436)
Observations	2110

# Conclusion

- ▶ Extension of the Roy model to account for household-based migration decisions
- ▶ Address the original issue of intra-household selection into migration
- ▶ Preliminary results suggest that the within-household selection decision depends on :
  - ▶ Expected remittances (from the migrant) - positively
  - ▶ Expected home earnings (of remaining members) - positively
  - ▶ Expected earnings abroad (of the migrant) - negatively