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

Isabelle Chort, Jean-Noël Senne

DOCUMENT DE TRAVAIL DT/2013-14

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

Isabelle CHORT
Jean-Noël SENNE

UMR DIAL 225
Place du Maréchal de Lattre de Tassigny 75775 • Paris Cedex 16 •Tél. (33) 01 44 05 45 42 • Fax (33) 01 44 05 45 45
• 4, rue d'Enghien • 75010 Paris • Tél. (33) 01 53 24 14 50 • Fax (33) 01 53 24 14 51
E-mail : dial@dial.prd.fr • Site : www.dial.prd.fr

INTRA-HOUSEHOLD SELECTION INTO MIGRATION: EVIDENCE FROM A MATCHED SAMPLE OF MIGRANTS AND ORIGIN HOUSEHOLDS IN SENEGAL

Isabelle Chort
PSL, Université Paris-Dauphine, LEDa, UMR DIAL, 75016 Paris, France
IRD, UMR DIAL, 75010 Paris
isabelle.chort@dauphine.fr

Jean-Noël Senne
Paris School of Economics (PSE),
UMR DIAL and CREST(INSEE)
PSL, Université Paris-Dauphine, LEDa
jn_senne@yahoo.fr

Document de travail UMR DIAL
Octobre 2013

Abstract

This paper fills the gap between individual selection models and collective approaches of migration. We build a theoretical model in order to account for household-based migration decisions and derive its implications on migrant selection. Assuming that the origin household maximizes a collective utility including earnings but also further remittances when choosing the one among its members who is to migrate, migrant selection in this case may differ from what is predicted by a pure individual decision model. Therefore, we specifically tackle the so far under-explored issue of intra-household selection into migration in order to identify what are the key determinants of household members' location choices. We derive our estimation procedure from an extension of the Roy-Dahl model and provide empirical evidence using a unique matched sample of 926 Senegalese migrants in three destination countries - France, Italy and Mauritania - and their origin household in Senegal. Our results show that expected remittances, along with earnings differentials, play a major role in shaping intra-household selection patterns, which stands in striking contrast with usual predictions from individual self-selection models.

Keywords : migration, remittances, intra-household allocation, selection.

Résumé

Ce papier, en se situant à l'interface des modèles de sélection individuelle et des approches collectives de la migration, apporte un éclairage nouveau sur la question centrale de la sélection des individus dans la migration. Un modèle théorique est tout d'abord proposé pour décrire le processus de décision de migration au niveau du ménage d'origine et, par là même, appréhender la complexité de la sélection du migrant lorsque la migration est envisagée comme une stratégie familiale. Le modèle fait l'hypothèse que le choix du membre du ménage en migration résulte de la maximisation de l'utilité collective du ménage d'origine, qui dépend non seulement des revenus mais aussi des transferts attendus de la part du migrant sélectionné. Dans ce cas, la sélection dans la migration peut différer de celle qui aurait prévalue dans le cadre d'un modèle de décision purement individuel. Ainsi, ce chapitre s'attaque à une problématique qui reste totalement inexplorée dans la littérature, à savoir la sélection *intra-ménage* dans la migration, et a pour ambition d'identifier les déterminants clés des choix de localisation des différents membres du ménage. Dans le prolongement du modèle de Roy, une procédure d'estimation novatrice est ensuite suggérée, permettant de tester les prédictions théoriques précédentes. L'analyse empirique se fonde sur des données uniques, issues du projet MIDDAS, et constituées d'un échantillon de migrants sénégalais dans trois pays de destinations (France, Italie et Mauritanie) appariés avec les non-migrants de leur ménage d'origine. Les résultats suggèrent que les différentiels de revenus, mais également les transferts attendus, jouent un rôle prépondérant dans la sélection du migrant *au sein* de son ménage d'origine, se posant ainsi en contraste avec les prédictions classiques des modèles de sélection individuels.

Mots Clés : migration, remises migratoires, allocation intra-ménage, sélection

JEL Code: F22, F24, D13, C51.

1 Introduction

The question of the characteristics that differentiate individuals who migrate and those who stay in their home country remains a vivid issue throughout the migration literature. Migrant selection has indeed been tackled by a large number of articles since the seminal paper by Borjas (1987) who applied to international migration the Roy model of self-selection. In this theoretical framework, location choices depend on individuals' comparative advantage based on both their observed and unobserved characteristics. All the papers derived from the Roy model of self-selection thus explicitly share an individualistic approach, in line with the first models of migration developed by economists who view migration as an individual income-maximizing strategy (Harris and Todaro (1970); Sjaastad (1962)). Indeed, in these models, individuals choose where to live and work according to their actual or expected earnings at each location, once migration costs are accounted for.

While the collective dimension of migration decisions has been acknowledged since the 1980s by a substantial strand of the literature, in particular Stark and Bloom (1985), this approach has exclusively been put forward to provide a rationale for remittances behavior that goes beyond mere altruism, especially in developing countries. Surprisingly, no paper has investigated the implication of a collective migration decision on migrant selection. Yet, if migration is rather a household welfare-maximizing strategy and then decided on collectively, so should the migrant member be selected *within* the household. Therefore, the selection of one or more migrants among household members may not be equivalent to individual self-selection into migration: indeed, future remittances to non-migrant members could well enter the collective decision process at the migration stage, jointly with comparative advantages in earnings. This paper is thus the first to explicitly model the implications of a household-based migration decision on migrant selection, by especially emphasizing the role of expected remittances. This issue is indeed ignored in the migration literature which is either focused on selection as a purely individual process or on remittances as the result of a collective strategy, while considering the migration decision as exogenous.

The question of intra-household selection into migration has received little interest to date, one reason being probably the lack of suitable data to empirically address this issue and the econometric challenges it raises. Indeed, in order to uncover the main factors that shape selec-

tion patterns within the household, one need to compare counterfactual allocations of household members across alternative locations accounting for the non-random double selection of who migrates and where. Therefore, information on the characteristics of both migrants and non-migrant members originating from a given origin household is required. This article thus contributes to fill a major gap in the migration literature by providing and estimating a household-based model for migrant selection using unique migrant-origin household matched data. *De facto*, we specifically tackle the issue of intra-household selection into migration and aim at answering the following questions: who, among household members, is more likely to migrate and what are the key components driving this collective decision?

First, we build a general theoretical model to account for a household's allocation decision of its members in different countries, including the home country. We assume that location choices result from the maximization of a household collective utility function that depends on both household members' earnings at each location and remittances received from abroad. We derive crucial implications in terms of migrant selection. We notably highlight how intra-household variations in remittances potential, along with earnings differentials across members, might play a role in the collective decision.

Second, we suggest an estimation procedure derived from an extension of the Roy model to a collective selection process with multiple alternatives. We then test the relevance of our model to explain household migration choices by providing an empirical application using unique survey data on a multi-sited and matched sample of Senegalese migrants and their origin households in Senegal. These data, collected in 2009-2010 as part of the MIDDAS project, provide information on migrants' characteristics in three of the top destination countries of Senegalese migrants (France, Italy, and Mauritania) as well as detailed information on all the remaining members of their origin household in Senegal. We collected in particular information on earnings of migrants in host countries and non-migrants in Senegal, and on remittances sent by migrants to their origin household.

The household-based framework that we adopt is relevant to the Senegalese migration case under study. Indeed, the data points out that 56% of surveyed migrants covered part or the totality of the costs of migration through family funding which indicates that migration is part of a collective investment. Furthermore, the data shows evidence of the strength of the links

between the migrants and their origin household. More than 80% of surveyed migrants send remittances to their origin households, most of the time on a regular (monthly) basis. The average amount of remittances represents between 15% and 30% of migrants' monthly income depending on the destination country and amounts to a large share of the resources of recipient households. In addition, remitted amounts are primarily used for daily consumption - 84% of money transfers - in order to cover the basic needs of all household members - 78% of money transfers are indeed targeted to the household as a whole for collective expenditures. Finally, around 60% of sampled migrants state their intention to return to Senegal, mostly in their origin household. These figures therefore indicate that, in spite of the geographical distance, both non-migrants and migrants remain part of a "transnational" household in which at least part of the resources are pooled. These empirical findings additionally suggest that remittances cannot be fully explained by a risk-sharing strategy.

Building on our theoretical model of intra-household selection into migration and estimating a three-step discrete model of location choices to recover the structural parameters in the household collective decision, we finally uncover which components mostly drive the allocation of household members across countries. Our results show that both earnings and remittances differentials play a role in shaping intra-household selection patterns. Interestingly enough, households select as migrants not only the members with higher comparative advantages in earnings at destination, but also those with higher remittances potentials, conditional on earnings. This very last feature stands in striking contrast with the usual predictions from individual self-selection models. This result is nevertheless consistent with our empirical findings that individuals with koranic education or who are the eldest among their siblings both have higher migration propensities and remit larger amounts of money, despite having no comparative advantages in earnings.

This paper thus reconciles the migrant selection literature, exclusively focused on self-selection of immigrants as an individual process (Chiswick (1999), Orrenius and Zavodny (2005), McKenzie and Rapoport (2010), Fernandez-Huertas Moraga (2011), Fernandez-Huertas Moraga (2013)), and other strands of the migration literature which have acknowledged the collective dimension of the migration decision, but mainly through the lens of remittances motives (see Rapoport and Docquier (2006) for a review). The first paper that intended to investigate the

impact of the family on migrant selection is Borjas and Bronars (1991). However, their model does not really depart from an individual selection approach. Since their theoretical results are based on the simplifying assumption that households do not split, they do not account for remittances and miss the largest part of the issue of collective migration decisions. For many developing countries, including Senegal, the assumption that the household does not split is overly restrictive.

The empirical part of the paper builds on Dahl (2002) who enriched the theoretical self-selection model inherited from Roy (1951) by dealing with selection based on multiple alternatives and providing an application to internal migration within the U.S. The same methodology is applied by De Vreyer et al. (2010) to the analysis of location choices in West African capital cities, and by Bertoli et al. (2013) to the migration of Ecuadorians to Spain and the U.S. However, Dahl (2002) and subsequent papers building on it all study individual location decisions. Our contribution is to adapt Dahl's individual theoretical framework to the modelling of collective location decisions. This article more indirectly relates to studies estimating conditional logit models with a number of alternatives varying across observations, with applications to marketing (Berry et al. (2004), Allenby and Rossi (1998)) or electoral choices (Yamamoto, 2012). Indeed, since we explore within-household allocation choices of members, the set of alternatives available to each household depends on the number of potential migrant members in the household.

Finally, this paper is one of the few empirical studies exploiting the specific information contained in matched data samples. We indeed surveyed both migrants and non-migrant members of their origin household. The resulting matched and multi-sited data set provides a unique opportunity to investigate original issues that are not or poorly tackled by the existing literature, such as the role of origin families in migrants' behavior and (intra-household) selection into migration for the case under study. It also helps bringing new insights on these topics through comparative analysis of migration in various contexts. The same survey design is found in Osili (2007) who studies the case of Nigerian migrants in the U.S., but the resulting matched sample is only made of 61 pairs of migrants and families of origin. A similar data structure is obtained by Abramitzky et al. (2012) and Ambler (2012), although constructed with very different methodologies. In the first case, the authors exploited data on individuals' names and

ages from the 1865 and 1900 Norwegian and the 1900 US censuses to link Norwegian migrants to their childhood household. In the second case, the author designed a controlled experiment to assess the role of information asymmetries between Salvadoran migrants in Washington and their origin household, during which family members back in Salvador were reached by phone and asked very specific questions related to the experiment.

The paper proceeds as follows. Section 2 presents a household-based model for migration and derives an extension of the Roy model of selection. Section 3 outlines the estimation procedure. Data and descriptive statistics are described in Section 4. Empirical specification and identification issues are discussed in Section 5. Estimation results are then presented in Section 6. Finally, Section 7 concludes.

2 Theoretical framework

In this section, we set up a structural model for household-based migration decisions. We consider that household members' location choices are decided on collectively within the origin household in order to maximize a collective utility depending on both earnings and remittances. We derive crucial implication in terms of migrant selection. We then build on Dahl (2002) to define an extended Roy model of intra-household selection that can be estimated to recover the key structural parameters driving the underlying collective decision.

2.1 A household-based model for migration

The origin household can be regarded as a "portfolio" of members whose geographical allocation is decided on collectively. Each member can be selected as a migrant in any possible destination country or as a stayer in the home country. Two preliminary comments have to be made. First, the model aims at investigating the rationality of migrant selection and location, conditional on the fact that households are yet selected into migration. Indeed, we choose to focus the analysis on intra-household selection, that is to say on the choice of the specific member who is to migrate and live abroad and the choice of the destination - this will be a simultaneous decision in our setting - once the decision to participate in migration has been taken[1]. Second, for simplicity matters, we ignore the fact that households can have several migrants and further

[1]Moreover, this will be consistent with the following empirical application and the structure of the matched data we use, which is exclusively composed of migrant households (see section 1.4).

focus the model on the case in which households send only one member abroad[2].

General setting

We consider an origin household h made of I members who can migrate to J possible destination countries. We note M_{hij} a dummy equal to one if member i of household h has migrated in destination country j and all other members have stayed in the home country[3]. We define the utility of each member in her relevant location as a function of individual earnings, at destination or in the home country, and remittances, sent or received. With regards to remittances from the migrant member, we first consider that they enter the $(I-1)$ utilities of non-migrant members positively. We additionally assume that the latter equally benefit from received amounts[4]. Conversely, we assume that remittances enter the migrant's own utility negatively, and that they are discounted by a positive and lower than one factor reflecting any indirect or deferred individual utility derived from sending money back to the origin household[5]. Therefore, depending on her migration status, the individual utility U_i of member i in household h writes:

$$U_i(M_{hij}) = U_i\left[(1-M_{hij})(Y_{is}+\frac{R_{kl}}{I-1})+M_{hij}(Y_{ij}-\delta_i R_{ij})\right] \qquad \forall i,j \text{ and } k\neq i \qquad (1)$$

where U_i is a concave and twice differentiable utility function. Y_{is} stands for earnings of member i in home country s and Y_{ij} for earnings of member i in destination country j, R_{ij} refers to remittances sent by migrant member i from host country j and R_{kl} to remittances received from migrant member k in host country l. Finally, δ_i is an individual-specific discount factor with $0<\delta_i<1$.

[2]More than 60% of origin households that were successfully tracked reported the surveyed migrant as the only member abroad. Unfortunately, the questionnaire does not record detailed information on other migrants from the same origin household. Yet, the following empirical results are robust, though less precise due to the small size of the resulting sample, to the exclusion of households with multiple international migrants.

[3]In the following empirical analysis, we restrict the pool of potential migrants to working-age household members.

[4]This might be too strong an assumption if remittances are targeted to specific recipients within the household for private use. It is nevertheless relevant in our context where remittances are mostly designed to the whole household to cover collective expenditures. This assumption is moreover consistent with the study by De Vreyer et al. (2009) on a representative sample of Senegalese households. Their results show that remittances used for daily consumption globally benefit all the members of the household. This is so even if they accrue to specific individuals or sub-groups within it.

[5]We could also consider the case of monetary transfers from the origin household to the migrant. However, this is not relevant in the specific context of migration from developing countries where this kind of transfers is almost never observed. Our data shows that only 0,4% of migrants received money from their family in Senegal.

The closer δ_i is to 0, the lower the negative effect of remittances in the migrant's direct utility. Intuitively, in the extreme case where δ_i would be equal to 0, the implied loss for the migrant in her direct individual utility would be totally compensated by the indirect utility derived from remittances to the home country, so that remittances would not affect the migrant's own welfare. In the polar case where it would be equal to 1, remitted amounts basically translate into lower disposable income at destination for the migrant. Note that δ_i may encompass any motive for remittances from which the migrant derives some positive but indirect or deferred utility. It might involve any exchange of services between the migrant and her origin household, the "warm glow" of taking care of those left behind through altruism or commitment to solidarity norms, or any form of deferred benefits and social prestige associated with migration and remittances upon return. This very last parameter will be crucial in the following analysis through its dual effect on transferred amounts and on the discounted loss in direct utility that it induces for the migrant member. To the extent that remittances potential plays a role in the migration decision, jointly with earnings differentials, it is an additionally relevant parameter in the selection process of the migrant within the household.

We then basically define the household's total utility U_h as an additively separable function of the weighted sum of each household member's individual utility plus a migrant-specific taste factor:

$$U_h(M_{hij}) = \sum_{i=1}^{I} \theta_i U_i(M_{hij}) + T_h(M_{hij}) \qquad \forall i, j \qquad (2)$$

where θ_i corresponds to the weight of member i in the household utility, with $\sum \theta_i = 1$. An alternative interpretation of these welfare weights is that they represent the bargaining power of each household member in the intra-household allocation process. The additional taste factor T_h aims at capturing the non-monetary determinants entering the total utility function. It includes in particular the specific costs of moving and any other non-monetary or psychic costs and benefits for household h of having a member i in country j. As such, note that this taste component is itself a weighted sum of costs and benefits for both the remaining household members in the home country and the migrant member in her chosen location[6].

[6]However, since our focus is on the relative role of earnings and remittances in the household decision, we leave the structural form of the taste component unspecified.

To describe the intra-household migration decision, we further adopt a simple unitary framework that can be characterized by a consensus model *à la* Samuelson (1956). Each member has specific preferences but these preferences are interrelated by a consensus that takes into account the welfare of other household members. In other words, household members agree on a unique collective objective and then act as if they were maximizing a well-behaved (Bergsonian) social welfare function. In our specific setting, we will therefore consider that migration and remittances decisions are interrelated and decided on collectively by each household member in order to maximize the above defined household utility[7].

Optimal amount of remittances

From the expression in equation (2), the utility for household h of having member i in country j and all other members in the home country s writes:

$$\begin{aligned} U_{hij} &= U_h(M_{hij}=1) \\ &= \sum_{k\neq i} \theta_k U_k \left(Y_{ks} + \frac{R_{ij}}{I-1} \right) + \theta_i U_i \left(Y_{ij} - \delta_i R_{ij} \right) + T_{hij} \qquad \forall i,j \end{aligned} \tag{3}$$

We consider that earnings and tastes at each location as well as parameters such as individual-specific bargaining powers and discount factors on remittances are exogenously given and known to all household members, or at least accurately expected conditional on the observed characteristics of members and locations[8]. Yet, the amount of remittances sent back to the origin household is the result of the migrant member's decision and therefore endogenously determined so as to maximize the household total utility including her own one. Hence, any optimal amount of remittances R^*_{ij} from migrant i living in country j should satisfy the following first order condition:

$$\frac{\partial U_{hij}}{\partial R_{ij}} = \sum_{k\neq i} \frac{\theta_k}{I-1} U_k^{'} \left(Y_{ks} + \frac{R^*_{ij}}{I-1} \right) - \theta_i \delta_i U_i^{'} \left(Y_{ij} - \delta_i R^*_{ij} \right) = 0 \qquad \forall i,j \tag{4}$$

[7]Although this unitary framework has been criticized first because the mechanism that leads to an agreement within the household remains unspecified and second because it somehow neglects household members' own rational preferences, as noted by Samuelson, it is particularly relevant in the case where household total resources are properly broken down into pre-specified shares so that the primary objective is to maximize the total (earnings) surplus. The only consensus decision to be made then relates to the allocation of the household surplus among members.

[8]We further assume that bargaining powers and discount factors are predetermined at the time when mobility decisions are taken. They are thus assumed not to be endogenous to migration.

By differentiating the previous equality with respect to each parameter (see Appendix 1.A), it can first be shown that the optimal amount of remittances R^*_{ij} is unsurprisingly an increasing function of migrant's earnings at destination Y_{ij} and a decreasing function of non-migrants' earnings in the home country Y_{ks}. More importantly, conditional on household members' total earnings, R^*_{ij} is also a decreasing function of the remittances discount factor δ_i. This last finding is quite intuitive: at the household level, when δ_i is lower than one and small enough, the marginal gain from each additional unit of remittances in non-migrants' utilities outbalances the concurrent marginal loss in the migrant's utility, therefore inducing an increase in the equilibrium amount of remittances. Overall, these predictions provide a first rationale for household-based migration choices to differ from pure individual self-selection based on earnings differentials. Indeed, conditional on members' individual earnings, intra-household variations in the δ_i parameter could well play an additional role in the collective decision through induced individual variations in the propensity to remit larger amounts of money.

Note that R^*_{ij} is also found to be a decreasing function of the migrant's bargaining power θ_i and an increasing function of non-migrants' bargaining powers θ_k, which is again quite intuitive since bargaining powers basically interfere through the differential weights attached to the welfare of each member at her relevant location. Nonetheless, a more important remark is that they do not alter the previous predictions. The specific role of bargaining powers in the selection process is more extensively discussed in the following sections.

Intra-household migration decision

From the above determined optimal amount of the remittances sent by migrant member i, the corresponding value of the collective utility function for household h of sending member i in destination country j and having all other members stay in the home country s writes:

$$V_{hij} = U_{hij}(R^*_{ij}) = \sum_{k \neq i} \theta_k U_k \left(Y_{ks} + \frac{R^*_{ij}}{I-1} \right) + \theta_i U_i \left(Y_{ij} - \delta_i R^*_{ij} \right) + T_{hij} \qquad \forall i, j \qquad (5)$$

Since every I household member can migrate in one of the J destination countries, the household's problem then boils down to choosing among $I \times J$ alternatives the geographical allocation of its members that maximizes the value of its collective utility. Household h thus

decides to locate member i in country j according to:

$$M_{hij} = \begin{cases} 1 & \text{if} \quad V_{hij} = \max(V_{h11}, ..., V_{h1J}, ..., V_{hI1}, ..., V_{hIJ}) \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

From the expression in equation (5) and by a direct implementation of the envelope theorem to allow for some comparative statics (see Appendix 1.B), it is easy to show that, conditional on welfare weights and migration costs and tastes, the value of the household utility at the remittances optimum is first an increasing function of the migrant's earnings at destination, Y_{ij}. Yet, in our household-based setting, it is also found to be an increasing function of non-migrant members' earnings in the home country, Y_{ks}. This last prediction indirectly reflects the fact that the opportunity cost of sending a member with high earnings at home is larger. Therefore, what precisely matters in the household decision is the comparative advantage in earnings across locations *among* potential migrant members, that is to say the difference between $(Y_{kj} - Y_{ks})$ and $(Y_{ij} - Y_{is})$ for two different members k and i. Put differently, the household thus chooses as its migrant the member with the highest earnings differential between host and home countries.

A second interesting feature for our matter at hand is that the optimal value of the household utility is a decreasing function of the remittances discount parameter δ_i. This is again intuitive, since for small enough δ_i, the marginal loss from remittances in the migrant member's welfare in the destination country is offset at the household level by the induced marginal gains in the non-migrant members' welfare in the home country. Together with the related and above stated effect of δ_i on remitted amounts, this result merely puts forward the fact that, conditional on earnings differentials, welfare weights and tastes, those members with a higher propensity to remit have a higher probability to be selected as migrants within the household.

Altogether, the household-based framework that we adopt to account for collective migration decisions allows us to derive crucial implications regarding migrant selection within the household with respect to both earnings and remittances. First, consistently with individual selection models that view migration as an income-maximizing strategy, household members' location choices are found to be primarily driven by individual comparative advantages in earnings across locations. At the household level, this finding is slightly more subtle since the decision

involves an additional intra-household level of comparison between potential migrant members, but basically comes down to maximizing the household total earnings surplus. Second, considering that part of this total surplus is further shared between migrant and non-migrant members through remittances, remittances potential together with factors influencing the marginal utility derived from remitted amounts play an additional role in the intra-household selection process through the optimal reallocation of welfare that it causes within the household. This last prediction stands in striking contrast with the usual predictions derived from individual self-selection models and is the key implication that we test in order to assess the appropriateness of our household framework in explaining intra-household migration patterns.

Two important points have to be raised. First, at the household level, the dual effect of δ_i on remittances and the propensity to migrate basically results from the induced discounted - an lower than one - marginal loss from transferred amounts in the migrant's utility. As a consequence, the interaction between δ_i and individual remitted amounts should precisely matter in the selection process, as implicitly shown in equation (5). This very last point is fundamental for further identification of the role of remittances potential in the following empirical analysis and is more precisely discussed in the next subsection.

Second, the optimal value of the household utility is finally found to be an increasing function of both migrant and non-migrants' bargaining powers $\{\theta_i, \theta_k\}$. Hence, the higher the relative weight of the member in the household utility, the higher the probability to be selected as a migrant. However, allowing bargaining powers to differ across household members does not challenge the above two main predictions from our theoretical model, since bargaining powers only affect the allocation of welfare within the household. Still, this last remark intuitively suggests that the respective role of individual earnings differentials and remittances potentials in the migration decision might vary according to differences in individual welfare weights within the household. In the following empirical analysis, we nevertheless could not find such an heterogeneity with respect to different proxy measures of bargaining powers within the household. For ease of presentation and considering that the effect of differential welfare weights is then negligible in our setting, we therefore assume that they are equal across household members in the next sections. We provide a detailed discussion in the empirical Section 1.6.

2.2 An extended Roy model of intra-household selection

The rest of the paper aims at testing the relevance of our household-based model to account for intra-household selection into migration. Building on the previous predictions and using detailed information on migrants and non-migrant members of the same origin household, we basically investigate the responsiveness of household members' location choices to both individual earnings differentials and remittances potentials, conditional on tastes.

First, we take a linear approximation of the above defined household utility function to allow tractable estimation of the underlying structural parameters driving location choices. Second, we assume that household members have accurate expectations about individual earnings, remittances and tastes, based on observable characteristics of each member and location[9]. Third, we consider that welfare weights are equal across household members. The collective value of the household random utility of locating member i in country j can then be written:

$$\tilde{V}_{hij} = \underbrace{\alpha(\sum y_{ks} + y_{ij})}_{\substack{\text{Earnings} \\ \text{component}}} + \underbrace{\beta(1-\delta_i)r_{ij}}_{\substack{\text{Remittances} \\ \text{component}}} + \underbrace{\gamma_j t_{ij}}_{\substack{\text{Taste} \\ \text{component}}} + \epsilon_{hij} \qquad \forall i, j \qquad (7)$$

with:

$$y_{ks} = E(Y_{ks}|x_k); \quad y_{ij} = E(Y_{ij}|x_i); \quad r_{ij} = E(R_{ij}|x_i); \quad t_{ij} = E(T_{hij}|z_{ij})$$

where x_k is a set of characteristics of non-migrant member k affecting home earnings, x_i is a set of characteristics of migrant i affecting destination earnings and remittances, z_{ij} is a vector of migrant i and destination j characteristics affecting tastes and ϵ_{hij} is an error term.

Following our theoretical framework, values for the remittances discount factor δ_i can equally be approximated by a subset x_{1i} of individual characteristics x_i that affect remittances amounts conditional on household earnings surplus. As such, note that the whole remittances component $(1-\delta_i)r_{ij}$ could be regarded as a simple reduced-form function $f(x_{1i})$ of those characteristics. Yet, to the extent that such characteristics also influence tastes t_{ij}, they would then stand for a mixed component of remittances and tastes. Therefore, further assessing their effect in

[9]This might be too strong an assumption since migrant earnings and remittances are only observed ex-post by the household. Moreover, information asymmetries may exist between the migrant and the origin household, due in particular to geographical distance. However, this assumption simplifies the setting and, as noted by Dahl (2002), adding in uncertainty to the Roy setting so that migration is based on expected utility maximization does not change the main insights from the model.

interaction with expected remittances amounts, as it appears in the structural form of the utility, will allow to disentangle and identify the relative role of remittances and tastes in the selection process, once earnings are properly taken into account[10]. Finally, note that allowing welfare weights to differ across household members would simply imply an additional heterogeneity in the structural parameters α and β with respect to relevant measures of individual bargaining powers. This potential heterogeneity is tested and ruled out in the empirical application that follows.

Utility thus comprises a deterministic mean component, which is a function of individual and locations' (observed) characteristics and a stochastic (unobserved) component which stands for household members' deviations from mean earnings, remittances and tastes[11]. The set of parameters $\{\alpha, \beta, \gamma_j\}$, which represents the relative weights of each factor in the above utility, is assumed to be identical across households. Moreover, while the γ_j parameters are location-specific to account for destination-specific costs or benefits of migration, the set of parameters $\{\alpha, \beta\}$ is further assumed to be homogenous across locations. Put differently, any increase in labor market earnings or remittances provides identical gains or losses in terms of utility, whatever the specific member's country of residence[12].

Considering that household members select among $I \times J$ alternatives the member's geographical allocation that maximizes the collective value of their random utility, the intra-household selection equations in (6) can alternatively be written as:

$$M_{hij} = \begin{cases} 1 & \text{if} \quad \tilde{V}_{hij} > \tilde{V}_{hkl} \qquad \forall (k,l) \neq (i,j) \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where M_{hij} is the indicator variable which is equal to one if member i from household h lives in destination country j and all remaining members k stay in the home country. The selection rule

[10] Intuitively, if the x_{1i} proxies for δ_i were only capturing a taste effect, they should play no role through remittances differentials. Hence, identifying the latter effect basically comes down to empirically investigate the heterogenous effect of remittances with respect to those proxies.

[11] Actually, this overall stochastic component is a complex sum of household members' individual-specific error terms. This point, as well as the choice of a functional form for expected earnings, remittances and tastes are further developed in Section 1.3.

[12] As noted by De Vreyer et al. (2010), this might be too strong an assumption if large differences exist between countries in the set of available goods and their prices (for instance public services), so that the living standards of individuals with equal incomes but residing in different country would be indirectly impacted. However, we can credibly assume that households are not in a position to take this dimension into account in their utility. Moreover, earnings will be converted into Purchasing Power Parity (PPP) units in the following empirical application, to allow relevant comparisons.

is such that non-migrant members' home earnings, migrant's earnings and remittances are only observed for the household's utility-maximizing allocation choice. In other words, the household can only locate each member in one specific destination, so that earnings and remittances are not observed for each member in every location but only if all $I \times J$ selection equations in (8) are simultaneously satisfied[13]. Equations (7) and (8) therefore define an extended Roy model of earnings, remittances and mobility, such as in Dahl (2002), but in which location choices result from a household utility-maximizing strategy. Hence, estimating this extended Roy-Dahl model directly derived from our household-based decision framework provides unique insights into the intra-household selection process of migrants.

2.3 Challenges to estimation

Since this article aims at investigating which component of the household utility mostly drives location choices, we are particularly interested in estimating the set of structural parameters $\{\alpha, \beta, \gamma_j\}$ in the latent structural utility from equation (7), which is equivalent to estimating a within-household discrete choice model of members' location depending on earnings and remittances. Such an estimation raises two main challenges.

First, following the above-described extended Roy model of selection, a fundamental identification issue stems from the fact that earnings and remittances are only observed at one location for each household member. To identify whether relative differentials in earnings and remittances determine the choice of the migrant member within the household, we therefore need to compute counterfactual earnings and remittances for each household member at each location, namely counterfactual earnings of migrants in the home country, *had they not migrated*, and counterfactuals earnings and remittances of both migrants and non-migrants in each possible destination country, *had they migrated* (for non-migrants) or had they migrated *elsewhere* (for migrants)[14]. However, a selection bias results from the fact that households choosing a specific utility-maximizing geographical allocation are not a random sub-sample of the population. In other words, selected migrants and non-migrants are likely to have specific observed and unob-

[13]Formally, each household h faces a $I \times J$ number of alternatives, so that $I \times J$ binary variables M_{hkl} can actually be defined, corresponding to $I \times J$ selection equations. M_{hij} equals one if alternative $\{ij\}$ is chosen and observed; consequently all the remaining M_{hkl} equal 0 since, by construction, only one allocation can be chosen. In other words, exactly one of the set of binary variables $M_{h11}, ..., M_{h1J}, ..., M_{hI1}, ..., M_{hIJ}$ is non-zero for each household h.

[14]More precisely for migrants, had they migrated in another destination country than the one in which they were surveyed.

served characteristics that simultaneously drive migration, earnings and remittances. "Naive" imputations based on earnings and remittances equations uncorrected for endogenous selection would then yield biased results. As a consequence, earnings and remittances for other locations must be imputed, taking into account the fact that location choices are not random but partially driven by observed and unobserved characteristics explaining earnings and remittances gaps. In the next section, we adapt the semi-parametric method suggested by Dahl (2002)[15]. We provide a simple three-step parametric estimation procedure to produce counterfactuals that correct for selection biases and finally test the consistency of our results with the household-based migration model that we develop.

An additional challenge inherent to our specific setting lies in the fact that, while the number of possible destinations for potential migrants is fixed, households are not necessarily of equal size. As a consequence, the number of potential migrant members varies across households. Each household is actually faced with a varying number of alternatives, each corresponding to the location of one specific member in one specific destination country. We thus need to estimate a within-household multiple choice model which takes into account variations in the size of the choice set across households. A few implementations of such non-standard multiple choice models can be found in the marketing literature, to estimate market shares of products' brands that are not available to every consumer from different regions (see Allenby and Rossi (1998) or Berry et al. (2004)), or in the political science literature, to analyze electoral choices within partially contested multiparty elections in which some parties do not run candidates in every district (see Yamamoto (2012)). We build on the latter papers to suggest an estimation procedure based on a (within-household) conditional logit model of location choices with a varying number of alternatives, which is extensively described in the following section.

3 Estimation strategy

We exploit the unique information contained in our matched and multi-sited dataset on individual earnings and remittances of migrants and non-migrants from the same household (see Section 1.4) to develop a three-step parametric estimation procedure of the extended Roy-Dahl model of intra-household selection into migration. We first estimate a reduced-form condi-

[15]Dahl's method is also applied in De Vreyer et al. (2010) and Bertoli et al. (2013) within the framework of individual self-selection models of migration.

tional logit model of intra-household location choices with a varying number of alternatives. Second, results from the first-stage estimation are used to estimate individual earnings and remittances equations corrected for endogenous selection. Third, we compute consistent counterfactual predictions to recover unbiased parameters on earnings and remittances components in the structural-form estimation of the conditional logit model of location choices.

Expected earnings, remittances and tastes

First, we assume that non-migrant household members k in the home country s face a Mincer-type earnings equation:

$$Y_{ks} = x_k^{'}\rho_s + \mu_{ks} \tag{9}$$

where x_k is a set of individual characteristics of non-migrant member k affecting (home) earnings and μ_{ks} is an individual-specific error term. The vector of parameters ρ_s identifies home country-specific returns to individual characteristics with respect to earnings.

Second, we assume that each migrant i living in destination country j faces the same Mincer-type earnings equation:

$$Y_{ij} = x_i^{'}\rho_j + \eta_{ij} \qquad \forall j \tag{10}$$

where x_i is a set of characteristics of migrant i affecting (destination) earnings and η_{ij} is an individual-specific error term. The vector of parameters ρ_j identifies destination country-specific returns to individual characteristics with respect to earnings.

Third, we similarly define a remittances equation for each migrant i living in destination country j:

$$R_{ij} = x_i^{'}\pi_j + \nu_{ij} \qquad \forall j \tag{11}$$

where x_i is a set of characteristics of migrant i affecting the amount of remittances sent back to the origin household and ν_{ij} an individual-specific error component. The vector of parameters π_j identifies destination-specific returns to individual characteristics with respect to remittances. Precisely investigating individual determinants of remittances amounts is of particular interest

in our setting. To this end, we will add to the vector of characteristics x_i both migrant and non-migrants' earnings in order to identify a subset x_{1i} of individual characteristics that (positively) influence remittances once earnings are accounted for. This latter subset will further be isolated as relevant proxies for (low) values of the δ_i parameter which essentially captures individual variations in the propensity to remit conditional on the household earnings surplus.

We finally specify tastes T_{hij} as a flexible function of migrant i and destination j characteristics, denoted z_{ij}. Many destination-specific variables may enter this taste component, some of them being potentially unobserved. We sidestep the estimation of this taste component by considering a country-specific dummy that accounts for any fixed differences in the costs or benefits of migration across destinations. It includes for instance moving costs, global standards of living or differences in public services, institutions and culture. We nevertheless assume that these costs and benefits might vary across individuals within a particular destination. Basically, the vector z_{ij} includes a destination dummy λ_j and a set of interactions with individual characteristics x_i and therefore stands for individual controls. We note $x_i'\phi_j$ this component.

Intra-household selection equation

We can substitute the above expressions of Y_{ks}, Y_{ij} and R_{ij}, together with the flexible specification of tastes T_{hij}, in equation (7) to get the household random utility in a reduced form, i.e as a function of household members' characteristics:

$$\tilde{V}_{hij} = \alpha(\sum_{k\neq i} x_k^{'}\rho_s + x_i^{'}\rho_j) + \beta(1-\delta_i)(x_i^{'}\pi_j) + x_i'\phi_j + \epsilon_{hij} \qquad \forall i,j \tag{12}$$

where $\epsilon_{hij} = \alpha(\sum_{k\neq i}\mu_{ks} + \eta_{ij}) + \beta(1-\delta_i)\nu_{ij} + \xi_{ij}$ and ξ_{ij} stands for individual deviations from mean tastes. The stochastic component of the utility is then a (weighted) sum of individual deviations from mean earnings y_{ks} and y_{ij}, remittances r_{ij} and tastes t_{hij}, respectively specified as deterministic functions of individual observable characteristics.

Equation (12), together with the selection rule in equation (8), depicts the general framework of an additive random utility model. Under the statistical assumption that error components ϵ_{hij} are i.i.d and have a type-1 Extreme Value distribution, it can be shown that the probability P_{hij} that household h locates member i in country j:

$$P_{hij} = P(M_{hij}=1) = P(\tilde{V}_{hij} > \tilde{V}_{hkl}) \qquad \forall (k,l)\neq(i,j) \tag{13}$$

can be written:

$$P_{hij} = \frac{exp[\alpha x_i^{'}(\rho_j - \rho_s) + \beta(1-\delta_i)(x_i^{'}\pi_j) + x_i^{'}\phi_j]}{\sum_{k=1}^{I}\sum_{l=1}^{J} exp[\alpha x_k^{'}(\rho_l - \rho_s) + \beta(1-\delta_k)(x_k^{'}\pi_l) + x_k^{'}\phi_l]} \tag{14}$$

P_{hij} writes as the usual conditional probability derived from a standard conditional logit model with a $I \times J$ fixed number of alternatives corresponding to each possible intra-household choice of member allocation[16]. A first specific feature is that the set of reduced-form parameters is destination-specific but alternative-invariant across choices of the member to be located at a given destination. Identification then relies on intra-household variations in individual characteristics of members. An additional specific feature is that the reduced-form probability for a member to be selected as a migrant appears to depend on both his own individual characteristics and the characteristics of all other potential migrant members within the household.

As previously noted, one important but non-standard issue for estimation is that households are not of equal size. As a consequence, the number of potential migrant members I varies across households so that each household h actually faces a varying $I_h \times J$ number of alternatives. If we further assume that the set of parameters to be estimated is identical across households, we can however easily write both conditional probabilities and the contribution to the log-likelihood function of a given household-level observation conditional on the specific number of alternatives available to that household, as follows:

$$\mathcal{L}_h = \ln(L_h) = \sum_{i=1}^{I_h}\sum_{j=1}^{J} M_{hij} \ln P_{hij} \tag{15}$$

where I_h is the number of potential migrant members in household h, J is the fixed number of possible destination countries for migrant member i, M_{hij} is the dummy equal to one if household h has a member i in country j and P_{hij} the associated conditional probability from equation (14) but whose denominator now depends on a household-specific $I_h \times J$ number of allocation choices.

[16]Note that $\tilde{V}_{hij} > \tilde{V}_{hkl}$ writes: $\alpha(\sum_{k\neq i} x_k^{'}\rho_s) + ... + \epsilon_{hij} > \alpha(\sum_{m\neq k} x_m^{'}\rho_s) + ... + \epsilon_{hkl}$ where sums on both sides of the inequality reduce: $-\alpha(x_i^{'}\rho_s) + ... + \epsilon_{hij} > -\alpha(x_k^{'}\rho_s) + ... + \epsilon_{hkl}$ and yields the simplified expression in equation (14). Note that components $\sum_{k\neq i}\mu_{ks}$ and $\sum_{m\neq k}\mu_{ms}$ also reduce in the household error term so that ϵ_{hij} (resp. ϵ_{hkl}) appears to be a function of individual i (resp. individual k) error terms only. This allows us to plausibly state the i.i.d assumption in equation (12).

The log-likelihood function for a sample of N households then writes as usual:

$$\mathcal{L}_N = \sum_{h=1}^{N} \mathcal{L}_h = \sum_{h=1}^{N} \sum_{i=1}^{I_h} \sum_{j=1}^{J} M_{hij} \ln P_{hij} \tag{16}$$

Equation (16) generalizes the sample log-likelihood function from a standard conditional logit model where choice sets are allowed to vary across observations. Standard maximization routines can then be applied to get consistent estimates of the set of reduced-form parameters.

Counterfactual earnings and remittances predictions

Considering that individual unobserved heterogeneity drives the intra-household probability to be selected as a migrant as well as earnings and remittances, observed samples of individuals at a given location are obviously not random. In other words, individuals with different propensities to migrate to a particular destination country are expected to systematically differ in their earnings and remittances realizations. Earnings and remittances equations thus need to be corrected for endogenous selection so that we can generate consistent counterfactual predictions.

We apply the selectivity-correction method implemented by Dahl (2002)[17]. The idea is to use the results of the above defined multiple choice model to compute, for each household member, a set of predicted location choice probabilities. A flexible function of these probabilities, denoted $\lambda(p_{hij})$, is then included as an additional set of regressors in equations (9), (10) and (11) to correct for selectivity biases. This scheme works as a control function procedure that corrects for differences in selection probabilities linked to both observed and unobserved individual characteristics which simultaneously affect the outcome under study. Theoretically, all choice probabilities could enter the control function. In practice, to avoid potential multicollinearity issues, Dahl (2002) suggests to use a high order polynomial of the first-best choice probability, i.e. the probability of the observed allocation, and a subset of other relevant probabilities. Implementation choices are discussed in Section 1.5.

A potential drawback to the conditional logit model is its Independence of Irrelevant Alternatives (IIA) property and its reliance on a parametric framework, so that a non-parametric

[17] For an exhaustive comparison of existing methods for selection bias correction based on a multinomial model, see Bourguignon et al. (2007). Resorting to Monte Carlo's simulations, they find that Dahl's approach is to be preferred to other commonly used methods such as Lee (1983) or Dubin and McFadden (1984).

estimation of choice probabilities should be preferred. However, this is only feasible with a large number of observations. Moreover, Bourguignon et al. (2007) pointed out that, even when the IIA assumption is severely at odds, selection bias correction based on multinomial models can be considered as a reasonable alternative when the focus is to estimate an outcome over selected populations. Therefore, our results should not be affected by the choice of the conditional logit model at this stage[18].

Structural model of intra-household location choices

In order to finally recover consistent estimates of the set of structural parameters $\{\alpha, \beta\}$ in the within-household model of location choices, a last step is needed. Using the unbiased estimates $\hat{\rho}_s$, $\hat{\rho}_j$ and $\hat{\pi}_j$ from the selectivity-corrected earnings and remittances equations, we can compute consistent earnings and remittances counterfactuals for each individual at each possible location, and then estimate the following structural conditional logit model with $I_h \times J$ alternatives:

$$\begin{aligned} P_{hij} &= \frac{exp[\alpha x_i^{'}(\hat{\rho}_j - \hat{\rho}_s) + \beta(1-\delta_i)(x_i^{'}\hat{\pi}_j) + x_i'\phi_j]}{\sum_{k=1}^{I_h}\sum_{l=1}^{J} exp[\alpha(x_k^{'}\hat{\rho}_l) - x_k^{'}\hat{\rho}_s) + \beta(1-\delta_i)(x_k^{'}\hat{\pi}_l) + x_k'\phi_l]} \\ &= \frac{exp[\alpha(\hat{y}_{ij} - \hat{y}_{is}) + \beta(1-\delta_i)\hat{r}_{ij} + x_i'\phi_j]}{\sum_{k=1}^{I_h}\sum_{l=1}^{J} exp[\alpha(\hat{y}_{kl} - \hat{y}_{ks}) + \beta(1-\delta_k)\hat{r}_{kl} + x_k'\phi_l]} \end{aligned} \tag{17}$$

In line with the main predictions of our household-based model for migration decisions, the probability to be located abroad first appears to respond to intra-household variations in expected earnings differentials $(\hat{y}_{ij} - \hat{y}_{is})$ between home and relevant destination countries. Nevertheless, conditional on earnings, a second dimension of selection further results from intra-household variations in remittances potentials. This additional selection channel can be captured through individual variations in relevant proxies x_{1i} for the δ_i parameter, since those characteristics simultaneously affect expected remittances amounts $\hat{r}_{ij}$[19]. Yet, proxy character-

[18]The non-parametric method suggested by Dahl (2002) consists in dividing the population into mutually exclusive cells according to observable characteristics such as gender, age or education. Migration probabilities are then estimated as the fraction of individuals in the same cell observed in a given country. The same approach is pursued in Bertoli et al. (2013). In this paper, we rely on the parametric alternative followed by De Vreyer et al. (2010).

[19]Remember that the lower δ_i, the higher the propensity to remit. In the extreme case where δ_i would be equal to one, note that remittances differentials should not play any role in the selection process. This is quite

istics for δ_i and individual controls x_i that enter the taste component credibly overlap, so that simple reduced-form parameters would confound both channels. Therefore, the introduction of interactions between (expected) remittances amounts and proxies for δ_i allows us to disentangle their potential joint effect through remittances and tastes. As a consequence, assessing the relative role of earnings and remittances in our within-household model of location choices by estimating the set of structural parameters $\{\alpha, \beta\}$ will be of crucial interest to shed light on the underlying rationality behind the intra-household selection process[20].

Two final comments have to be made. First, the violation of the IIA assumption could question the use of a conditional logit model at this stage, so that estimation methods that relax this assumption should be preferred[21]. Alternative estimation procedures are however hindered in our setting by the inherent challenges of dealing with a varying number of alternatives and by the limited size of the sample in the following empirical application. Second, some exclusion restrictions are needed for robust identification, in particular in steps two and three. Specification and identification issues are discussed in details in Section 1.5.

In the rest of the paper, we provide an empirical application using a unique matched sample of Senegalese migrants in three different destination countries - France, Italy and Mauritania - and non-migrant members of their origin household in Senegal. The next section presents the data.

4 Data

This article uses data from the surveys conducted between 2009 and 2010 within the framework of the MIDDAS project. Using Senegal as a case-study, this research project aims at documenting the links between migration, remittances and development. An important contribution of this research project is the collection of a unique matched and multi-sited data set, on which the subsequent analysis is based.

intuitive since it would imply that migrants would not derive any indirect utility from sending money back home and would therefore have no incentive to do so. Only earnings would then matter in the household decision.

[20]Note that we sidestep the precise estimation of the structural parameter γ_j related to the taste component whose exact functional form is not specified. Reduced-form parameters on individual characteristics entering this taste component will yet be indicative of additional differences in the non-earnings/remittances determinants of migration.

[21]For instance, Bertoli et al. (2013) resort to the estimation of a nested logit that allows for the correlation of individual unobserved heterogeneity in the propensity to migrate across possible destination countries. In our setting, we could consider an alternative correlation of unobserved heterogeneity in the propensity to choose a given destination across members of the same household.

4.1 Survey design

The main objective of the MIDDAS project was to build an original data set matching representative samples of Senegalese migrants in host countries with their origin household in Senegal, in order to collect accurate information on both "sides" of migration. The project was therefore phased in two successive stages.

First, surveys were conducted among representative samples of Senegalese migrants across four countries in two distinct receiving areas, namely France and Italy for the analysis of South-North migration to Europe, and Mauritania and Côte d'Ivoire for the analysis of South-South migration to Africa. These countries were selected as being the top-two destination countries in their respective area, and the top-four destination countries in the world for Senegalese migrants[22]. Second, migrants' origin households were tracked and interviewed in Senegal, thanks to the contacts provided by the migrants. The migrant questionnaire records in particular precise information on the migrant's socio-demographic characteristics, individual earnings and remittances. The same information was recorded for each resident member of her origin household through the tracking survey[23].

All origin households were tracked, except those of migrants residing in Côte d'Ivoire. We thus focus the following analysis on the French, Italian and Mauritanian migrant-household matched samples. A detailed description of the survey design, sampling methods and fieldwork procedures, as well as information on sample size and composition, can be found in Senne (2013).

4.2 Sample representativeness

Resorting to additional representative data sources, we can assess the performance of our design in achieving representativeness at different levels. Using French and Italian census data, we first show that migrant samples are fairly representative of Senegalese migrant populations in these two host countries. Unfortunately, we were not able to draw the same analysis for Mauritania for lack of reliable data. A more serious concern in our setting is the potential sample selection, in both migrant and household samples, resulting from imperfect matching. Yet, we find that no

[22] According to the 2012 United Nations Database on international migrants' stocks.

[23] As such, remittances were recorded twice. Importantly enough, we did not find any systematic differences between the sent amounts reported by the migrant and the received amounts reported by the resident members of her origin household (Seror, 2012).

systematic difference appears between the matched and unmatched migrant samples, especially regarding our main variables of interest, namely age, gender, formal and koranic education, link to the origin household head, earnings and remittance amounts. Furthermore, using data from the nationally representative PSF household survey conducted in Senegal in 2006 (De Vreyer et al., 2008), we find that our matched migrant households are quite similar to Senegalese migrant households according to their basic characteristics, and in particular their size and demographic composition.

Overall, although some sample selection issues may arise as a result of our survey design, we are nevertheless quite confident with the fact that they will not induce so much bias in the following empirical analysis. A detailed analysis of the representativeness of the matched samples can be found in Senne (2013).

4.3 Descriptive statistics

Household characteristics by migrant's location

Basic characteristics of origin households in Senegal depending on the location of the migrant are shown in Table 2. Observed features globally reflect both historical and current migration patterns to each destination[24]. Households with a migrant in Europe are mostly located around the Dakar region, whereas households with a migrant in Mauritania are unsurprisingly more likely to be settled in the region of the Senegal river valley bordering the country. The relatively high shares of migrants to Europe coming from the North-Eastern and Central parts of Senegal are explained by historical migration flows of Haalpulaar'en and Soninke to France for the former, and the more recent emergence of Louga and Touba as new emigration regions to Italy for the latter. Migrant households are mostly urban: the share of urban households ranges from 64.1% for households with a migrant in Mauritania, to 74.7% for households with a migrant in Europe, compared to 55.2% only for households without migrants, according to the above mentioned PSF survey. The Tijaniyyah brotherhood and the Wolof ethnic group are also dominant. The pooled figure for Europe however conceals the large prevalence of the Murid migration to Italy and the above mentioned and ancient Haalpulaar'en and Soninke migration

[24]In the following empirical analysis, we pool France and Italy due to the similarity in migrants' profiles in the two countries regarding our main variables of interest and particularly earnings and remittances. It further keeps estimation tractable since the size of matched sample for these two countries is smaller than for Mauritania due to lower matching rates in European surveys.

to France.

Heads of migrant households to both destinations are quite similar with respect to their age and gender, i.e they are mostly males around 60 years old. Nonetheless, some remarkable differences emerge with respect to education and earnings. Indeed, the average level of education of heads of household with a migrant in Europe is found to be significantly higher than the one observed among household heads having a migrant in Mauritania. At the same time, the latter are less likely to be employed and their monthly earnings are on average twice lower. This features confirms that households engaging in migration within Sub-Saharan Africa are on average poorer than households with migrants in Europe. The average size of origin households is around 12, with a share of active age adults around 60%. The pool of potential migrant alternatives is therefore large within the household, which further justifies our interest in intra-household selection. Finally, although origin households have on average two (international) migrants, note that 64% of them reported having only one member living abroad.

Individual characteristics by location

The above global picture however eclipses the observed differences in individual characteristics depending on household members' location, shown in Table 3. The non-migrant samples are made of all non-migrant members from migrants' origin households. Migrants are a few years older on average than non-migrants. They are predominantly male, even though we are closer to gender balance in the Mauritanian case. While migrants in France and Italy are much more educated on average than non-migrant members of their origin household, migrants in Mauritania are however more comparable to the members of their origin household with respect to schooling. Koranic education is also much more frequent among migrants at both destinations. More than three quarters of them are children or siblings of the origin household head, what is more, the eldest ones in a vast majority of cases. Finally, note that our survey mostly captures permanent migration since the average duration of stay ranges from 6 years in Mauritania to 12 years in Europe. Part or totality of the costs associated to migration were covered through family funding for respectively 60.5% and 52.1% of the migrants in Europe and Mauritania.

Table 4 provides more detailed information on earnings and remittances. Labour market participation is sensibly higher among migrants in their relevant destination country, reflecting

the high prevalence of labour migration in our survey. PPP earnings of working migrants are unsurprisingly much lower in Mauritania than in France and Italy, while still being higher than PPP earnings of non-migrant members of their origin household in all cases. 77% of migrants in Mauritania and 84% in Europe send remittances to their origin household in Senegal, most of the time on a regular monthly basis. While remittances from Mauritania are significantly lower than remittances from France and Italy, they amount on average to a larger share of migrants' income (around 30% in Mauritania and 15% in France and Italy). Interestingly enough, remittances amounts are quite substantial by comparison with average earnings in Senegal. They represent on average 24% of the monthly earnings of the origin household which is in 93% of the cases the main recipient of the migrant's money transfers. More than 80% of remittances are targeted to the household as whole in order to cover collective expenditures that overwhelmingly refer to daily consumption.

Overall, these features highlight the strength of the link between migrants and their origin household. Moreover, they emphasize the idea that remittances are far from being the mere result of a risk-sharing strategy in the Senegalese context, but rather represent a permanent and highly significant share of the household resources that are earmarked to cover the basic needs of all the members. In this setting, the migration of a member may credibly be a household investment in which remittances are part of a contract on total resources and might therefore be a relevant component entering the collective decision. It lends further credence to our household-based framework in which migration is viewed as a household welfare-maximizing strategy.

5 Econometric specification

We turn in this section to the implementation of our estimation of the extended Roy model of intra-household selection into migration. Following the theoretical framework and procedure described in Section 1.3, estimation proceeds in three main steps. We now discuss in more details empirical specifications and sources of identification at each stage.

5.1 Intra-household selection equation

The reduced-form estimation of the intra-household conditional logit model of location choices from equation (14) is conducted on the whole sample of migrants and non-migrant members from their origin households. The migrant sample is restricted to working-age individuals, i.e aged 18-59, at the date of migration[25]. We similarly define as potential migrants in their origin household all resident members that were aged 18-59 at the time of the surveyed migrant's departure[26]. Two possible destinations are considered: one broad European destination including France and Italy and one Sub-Saharan destination referring to Mauritania. One choice alternative for the household thus corresponds to one (migrant) member abroad, namely in France/Italy or in Mauritania, and all remaining (non-migrant) members in Senegal, so that each household is faced with a varying number of alternatives according to the number of potential migrants within the household.

The dependent variable is a dummy equal to one for the chosen (observed) household members' allocation among all possible (non-observed) allocations. The basic vector x_i of independent variables includes gender, age and education level in three categories: elementary, middle and high school and above, no schooling being the reference. We believe these variables to strongly affect earnings potentials at each location according to the standard Mincer framework. To further test the relevance of our intra-household decision framework resulting in a twofold selection involving earnings and remittances differentials, we enrich the specification with three additional dummies respectively equal to one if the individual attended koranic school, if she is the eldest child of the household head or the eldest sibling of the household head. We indeed expect those variables to have limited explanatory power on earnings but to be relevant determinants of remittances behavior. On the one hand, koranic schooling is expected to capture a higher commitment to prevailing solidarity norms conveyed by the islamic

[25]To focus the analysis on labor migration and ensuing selection within the origin household with respect to earnings and remittances outcomes, we also drop from the sample those migrants who were born in the destination country, those who migrated in order to study abroad and non-working women who migrated for family reasons. This amounts to 3.2% of the sample.

[26]This selection criterion bears some limitations due to potential changes in the household structure since the migration episode. Unfortunately, the data does not allow to reconstruct the exact composition of the origin household that prevailed at the time of migration. We therefore consider as relevant choice alternatives for the household all working-age members at that time, who are still household members at the time of the survey. Our empirical model could then be alternatively interpreted as a model of current location choices based on instantaneous utility comparisons.

religion, in a country where religion remains a strong vector of social control[27]. On the other hand, the other two eldest dummies can account for the fact that first-born children usually take on a greater responsibility in the welfare of the household, in a country where the hierarchy associated with age strongly defines social roles especially within the family[28].

Since all explanatory variables refer to the selected or potential migrant member, they are by definition alternative-specific. Identification is then fully achieved through within-household variations in members' individual characteristics. Yet, we allow parameters on these variables to vary across destinations by specifying a set of interactions with a destination-specific dummy. The reduced-form parameters that are estimated at this stage capture the overall effect of individual characteristics on intra-household selection through earnings differentials, remittances potentials and tastes. The implemented three-step procedure further aims at disentangling the relative role of these channels in the household-based decision.

These first-step estimation results are indeed used to compute appropriate choice probabilities p_{hij} that are added to second-step earnings and remittances equations in order to correct for endogenous selection in a given location. Robust identification at this second stage consequently relies on the inclusion in the first-step selection regression of at least one variable that explains location choices but does not affect earnings nor remittances. Following Munshi (2003) and Pugatch and Yang (2010) studies on Mexican migration to the U.S., we exploit rainfall data as an exogenous source of variation in emigration from Senegal. Indeed, precipitations in origin regions may well affect emigration flows, although their net effect is ambiguous: on the one hand, lower-than-average precipitations may damage local economic conditions and generate or increase incentives to emigrate; on the other hand, the induced negative shock on household income may also negatively impact propensities to emigrate if migration is costly and households are credit-constrained. Due to differential costs of migration to alternative destination countries, rainfall might then additionally affect the choice of a specific location. In the Senegalese context under study, migration to European countries is indeed much more costly than

[27]Using the same dataset,Chort et al. (2012) analyze the influence of solidarity norms conveyed by migrant networks on Senegalese remittances behavior. They point out the significant impact of koranic schooling on both the likelihood and amount of remittances.

[28]The eldest dummies more precisely refer to being the eldest among surveyed siblings in the origin household. Note that eldest dummies could capture the simultaneous effect of being a child or a sibling of the head, apart from being the eldest. The latter links are however reported by the vast majority of migrants (almost 80%). Thus, identification mostly relies on the variability in birth order among the observed part of siblings. On the issue of intergenerational relationships and the role of age and primogeniture in the Senegalese society, see Antoine (2007).

migration to Mauritania, which is a neighbouring country that imposes very few restrictions on Senegalese citizens' circulation. Note that we further contend that rainfall can affect migration from both rural and urban areas through respectively direct and indirect channels. These indirect channels might include for instance increases in food prices due to lower returns in the agricultural sector. Hence and whatever the channel considered and the direction of the effect, we yet expect rainfall to be a potential relevant determinant of the propensity to migrate from Senegal.

Rainfall data are derived from gridded datasets of monthly precipitations that we matched with our household-level data in Senegal thanks to registered GPS coordinates[29]. In practice, we computed variations in levels of rainfall by defining normalized yearly precipitation variables (z-scores) as observed precipitations minus a long term average (1970-2009), divided by the long-term standard deviation. We then include as an additional explanatory variable in the first-step selection equation the average z-score over the 5 years preceding the surveyed migrant's year of departure[30]. Identification is first achieved through both local and yearly variations in precipitations. Besides, since rainfall variations simultaneously affect all members of the same household, z-scores are interacted with individual characteristics to further identify differential effects on the intra-household probability to migrate. We can reasonably argue that rainfall in the home country has no impact on earnings at destination. Moreover, since estimation relies on past levels of precipitations at the time of realized migration, we additionally argue that the defined rainfall variables have no effect neither on *current* earnings in Senegal nor on *current* remittances from abroad[31].

5.2 Earnings and remittances equations

In a second step, we estimate earnings and remittances equations on the samples of migrants and non-migrant members of their origin households, using Mincer-type specifications (9), (10) and (11). We run separate OLS regressions for each of the three locations. Dependent variables are respectively the log of current monthly earnings in Senegal and the log of current monthly earn-

[29]We use data published by the Climate Research Unit of the University of East Anglia: `http://www.cru.uea.ac.uk/fr`. Worldwide and historical rainfall records (1901-2009) are provided at a 5° latitute/longitude resolution.

[30]We therefore additionally assume that the timing of migration is exogenous in our setting.

[31]This last assumption is nonetheless disputable in the case of very recent migration episodes. Our results are however robust to the exclusion of recent migrants - less than 3 years - from the sample.

ings and remittances to the origin household in destination countries[32]. Remittances amounts are expressed in the Senegalese currency (FCFA). For comparison purposes, earnings amounts are expressed in U.S. Purchasing Power Parity (PPP) dollars[33]. This conversion is also needed in the third-step estimation where predicted earnings differentials between locations are allowed to affect intra-household location choices. To expand the range of predicted earnings and remittances, we keep in the sample individuals who reported zero amounts[34].

Basic specifications include as independent variables x_i/x_k gender, age and its square, education level, koranic schooling and eldest dummies. Remittances equations include as additional regressors migrant's earnings at destination and non-migrant's (from her origin household) earnings in the home country. Results from these last specifications will be of particular interest. Indeed, they will serve as a basis to identify those individuals with a higher propensity to remit conditional on earnings, whose characteristics x_{1i} therefore stand as relevant proxies for low values of the δ_i parameter. These specific individual characteristics will further aim at capturing the additional role of remittances potential in the intra-household selection process.

As above mentioned, we finally add to the set of explanatory variables a function $\lambda(p_{hij})$ of choice probabilities obtained from the first step to correct for selection in a given location. In practice, we include a second order polynomial of the predicted first-best choice probability[35]. Since true selection probabilities are unknown, we correct standard errors with bootstrapping at this stage to account for the additional sampling variability induced by substituting estimates.

5.3 Structural model of intra-household location choices

Unbiased parameter estimates from the second step are finally used in a third step to identify the effect of expected earnings and remittances differentials on the probability to be selected as a migrant within the household. Practically, we impute counterfactual earnings and remittances

[32]Monthly earnings in Europe include labor income and social benefits.

[33]We use the conversion factors published by the World Bank in its World Development Indicators. PPP factors for private consumption in 2009 (country currency units buying the same amount of consumption goods as 1 USD in the U.S.) were 0.85 for Italy, 0.92 for France, 143.03 for Mauritania and 307.12 for Senegal (http://data.worldbank.org/indicator/PA.NUS.PRVT.PP).

[34]An issue raised by this sample definition is that we do not properly take into account additional selection on the labour market and into remittances. Bertoli et al. (2013) jointly model individual migration and working decisions. Within our household framework, dealing with both issues would nevertheless add to much theoretical complexity and is empirically hindered by the limited size of our sample.

[35]Note that in our setting, the first-best choice probability is similar to the selection probability for the movers and to a retention probability for the stayers, since it corresponds to the probability that the latter were not selected as migrants within the household.

for each member at each possible location, namely for migrants, had they not migrated or migrated elsewhere, and for non-migrants from their origin household, had they migrated abroad. Imputed earnings differentials between destination and home countries are first included as alternative-specific explanatory variables in the structural form of the conditional logit model from equation (17). Additional control variables include gender, age, koranic schooling, eldest dummies, previously defined rainfall variables and interactions with a destination-specific dummy. They aim at accounting for any non-wage determinants of location choices. As such, they first of all correspond to a flexible specification of tastes, including in particular individual and destination-specific costs or benefits of moving. However, they might equally capture the additional role of remittances potential in the intra-household selection process, so that we will pay peculiar attention to those characteristics that jointly affect the individual propensity to remit. To further disentangle the remittances and tastes channels, we add to the initial specification interaction terms between the latter characteristics and imputed remittances amounts. We again compute bootstrapped standard errors to correct for the extra sampling variability in imputed variables.

Identification at this stage is first achieved through within-household variations in earnings and remittances realizations. Yet, it more crucially depends on the exclusion from the structural model of selection of at least one variable that enters the earnings and remittances equations. We argue here that education level strongly affects earnings and remittances but not location choices, once earnings and remittances are accounted for. This might not be the case if households benefit directly from having an educated member at home, through externalities on other members for instance. Moreover, educated individuals could have strong preferences for moving in destination countries where the average level of education is higher. Finally, migration costs could vary across education levels. However, it is not clear whether the overall non-wage utility gains or losses from the migration of an educated member should be large. We can reasonably argue that the direct effect of education on location choices is negligible compared to its indirect effect through expected earnings and remittances. Both De Vreyer et al. (2010) and Bertoli et al. (2013) indeed find that education plays a limited role in shaping migration decisions once earnings are accounted for. They conclude that selection with respect to education is predominantly explained by expected wage differentials. Although some bias

might remain, the above-specified structural model allows us to consistently identify the set of structural parameters of interest α and β without relying on ad-hoc non-linear functional forms[36].

6 Results

We provide in this section a detailed presentation of estimation results from each step. We particularly focus the analysis on the relative role of earnings and remittances in order to uncover which component of the household structural utility plays a major role in the migration decision.

6.1 Step 1: Intra-household selection

Table 5 reports estimation results from the first-step conditional logit model of location choices in its reduced form, distinguishing both destination alternatives. Specification (1) includes the basic set of individual regressors. Alternative specification (2) includes rainfall z-score interactions as additional determinants of migration decisions.

Estimated coefficients on gender and age are respectively positively and negatively significant for both destination alternatives, though they are slightly larger for the European one. Being a men and relatively younger therefore increases the probability to be selected as a migrant within the household, whatever the chosen location. The lower coefficient of the male dummy for the Mauritanian alternative reflects the higher prevalence of independent female migration to this country. Some differences between locations however emerge with respect to education. Everything being equal, educated individuals (above the elementary level) have higher propensities to be the migrant members of their household in France and Italy than individuals who never went to school. This positive effect is also found to gradually increase with the level of education, a particularly high propensity being observed for household members who have at least a high school degree. Education is yet a weaker determinant of selection in Mauritania. The latter results are indeed true for individuals having a middle school degree and to a lesser extent for members having at least a high school degree, which is consistent with the fact that the average level of education among Senegalese migrants is noticeably higher

[36]Note that we nonetheless rely on the heterogeneity of the effect of remittances to separately identify the parameters related to the earnings, remittances and tastes components in the structural form of the utility.

in European countries. Overall, these findings are in line with the usual results derived from self-selection models. A more striking feature for the matter at hand is that koranic education and eldest dummies are important determinants of intra-household selection into migration. The positive coefficients on these variables are fairly large and highly significant. Another way of interpreting these results is the following: on average, surveyed households tend to favour allocations of their members in which male, eldest children or siblings, koranic and high school educated members migrate to France or Italy, and middle school educated members migrate to Mauritania.

Results from specification (2) show that rainfall deviations from a long term average are additional relevant determinants of migration decisions. Indeed, Wald tests for the joint significance of rainfall variables interacted with individual characteristics, reported at the bottom of Table 5, prove highly significant for both locations. Positive shocks on the level of precipitations seem to accentuate the above described patterns of intra-household selection according to gender and age, and to a lesser extent to education and eldest dummies in Europe. They also seem to foster migration of members with high school education to Mauritania[37]. To the extent that rainfall measured at the time of the migrant's departure can reasonably be excluded from current earnings and remittances equations, individual interactions can work as quite robust instruments for selection in the second step of our estimation procedure.

Note that the reduced-form parameters at this stage identify the joint effect of individual characteristics on intra-household selection through overall differentials in individual earnings, remittances and benefits or costs of migration. The next two steps thus aim at disentangling these channels to determine which component does indeed matter in the household's location decisions.

6.2 Step 2: Earnings and remittances

Tables 6 and 7 provide estimation results from individual earnings and remittances equations at each location. Uncorrected regressions refer to standard Mincer-type specifications, whereas corrected ones additionally include a correction term specified as a second order polynomial

[37]Note that since rainfall simultaneously affects all members of the same household, we cannot draw any clear conclusion about its overall effect on migration. It indeed cannot be identified through our within-household estimation procedure. Only its differential effect according to individual characteristics can be assessed and is anyhow relevant for the matter at hand.

function $\lambda(p_{hij})$ of the first-best selection probability obtained from step 1, specification (2).

Results show quite usual earnings profiles. Men tend to earn significantly more on average than women, with a wage premium of around 130% in each location. We also find positive but decreasing marginal returns to potential experience (for which age is a proxy), although it proves only marginally significant in Mauritania. We finally observe positive returns to education that are somewhat larger in destination countries than in the home country. They are yet markedly lower in African locations, especially in Senegal. A quite interesting result is that returns to education are particularly large in Europe for migrants who have at least a high school degree (around 130%) and in Mauritania for migrants who have a middle school degree (around 80%). To the extent that earnings differentials are taken into account in the household decision, this finding could partly explain why those very individuals are more likely to be selected as migrants within the household to each destination. Since overall samples include individuals with zero earnings, the large point estimates additionally point out the relatively lower participation to the labour market of women and individuals with fewer experience and low diplomas.

Estimation results from remittances equations provide further insights into the underlying rationale for household location decisions. Unsurprisingly, remitted amounts from each destination country are found to increase with migrants' earnings and to decrease with non-migrants' earnings. An increase of 100 PPP US dollars in migrants' income corresponds to an increase in remitted amounts of respectively 20% from Europe and 30% from Mauritania. Men tend to send back to their origin household larger amounts of money, gender differences ranging from 25% in France/Italy to 100% in Mauritania. We find an additional significant positive effect of age in the European sample. Everything being equal, only migrants with education above high school in France/Italy and migrants with middle education in Mauritania are found to remit significantly larger amounts, from respectively 28% to 39% more on average each month. These results are indeed consistent with our household-based framework that puts forward the additional role of remittances potential in the selection process and might additionally account for the observed patterns in household members' migration regarding education.

More crucial results for our issue of interest however emerge from the comparative analysis of earnings and remittances profiles with respect to koranic schooling and the eldest dummies. Indeed, conditional on earnings, the latter characteristics are found to be quite strong and

highly significant determinants of individual remittances behavior. Migrants who attended koranic school remit on average larger amounts than migrants without any koranic education from both France/Italy (+35%) and Mauritania (+53%). The same proves true for eldest children of the origin household head (+32% in Europe and +25% in Mauritania) and to a lesser extent for eldest siblings (+42% in Europe and +45% in Mauritania). However, these variables are simultaneously found to have no impact on individual earnings, whatever the location considered. These important findings are to be analyzed in light of step-1 results that put forward higher within-household migration probabilities to both locations among members having koranic education or being the eldest among children or siblings. They are a first indication that comparative advantages in earnings across members and locations are unable to fully explain observed patterns in intra-household selection into migration, and that expected remittances might be an additional key component in the household decision. In line with our theoretical framework, these results also suggest that koranic and eldest dummies stand for relevant proxies for low values of the δ_i parameter and therefore are appropriate characteristics to further uncover the role of remittances potential in our setting.

Finally, note that corrected and uncorrected coefficients from all regressions are very close in magnitude, suggesting that selection into migration does not induce such a high bias in earnings and remittances equations. This is consistent with the fact that Dahl's correction function is not significant in most specifications. We however reject the null hypothesis for home earnings and remittances from Europe, so that the following results are mainly based on counterfactual earnings and remittances predictions using parameters corrected for selection in each location[38]. Robustness using uncorrected predictions is nevertheless tested. In order to statistically confirm our intuitions from step-1 and step-2 reduced-form regressions, we now turn to the estimation results from the structural discrete model of location choices and assess their consistency with our household-based model for migration decisions.

[38] Quite similar patterns can be found in the previously mentioned studies by Bertoli et al. (2013) and De Vreyer et al. (2010). Robustness to alternative specifications of the correction function was tested. They yield similar results so that we finally kept the one that best fits the data, as suggested by Dahl (2002). Additionally note that all earnings and remittances specifications achieved quite reasonable goodness-of-fit. R^2 indeed range from 20% to 30%.

6.3 Step 3: Structural-form model of intra-household location choices

Table 8 reports the estimation results from the third-step conditional logit model of location choices in its structural form. Consistent with our theoretical framework and on counterfactual earnings and remittances predictions from the previous step, the main independent variables of interest include imputed individual earnings differentials between home and destination countries and imputed remittances amounts.

Specifications in columns (1) and (2) investigate the respective role of expected earnings and other individual characteristics that aim at capturing the non-wage determinants of migration in the household decision. Unsurprisingly, results show that earnings differentials play a major role in shaping intra-household selection patterns. The estimated effect is indeed positive and highly significant across the two specifications, using as independent variables either uncorrected or selectivity-corrected counterfactual earnings predictions. As above mentioned, this finding appears fully relevant to account for the higher migration propensities observed among members with middle and high education to respectively Europe and Mauritania where returns to education at these levels are comparatively larger than in Senegal. Consistent with our theoretical framework, this result therefore suggests that origin households support migration of the member with the highest comparative advantage in earnings across locations in order to maximize the total (earnings) surplus.

Yet, while controlling for earnings, most of the coefficients associated to individual controls are still found to be highly significant. These findings indicate that non-wage components are additional crucial determinants of intra-household migration decisions. First, men and relatively younger household members have higher probabilities to be located abroad. More importantly, the same pattern is observed for individuals with koranic education and who are the eldest among children and siblings of the origin household head. Besides, the point estimates of the effect of the latter characteristics are fairly large in comparison to the other dummy variables. In line with previously estimated remittances equations, this central result points out the fact that, conditional on earnings, those very individuals with a higher propensity to remit, reflecting low values of δ_i, have a higher probability to be selected as migrants within the household. This result is furthermore fully consistent with predictions from our household-based model for migration decisions.

However, two important limitations might challenge the latter interpretation. First, as extensively discussed in the previous sections, the reduced-form parameters on individual characteristics capture at this stage a potential simultaneous effect of tastes, which may confound the remittances potential channel we bring up. More concretely, a possible alternative interpretation of the observed patterns of selection with respect to koranic and eldest variables is that these individual characteristics basically account for differential costs or benefits of moving abroad[39]. For instance, religious networks at destination might well lower the cost of migrating and therefore foster migration of members having koranic schooling through the spread of information and the provision of assistance upon arrival or in the course of their stay. Besides, as eldest members usually take on a greater responsibility in the household decision, they might credibly "self-select" or be selected as the first link in the migration chain.

Hence, to further disentangle the relative role of remittances potentials and tastes in the selection process, we add to the set of explanatory variables imputed remittances amounts as well as interaction terms with koranic and eldest dummies. Indeed, as suggested by our theoretical model, to the extent that the latter variables stand as relevant proxies for low values of the remittances discount factor δ_i and therefore accurately capture the additional role of remittances potential, their effect positively interacts with remitted amounts. We should thus empirically observe some heterogeneity in the effect of remittances with respect to those individuals' characteristics: the lower δ_i, the higher the propensity to remit and the higher the role of remittances in determining location choices. Results are given in Table 8, columns (3) and (4) of . On the one hand, all the coefficients on remitted amounts and the relevant interaction terms are found to be positive and (highly) significant. On the other hand, point estimates of the direct effect of koranic and eldest dummies substantially drop and become marginally significant[40]. Overall, even if we cannot rule out that part of the effect of our variables of interest goes through differential tastes, these additional results lend further support to the hypothesis

[39]Indeed, Bertoli et al. (2013) point out the fact that the inverse taste component $-x_i\phi_j$ of the household utility could alternatively be interpreted as the *net cost* of migration to destination j, which is allowed to vary according to migrant characteristics. In our setting, it includes costs at both household and migrant levels. We do not report coefficients on rainfall variables since they present similar patterns to those from step-1 reduced-form specification.

[40]Note that conditional on remittances, gender and age still are significant determinants of migration to France/Italy. Moreover, their coefficients remain unchanged, which may reflect the lower implicit cost of young male migration to Europe. This also proves true for age in the Mauritanian alternative. The drop in size and significance of the male dummy may however be explained by the higher propensity to remit of male migrants in Mauritania. We nevertheless choose to focus the analysis on koranic and eldest dummies since they have no effect on earnings.

that the higher propensities to be located abroad of members having koranic education and being the eldest among children or siblings are mainly explained by their higher remitted amounts.

A second limitation refers to the potential confounding role of welfare weights in our setting. Indeed, additional predictions derived from our theoretical model suggest that migration propensities increase with bargaining powers within the household, so that the above stated patterns might alternatively be explained through this channel. This is particularly true regarding the eldest dummies since age is a fundamental determinant of social status in the Senegalese society. However, two elements credibly rule out this alternative interpretation. First, in our theoretical framework, expected remittances are expected to decrease with migrants' bargaining power. Therefore, if the koranic and eldest dummies were only proxy measures of individuals' bargaining power, they should negatively affect remitted amounts, which is not what we observe from the remittances equations estimated in Table 7. Second, as mentioned in the previous sections, differential welfare weights within the household translate into additional heterogeneity in the migration decision, especially according to the effect of earnings differentials across locations. Estimations shown in Table 9 formally test this hypothesis by adding interaction terms between predicted earnings differential and two relevant proxies for bargaining power, namely age and gender. Results show that no significant heterogeneity can be found along these two dimensions, allowing us to plausibly state that the effect of differential bargaining powers is quite negligible in our setting[41].

In line with the main predictions derived from our theoretical model, our results therefore support the idea that both earnings and remittances differentials play a major role in shaping household members' location choices. These findings have strong implications in terms of migrant selection within the household: households are not only to select as migrants the members with the highest comparative advantages in earnings across locations, but also those with the highest remittances potentials conditional on earnings. This very last implication stands in striking contrast with the usual predictions from individual self-selection models but appears fully consistent with the higher propensities to migrate observed among members having koranic education or being the eldest among children or siblings of the origin household head, despite no obvious comparative advantage with respect to earnings.

[41]To keep estimation tractable, we only test heterogeneity with respect to earnings differential. Indeed, additionally testing this hypothesis with respect to remittances potential would imply the inclusion of poorly identified triple interactions in the relevant specifications.

7 Conclusion

Although tackled by a large number of papers, migrant selection has always been modeled as the result of an individual income-maximizing strategy. However, individual selection models cannot account for most migration patterns observed in particular in developing countries where migration is part of a household welfare-maximizing strategy. Therefore, this paper aims at shedding a new light on the selection process of migrants by investigating the so far under-explored issue of intra-household selection into migration.

We first extend the seminal Roy model of self-selection to account for a household decision process for migration. In this framework, households base the location choices of their members on the maximization of a collective utility whose components include earnings of non-migrant members in the home country but also earnings and remittances from migrant members abroad. Using observed allocation choices of household members, we develop a three-step estimation procedure to estimate the weight on each component in the structural intra-household selection decision. We provide an empirical application using a unique matched sample of Senegalese migrants in France, Italy and Mauritania and their origin household in Senegal.

Our results show that together with earnings, remittances differentials play a significant role in shaping intra-household selection patterns. We find that, controlling for earnings differential, households are more likely to select into migration the ones among their members who have the highest remittances potential in order to maximize the collective welfare. These results stand in striking contrast with the results derived from usual individual-level selection models which do not account for the collective dimension of the migration decision and neglect the role played by future remittances in shaping migration patterns. Yet our framework appears relevant to explain the observed higher propensities to migrate among individuals having koranic education or being the eldest among siblings, despite no comparative advantages in earnings.

In our setting, the question of the exact benefits of migration for the migrants themselves however remains an open issue. More could indubitably be learnt from a dynamic analysis of migrants' individual trajectories in the long run.

Tables

Table 1: Sample size and composition by country

	France	Italy	Mauritania	Pooled
Stage 1: Migrant samples				
Number of eligible migrants	579	616	402	1,597
Refusal rate (%)	48.2	51.0	18.9	41.9
Number of surveyed migrants	300	302	326	928
...% of women	*24.3*	*22.9*	*36.5*	*28.1*
...% in capital/main cities	*72.3*	*48.0*	*73.0*	*64.0*
Stage 2: Origin household samples				
Number of provided contacts	158	114	266	538
Matching rate (%)				
... overall	*30.7*	*20.5*	*53.4*	*35.3*
... among provided contacts	*58.2*	*54.4*	*65.4*	*61.0*
Number of tracked households	92	62	172	326
...% in Dakar	*46.7*	*54.8*	*21.3*	*34.8*

Source: MIDDAS Survey, 2009 - 2010. Authors' calculation.

Table 2: Origin household characteristics by migrant's location

	France/Italy	Mauritania
Household characteristics		
Region (%)		
...Dakar	53.9	22.1
...North/East	18.1	45.3
...South	3.9	4.7
...Center	24.1	27.9
Environment (%)		
...Urban	74.7	64.1
...Rural	25.3	35.9
Composition (%)		
...Children (18-)	33.6	38.4
...Adults (18-60)	58.9	56.1
...Elderly (60+)	7.5	5.5
Size	12.8	11.2
Number of international migrants	2.19	1.59
Household head characteristics		
Age	58.3	58.4
Gender (%)		
...Male	62.9	64.0
...Female	37.1	36.0
Ethnic group (%)		
...Wolof	40.2	61.0
...Serere	10.4	9.3
...Peul	16.9	15.1
...Soninke/Mandinka	24.7	1.7
...Diola	4.5	9.3
...Other	3.3	3.5
Religion (%)		
...Murid	28.6	26.9
...Tijani	41.6	61.1
...Other	29.8	12.0
Schooling (%)		
...No schooling	45.2	64.0
...Primary	18.9	21.5
...Middle School	18.3	9.3
...High School and more	17.6	5.2
Labour status (%)		
...Unemployed/Non-working	41.3	51.3
...Working	58.7	48.7
Monthly earnings (XOF)	154,040.2	82,545.9
Monthly earnings (PPP)	501.6	268.8
Observations	146	164

Note: Earnings in XOF for Senegal (659 XOF = 1 euro). PPP refers to USD Purchasing Power Parity amounts, using the consumption conversion factor published by the World Bank (2009).
Source: MIDDAS Survey, 2009 - 2010. Authors' calculation.

Table 3: Individual characteristics by migrant's location

	France/Italy			Mauritania		
	Non-migrants	Migrants	t/χ^2	Non-migrants	Migrants	t/χ^2
Age	43.8	37.6	-6.2***	38.6	36.5	-2.1*
Gender (%)						
...Male	41.2	79.2	70.1***	41.9	65.7	31.6***
...Female	39.8	20.8		58.1	34.3	
Schooling level (%)						
...No schooling	40.7	14.9	64.5***	54.3	38.6	27.2***
...Primary	21.6	16.9		29.1	28.7	
...Middle School	17.5	18.4		8.5	19.8	
...High School and more	20.2	49.8		8.1	12.9	
Koranic schooling (%)						
...Only	10.1	13.6	3.5**	23.1	30.2	7.1**
...Some	52.4	63.1	10.7***	51.8	64.3	12.5***
Link to household head (%)						
...Son/daughter	24.7	62.9	134.8***	29.8	58.4	165.2***
...Brother/sister	6.3	13.6		4.2	22.7	
...Head/spouse	35.2	3.9		32.4	5.8	
...Other	33.8	19.6		33.6	13.1	
Eldest (%)						
...Son/daughter	35.8	64.2	28.4***	41.3	59.7	18.4***
...Brother/sister	27.2	72.8	45.6***	25.4	74.6	49.2***
Migration funding (%)						
...Family	/	60.5		/	52.1	
...Own savings only	/	18.9		/	23.6	
...Other channel only	/	20.6		/	24.3	
Migration duration	/	12.1		/	6.2	
Observations	568	146		716	164	

Notes: Sample restricted to individuals aged 18-59 at the time of migrant's departure. Samples of non-migrants are composed of non-migrant members from migrant households. χ^2 test for the equality of distributions for categorical variables, t-test for the equality of means for continuous variables between non-migrant and migrant samples.
Source: MIDDAS Survey, 2009 - 2010. Authors' calculation.

Table 4: Individual earnings and remittances by migrant's location

	France/Italy			Mauritania		
	Non-migrants	Migrants	t/χ^2	Non-migrants	Migrants	t/χ^2
Earnings						
Labour status (%)						
...Unemployed/Non-working	46.9	20.1	25.1***	44.8	15.7	49.4***
...Working	53.1	79.9		55.2	84.3	
Monthly earnings	92,690.6	1,255.7	/	59,048.7	78,326.4	/
Monthly earnings (PPP)	301.8	1,420.5	1,118.7***	192.3	407.8	215.5***
Remittances						
Propensity (%)						
...to any household	/	87.1		/	79.1	
...to origin household	/	84.4		/	76.7	
Frequency (%)						
...Monthly	/	63.8		/	59.7	
...Bimonthly/Quarterly	/	10.3		/	12.7	
...Less frequently	/	25.9		/	28.6	
Use (%)						
...Daily consumption	/	83.6		/	88.1	
...Education/Health	/	10.6		/	5.6	
...Other	/	5.8		/	6.3	
Targeted expenditures (%)						
...Collective	/	79.5		/	84.1	
...Private	/	11.1		/	11.8	
...Both	/	9.4		/	4.1	
Monthly remittances (XOF)	/	141,701.6		/	39,689.5	
Observations	568	146		716	164	

Notes: Sample restricted to individuals aged 18-59 at the time of migrant's departure. Sample of non-migrants are composed of non-migrant members from migrant households. χ^2 test for the equality of distributions for categorical variables, t-test for the equality of means for continuous variables between non-migrant and migrant samples. Earnings are expressed in euros for France and Italy, in XOF for Senegal (656 XOF = 1 euro) and in MRO for Mauritania (388 MRO = 1 euro). PPP refers to USD Purchasing Power Parity amounts, using the consumption conversion factor published by the World Bank (2009).
Source: MIDDAS Survey, 2009 - 2010. Authors' calculation.

Table 5: Intra-household location choices - Reduced-form conditional logit estimates

	Without rainfall		With rainfall	
	France/Italy	Mauritania	France/Italy	Mauritania
	(1)		(2)	
Male (d)	1.018***	0.418**	1.472***	0.615***
	(0.243)	(0.210)	(0.368)	(0.309)
Age	-0.026***	-0.038***	-0.029**	-0.047***
	(0.010)	(0.010)	(0.014)	(0.013)
Elementary school (d)	0.477	0.390	0.423	0.526
	(0.356)	(0.279)	(0.442)	(0.367)
Middle school (d)	1.420***	0.866***	1.252**	0.587**
	(0.375)	(0.321)	(0.504)	(0.281)
High school and more (d)	2.667***	0.610*	3.530***	0.912*
	(0.356)	(0.369)	(0.602)	(0.542)
Koranic school (d)	0.713**	1.754***	0.598**	2.785***
	(0.289)	(0.324)	(0.297)	(0.414)
Oldest child (d)	1.185***	1.033***	1.574***	1.007***
	(0.210)	(0.216)	(0.312)	(0.318)
Oldest brother/sister (d)	1.297***	1.690***	1.221***	1.733***
	(0.364)	(0.312)	(0.444)	(0.381)
Rainfall z-score x Male			0.854**	0.728**
			(0.425)	(0.361)
Rainfall z-score x Age			-0.013*	-0.026**
			(0.007)	0.011)
Rainfall z-score x Elementary			-0.209	-0.321
			(0.392)	(0.458)
Rainfall z-score x Middle			-0.315	-0.296
			(0.458)	(0.452)
Rainfall z-score x High			1.095*	1.737**
			(0.659)	(0.698)
Rainfall z-score x Koranic			0.436*	0.111
			(0.256)	(0.387)
Rainfall z-score x Oldest child			0.656*	0.122
			(0.374)	(0.383)
Rainfall z-score x Oldest brother/sister			-0.157	-0.281
			(0.492)	(0.424)
Destination dummy	yes		yes	
Observations	1,594		1,567	
Wald test for joint significance of rainfall variables			31.02***	19.22**
p-value			0.006	0.046

Notes: Sample is restricted to individuals aged 18-59 at the time of migrant's departure. Dependent variable is a dummy equal to 1 if member i of household h lives in country j. (d) stands for dummy variables. Reference category for education is no schooling. Rainfall z-scores refer to yearly deviations from the 1970-2009 trend period and correspond to the average z-score over the five years previous to the reported date of migration. Coefficients reported, standard errors in brackets.
*p<0.10; ** p<0.05; ***p<0.01
Source: MIDDAS Survey, 2009 - 2010. Authors' calculation.

Table 6: Individual earnings equations - OLS estimates

Dependent variable:	Senegal		France/Italy		Mauritania	
Log of monthly earnings, PPP	Uncorrected (1)	Corrected (2)	Uncorrected (3)	Corrected (4)	Uncorrected (5)	Corrected (6)
Male (d)	1.541***	1.579***	1.347**	1.258**	1.314***	1.283***
	(0.226)	(0.198)	(0.672)	(0.623)	(0.451)	(0.362)
Age	0.245***	0.297***	0.362**	0.394**	0.121*	0.136*
	(0.042)	(0.061)	(0.172)	(0.197)	(0.069)	(0.078)
Age squared (/100)	-0.286***	-0.322***	-0.483*	-0.496*	-0.067	-0.075
	(0.067)	(0.039)	(0.275)	(0.295)	(0.112)	(0.114)
Elementary school (d)	0.495***	0.568***	0.621	0.588	0.479*	0.385*
	(0.158)	(0.197)	(0.954)	(0.895)	(0.283)	(0.228)
Middle school (d)	0.426**	0.456**	0.665*	0.592*	0.934**	0.823***
	(0.212)	(0.221)	(0.387)	(0.346)	(0.469)	(0.309)
High school and more (d)	0.254*	0.174*	1.403***	1.250**	0.327	0.509
	(0.141)	(0.104)	(0.519)	(0.611)	(0.478)	(0.563)
Koranic schooling (d)	0.094	0.033	-0.301	-0.292	-0.326	-0.447
	(0.122)	(0.146)	(0.543)	(0.524)	(0.417)	(0.513)
Oldest child (d)	0.113	0.145	0.226	0.311	0.317	0.208
	(0.241)	(0.269)	(0.452)	(0.624)	(0.296)	(0.163)
Oldest brother/sister (d)	0.222	0.315	0.359	0.265	0.104	0.071
	(0.432)	(0.468)	(0.658)	(0.567)	(0.165)	(0.159)
Constant	-3.399***	-4.568***	-3.457**	-3.864**	-3.962**	-4.112**
	(0.758)	(0.871)	(1.612)	(1.804)	(1.917)	(2.025)
First-best probability		1.887**		-1.658*		-1.122*
		(0.947)		(1.006)		(0.677)
First-best probability^2		-1.492*		1.915		0.956
		(0.894)		(2.154)		(0.789)
Observations	1,248	1,248	141	141	160	160
R^2	0.27	0.30	0.27	0.29	0.21	0.24
Wald test for $\lambda(p_{hij})$		7.52**		3.25		4.21
p-value		0.03		0.17		0.12

Notes: Samples are restricted to individuals aged 18-59 at the time of migrant's departure. (d) stands for dummy variables. Reference category for education is no schooling. PPP refers to USD Purchasing Power Parity amounts, using the consumption conversion factor published by the World Bank (2009). Coefficients reported, bootstrapped standard errors in brackets (1000 replications) for corrected specifications.
*p<0.10; ** p<0.05; ***p<0.01
Source: MIDDAS Survey, 2009 - 2010. Authors' calculation.

Table 7: Individual remittance equations - OLS estimates

Dependent variable:	France/Italy		Mauritania	
Log of monthly remittances, FCFA	Uncorrected (1)	Corrected (2)	Uncorrected (3)	Corrected (4)
Migrant's earnings (/100)	0.174***	0.196***	0.278***	0.321***
	(0.000)	(0.000)	(0.001)	(0.001)
Origin household's earnings (/100)	-0.005**	-0.004**	-0.019*	-0.033*
	(0.001)	(0.001)	(0.011)	(0.017)
Male (d)	0.158*	0.247*	1.105**	1.092**
	(0.093)	(0.146)	(0.551)	(0.528)
Age	0.043**	0.040*	0.025	0.024
	(0.021)	(0.024)	(0.021)	(0.022)
Age squared (/100)	-0.102	-0.098	-0.154	-0.178
	(0.286)	(0.292)	(0.223)	(0.239)
Elementary school (d)	0.326	0.422	0.487	0.389
	(0.571)	(0.607)	(0.625)	(0.597)
Middle school (d)	0.431	0.396	0.362*	0.389*
	(0.582)	(0.457)	(0.217)	(0.212)
High school and more (d)	0.245*	0.283*	0.126	0.159
	(0.134)	(0.156)	(0.257)	(0.284)
Koranic schooling (d)	0.312**	0.348**	0.497**	0.526**
	(0.139)	(0.162)	(0.241)	(0.247)
Oldest child (d)	0.295*	0.321**	0.224**	0.248**
	(0.163)	(0.156)	(0.111)	(0.123)
Oldest brother/sister (d)	0.394	0.415*	0.436*	0.452*
	(0.229)	(0.247)	(0.259)	(0.271)
Constant	-2.128*	-2.156*	-1.057*	-0.894*
	(1.252)	(1.283)	(0.587)	(0.509)
First-best probability		-0.954**		-0.687
		(0.465)		(0.663)
First-best probability^2		1.257		0.879
		(1.356)		(1.102)
Observations	138	138	157	157
R^2	0.26	0.29	0.22	0.24
Wald test for $\lambda(p_{hij})$		5.18*		3.96
p-value		0.08		0.13

Notes: Samples are restricted to individuals aged 18-59 at the time of migrant's departure. (d) stands for dummy variables. Reference category for education is no schooling. Earnings are expressed in PPP and refer to monthly amounts. PPP refers to USD Purchasing Power Parity amounts, using the consumption conversion factor published by the World Bank (2009). Coefficients reported, bootstrapped standard errors in brackets (1000 replications) for corrected specifications.
*p<0.10; ** p<0.05; ***p<0.01
Source: MIDDAS Survey, 2009 - 2010. Authors' calculation.

Table 8: Intra-household location choice - Structural-form conditional logit estimates

	Uncorrected	Corrected	Uncorrected	Corrected
	(1)	(2)	(3)	(4)
Earnings differential, PPP (/100)	0.356***	0.236***	0.321***	0.215***
	(0.079)	(0.058)	(0.055)	(0.046)
Remittances, FCFA (/10,000)			0.198*	0.176*
			(0.115)	(0.103)
Remittances × Koranic			0.126***	0.118**
			(0.042)	(0.055)
Remittances × Oldest child			0.092**	0.076**
			(0.039)	(0.034)
Remittances × Oldest brother/sister			0.109*	0.096*
			(0.064)	(0.058)
France/Italy				
Male (d)	0.327**	0.395*	0.292**	0.338*
	(0.162)	(0.236)	(0.139)	(0.199)
Age	-0.012**	-0.016*	-0.010**	-0.013*
	(0.006)	(0.009)	(0.005)	(0.007)
Koranic school (d)	0.764**	0.628**	0.255**	0.276*
	(0.373)	(0.291)	(0.122)	(0.142)
Oldest child (d)	0.976***	1.271***	0.523*	0.607
	(0.374)	(0.485)	(0.311)	(0.456)
Oldest brother/sister (d)	1.078**	0.969**	0.651*	0.595*
	(0.532)	(0.478)	(0.394)	(0.307)
Mauritania				
Male (d)	0.222**	0.340*	0.116*	0.195
	(0.112)	(0.169)	(0.068)	(0.119)
Age	-0.035**	-0.042***	-0.031**	-0.046**
	(0.015)	(0.016)	(0.014)	(0.019)
Koranic school (d)	1.576***	2.058***	0.352	0.393
	(0.597)	(0.785)	(0.298)	(0.287)
Oldest child (d)	0.913***	1.302***	0.552*	0.619*
	(0.347)	(0.497)	(0.328)	(0.365)
Oldest brother/sister (d)	1.349***	1.413***	0.423*	0.486*
	(0.481)	(0.509)	(0.236)	(0.279)
Rainfall variables	no	yes	no	yes
Destination dummy	yes	yes	yes	yes
Observations	1,594	1,567	1,594	1,567

Notes: Sample is restricted to individuals aged 18-59 at the time of migrant's departure. Dependent variable is a dummy equal to 1 if member i of household h lives in country j. (d) stands for dummy variables. PPP refers to USD Purchasing Power Parity amounts, using the consumption conversion factor published by the World Bank (2009). Coefficients reported, bootstrapped standard errors in brackets (1000 replications).
*p<0.10; ** p<0.05; ***p<0.01
Source: MIDDAS Survey, 2009 - 2010. Authors' calculation.

Table 9: Intra-household location choice - Structural-form conditional logit estimates with unequal bargaining powers

	Uncorrected	Corrected	Uncorrected	Corrected
	(1)	(2)	(3)	(4)
Earnings differential, PPP (/100)	0.398***	0.264***	0.364***	0.245***
	(0.091)	(0.072)	(0.083)	(0.066)
Earnings differential × Age	0.031	0.022		
	(0.023)	(0.016)		
Earnings differential × Male			0.126	0.108
			(0.139)	(0.114)
Individual controls	yes	yes	yes	yes
Rainfall variables	no	yes	no	yes
Destination dummy	yes	yes	yes	yes
Observations	1,594	1,567	1,594	1,567

Notes: Sample is restricted to individuals aged 18-59 at the time of migrant's departure. Dependent variable is a dummy equal to 1 if member i of household h lives in country j. (d) stands for dummy variables. PPP refers to USD Purchasing Power Parity amounts, using the consumption conversion factor published by the World Bank (2009). Individual controls include gender, age, koranic schooling and eldest dummies. Coefficients reported, bootstrapped standard errors in brackets (1000 replications).
*p<0.10; ** p<0.05; ***p<0.01
Source: MIDDAS Survey, 2009 - 2010. Authors' calculation.

Appendices

A Optimal amount of remittances

Any optimal amount of remittances sent by the migrant member to her origin household to maximize the total household utility function should satisfy the following f.o.c:

$$\frac{\partial U_{hij}}{\partial R_{ij}} = \sum_{k \neq i} \frac{\theta_k}{n-1} U_k^{'} \left(Y_{ks} + \frac{R_{ij}^*}{n-1} \right) - \theta_i \delta_i U_i^{'} \left(Y_{ij} - \delta_i R_{ij}^* \right) = 0$$

To further determine how the optimal amount R_{ij}^* varies with other exogenous parameters in the model, we can simply differentiate the above equality with respect to each component:

$$\begin{aligned} & \sum_{k \neq i} \left[\frac{1}{n-1} U_k^{'}(.) \right] d\theta_k + \sum_{k \neq i} \left[\frac{\theta_k}{n-1} U_k^{''}(.) \right] dY_{ks} + \sum_{k \neq i} \left[\frac{\theta_k}{(n-1)^2} U_k^{''}(.) \right] dR_{ij}^* \\ = & \left[\delta_i U_i^{'}(.) \right] d\theta_i + \left[\theta_i (U_i^{'}(.) - \delta_i R_{ij}^* U_i^{''}(.)) \right] d\delta_i + \left[\theta_i \delta_i U_i^{''}(.) \right] dY_{ij} - \left[\theta_i \delta_i^2 U_i^{''}(.) \right] dR_{ij}^* \end{aligned}$$

For any concave and twice differentiable individual utility functions, implying $U^{'}(.) > 0$ and $U^{''}(.) < 0$, it is then straightforward to show that:

$$[+]dR_{ij}^* = \sum_{k \neq i} [+]d\theta_k \quad \sum_{k \neq i} [-]dY_{ks} \quad [-]d\theta_i \quad [+]dY_{ij} \quad [-]d\delta_i$$

so that R_{ij}^* can be depicted by the following function of all exogenous parameters:

$$R_{ij}^* = R_{ij}^*(\theta_k^+, \theta_i^-, Y_{ks}^-, Y_{ij}^+, \delta_i^-) \qquad \forall i, j \quad \text{and} \quad k \neq i$$

B Optimal value of the household utility function

The (remittance) optimal value of the collective utility function for household h of locating member i in destination country j and all other members in the home country s writes:

$$V_{hij} = \sum_{k \neq i} \theta_k U_k \left(Y_{ks} + \frac{R_{ij}^*}{I-1} \right) + \theta_i U_i \left(Y_{ij} - \delta_i R_{ij}^* \right) + T_{hij}$$

The envelope theorem states that marginal changes in the optimal value of a function with respect to exogenous parameters of that function can be accurately described by partially differentiating the objective function evaluated at its optimum. For any concave and twice differentiable individual utility functions, implying $U^{'}(.) > 0$ and $U^{''}(.) < 0$, and conditional on tastes, it is then straightforward to show that:

$$\begin{cases} \partial V_{hij}/\partial \theta_k = U_k(.) > 0 & ; \quad \partial V_{hij}/\partial \theta_i = U_i(.) > 0 \\ \partial V_{hij}/\partial Y_{ks} = \theta_k U_k^{'}(.) > 0 & ; \quad \partial V_{hij}/\partial Y_{ij} = \theta_i U_i^{'}(.) > 0 \\ \partial V_{hij}/\partial \delta_i = -\theta_i R_{ij}^* U_i^{'}(.) < 0 & \end{cases}$$

so that V_{hij} can be depicted by the following function of all exogenous parameters:

$$V_{hij} = V_{hij}(\theta_k^+, \theta_i^+, Y_{ks}^+, Y_{ij}^+, \delta_i^-) \qquad \forall i, j \quad \text{and} \quad k \neq i$$

C Sample representativeness tables

Table C.1: Migrant samples' representativeness by country - Comparison with OECD data

		France		Italy	
		Census	MIDDAS	Census	MIDDAS
Gender (%)	Men	54.7	75.5	88.1	77.3
	Women	45.3	25.5	11.9	22.7
Age (%)	20-29 year	20.1	27.6	15.4	23.4
	30-39 years	22.3	35.0	49.4	40.5
	40-49 years	25.0	21.6	29.2	30.4
	50-60 years	20.4	12.6	4.4	5.7
	60+ years	12.2	3.2	1.6	0.0
Duration of stay (%)	up to 5 years	17.5	14.8	29.2	18.9
	5 to 10 years	12.2	33.8	26.7	35.1
	10+ years	70.3	51.4	44.2	46.0
Citizenship (%)	National	58.6	25.5	1.6	2.3
	Other country	41.4	74.5	98.4	97.7
Education (%)	ISCED 0/1/2	45.1	54.6	83.8	48.5
	ISCED 3/4	26.9	20.3	12.3	20.1
	ISCED 5/6	28.0	25.2	3.9	26.4
Labor force status (%)	Employed	54.8	74.8	79.5	70.5
	Unemployed	12.6	14.1	9.1	21.2
	Inactive	32.6	11.1	11.4	8.3
Observations		93,076	286	28,030	299

Notes: OECD census data records information on all individuals born in Senegal, aged 20 and above and living in an OECD country. MIDDAS sample is restricted to this sub-population population for comparison purpose. ISCED refers to the International Standard Classification of Education of UNESCO. ISCED 0/1/2 corresponds to no formal education, primary and lower secondary education; ISCED 3/4 to upper secondary, vocational and technical education; ISCED 5/6 to tertiary education.
Source: DIOC 2005/06, OECD and MIDDAS Survey, 2009 - 2010. Authors' calculation.

Table C.2: Probit analysis of matching success

	France (1)	Italy (2)	Mauritania (3)	Pooled (4)
Migrant characteristics				
Age	0.013***	0.004	0.008*	0.009***
	(0.005)	(0.003)	(0.004)	(0.002)
Gender				
Male	0.016	0.043	-0.022	0.000
	(0.078)	(0.049)	(0.076)	(0.044)
Female	(ref)	(ref)	(ref)	(ref)
Ethnic group				
Wolof	-0.151**	-0.105	0.198**	-0.016
	(0.069)	(0.074)	(0.080)	(0.045)
Peul	-0.082	-0.020	0.063	-0.043
	(0.074)	(0.075)	(0.102)	(0.053)
Other	(ref)	(ref)	(ref)	(ref)
Religion/Brotherhood				
Murid	0.090	-0.064	0.109	0.014
	(0.154)	(0.143)	(0.127)	(0.080)
Tidjan	-0.015	-0.054	0.141	0.030
	(0.139)	(0.107)	(0.125)	(0.080)
Other muslim	0.015	0.206	0.166	0.062
	(0.133)	(0.264)	(0.127)	(0.081)
Other	(ref)	(ref)	(ref)	(ref)
Koranic schooling				
Yes	0.151	-0.025	0.063	0.085
	(0.091)	(0.067)	(0.089)	(0.063)
No	(ref)	(ref)	(ref)	(ref)
Formal schooling				
Elementary/Secondary	0.014	0.016	0.087	0.027
	(0.087)	(0.081)	(0.071)	(0.047)
Highschool/University	0.012	0.055	-0.064	0.027
	(0.093)	(0.084)	(0.120)	(0.056)
Vocational	-0.016	0.142	0.229*	0.056
	(0.138)	(0.172)	(0.134)	(0.085)
No formal schooling	(ref)	(ref)	(ref)	(ref)
Marital Status				
Married	-0.069	0.000	-0.062	-0.046
	(0.077)	(0.061)	(0.085)	(0.048)
Divorced/Widowed	-0.101	-0.099**	-0.068	-0.108*
	(0.095)	(0.041)	(0.132)	(0.062)
Single	(ref)	(ref)	(ref)	(ref)
Time since arrival	-0.010**	0.003	-0.006	-0.006**
	(0.004)	(0.004)	(0.005)	(0.003)
Place of residence				
Capital/Main cities	-0.052	-0.077*	0.068	-0.014
	(0.072)	(0.044)	(0.076)	(0.038)
Small cities	(ref)	(ref)	(ref)	(ref)
Labor status				
Working	0.105	0.346**	0.067	0.127
	(0.137)	(0.162)	(0.171)	(0.091)
Unemployed	0.221	0.125	0.266	0.240
	(0.176)	(0.112)	(0.182)	(0.187)
Non-working	(ref)	(ref)	(ref)	(ref)
Income				
2nd quartile	0.155	0.124*	0.149	0.139*
	(0.102)	(0.069)	(0.093)	(0.081)
3rd quartile	-0.052	0.125	0.110	0.086
	(0.110)	(0.127)	(0.101)	(0.065)
4th quartile	-0.010	0.091	0.252***	0.164**
	(0.110)	(0.132)	(0.095)	(0.070)
Missing	0.080	0.169	-0.336	0.025
	(0.201)	(0.207)	(0.298)	(0.119)
1st quartile	(ref)	(ref)	(ref)	(ref)

Table 4 (continued)

	France (1)	Italy (2)	Mauritania (3)	Pooled (4)
Origin household characteristics				
Environment				
Rural	0.089	-0.052	0.023	0.060
	(0.082)	(0.050)	(0.067)	(0.044)
Urban	(ref)	(ref)	(ref)	(ref)
Size	0.004	0.004	0.006	0.005**
	(0.003)	(0.003)	(0.005)	(0.002)
Missing size				
Yes	0.084	-0.060	-0.228	-0.046
	(0.224)	(0.068)	(0.294)	(0.115)
No	(ref)	(ref)	(ref)	(ref)
Resident spouse/child				
Yes	-0.098	-0.035	0.127*	0.016
	(0.075)	(0.048)	(0.072)	(0.042)
No	(ref)	(ref)	(ref)	(ref)
Wealth score	0.008	0.038***	0.010	0.029***
	(0.017)	(0.010)	(0.026)	(0.010)
Remittances in cash/kind				
Yes	-0.104	-0.036	0.033	0.006
	(0.111)	(0.065)	(0.082)	(0.048)
No	(ref)	(ref)	(ref)	(ref)
Remittances amounts (in euros)	0.002	0.001	0.001	0.001
	(0.002)	(0.001)	(0.005)	(0.001)
Country				
Italy				-0.117**
				(0.051)
Mauritania				0.255***
				(0.056)
France				(ref)
Observations	300	302	326	928

Note: Marginal effects at the mean for continuous variables, at 0 for dummy variables. Robust standard errors in brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: MIDDAS survey, 2009-2010. Authors' calculation.

Table C.3: Origin household samples' representativeness by migrant's location - Comparison with PSF survey

	PSF with migrants		MIDDAS				
	Out of Africa	In Africa	France/Italy	Mauritania	χ^2/Fisher/t		PSF
	(1)	(2)	(3)	(4)	(1)-(3)	(2)-(4)	without migrants
Household characteristics							
Size	10.9	10.3	12.8	11.1	1.75*	1.03	7.8
Dependency ratio	42.7	47.1	39.1	43.8	-1.55	-1.27	41.1
Proportion of (%)							
...*Male*	32.5	40.8	42.7	41.4	3.84***	0.22	47.5
...*Female*	67.5	59.2	57.3	58.6			52.5
Environment (%)							
...*Urban*	67.6	51.9	74.7	64.3	1.57	3.36	55.2
...*Rural*	32.4	48.1	25.3	35.7			44.8
Region (%)							
...*Dakar*	49.1	30.9	53.9	22.2	0.91	33.82***	35.2
...*North and East*	17.6	27.2	18.2	45.0			14.4
...*South*	4.6	27.2	3.9	4.7			9.8
...*Center*	28.7	14.8	24.0	28.1			40.5
Household head characteristics							
Age (in years)	52.8	51.8	58.2	58.1	2.81***	3.41***	49.9
Gender (%)							
...*Male*	46.3	60.5	62.3	62.0	6.62***	0.07	80.7
...*Female*	53.7	39.5	37.7	38.0			19.3
Religion (%)							
...*Murid*	31.5	14.8	28.6	26.3	0.26	4.06*	34.6
...*Other*	68.5	85.2	71.4	73.7			65.4
Education (%)							
...*None*	58.3	63.0	44.8	63.7	5.35*	0.36	62.8
...*Primary*	12.0	19.8	20.1	21.6			20.3
...*Secondary and higher*	29.6	17.3	35.1	14.6			16.9
Observations	108	81	154	174			1524

Note: χ^2/Fisher's exact tests for the equality of distributions for categorical and dummy variables, t-test for the equality of means for continuous variables, between columns (1)-(2)samples.
$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: PSF survey 2006 - 2007 and MIDDAS Survey, 2009 - 2010. Authors' calculation.

References

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, 102(5):1832–56.

Allenby, G. M. and Rossi, P. E. (1998). Marketing models of consumer heterogeneity. *Journal of Econometrics*, 89(1-2):57–78.

Ambler, C. (2012). Don't tell on me: Experimental evidence of asymmetric information in transnational households. Technical report, mimeo, University of Michigan.

Antoine, P. (2007). *Les relations intergénérationnelles en Afrique: approche plurielle*. CEPED.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy*, 112(1):68–105.

Bertoli, S., Moraga, J. F.-H., and Ortega, F. (2013). Crossing the border: Self-selection, earnings and individual migration decisions. *Journal of Development Economics*, 101(0):75 – 91.

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

Borjas, G. J. and Bronars, S. G. (1991). Immigration and the family. *Journal of Labor Economics*, 9(2):pp. 123–148.

Bourguignon, F., Fournier, M., and Gurgand, M. (2007). Selection bias corrections based on the multinomial logit model: Monte carlo comparisons. *Journal of Economic Surveys*, 21(1):174–205.

Chiswick, B. R. (1999). Are immigrants favorably self-selected? *The American Economic Review*, 89(2):pp. 181–185.

Chort, I., Gubert, F., and Senne, J.-N. (2012). Migrant networks as a basis for social control: Remittance incentives among senegalese in france and italy. *Regional Science and Urban Economics*, 42(5):858–874.

Dahl, G. (2002). Mobility and the return to education: Testing a roy model with multiple markets. *Econometrica*, 70(6):2367–2420.

De Vreyer, P., Gubert, F., and Roubaud, F. (2010). Migration, self-selection and returns to education in the waemu. *Journal of African Economies*, 19(1):52–87.

De Vreyer, P., Lambert, S., and Safir, A. (2009). Remittances and poverty: Who benefits in the household? *Paris School of Economics, mimeograph*.

De Vreyer, P., Lambert, S., Safir, A., and Sylla, M. (2008). Pauvreté et structure familiale, pourquoi une nouvelle enquête? *Stateco*, 102:261–275.

Dubin, J. A. and McFadden, D. L. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica: Journal of the Econometric Society*, pages 345–362.

Fernandez-Huertas Moraga, J. (2011). New evidence on emigrant selection. *The Review of Economics and Statistics*, 93(1):72–96.

Fernandez-Huertas Moraga, J. (2013). Migrant selection patterns in rural and urban mexico. *Journal of Development Economics*.

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *The American Economic Review*, 60(1):126–142.

Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica: Journal of the Econometric Society*, pages 507–512.

McKenzie, D. and Rapoport, H. (2010). Self-selection patterns in mexico-u.s. migration: The role of migration networks. *The Review of Economics and Statistics*, 92(4):811–821.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the us labor market. *The Quarterly Journal of Economics*, 118(2):549–599.

Orrenius, P. M. and Zavodny, M. (2005). Self-selection among undocumented immigrants from mexico. *Journal of Development Economics*, 78:215–240.

Osili, U. O. (2007). Remittances and savings from international migration: Theory and evidence using a matched sample. *Journal of Development Economics*, 83(2):446–465.

Pugatch, T. and Yang, D. (2010). The impact of mexican immigration on us natives: Evidence from migrant flows driven by rainfall shocks. *Working paper, University of Michigan*.

Rapoport, H. and Docquier, F. (2006). The economics of migrants' remittances. In Kolm, S. and Ythier, J. M., editors, *Handbook on the Economics of Giving, Reciprocity and Altruism*, volume 1, chapter 17, pages 1135–1198. Elsevier.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146.

Samuelson, P. A. (1956). Social indifference curves. *The Quarterly Journal of Economics*, 70(1):1–22.

Senne, J.-N. (2013). *Migration, Remittances and Schooling Decisions within the Household: Evidence from Innovative Surveys in Senegal and Madagascar*. PhD thesis, Ecole de Hautes Etudes en Sciences Sociales, EHESS.

Seror, M. (2012). Measuring information asymmetries and modeling there impact on senegalese migrants' remittances. Master's thesis, Paris School of Economics.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy*, 70(4):80–93.

Stark, O. and Bloom, D. E. (1985). The new economics of labor migration. *American Economic Review*, 75(2):173–78.

Yamamoto, T. (2012). A multinomial response model for varying choice sets, with application to partially contested multiparty elections. *Working Paper, Princeton University*.