Economic Drivers of Family Reunification in a Context of International Migration

Rémi Viné[1]

IHEID, Geneva, Switzerland

Abstract

This paper aims to analyze family reunification behavior of migrants whose spouses are still in the source country. Does s/he reunify? Where? When? It is based on a simple model of a utility-maximizing behavior of a representative household composed of two spouses. An illustration of South to North migration using MAFE database is provided with the use of survival analyses methods to test the timing of reunification.

Keywords: Sequential Migration, Family, Utility-Maximizing, Africa-Europe migration

Vains objets dont pour moi le charme est envolé ;
Fleuves, rochers, forêts, solitudes si chères,
Un seul être vous manque et tout est dépeuplé.
L'isolement. - Alphonse de Lamartine

[1]I am grateful to the participants and organizers of the YSEM (KOF, ETH Zurich, 2020), the Junior Economic Workshop on Migration Research (ifo CEMIR, Munich, 2019) ; the YSEM (Graduate Institute, Geneva, 2019); the 3rd International Conference: Understanding Forced and Voluntary Migration (Lille University, 2019). I would also thank Slobodan Djajić for his precious comments.

1. Introduction

The present article aims to contribute to the economic literature dealing with migration and especially the family migration when one migrant first moves and then wonders whether it is best to reunify and, if so, where. In developed countries family reunification is an important matter as it accounts to a significant share of the overall migration.

Definitions. It appears that the terms used are sometimes unclear. Then it looks best to try to settle simple definitions on the terms that fit to the scope of the present paper. The family-based migration literature is scarce and a significant part of migration literature is actually dealing with internal movers Sjaastad in [1], Harris and Todaro in [2], Mincer in [3], or Nakosteen and Zimmer in[4]. In this paper, “migrant” means an international migrant who is a foreign-born person who did not initially have the nationality of the country s/he is living in.

The second precision, or rather simplification, I am making deals with the concept of family. Obviously, this cannot be summarized properly in a couple of lines. The international family migration sociology is a rather new but dynamic field of study (see for instance Kraler *et al.* in 2012 [5] or Edo Korljan in Foppiani and Scarlatescu, 2018 in [6]). Traditional families are less frequent in Europe and the jurisdiction ought to adapt to better match with what a family nowadays is (whatever the final decision between backing more traditional or more liberal perceptions). However, I will stick to unrealistic and convenient concept: a family in this article is made of two spouses (if they have children we ignore them for simplicity).

How to define family-based migration, to define family reunification? Those concepts are not so obvious and sometimes cover different sets of people across countries. The present paper displays a model based on the neoclassical framework and thus the specific characterization of what family reunification means does not seem to be crucial. However, a key distinction arises by stating that family reunification supposes a separation at a point. Most of the family-based migration literature is dealing with *instantaneous* family migration (everyone moves together) while the reunification introduces a sequence of migration. Nevertheless, family reunification cannot be only taken as a subset of family migration since it also allows for the same individual to migrate and to return. Hence family reunification is meant to bridge the family migration literature on the one hand and the return migration literature on the other.

The empirical part is based on MAFE database which has three African countries as source countries and the host countries consist of with six European countries. Except for the UK the legislation regarding family reunification is similar across EU members. Conditions for family reunification concern the income, the size of the flat, in a nutshell the capacity the first migrant has to decently welcome (part of) the remaining family.

Why Does it Matter. Family migration accounts for more than a third of total migration in the European Union countries and for two thirds in the United States [2]. Among major OECD members, only Ireland, Korea ,and the US have a larger share of family reunification over total migration. Sweden, Finland, Portugal, and Belgium have rather similar patterns than France[3].

In most of the developed countries the flow of family migration has not changed much during the second part of the 20th century, except for the US where it more than tripled from 1966 to 2011. In general, family-based migration concerns the migrant's spouse alone and the spouse is more likely to be a woman. It should be noted that the lion's share of the remaining part of the family-based migration concerns either the child(ren) or both the spouse and the child(ren) of the migrant. Hence this present work would be enriched by including these two additional common patterns.

Since family-based migration accounts for an important part of the total inflow of migrants in most of the developed countries, it seems crucial to understand its roots. If all newcomers were to apply for a reunification procedure then the sheer number of migrants would theoretically explode. In one of the rare studies dealing with family reunification in the host country with an economic approach, Jasso and Rosenzweig in 1986 [7] assessed the value of the "family reunification multiplier" and they concluded that, in the case of the US it is lower than unity. Still, family-based migrants do not necessarily fit best with the needs of the host country in terms of skills, type of occupation, etc. Then it seems rational for the States to be able to better anticipate the flows of family-based migrants in the short or medium run.

[2]https://www.oecd.org/els/mig/Chaloff.pdf

[3]It should be mentioned, however, that the way to measure family-based migration changes across countries. Therefore those comparisons are mere illustrations but should not be taken as flawless data for international comparison. As an example, the US includes many more components in it than most of the European countries such as several cases that could have been included in the "accompanying family of workers" category (accompanying spouse/children of workers).

Another perspective that is less of an issue in economics *per se* concerns the question of family reunification and integration of migrants in a country when the reunification happens in the host country (Bonjour and Kraler - 2015 [8]). The reunified family is almost systematically an endogamous family and it appears that the second-generation migrants whose both parents come from the same country generally suffers from higher unemployment propensity and lower education. Endogamy in often deemed as an important brake of integration dynamic (see Todd in 1994 [9] or Algan *et al.* in 2013 who reviewed several European countries [10]) and it therefore matters to pay close attention to the family reunification factors since they are in turn key drivers to endogamous behavior.

One can see some more practical discussion about family migration in AppendixA. I present in some more details the European context and compare it with the MAFE data that will be used in the present paper.

Objective of the Present Article. I aim to use a simple utility-maximizing framework to better understand the channels that might push a migrant to reunify, where, and when. The temporality is a key aspect here as most of the related literature analyzes the simultaneous family migration and usually does so in an internal context in which the costs of migration (both financial and psychological) are extensively lower. The present article provides few insights about who is more willing to stay split longer. This analysis is backed empirically with the use of survival analysis methods. This paper is articulated as following: first, there is a literature review that tries to provide an understanding of the literature related to family migration. Second, I present a model based on very simple assumptions in order to only catch few aspects of the migrant's decisions. Third, an empiric illustration is done with the use of MAFE data, *i.e.* on the context of South to North migration. Fourth, a few lines are dedicated to conclude.

2. Literature Review

Interdisciplinarity. Broadly speaking, the literature dealing with migration studies has emerged lately, mostly after 1950. In essence, this topic is interdisciplinary because it is not trivial to completely separate perspectives such as economic aspects along with demographic, juridic, anthropological, sociological, historic ones (see for example Brettel and Hollifield's book, 2014 [11]). For most of the papers studied it should be highlighted that interdisciplinary methods are missing and a dialogue among disciplines appears to be scarce. However, it seems unlikely that family-based migration - and *a fortiori* family reunification - is entirely driven by economic factors. This present paper has a modest objective and do not intend to understand the full spectrum of the family reunification decision. As an example, the importance of the family structure itself seems to matter (Kraler *et al.* in 2012 [5]) but this is beyond the scope of this paper.

Cumulative Causation. Migration studies do not escape from the curse of causation issues. Migrants self-select for a wide range of reasons. They may be pushed to migrate because of higher expected gains (see the seminal Roy -1951 article [12] or the the famous application to migration by Borjas in 1987[13]) and gains can be both absolute or relative to other households in the source community (Stark and Taylor - 1989 [14]). On the other hand, part of them might be forced not to migrate, which is mostly due to liquidity constraints (Djajić - 2014 [15] - for a theoretical analysis). Alternatively, some authors have proposed other types of models that would better fit with the above-average skill levels of migrants stylized fact when dealing with South-North migration (*e.g.* Chiquiar and Hanson - 2005 [16], or Clemens, Montenegro and Pritchett - 2016 [17]). In any case, self-selection is a major issue that can bias empirical results and correcting from it generally requires longitudinal dataset to find proper instruments[4]. The cumulative causation process enlarges the range of the variables that might impact migration decision: as written in Massey (1990) [19]: macroeconomic shifts such as employment growth can push migration trends upwards, which can in turn lead to an even higher growth in employment and the cumulative causation impact varies with the size of the city (Fussell and Massey 2004 [20]). The present paper indicates that the reunifier is more likely to have entered the territory through a labor procedure. The different channels of migration themselves might describe a type of sequential migration.

Reunifying or Remitting. Dealing with family reunification issues requires to understand the different alternatives that a migrant is facing as well as the context of migration. As such, it happens that migration is almost purely

[4]But this is not always the case. Indeed, using the French dataset *Trajectoires et Origines*, Chabé *et al.* -2018 [18] - got instruments for remigration intentions using the desire to get buried in the source country and the feeling of discrimination in the host country.

driven by individual motives but international migration, especially from Southern to Northern countries may also include other components, among which family (Dustmann *et al.* - 2017 [21]) or the community. Then, whatever the reunification schedule, an interesting behavior of the migrant is the remittance pattern. In the model displayed in this paper remittances are likely to happen when the family splits so that, depending on the complementarity in terms of consumption paths, the spouses' consumption is smoothened. Yang -2008 [22] with the example of Philippines indicates that remittances have an impact on the remaining family's decisions such as the investment in housing or the child labor. Additionally, as explained in Dimova and Wolff - 2011 [23] with the case of European countries, remitting can also have a positive impact on the probability of reunifying in the host country. However, the macroeconomic impact of remittance on source countries is unclear (Clemens and McKenzie -2014[24]) or could even be detrimental in case of a *Dutch Disease* due to high amount of transfers (Acosta, Lartey, Mandelman - 2009 [25]).

Jasso and Rosenzweig wrote a fascinating paper in 2010 [26] that deals with the choice between reunifying and remitting. In their study the reunifiers are the parents and the children intially stayed in the source country. They found out that reunification through the migrants' children occurs for the most educated children whose gains from migration are likely to be the highest. On the other hand, low-educated children are less prone to migrate but will rather benefit from financial transfers.

Family migration. Since members of a household can all work on the labor market it matters to analyze how wages might affect the family-migration decision (Gemici 2007 [27]). It may happen that only one of the spouses is better off in terms of wage perceived in the host country and then the wage *premia* of the first migrant should outweigh the loss of the second migrant. Moreover, this subtraction ought to be net of the opportunity cost of coming back home for the initial migrant. Mincer - 1978 [3] looks at the conditions under which family migration can occur in terms of household net wage gains. Family migration might happen even if one spouse loses out on it as long as this loss is offset by the wage *premia* of the other spouse due to migration. Mont (1989) provides a similar analysis with the use of a search model that explains the selection of the migrating couples and in which setting the couples can be composed of double-tied stayers [28].

The thin thread of literature that followed bring insights about who in more likely to migrate given the expected gains from migrating. Borjas & Bronars (1991) [29] use the canonical Roy model in the case of a household composed of two migrants. Using immigration to the US over the 1970s they found out that family migration under this frame is more self-restrictive. Conversely, Munk, Nikolka, Poutvaara (2017) [30] obtained the opposite result with Danish emigration to other Scandinavian countries: self-selection is reduced. Eliasson *et al.* find similar result with Swedish internal migration [31]. Foged (2016) [32] includes both internal and international family migration and finds that the share of each spouses, in terms of education level, matters in deciding whether to migrate. In any case, family migration is understood as a *simultaneous* migration.

Migration occurs through 4 main channels based on 4 different motives to emigrate: (i) escape danger in the source country (asylum), (ii) study in another country, (iii) get a job in a new country, (iv) join the family in another country[5]. The two last categories generally dominate the absolute number of migrants flows. Hence it seems important to also analyse *sequential* migration. It is worth exploring whether there are differences between them concerning the economic performances. Jasso and Rosenzweig - 1995 [34] compares the performance of both groups in the US between 1977 and 1990. Labor migrants perform better than the family reunification migrants but the gap seems to decrease over time without vanishing completely.

A Bridge between Permanent and Temporary Migration. Permanent and temporary migrations are generally viewed as following completely different patterns. Temporary migration can be more intuitively explained through consumption and savings' maximization behaviors as in Djajić and Milbourne 1988 [35], Djajić - 1989 [36], Dustmann 2003 [37] and [38]. Indeed, the migrant decides to migrate in order to benefit from a higher wage (but higher price levels as well) that might allow her to save money that can be used in the future once back to the source country. A key question then is: when does the return migration take place in the migrant's lifetime. It is likely that temporary migrants are less risk averse than their non-migrants counterparts. On the other hand, permanent migrants, even if they might still

[5]Obviously those are the large categories and subtle separations can be added. As such, what it named "new theory of international migration" indicates that migrations can be driven by a risk-limiting behavior. Migration is understood as a diversifying process. See for example Rosenzweig and Stark - 1989 [33]

be applying a utility-maximizing approach, are likely to also introduce other factors than purely economic and individual characteristics such as family, network, political instability in the source country, etc.

In practice, except for the famous guest-worker programs (*e.g. Gastarbeiter* in Western Germany, and *Bracero* between the US and Mexico) which were major migration channels during the postwar era, return migration is not much studied empirically with few exceptions such as Dumont *et al.* - 2008 [39]. Return migration accounts for 20 to 50% of immigrants recently arrived (less than 5 years) in the host country. The authors observe a U-shaped frame of return frequencies and age: returns migrants are mostly young whose professional life remains to be set and retirees who seek to better benefit from their pensions.

This paper aims to analyze in what case a migrant prefers to return instead of trying to bring the family in the host country. Migrants whose spouses initially remained in the source country are implicitly facing a choice between temporary and permanent migration. De Coulon and Wolff - 2010 [40], indicated that a couple with child actually has one more choice: circular migration. They studied the desires of immigrants once retired and checked how the type of migration was impacted by the location of the child(ren). Djajić (2008) [41] provides a framework under which return migration of parent is a possibility, given that children remain in the source country. Even if it would be an interesting alternative as this might concern one fourth of the retiree migrants, I do not include this option in the model for several reason: I rather aim to analyze family reunification between spouses and this event is more likely to happen during the active period of life. Additionally, the MAFE dataset only deals with few distant pairs of countries, implying that the cost of migration would be prohibitive in the case of multiple migrations [6].

3. A Model

3.1. Basic Setting

Our aim is to provide an economic understanding of the family reunification phenomenon.This simple framework ought to shed some light on key

[6]On the other hand, that does not include transit migration as analyzed by Djajić in 2017 [42] Migration from Africa to Europe appears to hide "checkpoints", *i.e.* few transit countries.

determinants of the family reunification process.

3.1.1. Assumptions

The model is restricted to migrants who have relatives and more specifically spouses in the source country[7]. I assume that the migrant is the one taking all decisions in light of the spouse's preferences. Therefore, the migrant will look at the household's lifetime total utility and compare four cases. First, the family reunifies in the host country, which means that she brings the spouse there (Belgium, France, Italy, the Netherlands, Spain, the United Kingdom) at an optimally chosen point in time. Second, after some (optimal) time spent in the host country, the migrant decides to go back to the source country and reunify with the family there. Third, the migrant does not reunify and the family remains apart forever. Fourth, the spouses migrate simultaneously.

I also assume that there is neither uncertainty nor information asymmetry. It follows logically the hypothesis of a common worldwide interest rate, r. This will also be useful to drop any strategic arbitrage the migrant can be pushed to make in case of interest rate differential as in Djajić - 2010 [43]. Additionally, the household's wage can be spent without friction costs in any of the two countries when the family is separated. S/he decides to reunify or not according to the best situation possible in terms of household utility. According to the type of reunification the period of separation can differ.

3.1.2. Economic Variables[8]

Wages in the two countries are different as well as wages earned by the two spouses. Wages are denoted as following: for each period the household earns a sum of both spouses wages, $w_{\tilde{w},w_s}$ denotes the total wage obtained by the couple when the migrant works in the host country and and the spouse

[7]This is a restrictive assumption that allows some more coherence as we assume a specific household lifetime and that both agents are in working age - even if this second aspect can be dropped easily. Moreover, using the MAFE data presented later in this paper one can see that the most frequent decider of the migration, if the the migrant him or herself is the spouse and the partner, among the different people, is the most likely to have financially contributed to the migration.

[8]A practical comment: notation with s subscript indicates spouse, tilde notation indicates the host country. For the agent who does not move, there still are two situations, the one being separated and the one being reunified. Since separation occurs before, it will be denoted with the subscript "1" and the second period with the subscript "2".

works in the source country. It should be mentioned that individual wage, and most likely the spouse's wage(s) might be zero in the case of no activity on the labor market. This can be explained by gender inequality on the labor market, skill gap between the spouses, weekly number of hours worked, etc. The model includes prices of the consumption baskets of each spouse and prices are not necessarily the same in the two countries. It is more likely that the price index in the source country p would be lower than the price index in the host country $\tilde{p}$ but this is not obligatory (although I use this restriction for the calibration of the model, in line with the data I use in the present paper). I assume that migration from the source to the host country is costly and this cost is not necessarily the same for the two members of the family if it reunified in the host country, as the first migrant might have built a network that allow avoiding some costs.

Within the lifespan of the representative household there would be a decisive moment, τ, when it chooses to reunify. There are four possibilities at time τ: (i) the migrant goes back to the source country to reunify with her family there, (ii) s/he reunifies the family in the source country, (iii) the couple is split forever $\tau \geqslant T$, (iv) there is no separation at all ($\tau = 0$). It should be stressed that if a reunification occurs, τ can vary according to the country of reunification.

Additively separable household utility. To keep things as simple as possible, the household is assumed to maximize a Bergson-Samuelson joint welfare function. This is a subset of the CES framework where the substitution between the two spouses' utilities is perfect (it is equal to unity).

$$U(c_{1t}, c_{ts}) = \alpha u(c_t) + (1 - \alpha) u(c_{ts}) \tag{1}$$

Quite classically, α and $1 - \alpha$ accounts for the Pareto-weights accruing to each spouse. With households members fully altruistic among each others, the coefficient α is 1/2. The individual's utility functions are assumed to be well-behaved ($u'(c) > 0$ and $u''(c) < 0$). So this restriction erases the supermodularity property of the function which can partly explain the desire of both members of the household to increase each other's consumption level. Therefore, it would be interesting to compare the separable and Cobb-Douglas household utility functions, as the former does not include any complementarity between spouse consumption while the latter does. The results are quite similar.

Separation cost and migration psychological costs. It seems realistic that the utility function should add a "separation penalty" when agents are not living together: they are supposed to suffer from the distance between them. To my knowledge, literature of the functional from this penalty should take over is rather modest. A constant penalty allows to considerably simplify the derivations but does not seem to catch reality very clearly. An always increasing and convex penalty might be convincing in the case of temporary migration as in Vinogradova -2016[44]. Lastly, in the case of potential permanent migration the shape of the separation penalty might actually be either increasing but concave, or take the form of an inverse-U as it is implied in a Dutch case (Eurelings-Bontekoe *et al.*, 2000 [45]). For convenience, I stick to the easiest case where the separation generates a penalty on the household utility that is a function of the duration of separation. Let $\pi(\tau)$ be the remainder once the penalty has been taken in account with $0 \leq \pi(\tau) \leq 1 = \pi(0)$, $\pi'(\tau) < 0$ and $\pi''(\tau) > 0$. The household utility in while being separated has takes the form $\pi(\tau)U(c_{1t}, c_{ts})$.

A more classic cost of migration deals with the move from home. This does not only imply a (temporary) separation from the nuclear family, it also leads to a separation with the extended family, a network established locally, a climate, a culture, etc. Therefore, it will be assumed that for any given amount of consumption, a extra consumption will deliver a higher utility if consumed in the home country $\tilde{u}'(\tilde{c}) < u(c)$, $\forall\ \tilde{c} = c > 0$.

3.2. Model with reunification

When the family reunifies there must be an optimal moment for it and this will depend on the location of the reunification itself: source or host country. It clearly appears that one of the two agents would switch country, thereby switch wage earned as well as the price level she is facing. There is a break in the household utility function. Depending on the place of reunification the decision can appear at a different time as this moment is set endogenously in the model. Implicitly, a model without reunification implies (i) $\tau = 0$ in the case of the host country reunification, implying that spouses migrate together; (ii) $0 < \tau \leq T$ in the case of the source country reunification which implies no migration at all; (iii) $\tau > T$ in either case implying that, in the given scenario, the couple remains separated [9].

[9]This case isn't realistic but we stick to this possibility for completeness. Additionally, in the MAFE database there are few outliers who are still separated after an extremely

3.2.1. Host reunification

In this setting the spouse of the migrant who initially remained in the source country eventually migrates. The objective function of the household is the following sum of discounted utilities, with δ as the discount factor:

$$V_{frhost} = \int_0^\tau \pi(\tau) U(\tilde{c_{1t}}, c_{ts}) e^{-\delta t} dt + \int_\tau^T U(\tilde{c_{2t}}, \tilde{c_{ts}}) e^{-\delta t} dt \tag{2}$$

subject to its budget constraint:

$$BC_{frhost} : \int_0^\tau (w_{\tilde{w}, w_s} - \tilde{p}\tilde{c_{1t}} - pc_{ts}) e^{-rt} dt +$$

$$\int_\tau^T (w_{\tilde{w}, \tilde{w_s}} - \tilde{p}\tilde{c_{2t}} - \tilde{p}\tilde{c_{ts}}) e^{-rt} dt + A_0 - K_0 - K_1 e^{-r\tau} = 0 \tag{3}$$

The intuition is straightforward: the household aims to maximize its utility subject to the resources available and the fixed cost incurred by the (potentially) several migrations [10]. Then I derive the first order conditions that would imply to get the two different consumption levels of the spouse as a function of the migrant's consumption. Since the choice of τ is endogenous[11] there are five first order conditions:

$$\frac{\delta L}{\delta \tilde{c_{1t}}} = 0 \Leftrightarrow \frac{\delta U(\tilde{c_{1t}}, c_{ts}) e^{-\delta t}}{\delta \tilde{c_{1t}}} = \lambda_{frhost} \frac{\tilde{p}}{\pi(\tau)} e^{-rt} \tag{4}$$

$$\frac{\delta L}{\delta c_{ts}} = 0 \Leftrightarrow \frac{\delta U(\tilde{c_{1t}}, c_{ts}) e^{-\delta t}}{\delta c_{ts}} = \lambda_{frhost} \frac{p}{\pi(\tau)} e^{-rt} \tag{5}$$

$$\frac{\delta L}{\delta \tilde{c_{2t}}} = 0 \Leftrightarrow \frac{\delta U(\tilde{c_{2t}}, \tilde{c_{ts}}) e^{-\delta t}}{\delta \tilde{c_{ts}}} = \lambda_{frhost} \tilde{p} e^{-rt} \tag{6}$$

long period of time, up to three decades.

[10] One can assume that this fixed cost is much smaller than the first migration's fixed cost: it is likely that a network has been previously set by the initial migrant. However, some costs still exist such as the cost of moving, the cost of the procedure.

[11] In practice, it seems reasonable to state that τ is endogenous as long as it is not small enough. Indeed, migrants are generally required to have lived in EU member State for at least 18 months to be able to implement the reunification procedure. Alternatively, it could be zero and both spouses migrate together.

$$\frac{\delta L}{\delta \tilde{c_{ts}}} = 0 \Leftrightarrow \frac{\delta U(\tilde{c_{2t}}, \tilde{c_{ts}})e^{-\delta t}}{\delta \tilde{c_{ts}}} = \lambda_{frhost}\tilde{p}e^{-rt} \tag{7}$$

The first four Focs are straightforward: they provide the arbitrage among consuming one more unit - in each cases separately - against the cost it generates. Marginal utilities of the migrant and the spouse once s/he migrated are equal. Marginal utilities of the migrant before and after reunification are equal, once corrected for the cost of separation that only occurs before reunion. Therefore, if the costs of separation were zero, the consumption of the migrant might change at the reunification but the value of an extra consumption unit should remain the same. The marginal consumption of the spouse, before and after migration are the same, once corrected for the cost of separation and the price level encountered in both cases. Due to the preference for consumption in the source country, the real marginal consumption equality implies that, for equal prices in both countries and no separation costs, the amount of consumption in the source country is higher than both the consumption of the spouse once reunified and of the migrant before and after reunification.

Assuming for simplicity that $\delta = r$, then the for each spouse within the two situations is constant[12], the move occurs with a change of migration status:

$$\frac{\delta L}{\delta \tau} = 0 \Leftrightarrow [U(\tilde{c}_2, \tilde{c}_s) - \pi(\tau)U(\tilde{c}_1, c_s)] - \pi'(\tau)U(\tilde{c}_1, c_s)\int_0^\tau e^{-r(t-\tau)}dt = \\ \lambda_{frhost}[(w_{\tilde{w},w_s} - \tilde{p}\tilde{c}_1 - pc_s) - (w_{\tilde{w},\tilde{w}_s} - \tilde{p}\tilde{c}_2 - \tilde{p}\tilde{c}_s) + rK_1] \tag{8}$$

The meaning of the last condition, equation (8), is crucial in the present analysis: the left-hand side (LHS) indicates the extra-cost in terms of forgone utility brought by delaying the reunification of the spouses and the RHS assesses the utility value of extra wealth brought by being separated one more instant. Since the separation cost π only enters in the period before reunification, the utility of household consumption is reduced in this period only. However, the cost of separation itself is a function of the length of separation. Therefore, the foregone utility by staying split an extra instant

[12]Hence $U(\tilde{c_\tau}, c_{\tau s}) = U(\tilde{c}_t, c_{ts})$.

ought to be corrected by the change of the separation cost[13]. The Bergson-Samuelson functional frame given in equation (1) allows to simply the usual marginal utilities equivalences:

$$\pi(\tau)\alpha\tilde{u}'(\tilde{c_1}) = \alpha\tilde{u}'(\tilde{c_2}) = (1-\alpha)\tilde{u}'(\tilde{c_s}) = \pi(\tau)(1-\alpha)(\frac{\tilde{p}}{p})u'(c_s) \qquad (9)$$

From equation (9) one can easily reformulate $\tilde{c_2}$, $\tilde{c_s}$, c_s as functions of $\tilde{c_1}$ so that equation (8) is rewritten as:

$$G \equiv [U(\phi(\tilde{c_1},\tau),\varphi(\tilde{c_1},\tau)) - \pi(\tau)U(\tilde{c_1},\psi(\tilde{c_1}))] - \pi'(\tau)U(\tilde{c_1},\psi(\tilde{c_1}))\int_0^\tau e^{-r(t-\tau)}dt - \frac{\pi(\tau)\tilde{u}'(\tilde{c_1})}{\tilde{p}}[(w_{\tilde{w},w_s} - \tilde{p}\tilde{c_1} - p\psi(\tilde{c_1})) - (w_{\tilde{w},\tilde{w_s}} - \tilde{p}\phi(\tilde{c_1},\tau) - \tilde{p}\varphi(\tilde{c_1},\tau)) + rK_1] = 0 \qquad (10)$$

with $\phi(\tilde{c_1},\tau) = (\tilde{u}_{\tilde{c_2}})^{-1}[\tilde{u}'(\tilde{c_1})\pi(\tau)] = \tilde{c_2}$, $\varphi(\tilde{c_1},\tau)) = (\tilde{u}_{\tilde{c_s}})^{-1}[\tilde{u}'(\tilde{c_1})\frac{\alpha}{1-\alpha}\pi(\tau)] = \tilde{c_s}$, $\psi(\tilde{c_1}) = (u_{c_s})^{-1}[\tilde{u}'(\tilde{c_1})\frac{\alpha}{1-\alpha}\frac{p}{\tilde{p}}] = c_s$ where $(\tilde{u}_{\tilde{c_2}})^{-1}[.]$, $(\tilde{u}_{\tilde{c_s}})^{-1}[.]$, $(u_{c_s})^{-1}[.]$ is are respectively the inverse functions of the marginal utilities of the migrant once reunified, the spouse after and before reunification. If one assumes that the individual utility function is identical, or is linearly related as it is the case here given that there is a psychological cost of consuming in the host rather than in home, one can further simplify the last equation. In this case, all consumption levels are positively and linearly related to each other. Therefore, all the consumption levels are similarly related to the exogenous variables[14]. The budget constraint can be integrated and modified in a similar fashion:

$$B \equiv (\frac{1-e^{-r\tau}}{r})(w_{\tilde{w},w_s} - \tilde{p}\tilde{c_1} - p\psi(\tilde{c_1})) + (\frac{e^{-r\tau}-e^{-rT}}{r})(w_{\tilde{w},\tilde{w_s}} - \tilde{p}\phi(\tilde{c_1},\tau) - \tilde{p}\varphi(\tilde{c_1},\tau)) + A_0 - K_0 - K_1e^{-r\tau} = 0 \qquad (11)$$

The system of equations (10) and (11) has two endogenous variables, $\tilde{c_1}$ and τ. G assesses the gain from reunifying an extra instant before and B

[13]One can easily notice that this additional term on the LHS drops whenever π is constant.

[14]The Pareto weight being an obvious exception when comparing consumption between the two spouses.

simply evaluates the balances between accumulation during the separation period and overspending once reunified. Therefore, it is expected that an increase of B leads to more consumption $\tilde{c_1}$. An larger G will push down the duration of separation. Comparative statics can be classically obtained through the computations of partial derivatives of G and B with respect to both the two endogenous variables and the exogenous variables (I define $\mathbf{x}$ as the column vector with the exogenous variables for notation convenience). Here, the focus is put on wages, price levels, costs of migration and lifetime. One can apply the implicit function theorem (IFT) in the neighborhood of the equilibrium to obtain:

$$\begin{bmatrix} G_{\tilde{c_1}} & G_\tau \\ B_{\tilde{c_1}} & B_\tau \end{bmatrix} \begin{bmatrix} d\tilde{c_1} \\ d\tau \end{bmatrix} = - \begin{bmatrix} G_{\mathbf{x}^\mathbf{t}} \\ B_{\mathbf{x}^\mathbf{t}} \end{bmatrix} d\mathbf{x} \tag{12}$$

The lack of explicit formulas and the use of IFT restricts the comparative statics to minor changes of each exogenous variables one by one. This nevertheless allows to sketch the consumption behavior as well as the duration of separation. From equation (12) it is rather straightforward to derive the comparative statics. One needs to ensure that the determinants of the Jacobi matrix of the endogenous components (denoted Δ) is positive[15]. In the present paper the focus is not on the consumption pattern itself but rather on the timing of reunification (and the country of reunification).

CRRA individual utility. Since the procedure of this paper is based on a utility-maximizing approach, a key aspect is obviously the functional form of the utility function (based on Djajić models *e.g.* in [35], [36]) as the objective is to keep tractability in order to highlight the underlying mechanisms. Indeed, some of the comparative statics could be cumbersome and it highly facilitates the analysis to restrict the functional form of individuals' utility.

The migrant assesses the discounted utility all over the household's lifespan. Then, one can define each household member's utility based on a constant risk aversion function (henceforth CRRA) which is the following:

$$u(c_s) = c_s^{1-\theta}/(1-\theta) \tag{13}$$

$$\tilde{u}(\tilde{c_1}) = \gamma \tilde{c_1}^{1-\theta}/(1-\theta). \tag{14}$$

[15] See (AppendixB) for a more careful discussion.

with $< \gamma < 1$ the cost of being away from the source country. This allows to verify the aforementioned condition: $\tilde{u}'(\tilde{c}) < u(c), \forall \ \tilde{c} > 0$.

The higher θ the higher the risk aversion of the agent. The fraction $1/\theta$ is the elasticity of inter-temporal consumption substitution of the agent [16]. The utilities of both agents in the household are based on the similar pattern as θ, the fixed measure of relative risk, is deemed as exogenous. This can be viewed as a restrictive simplification, as migrants generally have a specific risk aversion pattern and Dustmann *et al.* -2017 [21] - showed that this not only applies to the individual but also to the household level.

Comparative statics for a host reunification is summarized in the following proposition.

Proposition 1. *In the case of South to North migration and the case of host reunification, with a price level in the host country relatively high (i.e. $\tilde{p}\theta > p$) and a not too large cost of separation the time being separated in the neighborhood of the optimal varies as:*
(i) $\frac{d\tau_h}{dK_0} > 0, \ \frac{d\tau_h}{dK_1} > 0, \ \frac{d\tau_h}{dA_0} < 0$
(ii) $\frac{d\tau_h}{dw_s} \gtreqless 0, \ \frac{d\tau_h}{d\tilde{w_s}} < 0, \ \frac{d\tau_h}{d\tilde{w}} < 0$
(iii) $\frac{d\tau_h}{dp} < 0, \ \frac{d\tau_h}{d\tilde{p}} > 0$
(iv) $\frac{d\tau_h}{dT} > 0$

Details can be obtained in (AppendixB). Results are rather intuitive. Larger costs of migration increases the time being split while a higher amount of initial savings reduces the duration of separation. This is intuitive: the more costly the first migration the longer the time to eventually afford the reunion and the higher the savings the easier it is to afford it. The time being separated is impacted by both spouses' wages. The spouse's earning before reunifying does not provide a clear answer as two forces are competing: a higher wage in the source country allows to reduce the pain from being split through a higher level of consumption but a higher wage in the source country also permits to afford reunification faster. Prices also refer the optimal choice

[16] Basic derivations indeed leads to

$$\frac{1}{\theta} = -u'(c_t)/u''(c_t) * c_t = -\frac{dln(c_t/c_t)}{dln(u'(c_t)/u'(c_t))}.$$

Then θ is the inverse of the sensitivity of consumption with respect to the change in marginal utility.

of the household: the higher the price level in the source country the more likely the couple will reunify quickly as the spouse does not benefit from cheap consumption and both suffer from the separation. On the other hand, a higher price level in the host country leads to postpone the reunification in the host country as the cost of separation will be lower relatively to the better consumption opportunities in the source country. Lastly, a longer life expectancy allows the couple to delay the reunification and still enjoy the reunified period.

3.2.2. Source reunification

The procedure in the case of source reunification is exactly the same, only few ingredients change and can be easily noticed in the objective function and the budget constraint. Now the utility in the second period is based on the consumption of both spouses in the source country and thus prices and wages must adapt accordingly. The objective function and the budget constraint only differ from the host reunification scenario in the after reunion phase. A core element is that the time of reunification does not need to be the same in the two scenarii (I set the problem directly assuming $r = \delta$).

$$V_{frsource} = \int_0^{\tau} \pi(\tau) U(\tilde{c}, c_{1s}) e^{-\delta t} dt + \int_{\tau}^{T} U(c, c_{2s}) e^{-\delta t} dt \tag{15}$$

subject to its budget constraint:

$$BC_{frsource} : \int_0^{\tau} (w_{\tilde{w},w_s} - \tilde{p}\tilde{c} - pc_{1s}) e^{-rt} dt +$$
$$\int_{\tau}^{T} (w_{w,w_s} - pc - pc_{2s}) e^{-rt} dt + A_0 - K_0 - K_1 e^{-r\tau} = 0 \tag{16}$$

The procedure is similar: the household decides the best time of reunification as well as the consumption patterns. Therefore, one can easily obtain the equivalent B and G equations that parallel the equation (10) and (11). In this case, the comparative statics will be derived from:

$$\begin{bmatrix} G_{c_{1s}} & G_{\tau} \\ B_{c_{1s}} & B_{\tau} \end{bmatrix} \begin{bmatrix} dc_{1s} \\ d\tau \end{bmatrix} = - \begin{bmatrix} G_{\mathbf{x^t}} \\ B_{\mathbf{x^t}} \end{bmatrix} d\mathbf{x} \tag{17}$$

with the vector $\mathbf{x}$ being almost identical[17]. Source reunification can be summarised by the following proposition:

Proposition 2. *In the case of South to North migration and the case of source reunification,*

(i) $\frac{d\tau_s}{dK_0} > 0$, $\frac{d\tau_s}{dK_1} > 0$, $\frac{d\tau_s}{dA_0} < 0$

(ii) $\frac{d\tau_s}{d\tilde{w}} \gtreqless 0$, $\frac{d\tau_s}{dw} < 0$, $\frac{d\tau_s}{dw_s} < 0$

(iii) $\frac{d\tau_s}{dp} > 0$, $\frac{d\tau_s}{d\tilde{p}} < 0$

(iv) $\frac{d\tau_s}{dT} > 0$

3.3. Comparison of the two models

The strategy here is quite simple: it aims to compare the objective functions of the two scenarii and how these change when exogenous variables are modified. In each scenario, when the optimal moment for reunification exceeds the household lifetime the "forever alone" option dominates. For $0 < \tau < T$ the household will reunify but the place of reunification remains to be defined.

The core idea is to check under which conditions one reunification is preferred over the other in terms of the exogenous variables of the model:

$$RF_{country} = \frac{V^*_{frhost}(w_{\tilde{w},w_s}, w_{\tilde{w},\tilde{w}_s}p, \tilde{p}, K_0, A_0, K_1, r, T, \pi, \theta, \gamma)}{V^*_{frsource}(w_{\tilde{w},w_s}, w_{w,w_s}p, \tilde{p}, K_0, K_1, A_0, r, T, \pi, \theta, \gamma)} \gtreqless 1. \quad (18)$$

Whenever the ratio exceeds unity the household will seek to reunify in the host country. Even if the model is dangerously simplistic it catches a decent bulk of different variables but this has a cost. As such, it becomes tricky to derive the comparative statics. This is why I stick to calibrations and display several graphs that give the intuitions that the model brings about. Therefore, I will stick to the cases that are of interest in this paper, *i.e.* when the source country happens to be much less developed than the host country. That allows me to assert few bold assumptions, such as the wages are much higher in host than in source country, so are the price levels.

One can see that the two types of reunification involve in most cases opposite signs. This simply relates to the fact that push factors in one reunion scenario often consists in a pull factor in the other scenario.

[17]The wage $\tilde{w_s}$ of the spouse once reunified is replaced by w, the wage of the migrant after returning. The cost of reunification might also be different as re-migration does not imply the same costs as host reunification.

4. Calibration

Equipped with of the tools to make the decision, it is now possible to calibrate the model and provide results driven by the model. It appears crucial to analyze under what conditions is one type of reunification preferred over the other.

Calibration. In the present context of South to North migration, it is fairly that the source country, compared to the host country, will have lower wages and much lower price level.

To further simplify, I choose ranges of wages so that the agent's wage in the host country does not overlap with the range of the agent's wage in the source country. I proceed similarly for prices ($\tilde{w} > w, \tilde{w_s} \geq w_s, \tilde{p} > p$). Then, in the calibration, I display results of migration from a poorer country to a richer country, or at least from a country in which the agent's wage is lower to a country in which it is higher, which is in line with the scenario of migration from Africa to Europe. The analysis is restricted to $\theta < 1$ implying that agents are risk averse which is the most common setting even though this can widely vary (see Chetty 2006 [46]). A obvious weakness of the calibration given in the table 1 is its arbitrary choice of default variables.

One can wonder whether one type of reunification would be privileged over another when, for instance, prices change or wages change. The values that τ takes is often either the minimum, *i.e.* $\tau = 0$ implying a simultaneous migration or the maximum value implying a full separation[18]. Depending on the situation, it can take value above the household lifetime, or below zero. In the first situation, this implies that there is no reunification at all; in the second case, that there is no separation at all (back to the initial model without reunification). Price level in host country diminishes consumption levels for all agents and in both scenarios. On the other hand, prices in source countries are positively related to the agents living in the host country in both scenarios.

In the host reunification case, the optimal time for reunification increases with the price level in the host country and decreases with the price level in the country of origin. There is a price threshold in host country at which it appears better not to reunify at all. Fundamentally, a higher price in host

[18]This unrealistic result deserves refinements of the model such as a non constant cost of separation.

Table 1: Calibration Setting

Variable	Default Value	Range for graph
$\tilde{p}$	2.5	[1.75 : 3.25]
p	1	[0.75 : 1.25]
$\tilde{w}$	8	[5 : 30]
w	1.5	[0 : 8]
$\tilde{w}_s$	5	[0 : 25]
w_s	1	[0 : 5]
r	0.05	[0.01 : 0.15]
A_0	80	[0 : 150]
K_0	80	[0 : 150]
K_1	60	[0 : 100]
θ	0.71	[0.5 : 0.9
T	40	[25 : 50]
γ	0.99	[0.9 : 1]

The price level in the host country is taken from the MAFE weighted data, the price ratio would be 2.49 and so is the price level in host as source price level is normalized. I simply assume that the first migration is not constrained by liquidity access while the second migration is until τ quite large (above 20 years of separation). The risk aversion coefficient simply is fixed at the value estimated by Chetty (2006) [46]. This is the value when utility is additive and where the cross derivative is zero.

country delays the reunification because the opportunity cost of being separated becomes relatively slower: they can wait more as this would otherwise imply to give up a significant purchasing power. Interestingly, below a given a price level there is not reunification at all because there is no separation either: spouses migrate together. Therefore, the sequential migration happens only in a window of cases for which simultaneous migration is too costly and in which migrants suffer from being split.

If the reunification is to take place in the country of origin, then the higher the price in the host country the sooner it happens and conversely for the price in the source country. In the setting provided here, migration is so profitable that the migrants are willing to stay separated instead of reunifying back home. This might relate to the data in which censored observations - those who have not reunified yet - are separated up to a much longer period of time. The figures 1, 2, and 3 provide an illustration of the model trends in the context of a migration from an South to North. In a context of much higher wages in the host country (by a factor of 4 or 5, as estimated for the US, see Clemens *et al.* (2016) [17]) and of prices around 2.5 higher in the host country, migration is, of course, attractive. In this setting this is so attractive that source reunion is not meant to happen and thus the couple remains split. That would be changed with a higher cost of separation. This specificity explains why the source reunion case sometimes appears more profitable than the host reunification. This would be clearly hampered by a lower cost of migration of the spouse. This would explain the sometimes confusing *source reunification* curves. On the other hand, it should be stressed that this scenario logically clogs the possible source reunion for economic reason if one only considers the raw value driven from the calibration. Very logically, another scenario, for example migration from Senegal to Ivory Coast would lead to much different predictions. In any case, what matter is more the dynamics than the specific values. One can notice that higher prices in host pushes the host reunification under the source reunification. The revert is true for the source country price levels.

Basically, this gives hints about the willingness to rather reunify in the host country, or even not to reunify at all. Then the household optimizes its situation by picking the most profitable scenario. Simultaneous migration is not rare ($\tau = 0$) but this happens in only certain cases.

Figure 1: Optimal Objective Functions

(a) Price level in host country (b) Price level in source country

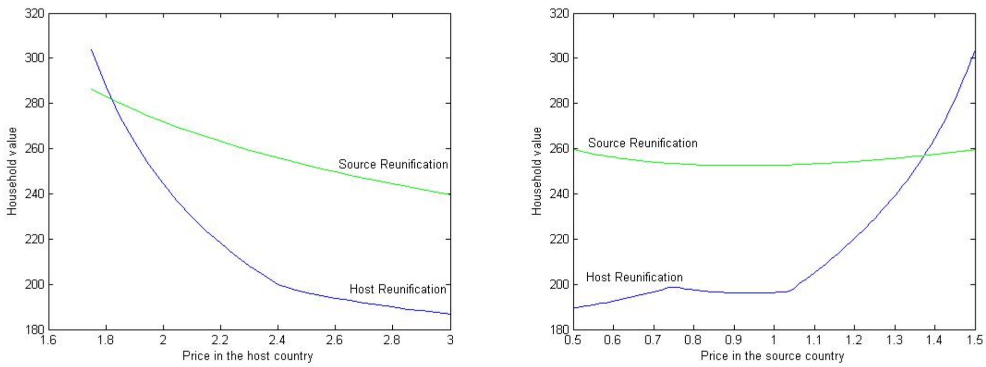


Figure 2: Optimal Objective Functions

(a) Wage of the first migrant in host country (b) Wage of the spouse in host country

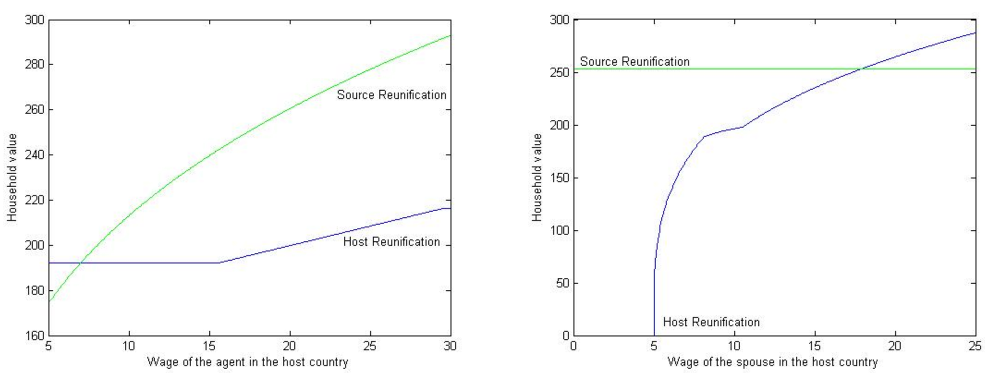
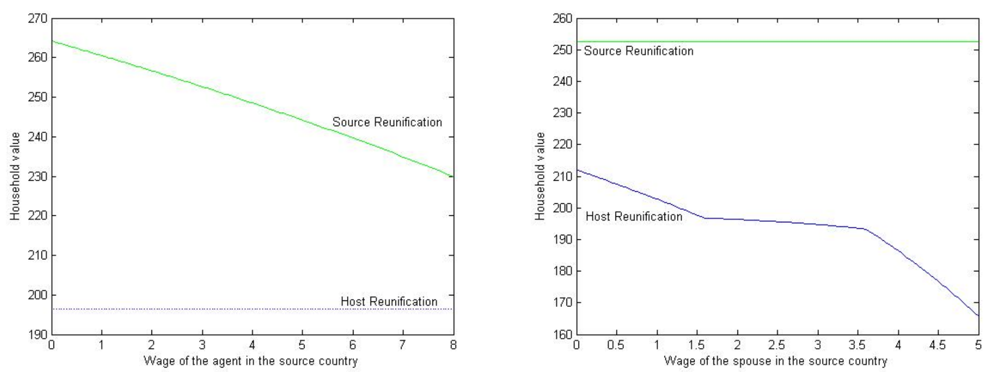


Figure 3: Optimal Objective Functions

(a) Wage of the first migrant in source country

(b) Wage of the spouse in source country



5. The Empirics

5.1. MAFE data

To empirically confront the insights of the model I proposed above, I use an empirical analysis with MAFE - Migration between Africa and Europe - data (for details about the database, please consult Beauchemin in 2012 [47])[19]. It covers three countries in Africa: Democratic Republic of Congo, Ghana, Senegal[20]; and six in Europe: Belgium, France, Italy, the Nether-

[19]The MAFE project is coordinated by INED (C. Beauchemin) in partnership with the Université catholique de Louvain (B. Schoumaker), Maastricht University (V. Mazzucato), the Université Cheikh Anta Diop (P. Sakho), the Université de Kinshasa (J. Mangalu), the University of Ghana (P. Quartey), the Universitat Pompeu Fabra (P. Baizan), the Consejo Superior de Investigaciones Científicas (A. González-Ferrer), the Forum Internazionale ed Europeo di Ricerche sull'Immigrazione (E. Castagnone), and the University of Sussex (R. Black). The MAFE project has received funding from the European Community's Seventh Framework Programme under grant agreement 217206. The MAFE-Senegal survey was conducted with the financial support of INED, the Agence Nationale de la Recherche (France), the Région Ile de France and the FSP programme 'International Migrations, territorial reorganizations and development of the countries of the South'. For more details, see: http://mafeproject.site.ined.fr/

[20]In the present study, the case of Senegal can be worrisome since polygamia is rather widespread - up to one third of the adults are in a polygamous relationship. See Gning and Antoine 2015[48] for details. In AppendixE I alternatively drop one source country from the analysis, results are quite similar. Moreover, I run the regressions with an extra

lands, Spain, the United Kingdom.Only few pairs of migration channels are available. The survey was performed in 2009 on both continent and consists of a total of 5399 people surveyed. Therefore, MAFE provides unique information on the two sides of the story, *i.e.* on reunifiers in the host country as well on reunifiers in the source country. Another advantage is legal: practices in terms of family reunion are very similar across European countries, the UK being slightly different under few aspects (European Migration network 2017 report [49]).Therefore, the legal constraint appears, at least in theory, to be similar in the host countries provided in MAFE data.

The questionnaire includes the entire pattern of the surveyed, implying that all the migration, all the different periods of activity are both reported. That provides a rich insight about the migration patterns between Africa and Europe. It should also be stressed that those frames are inherently right censored because several individuals will not have reunified at the moment of the study. The survival analysis will take this constraint into account. A major drawback of MAFE, though not unsolvable, is the absence of unified wage. The information is tricky as the researchers were asked to obtain the wages all over the life of the surveyed. Therefore, the currency could have changed during the period of activity. Since some country experience a change in currency (EU countries in MAFE except the UK) or a hyperinflation in the 1980s, that makes the computation of the wages per period of activity highly unreliable and too noisy to be rigorously included in the analysis.

dummy variable that is equal to one if the surveyed is in a polygamous household. This variable does not impact the results.

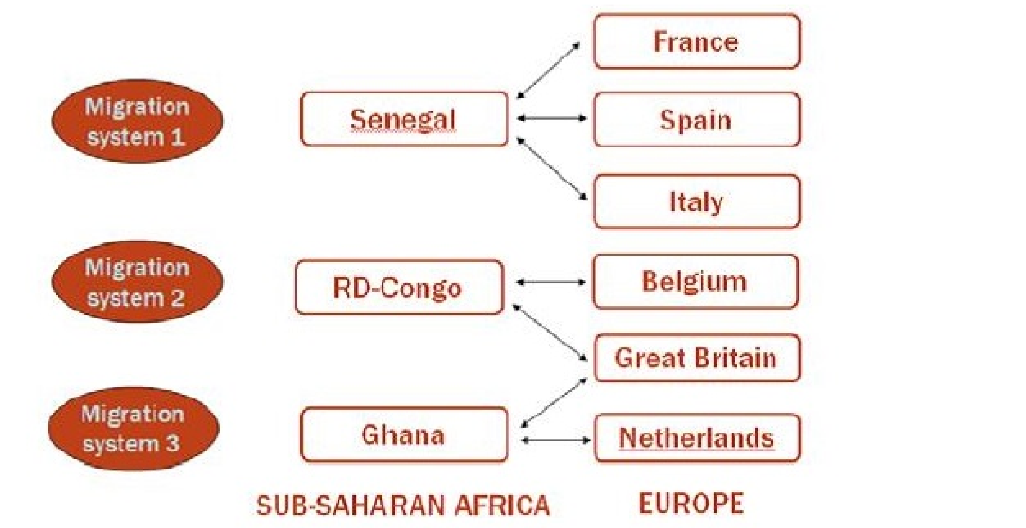


Figure 4: Countries involved in the MAFE Project

5.2. Descriptive Statistics

Few descriptive statistics are offered in table

The table (2) provides a hint about the share of sequential migration and hence justifies the analysis performed in the present paper: simultaneous migration is clearly less likely than sequential migration. Therefore, there must be some reasons why this is so. It can absolutely due specifically to legal constraints but it is worth analyzing whether economic reason are also involved. In the MAFE data, one can easily notice that being together in either host or source country seems to be rather the same, around half of the people choose either locations. This is especially true for the poorest of the three country, namely the Democratic Republic of Congo. Indeed, more than nine migrants out of ten migrate sequentially and reunify in the host country.

	DRC		Ghana		Senegal		Total	
	Freq.	Weighted	Freq.	Weighted	Freq.	Weighted	Freq.	Weighted
Simultaneous Migration	39	23%	41	23.9%	63	28.9%	143	26.1%
Source reunification	19	58.7%	49	50.3%	34	32.4%	102	42.8%
Host reunification	51	18.3%	67	25.7%	108	38.8%	228	31.2%

Table 2: Types of family migration

Table (3) illustrates few differences across the couples who experienced different migration patterns. There are four different groups: (i) those in which the spouse or the surveyed only is currently a migrant, (ii) those in which the spouse or the surveyed surveyed has returned, (iii) those in which both spouses are in the host country, (iv) those who migrated together. Therefore, the first group consists of the household still split; the second group characterises the source country reunification; the third display the spouses in the case of host reunification; the last one deals with simultaneous migration.

In general, migrants who have either returned or who reunified in host country migrated at a younger age. The age at migration is close across the migration patterns: migrants are likely to be in their late twenties. This is partly implied by the constraint that migrants should be separated, implying that they have a partner and hence are likely to be older. Indeed, if one included all the migrants the average age would be lower. In the surveys, most of migrants are males, expect in the last column in which the proportion is in favour of women [21]. Family migrants appear to usually be more educated than migrants who are still separated. The GDP per capita in the host country is particularly high for a single migration, in which the separation is ongoing, and it is particularly low in case of simultaneous migration. In the case of source reunification the GDP per capita is significantly lower and in the case of host reunification it is somewhere in between. This suggests that when migration occurs in a richer country return migration is less likely. Source country per capita GDP demonstrate few notable differences with the exception of simultaneous migrants. Sequential migration happens more frequently when there is a large gap between source and host countries. This is also true with the price level variable. Additionally, it seems that host reunification happens more frequently when the source price level is relatively high. When the host country was the former colonizer the migrants are more frequently staying there. This could be partly explained by the common language that the countries might still share but this might not be a sufficient explanation. There are institutional links among countries such as the *Commonwealth* or the *Francophonie*. Lastly, the mean number of years separated highly changes according to the groups. For the censored groups, *i.e.* when only one household member is currently in migration, the mean time being separated is much higher than for the groups of reunified. Among reunifiers

[21] This last result might be driven by the survey design itself in which men might have been less likely available to directly answer.

host reunification seems to take longer, but this is not significantly higher. Including simultaneous migration would simply push down the mean time of separation in the case of both spouses are together in host country.

Variables	**One migrant**	**Return migration**	**Both at migration**	**Simultaneous**
Age at separation	30.3	29.6	26.8	28
Gender of the first migrant		*Proportions*		
Male	94.3	78.7	88.6	58.6
Diploma	10	13.6	12	13.8
(ln) GDP per capita in host	10.36	10.14	10.01	9.74
(ln) GDP per capita in source	6.57	6.37	6.5	8.74
Price level in host	1.03	1.05	1.08	1.08
Price level in source	0.432	0.385	0.462	0.814
(ln)Distance	8.38	8.42	8.36	8.43
Is source a former colony?		*Proportions*		
	41.2	56.3	66.6	29.4
Do host and source share a language?		*Proportions*		
	43.6	63.5	68.2	85.4
τ (no simultaneous migration)	8.64	3.63	4.59	0

Table 3: Characteristics of different types of couples

5.3. Estimation of τ - When reunifying?

The MAFE data provides a unique opportunity to estimate τ - the time being separated - either in the source reunification scenario or in the host reunification scenario or eventually with non-yet reunified couples (see AppendixC).

Survival Analysis. Survival analysis appears to be a promising way of dealing with the research question for several reasons: (i) it is meant to study the waiting time until a defined even happens, here the reunification of the spouses; (ii) survival analysis has the key advantage to to into account censored data, which are the observations that have not encountered the fateful moment. Usually the fateful moment is death or marriage, but this is not necessary. In my analysis, that will be τ, the reunification moment. In MAFE data, there are migrants who are alone in the host country with a spouse in the source country at the time of the study. Those observations are right censored. (iii) Survival analysis eventually allows to study the impact of predictors. Hence I will be able to assess how τ is impacted by several variables included in the model.

Competitive Risks. The present study requires a slight complication of the basic survival analysis as the fateful moment can either be host or source reunification. Both cannot happen simultaneously. Therefore, it is as if the two possible reunification events were competing with each other until one happens. Therefore, I will apply the competing risks analysis (see Cleves *et al.* 2008 [50] or Digman *et al.* 2012 [51] or Austin *et al.* (2016) [52] for simple presentations). This requires a slight refinement of the casual hazard and survival functions and derive cause-specific hazard and cumulative incidence functions instead. The data are then composed of three distinct groups: (i) the censored observations for which the separation is still ongoing, (ii) the ones who reunified in the source country, (iii) those who reunified in the host country. Estimated survival functions indeed differ across the type of reunion. It appears that the migrants choosing host reunification usually remain split longer. Indirectly this specification is helpful to avoid biases that would emerge through potential informative censoring. Indeed, one can notice in table (3) that censored observations bear differences with the other groups of interest. It matters as Lin and Wei (1989) [53] stressed that inference in Cox proportional model can still be obtained even though the vector of regression coefficients are misspecified. Given the very few variables used in the present analysis this property is clearly appealing.

The cause specific hazard gives the probability of experiencing a failure for the reason i at a specific moment, provided that failure has not happened so far. It can be written in the following way:

$$h_l(t) = lim_{\Delta t \to 0} \frac{P[t < \tau < t + \Delta t, \text{failure from cause l } |t \leq \tau]}{\Delta t} \tag{19}$$

The only difference with classical hazard function consists in the additional constraint that failure has to happen for a specific reason. In this scenario, l is only a set of two components: $\{source; host\}$. Put differently, the subhazard accounts for the rate of failure for either reunifying in source or in host at a given instant and given that reunification has not occurred so far. In the case of competing risks the Survival function is of little help as it does not distinguish the types of reunifiers; instead, it is common to rather use the Cumulative incidence function (CIF):

$$\begin{aligned} CIF_l(t) = P(\tau \leq t \text{ \& failure from cause l}) \\ = \int_0^t h_l(x)S(x)dx = \int_0^t h_l(x)e^{-\sum_{j=1}^{2}\int_0^x h_j(u)du}dx \end{aligned} \tag{20}$$

The CIF counts the number of failure per units of time - year - for a given reason. Hence the CIF is generally not bounded by unity as are classic cumulative functions but the sum of $CIF_l(T)$ is[22].

Covariates. As covariates, I include the GDP per capita of both source and host countries as well as the price conversion factor with respect to the purchasing power parity. The first variable is meant to proxy the wage of the migrant, the second to approximate the price level. The data are taken from the World Development Indicators table. Importantly, I pick the values at the time of reunification and in 2009 for the censored observation as the MAFE data was built in 2009. This is in line with two strong assumptions of the model, (i) agents do not suffer from any uncertainty and (ii) wages as well as price levels are assumed to be fixed. Hence can make there decision with the reunification values.

Distance from host and source countries is a rough approximation of the cost of migration. I use the *GeoDist database* from Cepii (see Mayer and Zignago in 2011 [54]) to compute the geodesic distance from the capitals of both countries[23]. Moreover, it will be added two proxies, one dealing with common languages between source and host countries, the other with a potential link with colonial past. Form the MAFE data itself, I use the age of the surveyed at the separation. The smaller the age the higher the expected life expectancy of the migrant and, henceforth, of the couple. I also add the gender and a binary variable that take the value of one if the migrant's work is considered as intermediary or highly skilled (following isco taxonomy). The level of education, proxied by the number of year of schooling, can add some information about the wage of the surveyed. In contrast with the GDP per capita that is supposed to give a moment of location, education level would rather highlight a plausible impact of the spread of the wages, assuming

[22] There is only one failure per individual and both type of failures cannot happen in the scenario I set up. However, this is a restriction as it is possible that a migrant got married several times and did several reunifications. Another possibility is that the reunification happens in, say the host country, but then the couple decide that they are better off in the source country and hence go back in the same year. Since I cannot work with any finer time interval than year this can exist though it seems unlikely not to be a clear minority of cases.

[23] Having the precise location might be more accurate but this will only change distances marginally as the distance between, for instance, Madrid and Dakar is not so different from the distance between Valencia and Dakar, or Barcelona and Dakar. Moreover, the MAFE database purposefully sample individuals from large cities, mostly for the capital and the suburbs. Therefore, the distance is likely to be rather decently computed

as in a Mincer-type regression that education impacts wages. It could also have an impact in a Roy-Borjas model, involving that remuneration would be dependant on the level of education (see AppendixA for a discussion of the potential Roy-Borjas setting in the family-based migration). In the case of source reunification, the issue would rather deal with the transferability: do return migrant who are relatively highly educated suffer more or less from reintegration? The issue is not raised in the theoretical model, though this is clearly an important question potentially issuing future research. Lastly, I add gender because it seems that there is a strongly unbalance proportion of males among migrants.

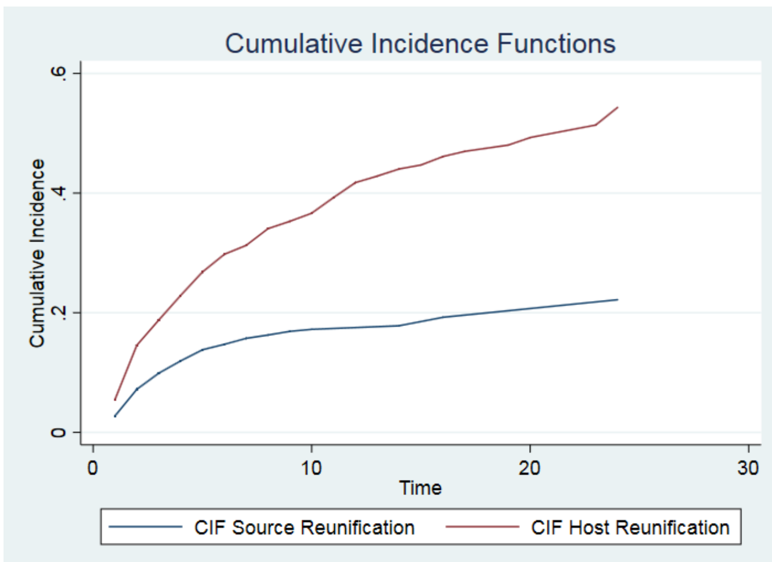


Figure 5: Non-parametric estimation of the Cumulative incidence Function

Cox Model. The estimation of τ is done according to the equation:

$$h_l(t|\mathbf{x}_i) = h_{0,l}(t)e^{\mathbf{x}_i\beta_x} \tag{21}$$

This denotes a simple Cox model with the refinement that hazards are specific to cause of reunification. The theoretical model highlights the need of the competitive risks frame, and the graph in figure (C.6) illustrates that the probability of reunifying by t differs across types of reunification and is more likely in the case of source reunification. Here, $\mathbf{x}_i$ accounts for the covariates presented above. Notice that I simply look at linear impact of the covariates on the probability of reunifying at time t for reason l. This clearly is for the sake of simplicity of interpretation and comparison with the model's intuitions[24]. Though very simple, the Cox model allows the baseline hazard function $h_{0,l}(t)$ to be unspecified (but nonnegative) and thus the baseline function can freely take any functions. It can be found in the appendix equivalent regression performed with parametric models, results are similar. The Cox model remains it this very simple setting and the restrictive proportional hazards assumption does not seem to be violated.

Subhazard Function to Model the CIF. In addition of the classic Cox model, I also display the subhazard function model derived by Fine and Gray (1999). This allows for different hazard patterns for the two competitive events. The hazard is slightly different than in equation (22) as it introduces the possibility of the other type of reunification given that the first one occurred. The sub-hazards are built according to the following limit:

$$\bar{h}_l(t) = lim_{\Delta t \to 0} \frac{P[t < \tau < t + \Delta t, \text{failure from cause l } | t \leq \tau \text{ OR } (\tau < t \text{ \& not l})]}{\Delta t} \tag{22}$$

Obviously, this does not bring additional light to the model presented in this context, but it still enriches the analysis by better bridging the explanatory variables to cumulative incidence as it is the case in a non competing structure. It is very clear that the Fine and Gray approach parallels the classic Cox model, both are semi-parametric and assuming proportional hazards. Only the baseline function might differ.

$$\bar{h}_l(t|\mathbf{x}_i) = \bar{h}_{0,l}(t) e^{\mathbf{x}_i \beta_x} \tag{23}$$

[24]One can go to AppendixE to have the results using a flexible parametric survival model that allows for more dimensions in terms of covariates' effects on the time being separated.

	SOURCE		HOST	
	(1)	(2)	(3)	(4)
	Cox Model	Sub-hazards	Cox Model	Sub-hazards
Age at separation	0.018	0.023**	-0.036***	-0.043***
	(0.012)	(0.011)	(0.011)	(0.011)
Gender of first migrant	0.266	0.226	-0.073	-0.114
	(0.252)	(0.243)	(0.202)	(0.199)
Years of schooling	0.047***	0.029*	0.076***	0.066***
	(0.018)	(0.017)	(0.012)	(0.011)
ln(GDP/capita) in host	-0.369***	-0.320***	-0.089	-0.012
	(0.111)	(0.113)	(0.094)	(0.088)
ln(GDP/capita) in source	0.265	0.304*	-0.385***	-0.361**
	(0.183)	(0.178)	(0.122)	(0.141)
Price level in host	1.490*	1.591**	-0.749	-0.955*
	(0.808)	(0.772)	(0.565)	(0.575)
Price level in source	-2.569***	-2.663***	-0.059	0.250
	(0.969)	(0.996)	(0.573)	(0.567)
ln(distance)	-0.223	-0.007	-1.049***	-0.963***
	(0.449)	(0.300)	(0.227)	(0.278)
Colony	-0.305	-0.389	0.743**	0.680*
	(0.443)	(0.428)	(0.358)	(0.355)
Language	1.472***	1.462***	-0.546	-0.640*
	(0.461)	(0.450)	(0.353)	(0.364)
N	3903	3903	3903	3903
Number of surveyed	651	651	651	651
Number of reunifiers	104	104	227	227
Number of competing events		227		104
Number of still separated		320		320
pseudo-R2	0.059		0.027	
chi2	74.324	60.630	71.793	74.789

Table 4: Estimation on hazards of separation

The results. The tables display the several models I exposed above. Both reunification types are depicted with the classic Cox model as well as with the subhazard model as it is stressed that both approaches should be included in the analysis (see Latouche *et al.* in 2013 [55] or Mozumder *et al.* in 2017 [56] for details). Differences are minor. Usually, it is said that cause-specific hazards better infer the rate of reunification while subdistribution hazards would better infer the risk of reunifying.

Given the small number of individuals, the results are not highly significant as the standard errors inflate. Indeed, only 104 surveyed reunified in the source country and 227 in the case of host reunion. There are 320 households that are separated at the moment of the survey.

It should be recalled that the results here concern the impact of the covariates on the hazards and not on τ. The impact on τ is reverted as an increase in one covariate will push down the survival function if the coefficient related to the variable is positive ($\frac{\delta S(t)}{\delta x_i} < 0 \; \forall \; \beta > 0$). This is straightforward: a positive impact on hazard decreases the probability of surviving. Obviously the impact on τ goes along the survival function and is then reverted to the one on hazard.

The table should be read in several manners. Firstly one can compare the results of the two suggested models. Cox and sub-hazards models provide fairly comparable results with very few differences. Secondly, a comparison between source and host reunification raises a key result: the variables often have reverted impacts (but coefficient are rarely significant for both types of reunification), as anticipated in the model, provided that the source country was much poorer than the host country. Thirdly, the table can be compared to the comparative statics provided above.

The model could not bring clear predictions about the variation of τ given an increase of the wage(s) in the host country. Indeed, in the case of host reunification, the spouses' wages are pushing τ in contradictory directions and it remains unclear in the case of source reunification. It appears that the richer the country per inhabitants the shorter both τs. The GDP per capita of the source country in the case of reunification in the host country has the anticipated sign: a higher wealth of the source country implies that the spouse is better off by staying more there, so separation lasts longer. The return migration does not seem impacted by the wealth of the source country *per se*. However, it appears that higher prices in the source country contributes to postpone the reunification: the migrant prefers to stay an extra-unit of time in the host country and benefit from *relatively* higher standards of living. This is confirmed by the impact of prices in the host country on the probability of reunifying at time t. On the other hand, prices do not seem to contribute to explain the time spent separated. Distance contributes to delay reunification in the host country. A colonial past between host and source countries or a common official language might also take part of the cost of migration. Indeed, inclusion on the labour market is related with the language skills and a colonial past often implies institutional agreements and a stronger network of migrants already in the host country. When studying the source reunification, only the variable dealing with a common language seems to have a significant impact and it is in line with what the model predicts for the migration costs: it should reduce the delay.

Without being fully convincing, the matching between the data and the model might seem surprisingly good, given the simplicity of the model that does not include any legal aspects. It is interesting to notice that gender does not impact the pace of reunifying. Foged (2016) showed that the gender is not a key determinant to family migration but the educational power of each spouse is. The result here seems to pursue this phenomenon as reunification appears to be independent from gender. Migrating is a cornerstone decision, not only of the individual, but also of the family and it seems that purely rational motives, in terms of economics, take a big part of the decision made.

5.4. Choice of reunification - Where reunifying?

A simple illustration. The final step requires to check whether value functions in the model lead to accurate prediction in terms of the choice of reunification. In order to do so, I simply use the dichotomous variable $y = 1 \; \forall$ *reunification in source* and $y = 2 \; \forall$ *reunification in host*. The setting is classic, one individual appears once, hence N = 333 which accounts for the sum of both types of reunifiers. The econometric model is extremely simple:

$$f(E[Y|\mathbf{x}_i]) = \mathbf{x}_i\beta \qquad (24)$$

where $f(.)$ is the linear function in the case of linear probability model, the logistic function for a logit model, the inverse normal distribution for a probit model, the logarithm for a Poisson model. The vector of variables is exactly the one used above but with τ included. Since the variables are weighted it is not possible to compute goodness of fit or several other basic post-estimation statistics. I simply put the unweighted results instead. It should be stressed that this is highly disputable as weighting changes significantly the results. Results are provided in table (5). A striking observation is that the results are barely significant for most of variables. On the other hand, τ has a large impact. As seen in figure (5) host reunification tends to take longer. On the other hand, a more optimistic interpretation would also have a consideration of the signs of non-statistically significant coefficients. There are in line with what is expected: the higher the standard of living in the host country, the more likely to settle there by reunifying and *vice versa* regarding the reunification in the source country. It it noticeable that the colony and language variables are impacting though both are slightly collinear[25].

	(1) Liner Probability Model b/se	(2) Logit b/se	(3) Probit b/se	(4) Poisson b/se
Time being separated	0.031***	0.159**	0.095***	0.050***
	(0.011)	(0.064)	(0.036)	(0.019)
Age at separation	-0.006	-0.030	-0.017	-0.012
	(0.006)	(0.030)	(0.018)	(0.016)
Gender of first migrant	-0.178	-0.807	-0.464	-0.551
	(0.118)	(0.632)	(0.359)	(0.434)
Years of schooling	-0.003	-0.010	-0.006	-0.008
	(0.007)	(0.034)	(0.020)	(0.014)
ln(GDP/capita) in host	0.087*	0.394	0.225	0.232
	(0.048)	(0.286)	(0.164)	(0.172)
ln(GDP/capita) in source	-0.026	-0.138	-0.085	-0.032
	(0.070)	(0.353)	(0.208)	(0.174)
Price level in host	-0.541*	-2.790*	-1.614*	-1.361
	(0.304)	(1.532)	(0.900)	(0.885)
Price level in source	0.560	2.638	1.676	0.854
	(0.364)	(1.842)	(1.071)	(0.745)
ln(distance)	-0.073	-0.613	-0.329	-0.284
	(0.160)	(0.997)	(0.582)	(0.394)
Colony	0.396***	3.161***	1.794***	1.550***
	(0.117)	(0.997)	(0.523)	(0.478)
Language	-0.248*	-2.381**	-1.360**	-1.166**
	(0.131)	(0.988)	(0.529)	(0.462)
Constant	0.888	4.526	2.414	1.092
	(1.737)	(10.804)	(6.273)	(5.242)
N	331	331	331	331
r2	0.165			
r2_p		0.139	0.140	0.029
chi2		57.286	57.585	18.225

Table 5: Where to reunify?

Future research. Without delivering fully satisfying results, the analysis here allows to envision some future research. Including the children is one promis-

[25]The variance inflating factor on an unweighted LPM would amount to 7.16 in the case of the colony variable. This might be troublesome but is not above the threshold usually used which is a VIF equal to 10.

ing alley of research. The model should also be tested in case of South-South or North-North migration. To do that, as well as to provide a more satisfactory data analysis, one should use micro-data for the wages in order to better compare the family migration plan with the model. A richer model would also take into account the selection of migrant couples and that will connect this model with several existing models (Mincer in 1978 [3], Borjas and Bronars in 1991 [57], Mont in 1989 [28], Eliasson *et al.* in 2014 [31], Foged in 2016 [32] or in a different setting Munk *et al.* in 2017 [30]).

6. Conclusion

This analysis has allowed to have a first and large approach of the family reunification issue when thought with an economic perspective. An empirical focus, backed with the model and a calibration, studied the South to North sequential family migration process with the use of the MAFE database that focuses on migration from Africa to Europe. Though not all, several intuitions exposed in the model are confirmed in the present empiric work. International family migration, depending on the context, can especially concern sequences of migration rather than a "one shot".

This paper offers a first step to highlight how wages, prices, length of migration impact the odds of reunifying. Future work, potentially bridging the discounted utility model and individual heterogeneity would clearly helpful in order to better understand the migration patterns.

If a broader conclusion were to be given, firstly, one can notice that even a very restrictive model which accounts for economic factors allow to derive rather appropriate conclusion about the family reunification in the host country. The source country family reunification appears to be somehow less dependent on economic factors or, at least, the reunifiers in the source country usually benefited from a very specific up-league in terms of social status, implying their wage to skyrocket there. I have not included this analysis as this is not reflected in the data. Secondly, a simple discounted utility-maximizing framework gives accurate insights, hence agents might behave quite rationally when decision are as important as the choice of migration which, presumably, will deeply affect the agent's life.

[1] Larry A Sjaastad. The costs and returns of human migration. *Journal of political Economy*, 70(5, Part 2):80–93, 1962.

[2] John R Harris and Michael P Todaro. Migration, unemployment and development: a two-sector analysis. *The American economic review*, 60(1):126–142, 1970.

[3] Jacob Mincer. Family migration decisions. *Journal of political Economy*, 86(5):749–773, 1978.

[4] Robert A Nakosteen and Michael Zimmer. Migration and income: the question of self-selection. *Southern Economic Journal*, pages 840–851, 1980.

[5] Albert Kraler, Eleonore Kofman, Martin Kohli, and Camille Schmoll. *Gender, generations and the family in international migration*. Amsterdam University Press, 2012.

[6] Oreste Foppiani and Scarlatescu Oana A. *Family, Separation, and Migration:An Evolution-Involution of the Global Refugee Crisis*. Peter Lang, 2018.

[7] Guillermina Jasso and Mark R Rosenzweig. Family reunification and the immigration multiplier: Us immigration law, origin-country conditions, and the reproduction of immigrants. *Demography*, 23(3):291–311, 1986.

[8] Saskia Bonjour and Albert Kraler. Introduction: Family migration as an integration issue? policy perspectives and academic insights. *Journal of Family Issues*, 36(11):1407–1432, 2015.

[9] Emmanuel Todd. *La destin des immigrés: assimilation et ségrégation dans les démocraties occidentales*. Éditions du Seuil, 1994.

[10] Yann Algan, Alberto Bisin, Alan Manning, and Thierry Verdier. *Cultural integration of immigrants in Europe*. Oxford University Press, 2013.

[11] Caroline B Brettell and James F Hollifield. *Migration theory: Talking across disciplines*. Routledge, 2014.

[12] Andrew Donald Roy. Some thoughts on the distribution of earnings. *Oxford economic papers*, 3(2):135–146, 1951.

[13] George J Borjas. Self-selection and the earnings of immigrants, 1987.

[14] Oded Stark and J Edward Taylor. Relative deprivation and international migration oded stark. *Demography*, 26(1):1–14, 1989.

[15] Slobodan Djajić and Alexandra Vinogradova. Liquidity-constrained migrants. *Journal of International Economics*, 93(1):210–224, 2014.

[16] Daniel Chiquiar and Gordon H Hanson. International migration, self-selection, and the distribution of wages: Evidence from mexico and the united states. *Journal of political Economy*, 113(2):239–281, 2005.

[17] Michael A Clemens, Claudio E Montenegro, and Lant Pritchett. Bounding the price equivalent of migration barriers. 2016.

[18] Bastien Chabé-Ferret, Joel Machado, and Jackline Wahba. Remigration intentions and migrants' behavior. *Regional Science and Urban Economics*, 68:56–72, 2018.

[19] Douglas S Massey. Social structure, household strategies, and the cumulative causation of migration. *Population index*, pages 3–26, 1990.

[20] Elizabeth Fussell and Douglas S Massey. The limits to cumulative causation: International migration from mexican urban areas. *Demography*, 41(1):151–171, 2004.

[21] Christian Dustmann, Francesco Fasani, Xin Meng, and Luigi Minale. Risk attitudes and household migration decisions. 2017.

[22] Dean Yang. International migration, remittances and household investment: Evidence from philippine migrants' exchange rate shocks. *The Economic Journal*, 118(528):591–630, 2008.

[23] Ralitza Dimova and François-Charles Wolff. Do downward private transfers enhance maternal labor supply? evidence from around europe. *Journal of Population Economics*, 24(3):911–933, 2011.

[24] Michael Clemens and David McKenzie. Why don't remittances appear to affect growth? 2014.

[25] Pablo A Acosta, Emmanuel KK Lartey, and Federico S Mandelman. Remittances and the dutch disease. *Journal of international economics*, 79(1):102–116, 2009.

[26] Guillermina Jasso and Mark R Rosenzweig. Remit or reunify? us immigrant parents, remittances, and the sponsorship of children. In *Research Conference on Remittances and Immigration, Federal Reserve Bank of Atlanta, Atlanta, GA*, 2010.

[27] Ahu Gemici. *Family migration and labor market outcomes.* University of Pennsylvania, 2007.

[28] Daniel Mont. Two earner family migration a search theoretic approach. *Journal of Population Economics*, 2(1):55–72, 1989.

[29] George J Borjas and Stephen G Bronars. Consumer discrimination and self-employment. *Journal of political economy*, 97(3):581–605, 1989.

[30] Martin Munk, Till Nikolka, and Panu Poutvaara. International family migration and the dual-earner model. 2017.

[31] Kent Eliasson, Robert Nakosteen, Olle Westerlund, and Michael Zimmer. All in the family: Self-selection and migration by couples. *Papers in Regional Science*, 93(1):101–124, 2014.

[32] Mette Foged. Family migration and relative earnings potentials. *Labour Economics*, 42:87–100, 2016.

[33] Mark R Rosenzweig and Oded Stark. Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of political Economy*, 97(4):905–926, 1989.

[34] Guillermina Jasso and Mark R Rosenzweig. Do immigrants screened for skills do better than family reunification immigrants? *International Migration Review*, pages 85–111, 1995.

[35] Slobodan Djajić and Ross Milbourne. A general equilibrium model of guest-worker migration: The source-country perspective. *Journal of international economics*, 25(3-4):335–351, 1988.

[36] Slobodan Djajić. Migrants in a guest-worker system: a utility maximizing approach. *Journal of Development Economics*, 31(2):327–339, 1989.

[37] Christian Dustmann. Children and return migration. *Journal of population economics*, 16(4):815–830, 2003.

[38] Christian Dustmann. Return migration, wage differentials, and the optimal migration duration. *European Economic Review*, 47(2):353–369, 2003.

[39] Jean-Christophe Dumont, Gilles Spielvogel, et al. Return migration: A new perspective. *Organization for Economic Cooperation and Development (OECD), International Migration Outlook, Annual Report*, 2008.

[40] Augustin De Coulon and François-Charles Wolff. Location intentions of immigrants at retirement: stay/return or go 'back and forth'? *Applied Economics*, 42(26):3319–3333, 2010.

[41] Slobodan Djajić. Immigrant parents and children: an analysis of decisions related to return migration. *Review of development economics*, 12(3):469–485, 2008.

[42] Slobodan Djajić. Transit migration. *Review of International Economics*, 25(5):1017–1045, 2017.

[43] Slobodan Djajić. Investment opportunities in the source country and temporary migration. *Canadian Journal of Economics/Revue canadienne d'économique*, 2(43):663–682, 2010.

[44] Alexandra Vinogradova. Illegal immigration, deportation policy, and the optimal timing of return. *Journal of Population Economics*, 29(3):781–816, 2016.

[45] Elisabeth HM Eurelings-Bontekoe, Evelien PM Brouwers, and Margot J Verschuur. Homesickness among foreign employees of a multinational high-tech company in the netherlands. *Environment and Behavior*, 32(3):443–456, 2000.

[46] Raj Chetty. A new method of estimating risk aversion. *American Economic Review*, 96(5):1821–1834, 2006.

[47] Cris Beauchemin. Migrations between africa and europe: Rationale for a survey design. Technical report, Institut National d'Etudes DÃ© mographiques, 2012.

[48] Sadio Ba Gning and Philippe Antoine. Polygamie et personnes âgées au sénégal. *Mondes en développement*, (3):31–50, 2015.

[49] Family reunification of third-country nationals in the eu plus norway: National practices.

[50] Mario Cleves, William Gould, Roberto Gutierrez, and Yulia Marchenko. *An introduction to survival analysis using Stata*. Stata press, 2008.

[51] James J Dignam, Qian Zhang, and Maria N Kocherginsky. The use and interpretation of competing risks regression models. *Clinical Cancer Research*, pages clincanres–2097, 2012.

[52] Peter C Austin, Douglas S Lee, and Jason P Fine. Introduction to the analysis of survival data in the presence of competing risks. *Circulation*, 133(6):601–609, 2016.

[53] Danyu Y Lin and Lee-Jen Wei. The robust inference for the cox proportional hazards model. *Journal of the American statistical Association*, 84(408):1074–1078, 1989.

[54] Thierry Mayer and Soledad Zignago. Notes on cepii's distances measures: The geodist database. 2011.

[55] Aurelien Latouche, Arthur Allignol, Jan Beyersmann, Myriam Labopin, and Jason P Fine. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *Journal of clinical epidemiology*, 66(6):648–653, 2013.

[56] Sarwar Islam Mozumder, Mark J Rutherford, Paul C Lambert, et al. A flexible parametric competing-risks model using a direct likelihood approach for the cause-specific cumulative incidence function. *Stata Journal*, 17(2):462–489, 2017.

[57] George J Borjas and Stephen G Bronars. Immigration and the family. *Journal of Labor Economics*, 9(2):123–148, 1991.

[58] Ted A Gooley, Wendy Leisenring, John Crowley, and Barry E Storer. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Statistics in medicine*, 18(6):695–706, 1999.

[59] Roberto G Gutierrez et al. Parametric frailty and shared frailty survival models. *Stata Journal*, 2(1):22–44, 2002.

AppendixA. Family Migration - Some more Discussion

Table (A.6) is derived from Eurostat data which provide yearly number of legal migrants entering each EU countries. This number is split by four main motives of migration: (i) education, (ii) family, (iii) occupation, (iv) other (which includes different motives such as health, asylum). One can notice that for all countries displayed here but the UK, the family-based migration accounts for a third up to a half of the immigration flow. The UK has a different legal frame, which justifies the role of outlier it plays here (see the European Commission report for more details [49]).

Table (A.7) is based on the MAFE data used in the core analysis. It is interesting to see how the data of the source countries corroborates the macro data from the host countries: indeed, around 40% of the migrants went to one of the European countries mentioned in table (A.6) through the family reunification channel. Mafe data is subjective in the sense that the interviewers directly asked people about the channel of migration without any means of controlling the veracity of the response. This is then more prone to measurement errors. Even though the comparison between the statistics from the two tables are highly disputable, it is striking to see how the objective measure from Eurostat data corroborates with the subjective measure from MAFE data.

	Belgium	France	Italy	Netherlands	Spain	UK
Education	12.9	30.3	7.7	18.8	11.3	40.5
Family	50.7	42.5	32.7	33.7	49.6	1.7
Occupation	10.1	9.4	39.8	18.8	26.8	20
Other	26.3	17.7	19.8	28.7	12.3	37.8

Table A.6: Share of reasons of migration per host countries

	DRC	Ghana	Senegal
Education	9.6	14.7	11
Family	39.6	37.9	43.3
Occupation	3.2	22.3	30.4
Other	47.2	25.1	15.2

Table A.7: Share of reasons of migration per source countries

A last check from Eurostat data is simply to run a excessively simple gravity model and to split the sample by the four categories described above. I cannot argue anything against all the biases simple gravity models suffer from because I do not deal with endogeneity issues. The objective of this illustration is merely to check whether the coefficients, biased similarly over the four sub-samples, would have similar magnitudes and statistical significance.

The econometric model can be written as such:

$$immigrants_{ijt} = \alpha + \beta LogDist_{ijt} + \gamma X_{it} + \delta Y_{jt} + \epsilon_{ijt} \tag{A.1}$$

with X_{it} and Y_{jt}, CGP/capita, Gini, Inflation and Population in source and host countries, respectively. Not suprisingly, the GDP and population coefficients have positive and significant coefficients, and distance has a negative impact. Interestingly, the category "other" also shares those coefficients, implying that the refugees, likely accounting for most of the immigrants in this section also take decisions partly in line with the basis gravity model.

More importantly, the family sub-sample is negatively impacted by inflation in both countries but the level of inequalities (as measured by the Gini coefficient) is not significant while they are significant for the other migration motives. This suggests that the Roy model of migration could fit for some of the migration channels but not for all of them, notably the family motive. The first migrant would potentially chose the host country according to the differences in the distribution of inequalities and his or her personal characteristic. On the other hand, that is unlikely to be the case for the migrant coming thereafter to reunify. This is another reason for the model I display in the present paper not to be further refined with more individual heterogeneity which could raise effects similar to the Roy model.

(Log) Number of Immigrants				
	(1)Education	(2)Family	(3)Occupation	(4)Other
LogDistance$_{ij}$	-0.8855***	-1.1042***	-1.0523***	-0.9124***
	(-16.35)	(-17.65)	(-16.62)	(-14.20)
GDP/capita$_{it}$	0.5440***	0.4230***	0.4511***	0.2861***
	(16.45)	(10.65)	(11.15)	(7.22)
Gini$_{it}$	0.0015	-0.0037	-0.0116**	-0.0090
	(0.32)	(-0.65)	(-2.07)	(-1.56)
Inflation$_{it}$	-0.0107***	-0.0105***	-0.0079**	-0.0047
	(-3.47)	(-2.70)	(-1.99)	(-1.01)
logPopulation$_{it}$	0.6202***	0.6145***	0.6721***	0.5142***
	(28.62)	(22.68)	(26.36)	(18.57)
GDP/capita$_{jt}$	1.7862***	1.8254***	1.5930***	1.6024***
	(16.31)	(17.50)	(13.49)	(11.76)
Gini$_{jt}$	0.0401***	0.0096	0.0466***	0.0966***
	(3.64)	(0.76)	(3.36)	(7.00)
Inflation$_{jt}$	0.0223**	-0.0328***	0.0127	0.0328***
	(2.50)	(-3.09)	(1.37)	(2.88)
logPopulation$_{jt}$	0.8189***	0.6515***	0.7243***	0.7339***
	(29.04)	(18.86)	(20.92)	(21.10)
Constant	-37.9803***	-31.1410***	-32.7909***	-31.9727***
	(-25.63)	(-17.29)	(-19.45)	(-18.24)
N	6543	7085	6230	5733

t statistics, clustered by pairs of countries, in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Immigration in Europe from RoW (2008-2016)

AppendixB. Comparative statics

AppendixB.1. Consumption statics in case of host country reunification

The comparative statics of most partial derivatives of G and B functions are straightforward. From equation (12) one simply uses classic analysis.

$$\frac{d\tilde{c_1}}{dx} = \frac{G_\tau B_x - B_{\tilde{c_1}} G_x}{\Delta} \tag{B.1}$$

The signs of G_τ, $G_{\tilde{c_1}}$, B_τ, $B_{\tilde{c_1}}$ are important to determine. One can easily obtain $G_\tau > 0$ as long as $|\pi''(\tau)| < |\pi'(\tau)|$ or $\pi''(\tau) < 0$. $G_{\tilde{c_1}} > 0$ is also straightforward. The change of the budget balance according to consumption is also simply derived: $B_{\tilde{c_1}} < 0$. The other term B_τ requires some more analysis. The impact of the length of separation on the budget constraint B_τ is also positive. One can see this by computing the derivative and noticing that $B_\tau > 0$ for $\tau = T$ and that $B_{\tau\tau} > 0$[26] so that $B_\tau > 0$ for a duration of separation close to the household's lifetime. Now the question deals with whether $B_\tau > 0$ also holds for smaller values of τ. Using a CRRA individual utility function, one can find that $B_\tau > 0$ holds for $\tau > 1$ if the cost of separation is not too confiscatory (here the cost of separation function would hold a shape in the form of $\pi(\tau) = (\frac{1}{1+\tau})^{1/10}$). Such a cost already implies a drop of almost 7% after one year being separated and over 10% after two years; a hypothetical separation all over the life would lead to a penalty of about a third. The data cannot provide results at a thinner scope than a year and therefore it sounds realistic to consider that a reunification that occurred within a year as a simultaneous migration rather than a sequential migration. To summarize, this leads to $G_\tau > 0$, $G_{\tilde{c_1}} > 0$, $B_{\tilde{c_1}} < 0$, $B_\tau > 0$ which logically implies $\Delta > 0$.

It is clear that $G_{\tilde{w_s}} > 0$, $G_{w_s} < 0$, $G_{\tilde{w}} = 0$, $G_{K_0} = 0$, $G_{K_1} < 0$, $G_{A_0} < 0$, $G_T < 0$, $B_{\tilde{w_s}} > 0 \; \forall \; \tau > 0$, $B_{w_s} > 0 \; \forall \; \tau < T$, $B_{\tilde{w}} > 0$, $B_{K_0} < 0$, $B_{K_1} < 0$, $B_{A_0} > 0$, $B_T < 0$[27]. The derivatives with respect to price levels are less obvious, $B_{\tilde{p}}$ is unambiguously negative but B_p is positive only for $p\psi_p(.) > |\psi(.)|$ with $\psi_p(.)$ the derivative with respect to the home price level. This condition, for example, is met in the individual CRRA case where $\tilde{p}\theta > 1 \equiv p$ (if one fixes, without loss of generality, p as a numéraire), which is realistic in the present South to North migration. Lastly, $G_{\tilde{p}}$ and G_p can be shown to be negative and positive, respectively, using the same restrictions as for the computation of G_τ and $G_{\tilde{c_1}}$. One should simply add that ψ with ψ_p or $\psi_{\tilde{p}}$ and notice that $\psi_{\tilde{p}} < \psi < -\psi_p$ and, make an additional assumption in the CRRA frame: $\frac{\pi'(\tau)}{\pi(\tau)} < \frac{r\theta}{e^{r\tau}-1}$. With the functional form of the separation cost taken here the partial derivatives are satisfactory up to a certain duration. For example, above 8 eight years of separation, with $r = 0.05$, the sign of the derivatives is reverted, implying an unclear result of the impact of prices

[26]A sufficient - but not necessary condition - for this implies that $\pi(\tau)$ is either linear or convex.

[27]This last result only stem for a household that uses the separated period of migration to eventually benefit from better consumption once reunified, which is implicitly the main characteristic of sequential migration

on the length of separation. However, eight years of separation is already an exceptional case, at least in the context of the MAFE data. It should be highlighted that those restrictions are not necessary. The use of CRRA is obviously more restrictive but also more convenient.

Once equipped with all the partial derivatives, it is intuitive to apply equation (B.1) to obtain the results displayed in Proposition 1.

AppendixB.2. Consumption statics in case of source country reunification

Comparative statics in case of source country reunification are derived through a very similar procedure so that it does not seem required to detail it, a parallel with the first case is sufficient.

AppendixC. Time being separated

The richness of the MAFE data allows to estimate the number of years during which the couples were split. One can easily notice that the number of couples being separated decreases exponentially and that the source reunification usually happens quicker than the host reunification.

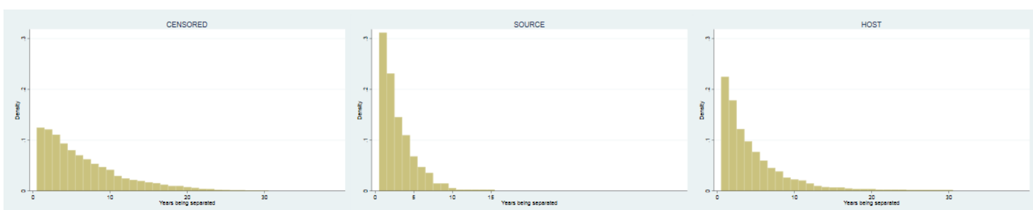


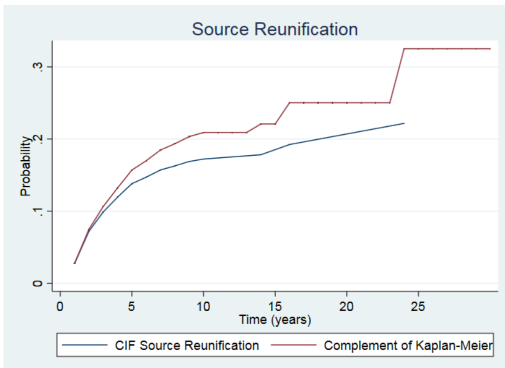
Figure C.6: τ in MAFE data for non censored individuals

AppendixD. Competing risks

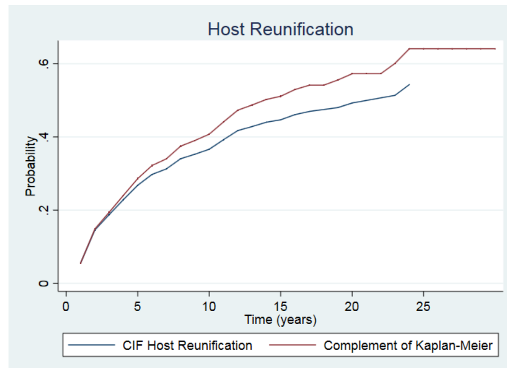
The use of classic tool for the competing risks survival analysis leads to biases. Following Gooley *et al.* in 1999 [58], one can see in the two graphs that, indeed, the usual non-parametric Kaplan-Meier estimation is bias. There is an upward bias (as explained by Austin *et al.* in 2016 [52]). It overestimates the reunifications in both cases.

Figure D.7: CIF versus (complement of) Kaplan-Meier

(a)



(b)



AppendixE. Robustness checks

Flexible parametric survival models. In order to check whether adding dimensions on the regression equations would have an impact, one can run a flexible parametric survival model. This introduces restricted cubic splines to gain flexibility. On the other hand, this implies to drop the non-parametric baseline function and instead use a Weibull distribution (see E.11 for parametric models' results). The cause-specific log-cumulative is written as:

$$ln(H_l(t|\mathbf{x}_i)) = s_i(ln(t)|\gamma_i; m_{0i}) + \mathbf{x}_i\beta_x \tag{E.1}$$

with γs giving the baseline of the cumulative hazards and m_{0i} being the cause-specific number of knots to be included. The bridge from the cumulative to the hazards is not problematic though it implies to include the baseline. The coefficients remain similar to what what obtained in table (4).

	(1) SOURCE b/se	(2) HOST b/se
xb		
Age at separation	0.019	-0.034***
	(0.012)	(0.011)
Gender of first migrant	0.288	-0.083
	(0.253)	(0.203)
Years of schooling	0.053***	0.081***
	(0.018)	(0.012)
ln(GDP/capita) in host	-0.400***	-0.115
	(0.109)	(0.092)
ln(GDP/capita) in source	0.229	-0.438***
	(0.182)	(0.121)
Price level in host	1.510*	-0.845
	(0.804)	(0.565)
Price level in source	-2.698***	-0.165
	(0.964)	(0.580)
ln(distance)	-0.279	-1.133***
	(0.439)	(0.226)
Colony	-0.296	0.815**
	(0.445)	(0.356)
Language	1.491***	-0.575
	(0.463)	(0.351)
restricted cubic spline 1	1.006***	1.154***
	(0.086)	(0.068)
restricted cubic spline 2	0.363***	0.309***
	(0.063)	(0.049)
restricted cubic spline 3	-0.025	-0.082***
	(0.038)	(0.029)
Constant	0.030	12.936***
	(4.426)	(2.683)
dxb		
derivative of restricted cubic spline 1	1.006***	1.154***
	(0.086)	(0.068)
derivative of restricted cubic spline 2	0.363***	0.309***
	(0.063)	(0.049)
derivative of restricted cubic spline 3	-0.025	-0.082***
	(0.038)	(0.029)
N	3903	3903

Table E.9: Flexible Parametric Survival Model

Shared Frailty. A last refinement of the Cox model introduces shared frailty, for details see Gutierrez (2002) [59]. This will take into account random effects among groups of migrants according to the pair of country that exist in the present paper. It should be highlighted that few returned migrants also went in another country, even though the vast majority went to the six European countries presented above. The model including shared frailty is written in the following way:

$$h_{lij}(t|\alpha_j) = \alpha_j h_{lij}(t) = \alpha_i h_l(t|\mathbf{x}_{ij}) \tag{E.2}$$

The subscript j refers to the pairs of countries group. In the shared frailty frame, individuals of different groups are assumed to be independent while they are not within groups. One can see that the model loses most of its significance, which makes sense as the variability of macro-variables are not large compared with the between variability. Therefore, this suggests that the differences are especially important between countries. Of course, with a better approximation of wages the results might be less striking. Nevertheless, this result is useful to have a broad estimation of the willingness of migrants to reunify in host country rather than in source country and to call for more micro-data to assess the choice of reunification.

	(1) SOURCE b/se	(2) HOST b/se
Age at separation	0.010	-0.013
	(0.013)	(0.011)
Gender of first migrant	0.327	0.190
	(0.267)	(0.219)
Years of schooling	0.049**	0.070***
	(0.021)	(0.015)
ln(GDP/capita) in host	-0.354***	0.070
	(0.120)	(0.113)
ln(GDP/capita) in source	0.062	-0.134
	(0.249)	(0.176)
Price level in host	0.928	-1.070
	(0.944)	(0.801)
Price level in source	-1.687	-0.095
	(1.323)	(0.765)
ln(distance)	0.379	-0.658*
	(0.919)	(0.387)
Colony	-0.341	-0.248
	(0.823)	(0.584)
Language	0.666	0.382
	(0.814)	(0.608)
N	3903.000	3903.000
N_sub	651.000	651.000
N_fail	104.000	227.000
pseudo-R2		
chi2	19.103	30.475

Table E.10: Cox model with shared frailty per paris of countries

Different parametric Survival Models. Table (E.11) simply displays the results of parametric models instead of Cox or Fine and Gray setting. The the baseline hazard is assumed to follow, respectively, a Weibull, an exponential, or a Gompertz distribution. The results do not seem to be heavily affected by a change in the shape of the baseline. For some more flexibility, and to stick with the most common model, I will simply use the Cox model as the preferred model.

	SOURCE			(HOST)		
	(1)	(2)	(3)	(4)	(5)	(6)
	Weibull	Exponential	Gompertz	Weibull	Exponential	Gompertz
	b/se	b/se	b/se	b/se	b/se	b/se
analysis time when record ends						
Age at separation	0.027**	0.027**	0.021*	-0.028***	-0.031***	-0.033***
	(0.012)	(0.012)	(0.012)	(0.010)	(0.010)	(0.011)
Gender of first migrant	0.257	0.257	0.261	-0.148	-0.124	-0.098
	(0.254)	(0.254)	(0.252)	(0.202)	(0.201)	(0.202)
Years of schooling	0.063***	0.063***	0.051***	0.086***	0.082***	0.080***
	(0.018)	(0.018)	(0.018)	(0.012)	(0.011)	(0.012)
ln(GDP/capita) in host	-0.409***	-0.408***	-0.377***	-0.124	-0.112	-0.105
	(0.109)	(0.108)	(0.111)	(0.091)	(0.092)	(0.092)
ln(GDP/capita) in source	0.230	0.230	0.255	-0.491***	-0.454***	-0.425***
	(0.185)	(0.185)	(0.184)	(0.123)	(0.121)	(0.123)
Price level in host	1.697**	1.696**	1.575*	-0.773	-0.742	-0.736
	(0.798)	(0.798)	(0.804)	(0.565)	(0.565)	(0.565)
Price level in source	-2.908***	-2.904***	-2.628***	-0.164	-0.124	-0.097
	(0.979)	(0.975)	(0.970)	(0.582)	(0.577)	(0.575)
ln(distance)	-0.413	-0.412	-0.290	-1.260***	-1.200***	-1.145***
	(0.433)	(0.432)	(0.446)	(0.227)	(0.225)	(0.229)
Colony	-0.419	-0.418	-0.354	0.815**	0.779**	0.763**
	(0.447)	(0.447)	(0.444)	(0.362)	(0.359)	(0.357)
Language	1.595***	1.594***	1.516***	-0.604*	-0.576	-0.561
	(0.468)	(0.468)	(0.464)	(0.358)	(0.355)	(0.353)
Constant	-0.564	-0.570	-1.365	12.729***	12.131***	11.556***
	(4.448)	(4.447)	(4.507)	(2.699)	(2.674)	(2.704)
/						
ln_p	0.004			0.103**		
	(0.078)			(0.052)		
gamma			-0.090***			-0.021
			(0.030)			(0.015)
N	3903	3903	3903	3903	3903	3903
Number of surveyed	651	651	651	651	651	651
Number of failure	104	104	104	227	227	227
chi2	88.714	91.584	77.780	85.571	82.043	76.553

Table E.11: Survival Analysis with parametric hazards models

What if we only look at post-1990 since WDI not available before for lots of countries?. For most of countries, the WDI database does not provide any information for the price level by 1990. Therefore, I simply put the last value available, which is obviously terribly restrictive. Then it is worthwhile checking whether the results could be driven by this fact. It appears, on the table (E.12), that they are not.

	(1) SOURCE b/se	(2) HOST b/se
Age at separation	0.018	-0.043***
	(0.014)	(0.013)
Gender of first migrant	0.357	0.039
	(0.310)	(0.244)
Years of schooling	0.052**	0.096***
	(0.022)	(0.015)
ln(GDP/capita) in host	-0.330**	-0.073
	(0.150)	(0.127)
ln(GDP/capita) in source	0.686***	-0.450***
	(0.246)	(0.153)
Price level in host	1.282	-1.083
	(0.966)	(0.674)
Price level in source	-5.006***	0.025
	(1.370)	(0.736)
ln(distance)	-0.162	-1.225***
	(0.463)	(0.294)
Colony	-0.766	0.825*
	(0.563)	(0.441)
Language	1.909***	-0.797*
	(0.572)	(0.435)
N	3017	3017
Number of surveyed	541	541
Number of reunifiers	75	165
pseudo-R2	0.071	0.036
chi2	62.999	68.543

Table E.12: Cox model with post 1990 observation only

Separate countries. Another interesting check deals with whether one source country only explains everything. It appears that results barely change or, at least, signs remain unaffected[28].

[28] There is one exception: gender level once Senegal has been dropped but this is not a key variable in the present study.

	SOURCE				(HOST)	
	(1)	(2)	(3)	(4)	(5)	(6)
	No DRC	No Ghana	No Senegal	No DRC	No Ghana	No Senegal
	b/se	b/se	b/se	b/se	b/se	b/se
Age at separation	0.016	0.010	0.031**	-0.056***	-0.023*	-0.027*
	(0.014)	(0.018)	(0.014)	(0.013)	(0.012)	(0.014)
Gender of first migrant	0.444*	0.291	0.030	0.211	-0.427	-0.084
	(0.268)	(0.408)	(0.302)	(0.233)	(0.298)	(0.238)
Years of schooling	0.055***	0.022	0.033	0.079***	0.081***	0.070***
	(0.020)	(0.024)	(0.032)	(0.014)	(0.014)	(0.024)
ln(GDP/capita) in host	-0.341***	-0.306*	-0.434***	-0.023	-0.197**	-0.033
	(0.124)	(0.174)	(0.130)	(0.115)	(0.100)	(0.139)
ln(GDP/capita) in source	0.106	-0.035	0.334	-0.440**	-0.555***	-0.262*
	(0.276)	(0.289)	(0.210)	(0.176)	(0.176)	(0.138)
Price level in host	1.330	0.645	1.814*	-1.627**	-0.448	-0.371
	(0.937)	(1.107)	(1.011)	(0.696)	(0.653)	(0.822)
Price level in source	-1.158	-1.772	-3.108***	-0.534	0.008	0.337
	(1.459)	(1.427)	(1.168)	(0.835)	(0.790)	(0.704)
ln(distance)	-0.018	-0.359	-0.469	-1.127***	-1.707***	-0.688***
	(0.708)	(0.676)	(0.443)	(0.285)	(0.361)	(0.264)
Colony	-0.013	-0.308	-0.217	0.505	1.280***	0.502
	(1.185)	(0.481)	(0.506)	(0.701)	(0.475)	(0.369)
Language	1.233	1.790***	1.121**	-0.155	-0.985**	-0.467
	(1.164)	(0.528)	(0.535)	(0.693)	(0.451)	(0.368)
N	3374	2929	1532	3374	2929	1532
Number of surveyed	537	450	319	537	450	319
Number of reunifiers	86	53	69	176	160	120
pseudo-R2	0.069	0.058	0.049	0.035	0.034	0.018
chi2	70.235	35.864	35.642	70.570	59.807	21.932

Table E.13: Cox model without one source country each time

No log of GDPs. Table (E.14) provides the same results as above but without taking the logs of GDPs. Signs are not affected.

	(1) SOURCE b/se	(2) HOST b/se
Age at separation	0.022805*	-0.038376***
	(0.012173)	(0.010385)
Gender of first migrant	0.234346	-0.058971
	(0.254258)	(0.202061)
Years of schooling	0.059885***	0.057005***
	(0.019052)	(0.012451)
GDP per capita in host per year	-0.000034***	0.000008
	(0.000008)	(0.000008)
GDP per capita in source per year	-0.000075	0.000067***
	(0.000073)	(0.000016)
Price level in host	1.707795*	-0.549359
	(0.872039)	(0.642612)
Price level in source	-1.593259*	-2.353925***
	(0.944032)	(0.622644)
ln(distance)	-0.934453	-0.269465
	(0.599846)	(0.220475)
Colony	-0.332377	0.350539
	(0.425943)	(0.339414)
Language	1.539782***	-0.169829
	(0.454319)	(0.329740)
N	3903	3903
Number of surveyed	651	651
Number of reunifiers	104	227
pseudo-R2	0.062	0.030
chi2	79.214	80.593

Table E.14: Cox model without logs for GDPs

Post 23 years old people only. Table (E.15) provides the same results as above but it drops the couples whose separation started before the age of 24. This is likely to drop those who might have migrated for educational motives as well, in which case the causality might be reversed. Here also, results are barely modified.

	(1) SOURCE b/se	(2) HOST b/se
Age at separation	0.013	-0.036**
	(0.015)	(0.015)
Gender of first migrant	0.418	0.080
	(0.258)	(0.228)
Years of schooling	0.057***	0.080***
	(0.020)	(0.015)
ln(GDP/capita) in host	-0.436***	-0.123
	(0.120)	(0.122)
ln(GDP/capita) in source	0.368*	-0.424***
	(0.204)	(0.144)
Price level in host	1.752**	-0.488
	(0.863)	(0.642)
Price level in source	-3.667***	0.973
	(1.148)	(0.747)
ln(distance)	-0.232	-0.923***
	(0.475)	(0.288)
Colony	-0.720	0.444
	(0.483)	(0.449)
Language	1.805***	-0.448
	(0.494)	(0.434)
N	2864.000	2864.000
Number of surveyed	515.000	515.000
Number of reunifiers	89.000	155.000
pseudo-R2	0.074	0.030
chi2	76.590	52.325

Table E.15: Cox model without less than 23 individuals when separation occurred