



The structure and dynamics of migration patterns in 19th-century northern France

Claire Lemerrier, Paul-André Rosental

► To cite this version:

Claire Lemerrier, Paul-André Rosental. The structure and dynamics of migration patterns in 19th-century northern France. 2009. halshs-00450035v1

HAL Id: halshs-00450035

<https://shs.hal.science/halshs-00450035v1>

Preprint submitted on 25 Jan 2010 (v1), last revised 4 May 2010 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The structure and dynamics of migration patterns in 19th-century northern France

This is the text of a presentation given by one of the authors in Columbia University in October, 2009 (thanks to the organizers!). It is barely a draft and hence should not be cited or quoted without permission – and it unfortunately includes a number of typos, etc. It however describes a large part of our current results, so that comments are welcome while we are writing the “real” draft.

Abstract: Using empirical data from 19th century Northern France, an area of quick urbanization and industrialization, and techniques derived from dynamic network analysis, this paper first suggests a method to identify invisible spatial patterns, ie hidden channels of preferential relationships between spatial units (in our case the *communes*, which are the administrative equivalent of religious parishes). Second, it studies whether and to what extent these channels were instrumental into migration processes. Did migrants who left to the cities borrow ancient paths that irrigated their village? In that case, urbanization process, which is a traditional symbol of “modernization”, went through old patterns, possibly directly inherited from *Ancien Régime*. Or on the reverse, have they been uprooted by a macroeconomic process that has left no space to previous structures and determinations? Raising this question as such is the only way to check whether the “rural exodus” model has any relevance. In order to test it, we will have to introduce time within our demonstration and statistical tests. To what extent is a spatial pattern at moment B contained within (or at least compatible with the possible evolutions of) a spatial pattern at moment A? Without reducing our hypotheses to a fight between economic and spatial determinations, our empirical work will help us to study precisely how those two dimensions interfere in social processes.

In the last decades, social sciences have been paying a growing attention to the relational aspects of migration – both quantitative (e.g. Palloni *et al.*, 2001, who review the literature on family, social capital and migration) and qualitative (e.g. Fertig, 1998, who contrasts “chain migration” and “over-population” effects, Rosental, 1999, 2006 or Takai, 2001). The decision to migrate has especially been explored in terms of collective dynamics. This approach emphasizes the role of information on distant opportunities, that can be provided by correspondence with family members or former neighbors (Hvidt, 1980).

This line of research has been enhanced by new methodological opportunities such as panel studies, genealogical reconstitution, complex logistic regressions, event history analysis and multi-level modeling (see e.g. Dribe & Lundh, 2005, Courgeau, 2007, Bonneuil *et al.*, 2008). It had been pioneered in the 1950s by the Swedish geographer Torsten Hägerstrand (1916-2004), whose work is however generally not well-known among historians (Rosental, 1997).

Previous models of migration, such as those developed by Ravenstein at the end of the 19th-century, and later by Zipf and Stouffer (see e.g. Taylor, 1975), generally relied on aggregate data. Zipf may be considered as the pioneer in mechanical, so-called “gravity” models, which to some extent are similar to the “push and pull model” developed from the

¹ CNRS-ENS, Institute for Modern & Contemporary History, Paris (Claire.Lemerrier@ens.fr); Sciences Po & INED, Paris (rosental@sciences-po.fr). The data-collection was funded by DATAR (contract « Prospective et territoire »).

1920s onwards. He explained migration flows as a combination between geographic distance (seen as a cost) and total populations (at best seen as a proxy for job opportunities). Stouffer introduced a more realistic conception of space through the notion of “intervening opportunities”, ie the amount of economic opportunities between place of origin and destination. Both opened up a long tradition of migration studies, and searched for a general “law of migration”, a concept which is still heuristically useful but only describes part of the phenomenon.

Hägerstrand, on the contrary, endeavored to build a model which would integrate the analytical and historical perspectives. His vision of migration was derived from his more general studies on diffusion processes. Taking empirical refutations seriously, he tried to understand why pairs of cities or countries did actually exchange more (or fewer) migrants than predicted by Zipf’s or Stouffer’s models. He insisted on “deviations from the ‘inverse-distance rule’” (1957, 126), even arguing that this rule could be nothing more than a limit-case of the exceptions. Long before social network analysis was applied to migration, and in a more systematic way, he emphasized personal contacts and the circulation of information more than macro-phenomena such as urbanization. More precisely, he considered that macro-economic processes, as formalized by Stouffer in a more general way than Zipf, were only a background, which allowed to understand why, globally speaking, poor areas produced migrants towards wealthy ones, but which was unable to explain why some expected flows of mobility did not happen. Migration models had to explain why, all things being equal, migrants starting from a town or country A did or did not go to such or such town or country B located at a given distance and providing the same opportunities. Hägerstrand’s unit of analysis was thus the set of destinations from a given place of origin, which he called “migration field” or “migration area”.

Using the exceptional Swedish nominative data, Hägerstrand empirically demonstrated the relative enduring patterns of migration fields over time, sometimes over almost two centuries, from the 1780s to the 1950s. While describing migration fields as “a chain of connected events” (*ibid.*, 131), not a static structure, he also insisted on their relative inertia, explaining it by a path-dependency mechanism rooted in his relational view of migration:

“It seems as if irregularities, resulting partly from transportation conditions of former times and partly from other ‘historical’ factors, have created a network of social contacts, which tend to conserve a ‘bias’ in the migration frequencies even when changed conditions no longer limit travel. It seems not unlikely that once they have arisen, irregularities in the shape of migration fields have a tendency to perpetuate themselves because migrations at any given time are *dependent* on preceding migrations.” (Hägerstrand, 1957, 130)

Hägerstrand used ingenious maps and sometimes even hand-made simulation to identify the precise chronology and shape of migration fields. He stressed their directional asymmetries. The migration process was cumulative: the more a given rural village produced migrants to the urban world, the more those migrants accumulated into the cities into which this village had previously sent migrants. Hägerstrand demonstrated that it was possible to simulate this empirical evolution. Formalizing it into a model required two conditions: a) treating migration as a diffusion process initiated by pioneers (which in Hägerstrand’s model were randomly simulated, but that historians inspired by his work managed to retrieve in the astonishing Swedish nominative registers²) b) treating space as “heterotropic” rather than “isotropic”. For Zipf or even Stouffer, space was reversible because it was in fact reduced to

² See Akerman *et al.*, 1977, or Rice and Ostergren, 1978, who identified the first migrant sent to the U.S from a Swedish micro-region in the mid 19th century, and demonstrated how the subsequent snowball process was initiated by this pioneer’s relatives and close neighbors.

economic and demographical variables: if a destination place B suddenly gained (resp. lost) half of its population, its attractivity increased (resp. declined) proportionately. For Hägerstrand, each place in a migration field was part of specific patterns shaped by migration history. All things being equal, a migrant from A would go to B rather than C according to the migratory history between the two pairs (A,B) and (A,C). Directionality, in other words, became a part of causal processes.

Additionally, Hägerstrand and his disciples integrated micro and macro processes in a very impressive way: pioneers from village A chose a preferential destination B; they developed human and economic relationships between A and B (marriage, commercial exchanges). Over time, those relationships became institutionalized, e.g. they influenced the creation of roads or railroads. Institutions and personal networks combined their influence to intensify the flows of mobility between A and B (Dahl, 1957).

Studying the migration field of one or a few parishes finally helped to uncover the fact that many if not most of the migratory relationships between places appeared as less hierarchical than symmetrical (in-migration more or less balancing out-migration, and net fluxes being comparatively negligible).

One may object that Hägerstrand's model could only give hints on some more general processes, since the representative character of cases could always be questioned. In addition, it focuses only on the place of origin or destination. Treating both symmetrically is a progress (as in most cases, data are more available or precise for in- rather than out-migration), but this concentric approach fails to consider the interaction between neighboring migration fields. One knows with which towns or countries a place A exchanges migrants, but without having any idea of the aggregate structure of A, B, C, ..., N migrations areas, (A, ..., N) being places in the same region or continent.

Conversely, the global study of flux matrices between towns, regions or countries has its own limits and has not, up to now, incorporated most of Hägerstrand's ideas. While they have benefited from increased computing power and spatial auto-correlation techniques (see e.g. Anselin, 1995, Tobler, 1995), these studies still often seem confined to taking into account the effects of physical distance (or transportation length), population, and, in the best cases, of borders (political, linguistic... see e.g. Cattán *et al.*, 1996) or clustering effects. The "residuals", that are extremely interesting if we want to reconstruct the history of migration fields in the way suggested by Hägerstrand, are not systematically analyzed, and specific relational phenomena between places, such as the tendency to reciprocate migration fluxes, are not included in the models. Finally, the studies of the decision to migrate that we first mentioned, while they now often incorporate a relational dimension, fail to precisely take into account the spatial elements of the choice. At best, some differentiate between short- and long-distance migration; but if migration is considered as an option, how does the actor choose to go north or south, or to a specific city?

The "spatial" and "statistical" logics in the study of migration still seem to be often separated (Rosental, 1997), even if the latter more or less includes relational and not only economic factors. We would therefore like to complement these different lines of research by presenting a method that could quite easily be applied to many regions (with contemporary or historical data). It uses several variants of network analysis to make sense of the residuals of a simple model (based on distance and population), thus providing a multi-faceted description of the structure of a migratory field including 75 different places and of its evolution between three time-periods. This research strategy allows us to simultaneously test macro, historical explanations (rural migration, industrialization...), specific, local cultural factors (here a linguistic border), classical hypotheses on determinants of individual migration (gender, literacy...) and more structural phenomena linked to Hägerstrandian hypotheses on the

relational components of migration (tendencies to reciprocation or transitivity for preferential migration between places). In addition, considering these various scales of the migration phenomenon allows us to discuss the alleged long-term stability of migration fields: some of our indicators seem to point to massive inertia, while others show multiple alterations. We will try to figure out what global pattern stands out of these various possible measures of change.

To our knowledge, network analysis has never been used in this way to study migration relationship between places³. Although Hägerstrand himself was aware of Moreno's sociometric research, he criticized structural network analysis as too static; most geographers and economists since have not paid much attention to social network analysis as developed by sociologists. Our underlying idea is close to that network studies of other fluxes, such as those of international trade (Smith and White, 1992), but we do not limit ourselves to the blockmodeling of a migration field, allowing to describe its structure, as we did it in a previous paper (Lemercier & Rosental, 2000). Treating spatial units as individuals, we use recent techniques developed to model static and dynamic network data, that have up to now mainly been used to study interpersonal relationships⁴; they allow to take simultaneously into account micro and macro, general and relational hypotheses.

Our empirical field of study is 19th-century Northern France, that we have chosen for its “experimental” value. First, it was one of the only areas in 19th-century France that was exposed to brutal industrialization and urbanization, as has been the case in England before or in Germany thereafter. Urbanization towards booming towns or cities of this area has been often studied (e.g. Pétillon, 2006), but there has been almost no research on what went on in the countryside during this process, as if it had been entirely passive and subject to urban attraction. How important were spatial relationships between villages in this area during the rise of urbanization? Did immigration flow go through pre-existing channels (which Hägerstrand’s model would predict), or did it suddenly create new paths of mobility? Did it shatter intra-rural preferential relationships, or did it leave them unchallenged?

Second, there was not one point of destination within the region, but several: three booming cities (Lille, Roubaix, Tourcoing, the former also being an administrative capital city), plus several immigration towns. How were the flows segmented between those destinations? Third, this region was among the few ones in France that were divided by a linguistic boundary: the Eastern, more urban half of it spoke French, the Western, more rural half, spoke Flemish. Fourth, the North was one of the first two massive regions of foreign immigration (with the South East which attracted many Italian migrants). We will not precisely deal here were the articulation between internal and external migration – an understudied question – but we have used the same dataset to explore this question (Rosental, 1996).

We will first more precisely present our case study and hypotheses (1.) and our data (2.), then discuss various results on the general, seemingly quite stable structure of the migration field (3.) and its dynamics (4.), before trying to map the network structure (5.) and briefly discussing avenues for future research (6.).

³ See an application of factor analysis in XXXX.

⁴ For an updated list of applications, see the webpage http://stat.gamma.rug.nl/siena_applications.htm.

1. The case study

Location

As we were especially interested in “routine” migration between villages and wanted to qualify the importance of macro-phenomena such as rural migration and industrialization, we deliberately chose to study one of the French regions in which they were the strongest in the 19th century. Our sample consists of 75 adjacent *communes*⁵ (the finest administrative unit, established during the French Revolution and only slightly modified since) situated in a ca. 50kmx20km zone west of Lille, along the French-Belgian border⁶ (Map 1). It includes Lille (60,000 inhabitants in 1800, 200,000 in 1890) as well as two of the fastest-growing 19th-century towns, due to their manufacturing activity (especially in the wool industries): Roubaix (that grew from 9,000 to 115,000 inhabitants) and Tourcoing (from 11,000 to 65,000), along with many smaller towns and villages; 17 *communes* had less than 1,000 (and more than 200) inhabitants in 1851 (Map 2). In order to test both the stability of migration fields and the effects of urban growth and economic changes, we collected data for three different time periods, around 1825, 1860 and 1880.

Another peculiarity of this region that was interesting for us (and made us choose the west of Lille) is that it was divided by a linguistic border: in its Western part, most of the inhabitants spoke Flemish, not French, in their daily life. This fact was however not documented by official sources at that time: the linguistic map (Map 3) is based on work by scholars who were generally themselves regionalist activists. It is all the more interesting to test if we find patterns correlated with their reconstructions in our data.

Hypotheses

This brief presentation of the region allows us to sum up the hypotheses on the structure and dynamics of the migration field that we will try to test, using several different network analysis techniques in order to get a multi-faceted and multi-scalar view of this structure.

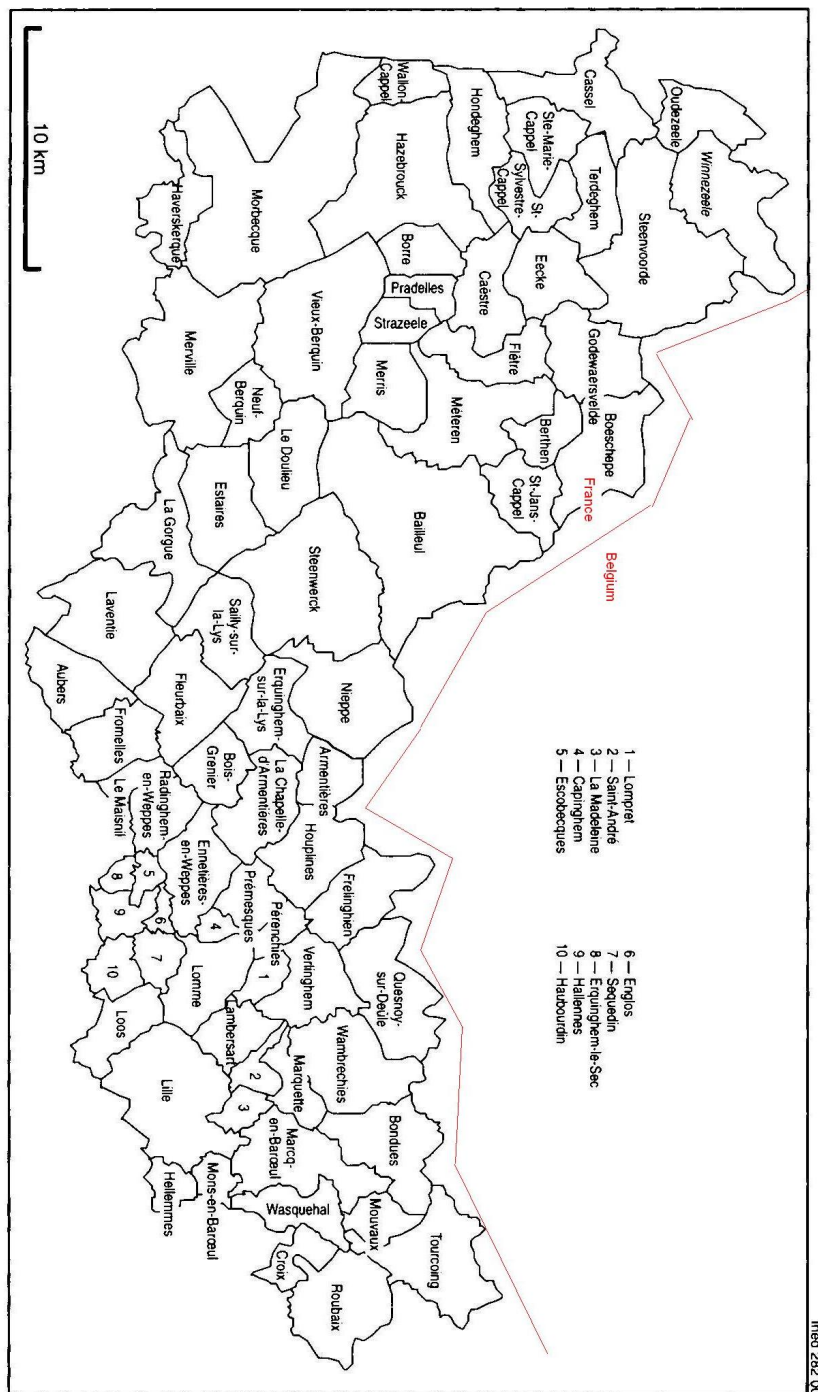
1. It is not unreasonable to suspect that, other things being equal, migration was more common between *communes* that were close to each other and that had many inhabitants. This is summed up by the classical “gravitational model” that states that the total number of migrations between two places (their direction is not considered here) is equal to a coefficient A , multiplied by the populations of the two places, and divided by the distance between them elevated to the power B (as an undue, but now classical extension of the word “gravitational”, which presupposes a power 2). It is quite easy to use a linear regression on the logarithms of the data in order to estimate these coefficients (see e.g. Taylor, 1975). We did this⁷ and only considered the residuals of this simple model in our network study. As the R^2 in our three regressions lies between 0.23 and 0.26, indicating some adjustment but also quite an amount of dispersion, it seems reasonable both to control for distance and population and to have a closer look at residuals.

⁵ Data for Le Douliou are missing, as are those for Haverskerque in the first period and Bois Grenier in the first two periods (this last *commune* was created just before the second period, so that we could only use the third-period data, as we wanted to consider it both as a place of birth and place of dwelling).

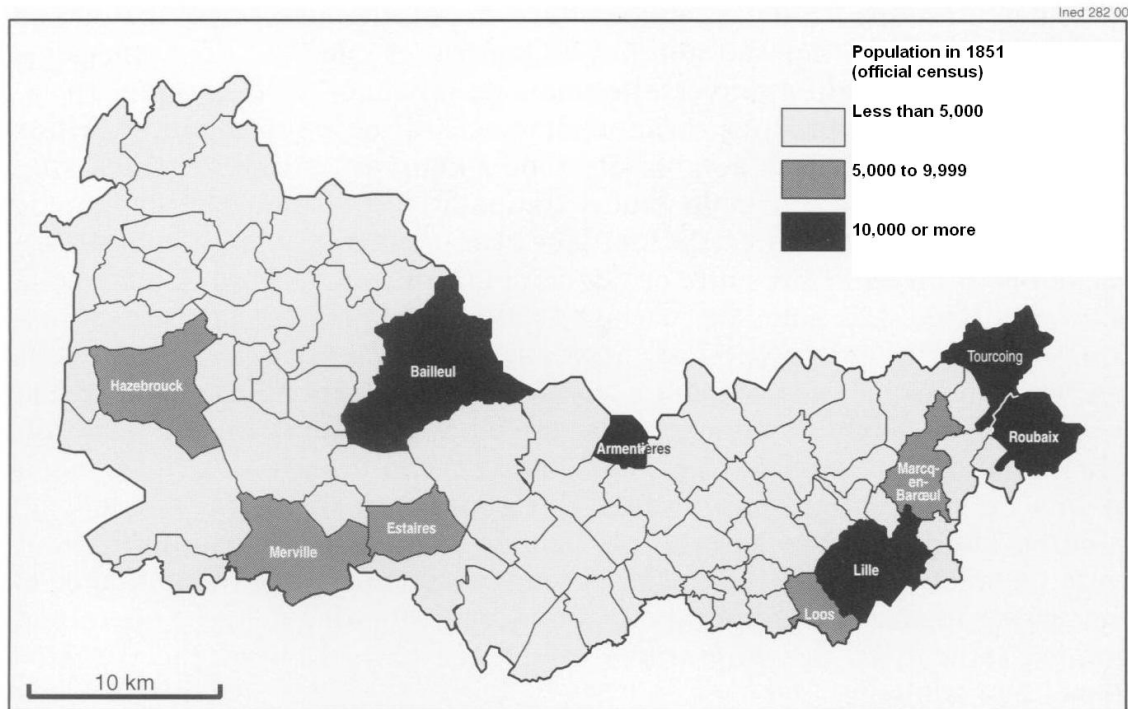
⁶ Belgian migration to our region, which was important and generally shaped by the linguistic border, has been studied in a paper using the same dataset (Rosental, 1996). We could however not include Belgian villages in the study of the migratory field, as marriage records were a bit different and, more importantly, administrative units were much larger, thus giving a quite different view of local migration.

⁷ We used geodesic distance between the city halls of *communes* because they were easily available. There were no obvious physical barriers and many roads in our region, and most of the *communes* were quite small, so that this is probably a reasonable proxy for traveling distance.

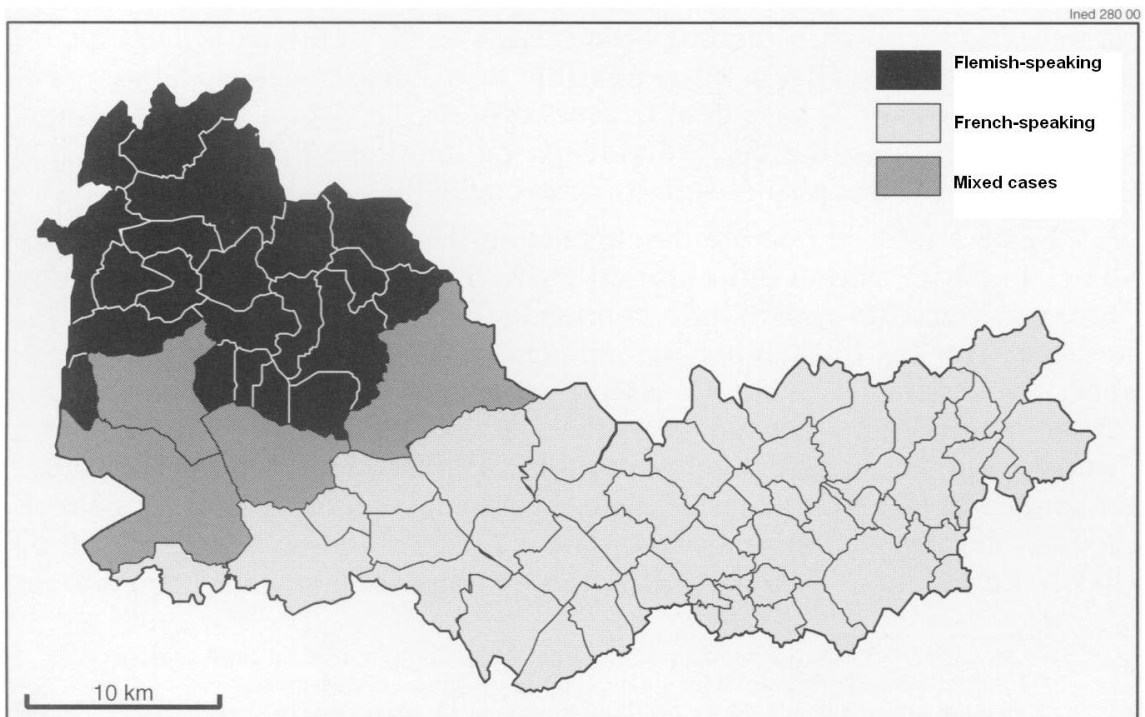
Map 1



Map 2



Map 3



Reconstructed from Kurth, 1896 and Coornaert, 1970.

We have chosen this strategy because if we would have directly processed data on the original network, some of our findings could have been suspected to simply reflect effects of population and distance. It should however be noted that what we consider as “effects of population and distance” are in fact proxies for other social mechanisms, some of which might be quite similar to those that we study below. We can only conclude that some other effects that we will find can be underestimated in this treatment, if some of them are correlated with distance or population.

We have treated the residuals in a simple, binary way. We have considered that there was an “over-attraction” from *commune* A to *commune* B if the observed number of migrants was above the number predicted by our regression for the period: this was coded as a (directed) tie in our network; the opposite case (“under-attraction”) was coded as an absence of tie⁸. This produced matrices with densities between 9 and 10% in each period, meaning that each *commune* had, on average, seven different “over-attraction” ties with others – the number of actual preferential “partners” for in-migration or out-migration actually being in a range of 0 to 19 for each *commune*, period and direction of migration.

2. In the matrix of residuals, we expect to find over-attractions to Lille, Roubaix and Tourcoing, and perhaps more generally to the largest and/or most industrial *communes*. It must be noticed that a large “in-degree” in our network represents a wide, diverse migration field, not necessarily a high number of total “real” migrations. The most important cities are nevertheless likely to have had such a wide field.

We expect this effect to be more and more important in the second and third periods. We have distinguished three periods of analysis⁹. Around 1825, there was no massive mobility towards cities. Around 1860, the urbanization process started, with massive outflows from the villages to the cities and industrializing towns. Around 1880, the intensity of migration was dramatic. It is one of the only cases in which, in 19th century France, one could refer to the traditional vision in terms of “rural exodus”.

In addition, we used data on the presence of industries and of some public equipments in the *communes*, in order to test an approach in terms of opportunities, if not directly intervening opportunities (Stouffer's model being quite difficult to test empirically). Another version of the rural migration argument would state that places offering these jobs or services were attractive to more different *communes* than others.

3. As for migrations between small, rural *communes*, they have been considered by some scholars as a quasi-random (thus insignificant, despite of their weight) “micro-mobility” (Poussou, 1970, 2002). On the contrary, studies of individual migration fields that appeared to be very asymmetrical revealed preferential links. Merely showing the presence of such links is therefore interesting. In addition, there are alternative hypotheses on their patterns and causes. On the one hand, they could form closed clusters of adjacent villages (well-defined subregions, or cliques, in network terms). Else, they could be organized in a more complex way (purely bilateral preferences, non-adjacent places...). What is at stake here is to identify the “invisible” structures of a rural (resp. urban) region, and to check whether they were homogeneous (with similar patterns everywhere) or heterogeneous (cliques here, random short-distance movements there, etc.).

On the other hand, couples or clusters of linked villages could be characterized by a common feature present at the time of observation and thus “explaining” preferential

⁸ This dichotomizing might create unwanted effects, e.g. “changes” in the network due to minor variations around the threshold. We are planning to test the robustness of some of our analyses for a different threshold.

⁹ The exact time span varies according to the demographic importance of each *commune* : the smaller it is, the longer the period considered. For instance, one year is largely more than enough to collect all the required marriage certificates (see below) from Lille, but 20 years may be necessary in a little village.

migration, like being in the same linguistic zone, in the same administrative canton or on the same river (rivers could be used for travel and some, like the Lys for line-weaving, also defined industrial territories). Else they could, as stated by Hägerstrand, be only linked by the personal relationships created by previous migrations (themselves possibly based on past common features). Network modeling can help to decide between these alternative explanations, as it allows to test the effect of attributes such as language as well as that of more structural, general phenomena (that might happen between non-adjacent places). Several of these structural effects (Snijders *et al.*, 2009a) have interesting interpretations in the case of migration.

A tendency to find *reciprocity* in “over-attractions”, and even more attractions that become reciprocal between periods (when more people than expected migrate from A to B, more people than expected tend also to migrate from B to A in the next period), point to a non-hierarchical view of migration, therefore less to economic causes¹⁰ than to personal contacts provided by earlier migrants and possibly leading, for example, to migration for marriage.

Simple chain migration – either individuals moving step by step to the big town or different sorts of workforce coming from the country to intermediary towns and from them to big cities (Hägerstrand, 1957) – would result in asymmetrical ties and low *transitivity*, if we define transitivity as follows: when more people than expected migrate from A to B and from B to C, more people than expected tend also to migrate directly from A to C in the next period. The role of intermediary towns in this pattern of chain migration could also, in network terms, lead to an attractive effect of *betweenness*: places acting as hubs for migrants would in this case become more and more attractive, as the outcome of a cumulative process.

We could however think of another sort of chain migration, with information moving backwards to places of emigration, that could eventually lead to *transitivity*. Short-distance moves in the older generations would thus enable the younger to migrate further.

Finally, combining the non-hierarchical and relational view of migration could even lead us to expect a *3-cycle* network effect, i.e. when more people than expected migrate from A to B and from B to C, more people than expected tend also to migrate directly from C to A in the next period.

Using network concepts to devise models proved useful here as it compelled us to finely specify our structural hypothesis, thus the concrete mechanisms that we imagined under the general idea of “the importance of personal relationships for migration fields”. Hägerstrand, who underlined the non-hierarchical, path-dependent patterns, would probably have hypothesized strong reciprocal and 3-cycle effects; his description of chain migration however points more in the direction of transitivity, or even betweenness.

4. Our method allows us to move from macro- to microscopic scales of analysis. At the individual level, we could expect some categories of migrants to depart from the “normal” behaviors as defined by the structures of the migration field. They would more easily act as pioneers (a typically Hägerstrandian notion) or otherwise make atypical migration choices. For example, migrants crossing the linguistic border, or more generally choosing atypical places, could be more educated than the others (hence more aware of opportunities unknown in their personal network and/or benefiting from a more extended network and/or being driven by another spatial distribution of opportunities). They could also be early followers of a pioneer (e.g. in their close kin). The literature helps us to identify attributes of long-distance migrants (“long-distance” meaning more than 20km) as compared to short-distance migrants: being male, literate, and born in a town rather than a village (Bonneuil *et al.*, 2008). But we lack specific hypothesis on “countercurrent” migrants, who in fact represent a different case.

¹⁰ At least to simple economic causes. Economic complementarity could be involved if each of the fluxes implied a different sort of workforce, e.g. blue-collars vs. white-collars.

2. Data

Marriage records and migration

Our aim was to present a replicable research strategy, and our results will be far more interesting if we are able to compare them with others from different regions and periods. We therefore use materials that should be reasonably easy to connect for many other places and periods. As 19th-century France, contrary to other countries such as Belgium or Sweden, did not record migration *per se* in “population books”, we use an admittedly rough proxy, only counting as “migration” the differences between place of birth and place of dwelling at the time of marriage, as recorded by (civil) marriage certificates (to be legal, a marriage had to be recorded at the mayor's office)¹¹. The median age at marriage in our sample was 29 for grooms (interquartile range: 25-34) and 26 for brides (22-31). This leaves aside several types of migration: migration after marriage, by non-married persons (possibly on average more mobile), and first migration(s) followed by another or by a return migration before marriage. On the other hand, this includes both young people who migrated alone and children accompanying their parents. As many migrations occurred during the first decades of life (this pattern was already identified by Ravenstein, 1885), it is however not an unreasonable proxy of migration, and it is certainly one of the most accessible ones. Our proxy thus includes people who moved to take one of their first jobs and/or to marry – the choice of the place, in both cases, being likely to be influenced by information obtained from social networks. Marriage records, in addition to the places of birth and dwelling of the spouses, also include their dates of birth, occupation, signature or lack thereof¹² and the same data on their parents (when they were alive) and on (generally four) witnesses. Up to now, we did not use information on parents and witnesses in this research, but they could provide useful complementary elements.

In each *commune*, ca. 70 marriage records were coded for each period. For example, in Séquedin (ca. 600 inhabitants), we had to include all records from 1877 to 1887 to gather 71 records; on the contrary, five months of 1880 were sufficient for Armentières. We thus used different sampling rates in each *commune*. The records therefore had to be weighed accordingly, weights being a function of the time-span and the exact number of records coded. For example, the trajectory of one groom in Séquedin was considered as representing 0.09 in terms of migration in the third period, while that of one groom in Armentières weighed 2.4. These are not “real” numbers of migrations, but they should be proportional to the number of moves made by new spouses marrying around 1879 from their place of birth to their place of dwelling¹³. Our database comprises 15,437 marriage records (ca. 70 records x 75 *communes* x 3 periods), each including information on two “migrations” (bride and groom), so that there are ca. 10,000 “migration trajectories” per period.

¹¹ We did not exclude marriages that were not the first one for one or both of the spouses – they anyway represent a small proportion of our sample.

¹² A rough indicator of literacy. 49% of grooms in our first period signed, 60% in the second and 81% in the third; the figures for brides were 39, 47 and 69%.

¹³ 1825, 1860 and 1880 were our original starting dates, but 1826, 1860 and 1879 are closer to the median dates of marriages actually coded by our research team.

Attributes of the communes

Testing our hypotheses required information on each of the *communes* in the sample – even more so because we claim that such attributes do not explain all the observed data, an hypothesis that requires serious testing. Some of these attributes could be derived from the marriage records themselves, while others were found in different external sources. We coded the following ones:

A. attributes that might have an “homophily” effect in network terms:

- main spoken language
- location on one of these rivers: Borre, Deule, Lys, Marcq;
- belonging to two sorts of administrative units: *canton* (a small unit) and *arrondissement* (a larger one);
- indicators of industrial activity in the 1860s (from Joanne, 1869). The source records plants, mills, and the main textile activities in a qualitative and probably not very consistent way. We created dummies for the presence of textile industry, chemistry/pottery and agro-industry and an index based on the addition of the dummies;
- occupational data, that deserve a longer comment (see Appendix). Suffice it to say here that we created three indexes for each period, based on the percentage of workers and cultivators among grooms and on a synthesis of literacy and the most common occupations of brides and grooms – and that describing socio-economic features of our *communes* appeared much more complicated than expected, far from our prenotions of a rural, agricultural Flemish-speaking west and a more urban, textile French-speaking east.

B. attributes that might have a more “hierarchical” effect on migration:

- indicators of industrial activity (see above);
- occupational data (see above);
- the fact of being the *canton* seat;
- population, from the official censuses (1806; average of 1826 and 1831; average of 1861 and 1866; average of 1876 and 1881);
- a dummy for Lille, Roubaix and Tourcoing;
- taxes paid per inhabitant at the beginning of the century (from *Statistique*, 1804);
- an index of public services at the beginning of the century (from *Statistique*, 1804; based on churches, markets, justices of the peace, etc.) and in the 1860s (from Joanne, 1869; based on markets, notaries, post offices, customs, tax offices, charity services);
- an index of social services in the 1860s (from Joanne, 1869; based on municipal charity offices, societies of mutual help and hospices);
- data on postal traffic (letters, newspapers, printed material, etc. sent and received per inhabitant) from *Relevé*, 1847.

It must be reminded that “hierarchical effects” here describes the fact that some *communes* might have had a wider migration field than others due to these attributes – not necessarily a higher number of in-migrants.

C. other attributes with different possible effects:

- geodesic coordinates of the city hall;
- adjacency between *communes*¹⁴;
- type of agglomeration (more or less grouped population) from *Relevé*, 1847.

¹⁴ Adjacency (a proxy for the absence of “intervening opportunities” might have a specific effect, different from that of distance.

3. General features of migration and over-attractions: a stable structure?

Types of migration, types of migrants

We concentrated our study on migration between the *communes* of our sample, in order to understand the structure and evolution of this migration field. We nevertheless have to mention the other types of migration that are present in the data.

In our weighted observations of places of birth and dwelling, we found 42% of sedentary spouses (40% of grooms, 44% of brides), 28% of spouses born outside of the sample (31% of grooms) and less than 0.5% of spouses living outside of the sample (although they married there). In the first period studied, sedentariness was a bit higher (46%) and in-migration a bit lower (21%) than afterwards (40% and 29%) – but these changes do not appear drastic for a region that was so fastly industrializing and urbanizing.

Our investigation thus concentrates on only 29% of the trajectories that we could reconstruct. It is of course important to keep these numbers in mind: our sample is by no means an actually closed network, and some of the *communes* that do not seem to have many ties within the sample probably have their own migration fields elsewhere. Defining boundaries for a network study is always difficult (Laumann *et al.*, 1992). However, we do not want to show here e.g. which *commune* is the most central in the group: our investigation is centred on dyadic and triadic patterns, that might be less sensitive to the boundary problem. In addition, it happens that few of the in-migrants actually came from French *communes* adjacent or very close to the sample, as a majority was born in Belgium.

Inside the sample, if we begin with a rough distinction between French-speaking, Flemish-speaking *communes* and mixed cases (as in Map 3), which seemed of central importance in our preliminary study (Lemercier & Rosental, 2000), and if we single out movements towards Lille, the weighed observations can be divided into 58% between French-speaking places other than Lille, 26% from any place towards Lille (29% for women), 8% involving mixed cases, 5% between Flemish-speaking places, 1.8% from a Flemish-speaking to a French-speaking place other than Lille, and 0.3% in the opposite direction. These figures are quite stable across periods, except for an increase of the first type (51% in the first period, 63% in the third) at the expense of intra-Flemish migrations and movements towards Lille. This might point to the growing attraction of Roubaix and Tourcoing, but is again not a very large change.

These results already point to the existence of a linguistic boundary. In addition, they give the impression of a higher mobility in the Eastern part of our region, that must however be qualified. As Flemish-speaking *communes* were on average much less populated, it is no wonder that few migrants, in absolute numbers, moved between them. Regressing migrations on populations and distances will thus help us to better understand attractions even between the smallest villages.

What we can do with the original figures is look for specific patterns of migration among some categories of individuals, as defined by gender, literacy and occupation. We isolated the 95 most frequent (in weighed data) combinations of period, gender, literacy and occupation (e.g. “male literate butchers, 3rd period”). We then looked for the over-representation of each type of migration among them¹⁵, and tried to cluster the 95 categories according to their migration patterns (% of each type of migration – general or intra-sample).

¹⁵ We wanted to use a logistic regression linking each type of migration to these attributes, but it appeared difficult to use a regression on a dataset where individual records are weighed. Suggestions in this respect are welcome.

Results on sedentarity and in-migration confirm expected patterns. 4 out of the 6 most sedentary categories (more than 75%) comprise female cultivators, and 7 out of the 14 most sedentary categories include the occupation “cultivator”. 6 of the 12 least sedentary profiles (less than 20%) are those of servants or female cooks; these occupations, along with housewives, are also among those with many spouses (more than 44%) born outside of the sample. Female cultivators were almost never born outside of the sample, but male cultivators were only a bit below the mean in this respect. Movements between French-speaking places clearly over-represented weavers and some other textile occupations, while it was overwhelmingly servants and especially farm servants (and illiterate spouses) who migrated between Flemish-speaking *communes*: this is both a reflection of different occupational structures (see Appendix) and of patterns of economic and life-cycle migration. People who migrated to Lille generally were not spinners or cultivators; employees (in retailing or railways) and seamstresses were over-represented among them, as were literate spouses. Finally, only a handful of migrants crossed the linguistic border, so that it is difficult to draw any conclusion about them. There were not only civil servants, as could have been expected, moving in both directions, but also, towards the Flemish-speaking zone, bakers, day-laborers, female servants and even male and female cultivators; patterns for the opposite migration are even less clear, with perhaps more servants and textile workers.

In a region with an admittedly complex socio-economic structure (see Appendix), it is therefore not easy to find individual attributes that clearly constrained migration, apart from the most expected ones (sedentary cultivators, servants coming from Belgium, literate spouses moving to Lille). Women and men, in particular, seem to have had very similar behaviors at this scale, and “countercurrent migrants” do not seem to have had much in common. This points to the need of more contextual information on migrants, especially pioneers, and to the fact that aggregate individual attributes do not explain much in terms of migratory patterns. Let us therefore turn to the study of the structures of our migratory field.

Distance, population and migration

As already mentioned, we used linear regression to determine coefficients in the classical “gravitational model” of migration, which gave a R^2 between 0.23 and 0.27 in each period – a reasonable, but not excellent adjustment. More interestingly, the coefficients A (multiplying population) and B (the power that distance is elevated to) did not vary much either, A fluctuating between 14.3 (first period) and 14.8 (third) and B between 1.06 (second period) and 1.17 (first). If there was a tendency to a growing effect of population, it was at best limited, and nothing much seems to have changed as for the effect of distance. The available means of transportation arguably did not evolve much inside our small region, but we could have thought that urbanization and industrialization would have made long-distance migration and migration to the largest centers more likely, thus changing these coefficients. It is not the case: at the most general level, we once again find stability.

Coefficient B can be compared to those found in different times and places when estimating the same equation (Hägerstrand, 1957, 115-7 [other references needed]). They generally range from 0.4 to 3, our case being in this respect closer to 20th-century urban Scandinavian areas than to 19th-century Sweden – indicating a certain degree of modernity, or, more probably, the lack of physical obstacles to migration inside the sample¹⁶.

¹⁶ Some if not most of Hägerstrand's data also include all migrations to/from a place, whereas ours are truncated for long distances, which might have an impact on the coefficient.

Attributes of the communes, homophily and hierarchy

As we had many data on attributes of the *communes* (and sometimes several indicators for the same characteristic) that could have had an impact on migration, but could not all be included in multivariate models, we first tried to measure their possible influence by dividing the matrix of over- and under-attractions according to various groupings, looking at the inter-group and intra-group densities and trying to find some possibly non-random patterns¹⁷. In addition to visual inspection of the matrices, this was done using a procedure of network autocorrelation available in Ucinet (Borgatti, Everett & Freeman, 2002)¹⁸ – and QAP-regression [ref] in the case of adjacency between *communes*. When the variables changed across periods (e.g. percentages of workers), we generally grouped network data of period N according to attribute data of period N-1 (although we sometimes tested data of period N), as the migrations that we study mostly occurred in the interval between periods and can be considered as more influenced by the past than the present conditions.

In some cases, such as attractions, we found no such pattern (no really important and consistent discrepancies), so that an influence of this variable, even in a multi-variate model, seemed quite unlikely. This was the case for the following:

- hypothesized hierarchical effects: being the seat of a *canton* and all indicators of public or social services, taxes and postal traffic.
- hypothesized other effects: almost all indicators of industries, percentage of workers, type of agglomeration.

Most indicators of the importance of *communes* that could play a different role than population alone therefore do not seem to have influenced over-attractions. The same is true for industrial activity, except for a possible attraction to *communes* with chemical or pottery industries in the second period. As for population and big cities (in addition to the effects already captured by the gravitational model), there was a specific over-attraction from the towns with more than 5,000 inhabitants to Lille in the first period, from all *communes* to Lille in the second one and from towns with more than 5,000 inhabitants to Lille, Roubaix and Tourcoing in the third.

Homophily effects seem to have been more present, as might have been expected for the residuals of a gravitational model. In our univariate treatment, the following groups appear to follow this tendency: linguistic groups; *communes* located on the Borre, on the Lys (all periods) and on the Deule (period 1); belonging to the same *arrondissement* or to the same *canton*¹⁹; *communes* with a high percentage of cultivators; and those belonging to the same socio-economic cluster (except for cluster 1 of period 2 for migration in period 3). In addition, the coefficient of the QAP-regression on adjacency is highly significant.

Of course, many of these indicators are in fact highly correlated (e.g. is it likely that two adjacent, Flemish-speaking *communes* had a similar percentage of cultivators and socio-economic profile, were situated in the same *arrondissement* and perhaps on the same river or in the same *canton*). Language and *arrondissement* are particularly correlated, but language homophily is quite stronger, which led us not to consider *arrondissements* in our other analyses. Specific *canton* and river effects (these attributes being quite distinct from each

¹⁷ This method required to create classes for continuous variable. We generally begun with more or less equal-frequency classes, but in ambiguous cases (e.g. percentage of in-migrants from outside of the sample), we sorted the matrix according to the continuous variable in order to find thresholds that would make the patterns clearer.

¹⁸ Network autocorrelation with categorical attributes, ANOVA density models, with the options “variable homophily” and/or “structural blockmodel” (that helps to capture either homophily or more or less complex hierarchical effects).

¹⁹ Two *cantons* (Vieux-Berquin and Hazebrouck) showed particularly high internal densities for each of the three periods; among the ten *cantons* including at least four *communes* of the sample that were tested for homophily, six had significant rates for all periods and the remaining three for all but the second period.

other) were kept to be tested, because they defined smaller areas. As the percentage of cultivators was important in the definition of socio-economic clusters, but the latter also included additional information, we chose to test the effect of clusters in the multivariate models.

This leaves us with quite a number of non-structural effects to test in our further analysis. It is already worth noticing that, apart from the exact pattern of attraction to the biggest cities, these effect do not seem to change much between periods (although some of the attributes themselves do): that these “non structural” effects (in terms of network analysis) are in fact quite stable.

Structural network tendencies

Our hypotheses on endogeneously evolving spatial patterns are better tested using dynamic actor-oriented models (Snijders *et al.*, 2009a), but preliminary more static tests, based on triadic census (as implemented in Pajek (Batagelj & Mrvar, 1998)) and ERGM models (as implemented in Siena (Snijders *et al.*, 2009b)) for each period are useful, as they give a first idea of the micro-structures of our network, at the dyadic and triadic scale.

As for triads, Table 1, from the Pajek manual, defines all the possible types and Table 2 compares the observed and expected frequencies of each one.

Table 1

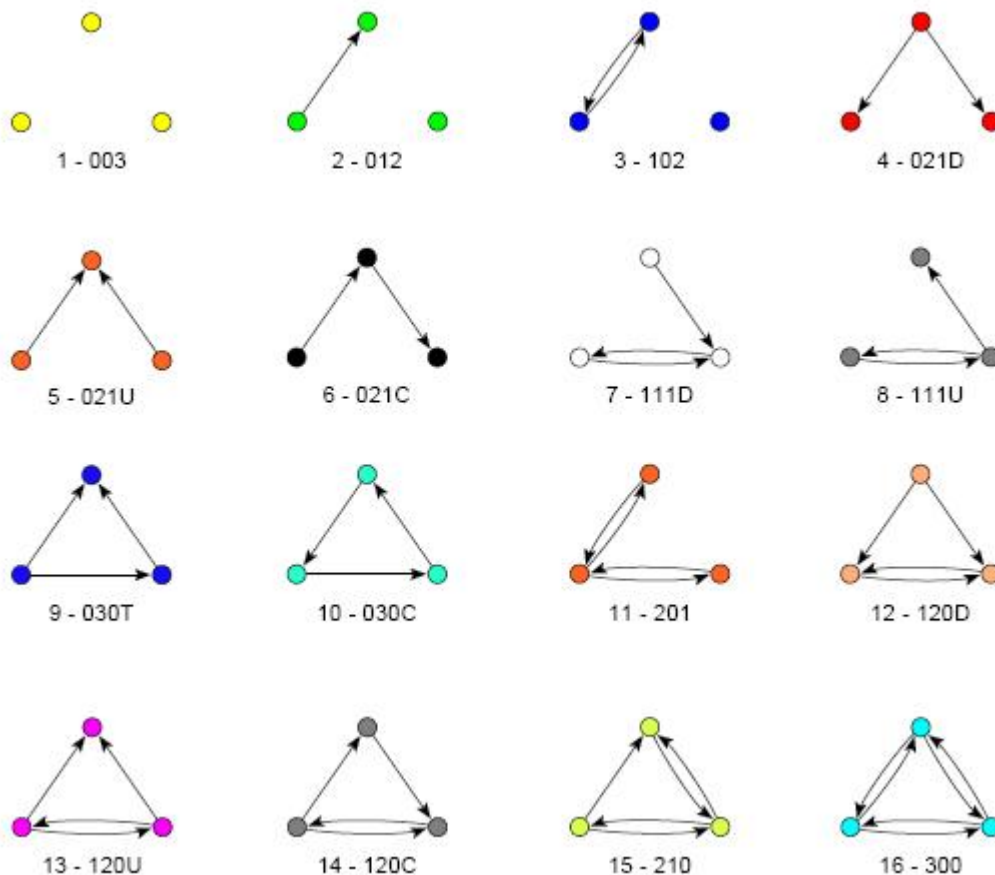


Table 2: Triad counts

| | Period 1 (observed – Type observed expected)/ expected | | Period 2 (observed – observed expected)/ expected | | Period 3 (observed – observed expected)/ expected | |
|----|---|--------|--|---------|--|--------|
| 1 | 37201 | 0.08 | 38629 | 0.11 | 39701 | 0.08 |
| 2 | 17275 | -0.19 | 17170 | -0.25 | 19278 | -0.18 |
| 3 | 4080 | 2.71 | 4933 | 2.97 | 4163 | 2.35 |
| 4 | 580 | -0.47 | 683 | -0.45 | 871 | -0.3 |
| 5 | 655 | -0.4 | 590 | -0.53 | 647 | -0.48 |
| 6 | 828 | -0.62 | 819 | -0.67 | 1080 | -0.57 |
| 7 | 394 | 0.74 | 371 | 0.37 | 356 | 0.35 |
| 8 | 481 | 1.12 | 676 | 1.5 | 626 | 1.37 |
| 9 | 235 | 0.04 | 229 | -0.15 | 271 | 0.03 |
| 10 | 35 | -0.54 | 29 | -0.68 | 33 | -0.62 |
| 11 | 85 | 6.28 | 136 | 8.21 | 87 | 5.22 |
| 12 | 21 | 3.37 | 61 | 3.13 | 62 | 3.43 |
| 13 | 76 | 5.51 | 123 | 7.33 | 105 | 6.51 |
| 14 | 107 | 3.58 | 126 | 3.27 | 93 | 2.33 |
| 15 | 89 | 35.98 | 181 | 55.24 | 122 | 40.15 |
| 16 | 24 | 579.62 | 68 | 1161.88 | 30 | 571.72 |

The relative weight of each sort of triad does not change much from period to period, with a few exceptions often involving period 2 and that would not be easy to interpret²⁰. The general structures of migration in our region once again appear stable.

Although the network of over-attractions is overall not very dense, as there was a large number of null migrations, the densest local patterns (types 11-16) are clearly over-represented. Along with type 3, these are the triads that involve forms of reciprocity, hence probably an important phenomenon in our data. On the contrary, clearly hierarchical types (4 and 5) are under-represented; but type 13 could also be viewed as hierarchical, and it is quite present. The most important tendency that can be noticed is therefore reciprocity, whereas the tendency to “egalitarian” patterns (that could be derived from a “micro-mobility” vision of rural migration) is not always present. Hence, probably, the under-representation of generalized exchange (3-cycles) without reciprocation (type 10), while the quite similar type 14 is more often found. “Pure” transitivity (type 9) is not more present than in random networks, but transitivity with partial reciprocation (type 13) is. Apart from this tendency towards reciprocity and even clique-formation (type 16)²¹, these static patterns do not help us to decide between our hypotheses, especially on chain migration. For example, types involving betweenness (4 to 8) are often under-represented, but the scarcity of betweenness position does not imply that *communes* in such positions will not become especially attractive in a later period.

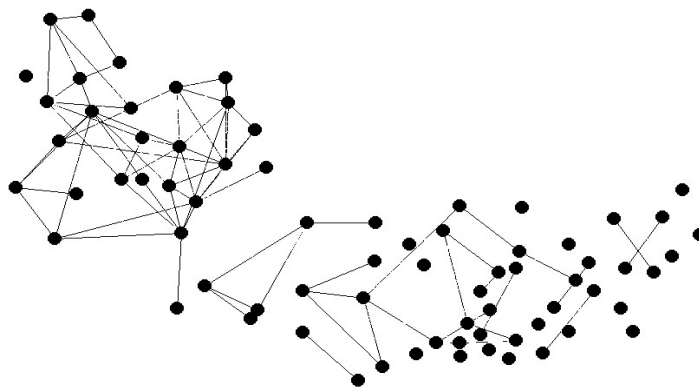
The specific, but here were important case of reciprocity can be better interpreted thanks to maps (Maps 4 to 6) for each period (that we were not able, as for now, to draw for specific

²⁰ There might be a time-tendency for type 7, but an important problem with triad census, that to our knowledge no software has solved, is that we do not know exactly which triplets of actors (here *communes*) are involved in triad of each type. It does not allow e.g. to follow what happened to former triads of type 7 after period 1, or to know if some types were typical for adjacent *communes* and others for long-distance ties. Siena models allow to partly answer such questions, but not with this degree of precision.

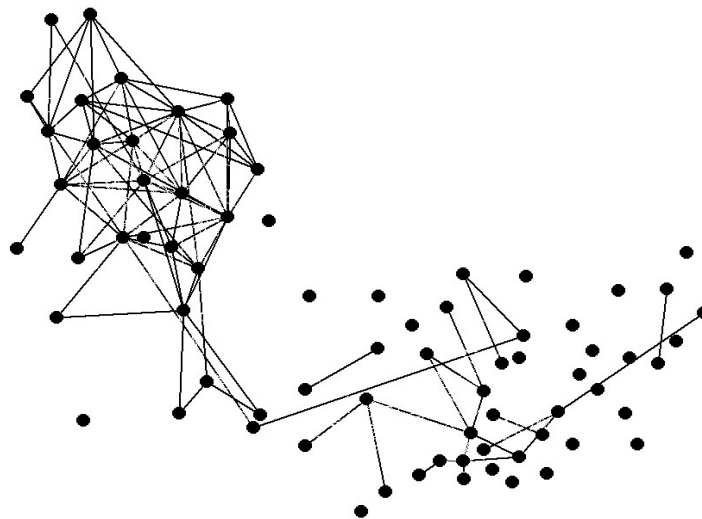
²¹ Even without symmetrizing the network, 15 highly overlapping 4- or 5-cliques are found in the second period, 6 in the third and one in the first (along with 21 3-cliques in this first period).

triad types)²². Out of a total of 500 to 530 over-attractions in each period, there were 80 to 100 reciprocal ties, thus representing one third of the over-attractions – slightly more in the second period than before or after. The maps clearly show that this tendency to reciprocation, as well as to clique formation, was more typical for the Western, Flemish-speaking part of the sample – while reciprocation without large cliques was found in the central part of the French-speaking region, not in the more urban and industrial East. In addition, it was generally short-distance preferential ties that were reciprocated: the Pearson correlation (controlled by QAP-correlation) between adjacency matrices and reciprocated over-attraction matrices is around 0.4 for each period. These general reciprocation pattern point to a quite simple model of rural, local, non-directed migration, perhaps complemented by more asymmetrical attractions to Lille, Roubaix or Tourcoing. It is however noticeable that it seems to work differently in each of the two linguistic subregions.

Map 4: Reciprocal over-attractions, first period

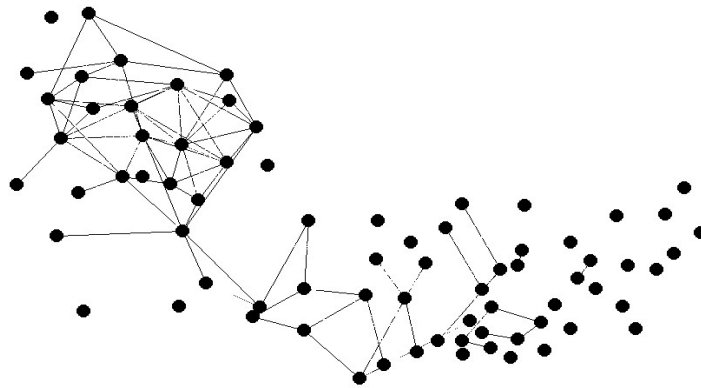


Map 5: Reciprocal over-attractions, second period



²² The circles representing *communes* are placed according to their geodesic coordinates.

Map 6: Reciprocal over-attractions, third period



As a less straightforward, but useful complement to triad counts, exponential random graph (ERG) models allow to check the relative weight of structural network tendencies (e.g. reciprocity or closure) and independent attribute effects (e.g. the attraction to Lille) for each period. In our case, both structural and independent effects seem to have played a significant role. The best models, with a very good convergence, are described in Table 3. For structural parameters, we have chosen to test some the generic “alternating” parameters (Snijders *et al.*, 2006, Robins *et al.*, 2007), as they seem to be statistically more robust, although their interpretation is still somewhat experimental, due to the lack of empirical studies having used them.

Table 3: ERG models

| Parameter | Period 1 | | Period 2 † | | Period 3 | |
|---|----------|-----------|------------|-----------|----------|-----------|
| | Estimate | Std error | Estimate | Std error | Estimate | Std error |
| Reciprocity | 0.44 | 0.22 | 0.54 | 0.19 | 0.32 | 0.21 |
| Alternating out-k-stars | 0.29 | 0.20 | 0.42 | 0.17 | 0.66 | 0.17 |
| Alternating in-k-stars | -0.04 | 0.24 | -0.21 | 0.23 | -0.36 | 0.32 |
| Alternating k-triangles | 0.75 | 0.07 | 0.82 | 0.06 | 0.71 | 0.07 |
| Alternating independent twopaths | -0.18 | 0.01 | -0.16 | 0.01 | -0.14 | 0.01 |
| Same socio-economic cluster* | 0.24 | 0.07 | 0.06 | 0.06 | 0.18 | 0.08 |
| Same <i>canton</i> | 0.26 | 0.10 | 0.13 | 0.08 | | |
| Same language | 0.26 | 0.09 | 0.17 | 0.09 | 0.24 | 0.09 |
| Same position on the river Deule (or not) | 0.36 | 0.12 | | | | |
| Adjacent <i>communes</i> | 1,66 | 0.13 | 1,52 | 0.12 | 1,65 | 0.12 |
| Towns to large cities** | 1,79 | 0.79 | 1,5 | 0.44 | 1,71 | 0.50 |
| Flemish to large cities*** | | | 0.45 | 0.15 | 0.48 | 0.19 |

* clusters as defined for period 1 is used for periods 1 and 2, clusters as defined for period 2 are used for period 3.

** period 1: attraction from towns with more than 5,000 inhabitants (in period 1 as well as in 1806) to Lille. Periods 2 and 3: attraction from the same towns to Lille, Roubaix and Tourcoing²³.

*** attraction from Flemish-speaking *communes* to Lille, Roubaix and Tourcoing. This effect was included after a closer examination of the networks allowed by blockmodeling (see below).

²³ Two *communes* in the vicinity of Lille passed the 5,000 inhabitants bar between periods 1 and 2, but were not especially attracted to Lille, Roubaix and Tourcoing: hence this specification that concentrates on older medium-size towns.

† Results for Period 2 should be taken with caution, as convergence was only reached with difficulties and thanks to the inclusion of non-significant parameters. Cluster, language and *canton* were significant in some slightly different, but otherwise less satisfactory specifications.

Some of the candidate effects (migrations following valleys, attraction to *communes* with chemical industry and general attraction to Lille in the second period) were not significant in any multivariate model. Several effects of characteristics of the *communes* however were, even when considered along with more structural tendencies.

As for the latter, and suprisingly if we consider the triad counts, reciprocity was barely significant (and not at all in the last period). It may have been absorbed by stronger attractions between adjacent *communes* or by some of the homophily effects²⁴. However, the strong positive estimate for alternating k-triangles, along with the significant negative effect for alternating independent twopaths, are more consonant with the view given by triad counts and confirm. The purpose of the alternating k-triangles effect is to model transitivity and clustering; along with a negative alternating independent twopaths effect, it is supposed to describe a segmented network of multiple (but small) dense regions connected by low density paths. This confirms the importance of clique-like phenomena in our sample that could have structural causes – i.e. be related to the past history of migration between *communes*, as hypothesized in part 1. – and not only be an indirect consequence of homophily.

The estimate for alternating out-k-stars, that becomes higher in the last periods, point to an important and growing variance in the size of migration fields for out-migration (whereas it is not true, perhaps suprisingly, for in-migration). It might be connected to growing differences between the Flemish-speaking and French-speaking parts of the sample, *communes* in the former exhibiting more diverse local attractions and increasing propensities to migrate also to the French-speaking *communes* (see part 5. below).

Adjacency lies somewhere between “structural” and “attribute” effects; as we deal here with residuals of a model already including distance effects, the high level of this parameter points to the importance of the most local connexions *per se*, even when the adjacent *communes* are otherwise not similar (as it does not disappear when included along with several homophily effects). We thus are able to distinguish between several variants of a general intuition on “Brownian motion between similar, close places”. Over-attraction occurred more often between adjacent *communes*, between *communes* that shared the same language, socio-economic features (in the first and third period), or, in the first period, that belonged to the same *canton* or were located in one of the valleys. In addition, *communes* that shared common over-attractions probably tended to exchange more migrants than expected (in the transitive/clique-like effect described by the structural parameters). Among these different variants, homophily might have had weaker effects, but was still present, especially in the first period and with respect to language. There were thus several different ways to move to a “close” place, and all of them seem to have influenced individual strategies.

This general picture of migration between more or less equal places, creating a landscape of cliques, must however be qualified. The two last parameters in the models remind us that Lille, then Lille, Roubaix and Tourcoing were even more attractive than their population alone would have explained. They however did not attract migrants from all *communes* indifferently, but preferably from the middle-size towns and from the Flemish-speaking region (although they were located on the other side of the sample). These two different effects might be related to two different sorts of migrants, and/or middle-size towns could have played the role of stopover for some of the migrants (as our blockmodeling below seems to indicate). They anyway point to some very specific forms of asymmetry in the network of

²⁴ We tested the interaction between reciprocity and adjacency, but the estimate was negative for this parameter.

over-attractions: they are the only hint of modernization as generally envisioned in the ERG models, and as such they show a subtle image of “rural exodus”.

4. Dynamic modeling

The results that we have given often tended to exhibit a remarkable stability between periods, be they structural or related to characteristics of the *communes*, the migrants or types or migration. A lower homophily and growing attraction of Flemish-speaking migrants to Lille, Roubaix and Tourcoing after the first period almost appeared as the only exceptions.

However, if we look closely at the egocentric network of over-attraction (or migration field) for each *commune*, a lot of changes appear. Whereas there were 491 to 532 over-attractions in the network in each period, with a quasi-stable density, 305 new over-attractions in fact appeared between the first and second periods, whereas 265 disappeared and 226 only remained stable. The figures for changes and stability between the second and third period were respectively 290, 289 and 242²⁵.

This implies that the sort of stable migration fields at a micro scale that Hägerstrand and his followers found in Scandinavia were not really common in our region – for reasons that remain to be investigated. The fact that so frequent micro-changes gave birth to a more or less stable regional structure is nevertheless quite interesting, and in line with other findings by Hägerstrand, who actually was interested in innovation-spreading, not in immovable structures. The result is similar to what Lazega *et al.*, 2006, termed a “spinning-top model”: multiple micro moves or changes at the individual level are what allows the general structure to be conserved. Snijders *et al.*, 2009a, similarly noticed that:

“If one has observed a longitudinal network data set of which the consecutive cross-sections have similar descriptive properties – no discernible trends or important fluctuations in average degree, in proportion of reciprocated ties, in proportion of transitive closure among all two-paths, etc. –, then it would be a mistake to infer that the development is not subject to structural network tendencies just because the descriptive network indices are stationary. For example, if the network shows a persisting high extent of transitive closure, in a process which is dynamic in the sense that quite some ties are dissolved while other new ties appear, then it must be concluded that the dynamics of the network contains an aspect which sustains the observed extent of transitive closure against the random influences which, without this aspect, would make the transitive closure tend to attenuate and eventually to disappear. [...] Given a sequence of consecutively observed networks, if one were to make an analysis of the first one by an ERGM and of the further development by an actor-based model, then in theory it is possible to obtain opposite results for these two analyses, and this would point toward a non-equilibrium situation.”

It is exactly what we observe here; this notion of non-equilibrium with seemingly stable indicators might probably yield interesting results in the study of migration. From a more empirical point of view, it led us to use dynamic actor-oriented modeling as implemented in Siena (2009b), that seemed to make sense in the case of a such an amount of changes and, more importantly, offered parameters that fit exactly with our questions on patterns of future migration “contained” in those of past migration – to put it differently, we already had a Markovian model in mind. As their name indicates, dynamic actor-oriented models have however been developed to deal with individuals making (constrained) choices on their ties

²⁵ This is just the acceptable amount of change for Siena modeling, as measured by the “Jaccard index” (2.9 and 3.0), in that our periods are not too further apart from each other to make dynamic modeling artificial, according to Snijders *et al.*, 2009a.

with other individuals, while taking into account information on the local structure of the network around them and on attributes of their potential partners. Does it make sense to use such a model to describe “over-attractions” between places? We considered that it was not absurd, not only for practical reasons (the statistical operations practically performed by the software described what we wanted them to describe), but also because our vision of migration is not at odds with this model. Of course “places” did not decide on their ties, nor did the inhabitants gather to deliberate on migration. But our hypotheses are based on the idea that the individual migrants collectively created, then followed preferential channels due to their information both on places and on previous migrations. Although nobody really computed an “objective function”, thinking of one might help to understand what happened; it can be understood in terms of “rules for network behavior” (Snijders *et al.*, 2009a) – here, implicit social rules of migration.

We tested parameters related to our structural hypotheses as described in part 1.²⁶, along with others related to the attributes of the *communes*. The latter were chosen after an inspection of the previous results, where we looked for hints of changes. All the tested parameters actually proved significant (with rapidly converging models), except for the disappearance of homophily along the Deule river and, more interestingly, the 3-cycles parameter. Results are shown in Table 4.

This particular version of the historicity of migration (if more people than expected move from A to B and from B to C, then others will eventually move from C to A) was neither especially present nor clearly absent from the data (it appeared more or less as often as in random networks).

Table 4: Dynamic actor-oriented models

| Parameter | Period 1 to Period 2 | | Period 2 to Period 3 | |
|--------------------------------------|----------------------|-----------|----------------------|-----------|
| | Estimate | Std error | Estimate | Std error |
| Rate parameter | 17,7 | 1,33 | 16,22 | 1,17 |
| Outdegree (density) | -1.95 | 0.15 | -1.43 | 0.13 |
| Reciprocity | 0.26 | 0.15 | 0.25 | 0.13 |
| Transitive triplets | 0.20 | 0.02 | 0.12 | 0.01 |
| Betweenness | -0.06 | 0.02 | -0.11 | 0.02 |
| Towns to large cities* | 1,23 | 0.48 | 1,48 | 0.52 |
| Flemish to large cities* | 0.97 | 0.18 | 0.89 | 0.24 |
| Same socio-economic cluster* | 0.32 | 0.08 | 0.21 | 0.08 |
| Same <i>canton</i> | 0.80 | 0.17 | 0.40 | 0.13 |
| Same language | 0.39 | 0.10 | 0.42 | 0.11 |
| Migration from cluster 3 of period 1 | 0.19 | 0.08 | | |
| Migration to cluster 2 of period 1 | 0.30 | 0.10 | | |

* See Table 2. Towns to large cities, periods 1 to 2: to Lille. Periods 2 to 3: to Lille, Roubaix and Tourcoing. Periods 1 to 2: cluster defined in period 1. Periods 2 to 3: cluster defined in period 2.

The three other structural effects that we considered exhibited clearer features. We had already noticed that reciprocity was not significant in our ERG model for the third period; reciprocation was in turn barely significant in dynamic models. Although quite many reciprocated over-attractions were found in the data, at least in the Flemish-speaking part of the region, they seem to have been driven by other forces than reciprocation *per se*. On the contrary, transitive triplets had significant estimates, positive for the former and negative for the latter. We thus do not find chain migration in the sense that stopover towns, playing the

²⁶ Structural “popularity” and “activity” parameters could have been included in alternative specifications, but we chose not to test them as we had no substantive hypotheses on migration that would have explained, e.g., why *communes* with an already diverse in-migration field tended to attract even from new *communes*.

role of hubs for migrants (thus characterized by a high betweenness) would become more and more attractive. The sort of chain migration implied by transitivity phenomena (migrations from A to B and from B to C leading to migrations from A to C) seem to point to a specific social phenomenon: information moving backwards to places of emigration, thus allowing more direct trajectories to some opportunities for the youngest generations.

Whereas the views of our network given in part 3. seemed consonant with some of Hägerstrand's depictions of non-hierarchical, path-dependent, egalitarian and/or clique-like patterns, the dynamic modeling, that seems even more suitable for testing hypotheses on “Markovian migration”, gives a different, more oriented view of chain migration. Other case studies would be needed to assess if this result is specific for ours or describes a more general feature of migration phenomena. It anyway shows the importance of structural network evolution, along with the weight of some attributes of the *communes*.

As in the ERG models, and even more clearly, we both found:

- tendencies to homophily according to three different criteria (that were not as correlated as it could be thought) referring to cultural (language), socio-economic and political/administrative homogeneity. Each of these similarities thus regularly opened new roads for migration (although some older roads were closed at the same time), perhaps for different sorts of migrants – an hypothesis that might be investigated by returning to our individual-level data. Consistently with ERGM results, *canton* homophily seems to have been a decreasing driving force, but it did not disappear; contrary to simple versions of modernization, the effect of language homophily did not fade when literacy increased;
- and specific increasing attractions to the largest cities from middle-size towns and Flemish-speaking *communes*.

In addition, a close look at the density patterns investigated in part 2. has led us to test two candidate effects that proved significant only in the transition between the two first periods, when industrialization began. *Communes* that were part of cluster 3 in the first period were mainly situated in the center of our region; they had many linen spinners and cultivators. Between the first two periods, their migration fields tended to become more diverse as regards out-migration, with many different new destinations. This might be linked to the crisis in linen-spinning, although this hypothesis deserves further, micro investigation. On the contrary, *communes* that were part of cluster 3 in the first period were generally urban and had few day-laborers and cultivators. Their in-migration fields became more diverse, as they received more migrants than expected from several different *communes* of the sample (an, in fact, outside of the sample). These tendencies seem to point to the effect of economic changes, leading to the diversification of some migration fields as older channels possibly did not prove sufficient to meet new needs. However, this is not a straightforward effect of the size of towns or the presence of industries (that were directly tested above): we have been able to describe it more precisely thanks to our clustering and modeling. This initial effect anyway seems to have faded when industrialization continued.

5. Networks, blocks and maps

In a preliminary stage of this research (Lemerancier and Rosental, 2000), we had used blockmodeling for the second period in order to describe the general patterns of attraction in our region – mainly showing the significance of the linguistic boundary and the differences between local structures in the Flemish-speaking and French-speaking zones, that we already mentioned here when discussing reciprocal attractions. Our attempt was limited by the algorithms available, both for blockmodeling *per se* and for the modeling of structural local phenomena – let alone dynamic modeling.

Although the models showed in part 4. answer to most of our initial questions about the historicity of local migration patterns, they do not give a precise sense of what the migration field in our region looked like, especially in spatial terms. For example, we know that there was a tendency to transitivity, but we do not know where it was found, or if it was “local” in terms of distance, not only of networks. This precise question seems to have no proper answer, as it would be tedious to manually identify triads that became transitive, and we do not know of software allowing to automatize this task. However, blockmodeling helped us to return to maps and to get a better sense of what changed at the scale of the region and its main parts, not only in a series of network triads. In fact, it even led us to discover a parameter that proved significant (the attraction from Flemish-speaking *communes* to Lille, Roubaix and Tourcoing) and to better understand another (the role of middle-size towns). We thus consider that this research strategy gives an important complementary view of the network structure and should not be abandoned²⁷, at least for migration – or, more generally, exchange – studies.

We used the software Blocks (Snijders & Nowicki, 2007), that has three distinct advantages: it provides indicators that help to choose an appropriate number of blocks; it points to cases that are difficult to cluster (we have chosen to exclude them from image matrices and to show them in white on the graphs); and, quite importantly in our case, it specifically takes into account the orientation of ties – and it provide separate image matrices for reciprocated and non-reciprocated ties. “Maps” 7 to 9 show our results: in each case, the first graph uses the geodesic coordinates of *communes*, thus giving a “map” of the network, while the second is a more classical spring-embedding graph allowing to better understand the positions in the network, e.g. for the red blocks in the second and third period. Like the dynamic modeling, these maps show a changing image of our network, although there are also obvious elements of stability.

In the first period, three groups were clearly distinct, geographically as well as in network terms. These sub-regions were not much connected and the presence of large cities did not distort their features (apart from inducing a few migrations from the red to the black block). It is interesting to notice that, whereas there was a clear boundary in attractions, it did not exactly follow the linguistic boundary as depicted in map 3. The practice of Flemish might have been more extended at that time (as our map is based on later surveys), or *communes* situated at the geographic center of the sample might have been less prone to linguistic homophily at that time. In addition – and this is true for all the three periods –, the blocks defined by migration patterns are not significantly correlated with our socio-economic clusters: although there was a statistical tendency to homophily inside these clusters, it was not strong enough to appear here – adjacency, *canton* and language homophily were more likely combined to create these blocks.

The structure was both less clearly defined (according to Blocks) and more complex in the second and third periods, when large cities seemed to distort the previous patterns. This is consonant with our observation that dynamic modeling did not as strongly support egalitarian, clique-like tendencies as transitivity.

In the second and third period, there was still a relatively cohesive and closed Flemish-speaking group, but it was smaller, now restricted to the region depicted in map 2 (which might explain the growing language homophily found in dynamic modeling); inside this group, attractions were more often reciprocated than in the first period. Lille, along with four middle-size towns, was part of an intermediary (red) block receiving migrants from the two linguistic zones. In the second period, Roubaix and Tourcoing were still not very much connected to the rest of the sample (migrants there came from other zones, especially from Belgium), while they tended to play a role more similar to that of Lille in the last period. Finally, a relatively cohesive (green) block of French-speaking *communes* appeared and was

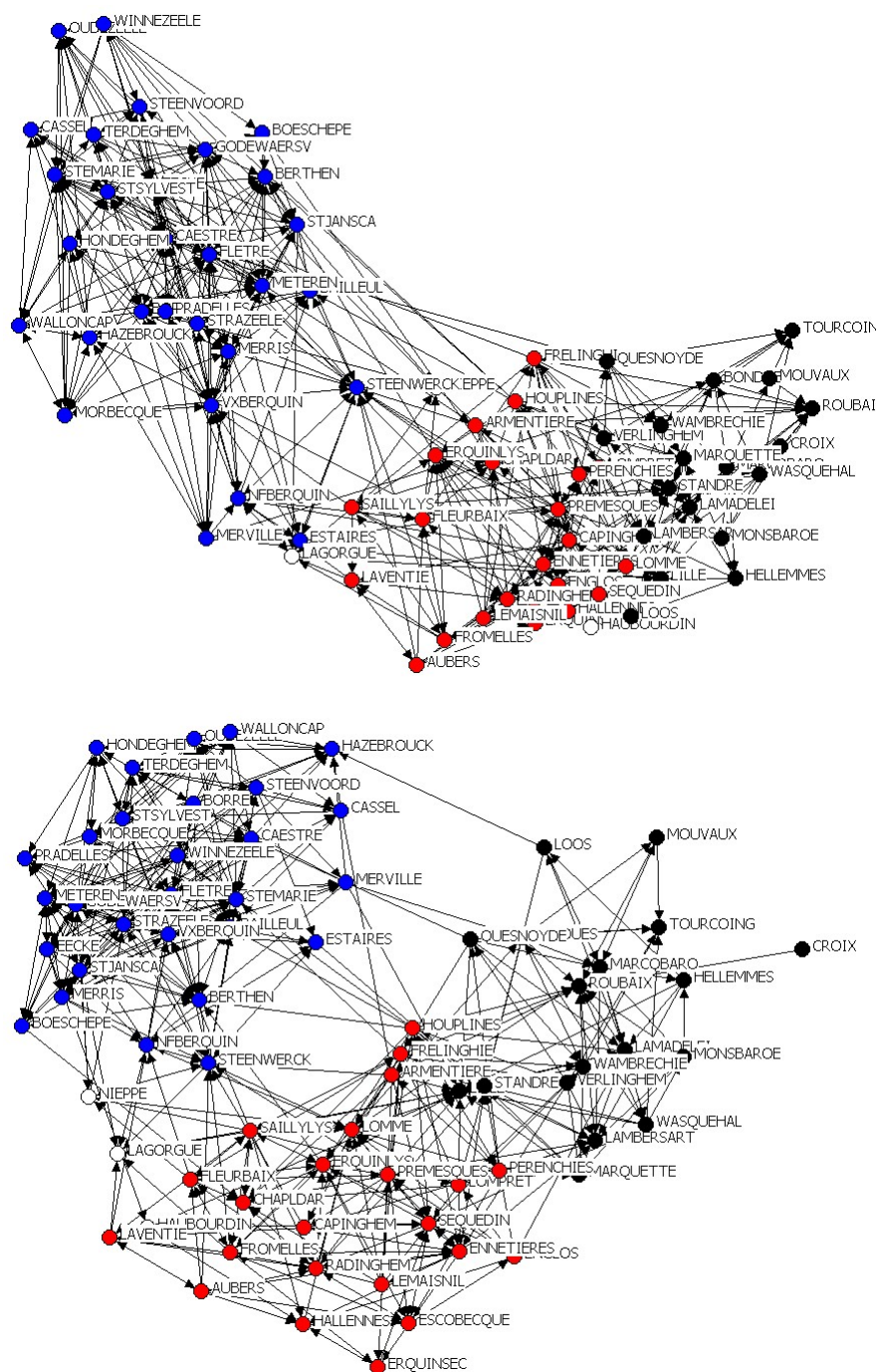
²⁷ On this question, see also Lazega, 2007.

again found in the third period. As far as we know, these *communes* did not have much in common in terms of size, activities or even *cantons* or rivers. Their case remains to be investigated.

Map 7: Results of blockmodeling, first period

| Observed mutual over-attraction frequencies | | | | Observed non-mutual overattraction freq. | | | |
|---|----|-----|----|--|-----|-----|-----|
| Birth \ Marr. | | | | Birth \ Marr. | | | |
| | | | | | | | |
| | 7% | 0% | 0% | | 17% | 2% | 6% |
| | 0% | 13% | 0% | | 0% | 16% | 1% |
| | 0% | 0% | 4% | | 2% | 0% | 16% |

Communes that could not be robustly clustered appear in white below



Map 8: Results of blockmodeling, second period

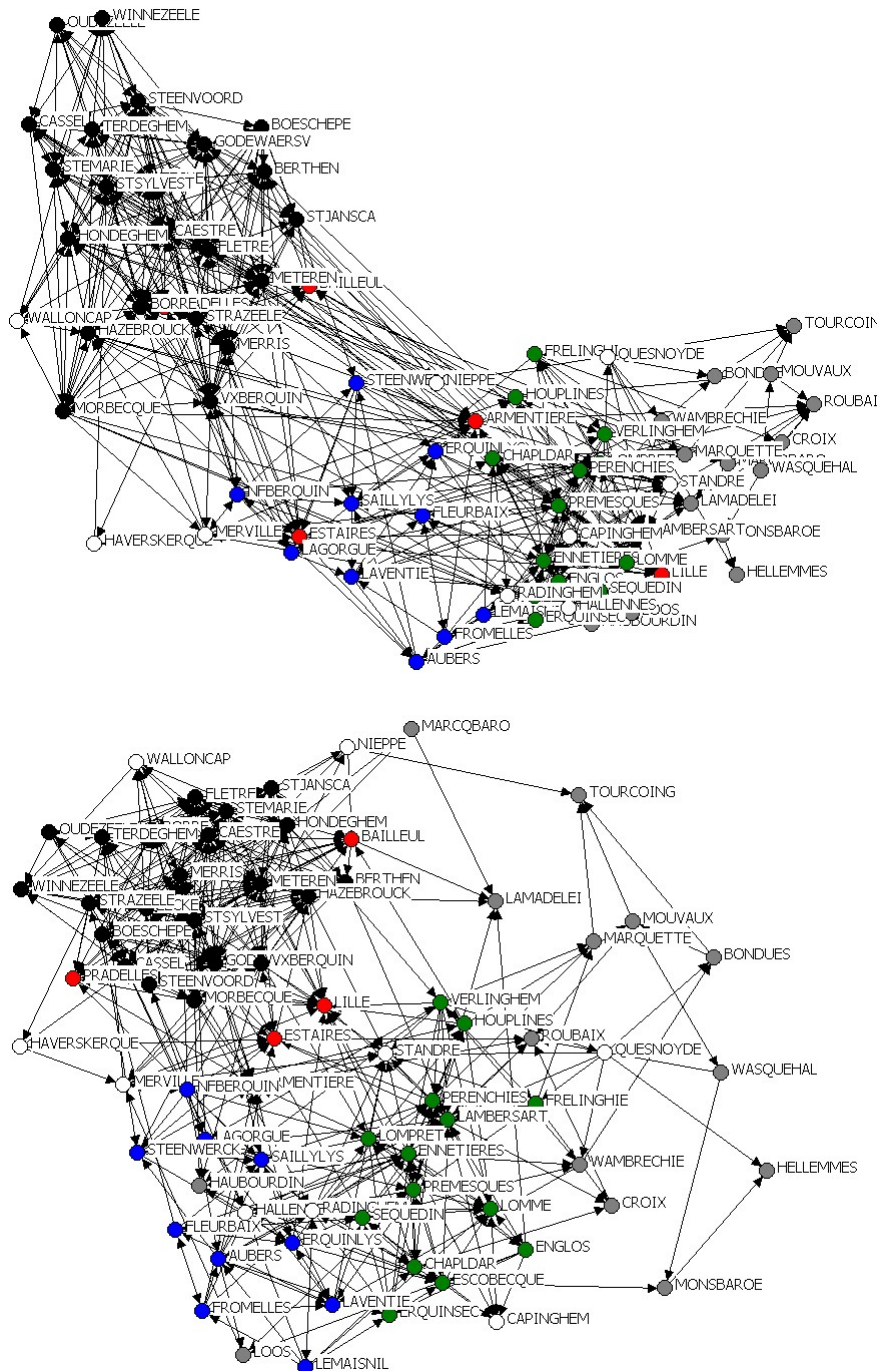
Observed mutual over-attraction frequencies

| Birth \ Marr. | | | | | |
|---------------|-----|------|-------|------|-------|
| | Red | Blue | Black | Grey | Green |
| Red | 0% | 2% | 1% | 0% | 0% |
| Blue | 2% | 7% | 1% | 0% | 1% |
| Black | 1% | 1% | 30% | 0% | 0% |
| Grey | 0% | 0% | 0% | 1% | 1% |
| Green | 0% | 1% | 0% | 1% | 13% |

Observed non-mutual overattraction freq.

| Birth \ Marr. | | | | | |
|---------------|-----|------|-------|------|-------|
| | Red | Blue | Black | Grey | Green |
| Red | 5% | 0% | 4% | 4% | 1% |
| Blue | 16% | 22% | 2% | 1% | 11% |
| Black | 31% | 1% | 17% | 2% | 1% |
| Grey | 0% | 1% | 0% | 6% | 0% |
| Green | 13% | 3% | 0% | 10% | 20% |

Communes that could not be robustly clustered appear in white below



Map 9: Results of blockmodeling, third period

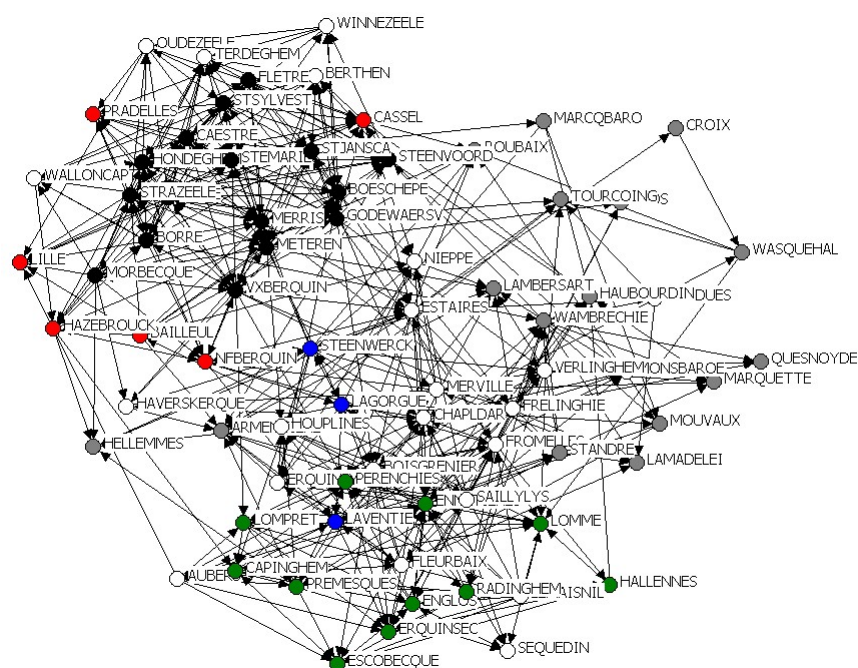
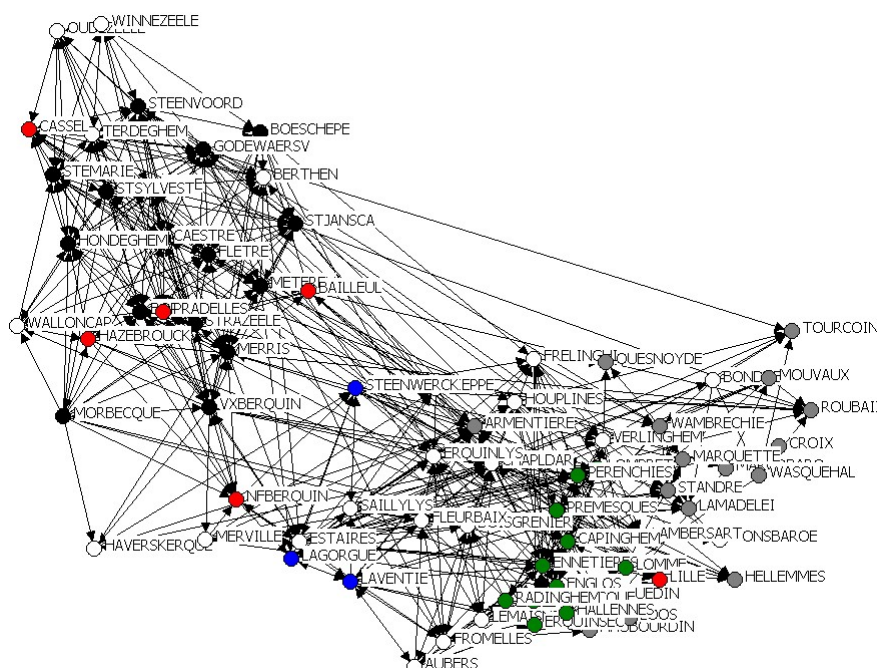
Observed mutual over-attraction frequencies

| Birth \ Marr. | | | | | |
|---------------|----|-----|-----|----|-----|
| | | | | | |
| | 3% | 0% | 0% | 0% | 0% |
| | 0% | 33% | 0% | 0% | 0% |
| | 3% | 0% | 32% | 0% | 0% |
| | 0% | 0% | 0% | 1% | 0% |
| | 0% | 0% | 0% | 0% | 15% |

Observed non-mutual overattraction freq.

| Birth \ Marr. | | | | | |
|---------------|-----|-----|-----|-----|-----|
| | | | | | |
| | 13% | 17% | 3% | 6% | 2% |
| | 0% | 17% | 4% | 19% | 12% |
| | 29% | 2% | 20% | 4% | 1% |
| | 0% | 0% | 0% | 6% | 1% |
| | 2% | 0% | 0% | 6% | 24% |

Communes that could not be robustly clustered appear in white below



6. Discussion and future research

We tried to use network methods to produce different views of a regional migration field: static and dynamic, local and global, investigating individual moves and relationships between places, structural effects and modernization theories. This cubistic strategy might sometimes make the reader dizzy. We however did not simply try all the possible features of network software: each of our analyzes was driven by a specific research question on a highly multifactorial phenomenon. We hope that the set of tools that we discussed will be used in other case studies – as our strategy to build data on arguably very simple migration trajectories is not very time-consuming. Comparing results would of course help to better understand what can be generalized in our findings and what is specific for a linguistically segmented and fastly modernizing region.

Important general results include the fact that a high amount of local changes, partly visible on maps, could lead to overall stable aggregate results. This helps to understand that social features of migration that might be described as form of path dependency, thus connected to an idea of historicity in the sense of determination by the past, or slow change, could be active in a rapidly changing economic context and not prevent important modifications in global fluxes.

As mentioned at several points in the text, some of our results gave birth to new hypotheses on local history and/or the interaction between types of migrants and types of migrations that could be tested in the future by returning to our individual micro data. The same data would also allow to test the consistency of several different patterns of relationships between places, e.g. our “migration” pattern and the relationship between the places of birth of grooms and brides (an “alliance” pattern) or between places of dwelling of the spouses and witnesses (capturing arguably heterogeneous, but also more diverse, social relations between places).

References

- Sune Akerman, Bo Kronborg & Thomas Nilsson, "Emigration, Family and Kinship", *American Studies in Scandinavia*, vol. 9, 1977, pp. 105-122.
- Anselin, Luc, (1995), "Local indicators of spatial association – LISA", *Geographical Analysis*, Vol. 27, n°2, pp. 93-115.
- V. Batagelj, A. Mrvar: "Pajek – Program for Large Network Analysis". *Connections*, 21(1998)2, 47-57.
- Bonneuil, N., Bringé, A. and Rosental, P.-A. (2008), "Familial Components of First Migrations after marriage in nineteenth century France", *Social History*, 33, 1, pp. 36-59.
- Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002. *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies
- Cattan N., Grasland C., Rehak S., 1996, "Migratory flows between Czech and Slovak republics in 1992: Which forms of transition? ", in : Carter F., Jordan P., Rey V. (eds), *Central Europe after the Fall of the Iron Curtain : Geopolitical Perspectives, Spatial Patterns and Trends*, Wien, Peter Lang, Wiener Osteuropa Studien :4, 319-336.
- E. Coornaert, *La Flandre française de langue flamande*, Paris, Editions ouvrières, 1970.
- Courgeau D., (2007) *Multilevel synthesis*, The Springer Series on Demographic Methods and Population Analysis, Springer, approx. 225p.
- Cécile Dauphin, Pierrette Lebrun-Pezzerat et Danièle Pouban avec la collaboration de Michel Demonet, « L'enquête postale de 1847 », in *La correspondance*, ed. Roger Chartier, Paris, Fayard, 1991.
- Martin Dribe, Christer Lundh, "People on the Move. Determinants of Servant Migration in Nineteenth Century Sweden", *Continuity and Change*, Vol. 20, No. 1, pp 53-91. 2005.
- Georg Fertig, "Balancing, Networking, and the Causes of Emigration: Early German Transatlantic Migration in a Local Perspective, 1700-1754", *Continuity and Change*, 13:3 (1998), 419-442.
- Hägerstrand T., "Migration and area. Survey of a sample of Swedish migration fields and hypothetical considerations on their genesis", in D. Hannerberg, T. Hägerstrand et B. Odeving éd., *Migration in Sweden. A symposium*, Lund, Lund studies in geography, series B, n°13, 1957, pp. 27-158.
- Adolphe Joanne, *Dictionnaire géographique de la France*, Paris, Hachette, 1869.
- G. Kurth, *La frontière linguistique en Belgique et dans le nord de la France*, Bruxelles, Société belge de librairie, 1896, t. 1.
- Laumann, E. O., Marsden, P. V. and Prensky, D. (1992) "The Boundary Specification Problem in Network Analysis", L. C. Freeman, D. R. White and A. K. Romney (eds.), *Research Methods in Social Network Analysis*, Fairfax, pp. 61-87.
- Emmanuel Lazega, *Réseaux sociaux et structures relationnelles*, Paris, PUF, 2nd ed. 2007.
- Emmanuel Lazega, Claire Lemerrier & Lise Mounier, "A Spinning top model of formal organization and informal behavior: Dynamics of advice networks among judges in a commercial court", *European Management Review*, 3 (2006), 113-122.
- Claire Lemerrier & Paul-André Rosental, « 'Pays' ruraux et découpage de l'espace. Les réseaux migratoires dans la région lilloise au milieu du XIX^e siècle », *Population*, 55, 4-5, 2000, p. 691-726.
- Palloni, A., Massey, D. S., Ceballos, M., Espinosa, K. and Spittel, M. (2001), "Social Capital and International Migration: A Test Using Information on Family Networks", *American Journal of Sociology*, 106, 5, pp. 1262-1298.
- Chantal Pétilon, *La population de Roubaix. Industrialisation, démographie et société 1750-1880*, Villeneuve d'Ascq, Presses universitaires du Septentrion, 2006.

Poussou J.-P., " Les mouvements migratoires en France et à partir de la France de la fin du xve siècle au début du xixe siècle : approches pour une synthèse ", *Annales de Démographie historique*, 1970, p. 11-78.

Relevé par département du nombre de communes et autres localités ayant une appellation propre en France, d'après le résultat de l'enquête générale faite au mois de novembre 1847 par les soins de l'administration des Postes, Bibliothèque nationale, rue Richelieu, Ms. Fr. 9787-10129. The source and the variables that it records are described in Dauphin *et al.*, 1991 and <http://crh.ehess.fr/docannexe.php?id=1155>

Robins, G., Snijders, T.A.B., Wang, P., Handcock, M., and Pattison, P. 2007. Recent developments in Exponential Random Graph (p*) Models for Social Networks. *Social Networks*. 29 (2007), 192-215.

Paul-André Rosental, « Scomposizione spaziale di una migrazione internazionale : l'integrazione dei belgi nel nord della Francia nel XIX secolo », *Memoria e Ricerca*, 8, 1996, p. 33-56.

Paul-André Rosental, « Les formalisations spatiales de la mobilité : fragments pour l'histoire longue d'une non-réception », *Genèses*, 4, 1997, p. 75-98.

Paul-André Rosental, *Les sentiers invisibles. Espace, familles et migrations dans la France du XIX^e siècle*, Paris, éditions de l'EHESS, 1999.

Paul-André Rosental, « Liens familiaux, position sociale et mobilité des femmes dans la France du Nord au milieu du XIX^e siècle », *Cahiers du MAGE. Marché du Travail et Genre*, 3, 2000, p. 27-40.

Paul-André Rosental, « Between Micro and Macro: Theorizing Agency in Nineteenth-Century French Migrations », *French Historical Studies*, 29, 3, 2006, p. 457-481.

David A. Smith and Douglas R. White, "Structure and Dynamics of the Global Economy: Network Analysis of International Trade 1965-1980", *Social Forces*, vol. 70, n°4, June 1992, p. 857-893.

Tom A. B. Snijders & Krzysztof Nowicki, *Manual for Blocks version 1.8*, June 2007.

Snijders, Tom A.B., Pattison, Philippa E., Robins, Garry L., and Handcock, Mark S., New specifications for exponential random graph models. *Sociological Methodology*, 2006, 99-153.

Snijders, T.A.B., Steglich, C.E.G., and van de Bunt, G.G. (2009a). « Introduction to actor-based models for network dynamics » (preprint). *Social Networks*, in press, <http://dx.doi.org/10.1016/j.socnet.2009.02.004>.

Tom A.B. Snijders, Christian E.G. Steglich, Michael Schweinberger, Mark Huisman, *Manual for SIENA version 3.2, Provisional version*, University of Groningen: July, 2009b.

Statistique du département du Nord du préfet Dieudonné, Douai, Marlier, 1804.

Takai, Yukari (2001), « The Family Networks and Geographic Mobility of French-Canadian Immigrants in early-twentieth-century Lowell, Massachusetts », *Journal of Family History*, 3, 373-394.

Peter J. Taylor, *Distance decay models in spatial interactions*. Norwich : Geo Abstracts, 1975

Tobler, Waldo "Migration: Ravenstein, Thorntwaite, and Beyond." *Urban Geography*, Vol. 16, No. 4 (1995), pp. 327-343.

Appendix: Occupations

It is common knowledge that the Western (Flemish-speaking) part of our sample was generally more rural and agricultural than the Eastern part. We however lack precise data on each *commune*, apart from the somewhat vague information given in Joanne, 1869. We could look for local monographs in the future, but as for now we have instead chosen to use information from marriage records, namely the occupations of spouses living in each *commune*. They are however not easy to code: we would have liked to have an index of industrial and agricultural activities, but it is impossible to accurately divide *journaliers* (day laborers), *domestiques*, *servantes* (servants) and probably even *ouvriers* (workers, usually with an artisanal or industrial connotation) into these categories. We chose to build three indexes for each *commune* and period:

- percentage of workers among grooms, including all occupation names beginning with *ouvrier*, e.g. *ouvrier mécanicien* (mechanical worker), as well as occupations that reasonably could be included in a category of artisanal or industrial workers – mainly spinners, weavers, sawyers and workers in brewery (*garçons brasseurs*);

- percentage of farmers among grooms, including combinations based on the words *cultivateur* and *fermier* and a few gardeners;

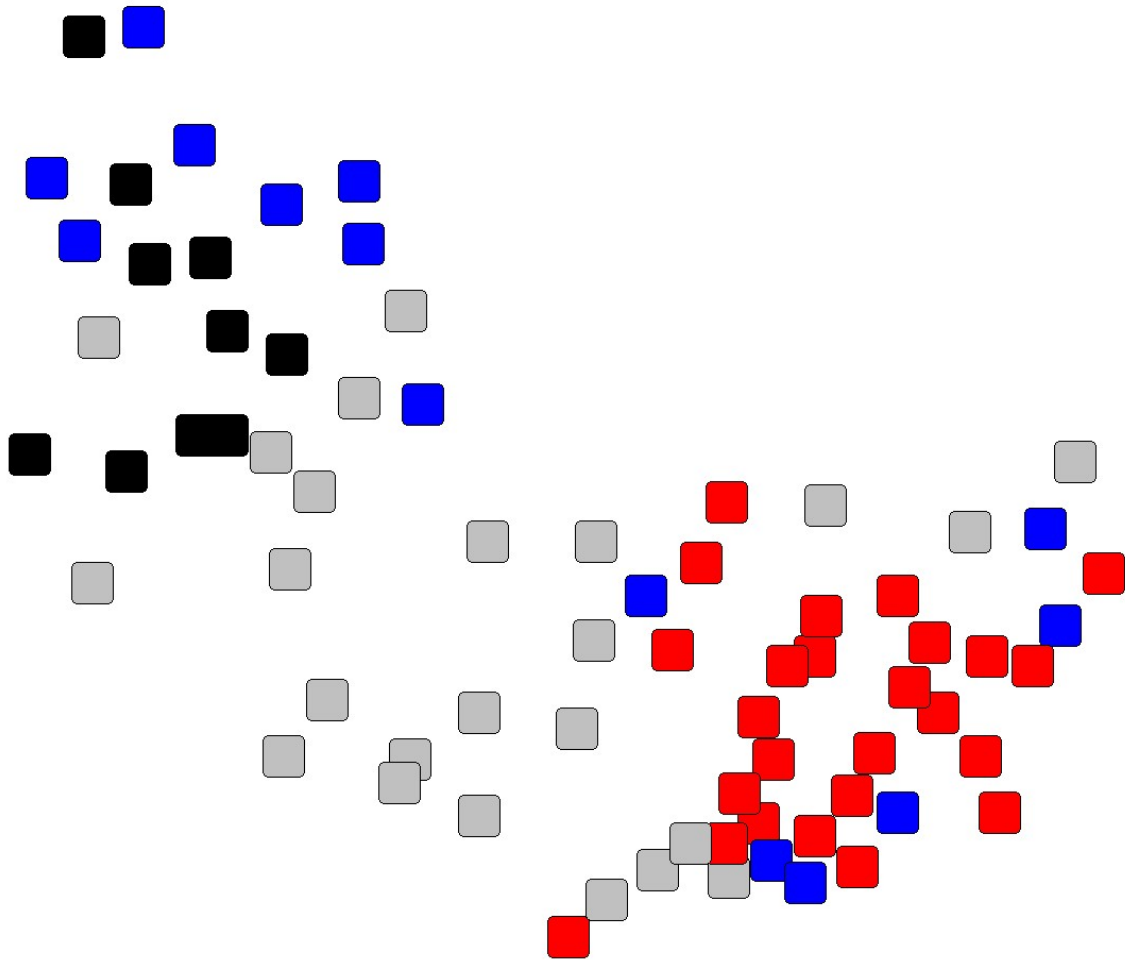
- a more complex index based on the percentage of literate grooms and brides and of the following exact occupations²⁸ among grooms and brides: bakers, carpenters, carters, cordwainers, cultivators, day laborers, farm servants (*domestiques de ferme* and some very close variants), joiners, masons, servants (*domestiques* and *domestiques à gages*), weavers and workers (*ouvriers* without precisions); cultivators, day laborers, farmers, farm servants (*servantes de ferme* and some very close variants), housewives (*ménagères*), ironers, lacemakers, linen spinners, maids (*domestiques* and *domestiques à gages*) seamstresses, servants (*servantes*), spinners, weavers, and “without occupation” (*sans profession*). The main occupations for men in the whole sample were cultivator (14 to 18% in each period), day laborer (12-21%) and weaver (9-18%), and for women cultivator (10-20%), day laborer (13-18%), spinner in the first period (10%; there was a crisis in linen hand-spinning in the 1840s), and “without occupation” in the last two periods (11-18%).

We used principal component analysis, then hierarchical clustering (Euclidean metric, Ward criterion, on coordinates on the first 5 axes)²⁹ to build clusters of *communes*. These clusters exhibit much clearer spatial patterns (see Maps A-C) than the indexes based on the percentage of cultivators or workers. They are also more correlated to attributes such as total population and percentage of sedentarity or in-migration from out of the sample. While there is no perfect indicator of the socio-economic structure of our *communes*, we therefore chose to include these clusters in some of our calculations. In addition, they are interesting *per se* in that they show, for example, changes in rural industrial occupations (linen spinners, lacemakers...) and literacy (see the last cluster for each period). Clusters, especially but not only for period 1, often seem at odds with classical depictions of archaic vs. modern or rural vs. urban – the opposition between clusters 1 and 2 in the last period is probably the only one that really fits in such antinomies.

²⁸ These occupations were by far the most common ones, together including 62% of mentions for men and 82% for women. We based our list on the most common occupations, only checking for variants in spelling, but keeping separate labelings of probably similar occupations, e.g. cultivators and farmers, maids and servants. It is therefore possible that some of these differences were more cultural (similar occupations were called in a different way in different parts of our region by the spouses and/or city hall clerks) than economic.

²⁹ Using the R package FactoMineR: <http://factominer.free.fr/>

Map A: Clusters in period 1



Each square represents a *commune* and they are placed according to geodesic coordinates. Criteria with v-tests above 2 or below -2 are listed below. Labels are very tentative.

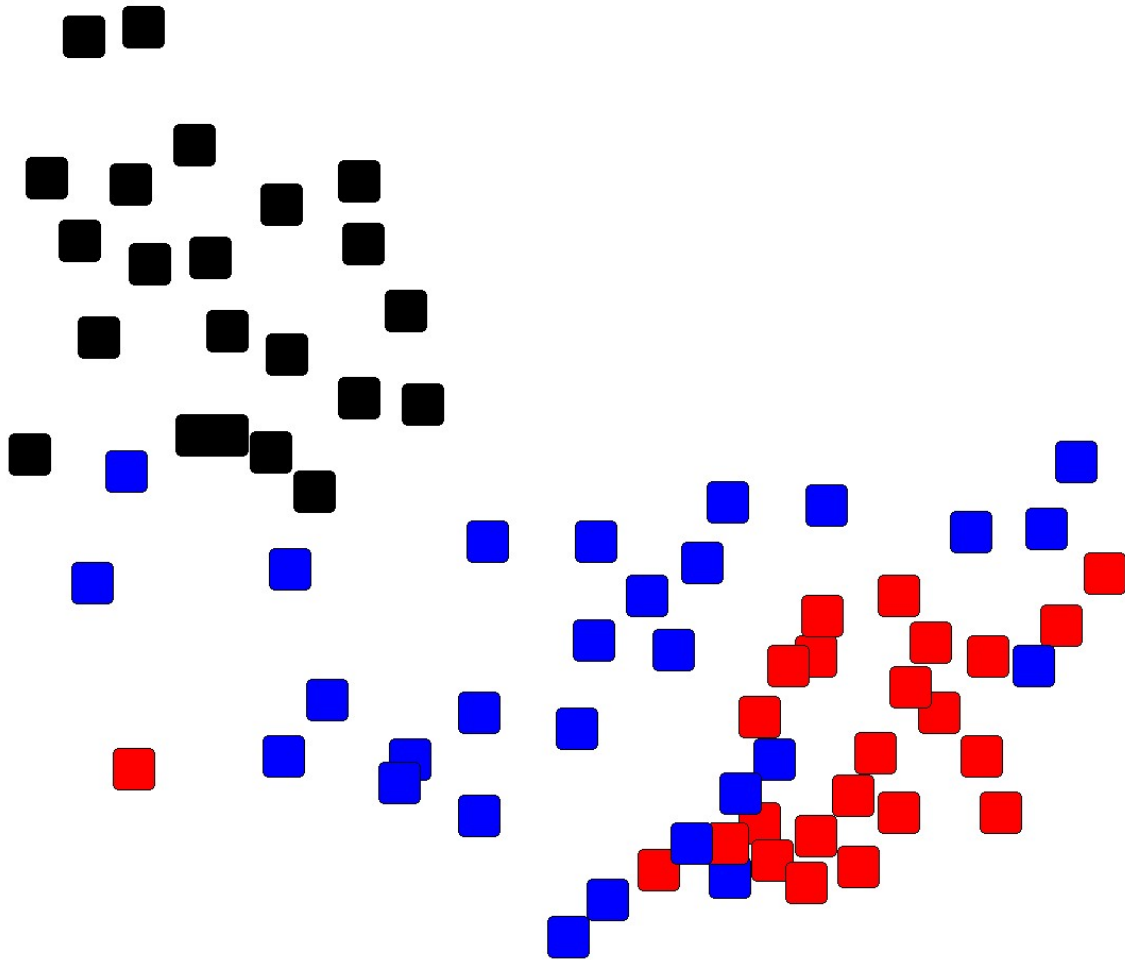
Cluster 1 (red) – *day-laborer villages* – over-represents day-laborers (43% of women, 37% of men) and masons (4% of men) and under-represents literate spouses (34% of women, 42% of men), female spinners (0%) and linen spinners (1%), maids and male servants (2%, 3%) and farm servants (0%, 1%), as well as housewives, cordwainers and workers. It is also totally French-speaking and over-represents small villages.

Cluster 2 (blue) – *urban* – over-represents farm servants (6% of women, 9% of men), bakers (2%), lacemakers (4%), seamstresses (9%), female weavers (5%) and literate men (54%) and under-represents day-laborers (7% of women, 12% of men), cultivators (12%, 13%) and male servants (3%). It also over-represents towns and includes many spouses not born in *communes* of the sample.

Cluster 3 (grey) – *agricultural* – over-represents linen spinners (11% of women), servants (7% of women, 10% of men), cultivators (24%, 21%) and under-represents female day-laborers (9%), ironers, carpenters, masons and bakers (0 to 1%). It is also distinctly French-speaking, exhibits high sedentariness rates and includes few spouses born outside of the sample.

Cluster 4 (black) – *worker-servant villages* – over-represents workers (6% of men), maids and male servants (22%, 18%), literate spouses (50%, 57%), as well as female farmers, spinners, ironers and housewives and under-represents day-laborers (0% of women, 9% of men) and female servants (0%). It is also distinctly Flemish-speaking, over-represents small villages, exhibits low sedentariness rates and includes few spouses born outside of the sample.

Map B: Clusters in period 2



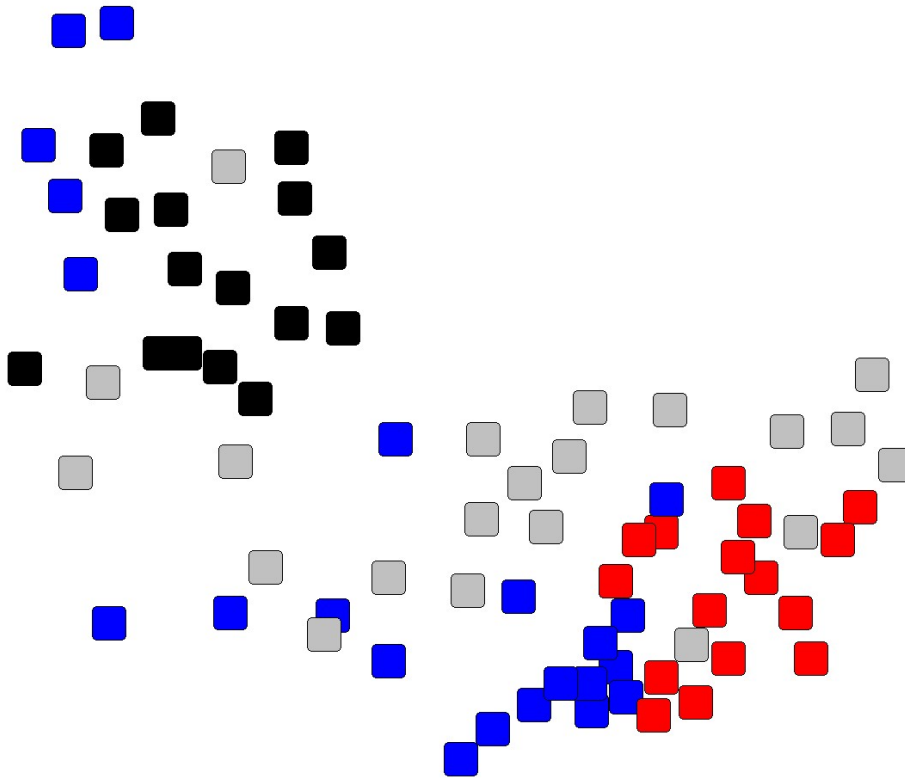
Each square represents a *commune* and they are placed according to geodesic coordinates. Criteria with *v*-tests above 2 or below -2 are listed below. Labels are very tentative.

Cluster 1 (red) – *day-laborers, urban* – over-represents day-laborers (27% of women, 21% of men), seamstresses (9%), as well as joiners, ironers, linen spinners and masons and under-represents cultivators (9% of men, 8% of women), weavers (6%, 2%), maids (2%) and servants (5%, 3%), as well as cordwainers and lacemakers. It is also totally French-speaking, with low sedentarity and many spouses born outside of the sample.

Cluster 2 (blue) – *rural industry* – over-represents weavers (15% of women, 27% of men) and male cultivators (19%) and under-represents male day-laborers (9%), seamstresses (4%), farm servants (1% of women, 3% of men), as well as lacemakers, spinners and workers. It is also distinctly French-speaking, with high sedentarity rates.

Cluster 3 (black) – *rural* – over-represents lacemakers (15%), maids (11%) and servants (12% of women, 12% of men), as well as cordwainers, carpenters, male workers and farm servants and female spinners. It under-represents weavers (1% of women, 9% of men), female day-laborers (4%), carters, joiners and ironers. It is also distinctly Flemish-speaking, over-represents small villages and includes few spouses born outside of the sample.

Map C: Clusters in period 3



Each square represents a *commune* and they are placed according to geodesic coordinates. Criteria with v-tests above 2 or below -2 are listed below. Labels are very tentative.

It is worth noting that clusters 1 and 4 are defined in a way that is extremely similar to clusters 1 and 3 in period 2. Cluster 3 in this period, while quite close geographically to cluster 2 in period 2, exhibit interesting differences (e.g. less cultivators).

Cluster 1 (red) – *day-laborers, urban* – over-represents day-laborers (27% of women, 24% of men), joiners (5%), housewives (6%), as well as ironers and linen spinners, and under-represents cultivators (7% of men, 3% of women), weavers (3%, 2%), servants (1%, 4%), as well as carters. It is also totally French-speaking, with low sedentarity and many spouses born outside of the sample.

Cluster 2 (blue) – *agricultural* – over-represents literate spouses (82% of women, 86% of men), cultivators (15% of women, 26% of men), as well as carters and “without occupation” (22% of women). It under-represents housewives (1%), male weavers, bakers, lacemakers and spinners. It is also distinctly French-speaking, with many small villages and high sedentarity rates.

Cluster 3 (gray) – *textile* – over-represents weavers (13% of women, 20% of men), female spinners (5%) and bakers (2%) and under-represents male day-laborers (9%), cultivators (7% of women, 12% of men), literate brides (65%), as well as cordwainers and seamstresses. It is also distinctly French-speaking.

Cluster 4 (black) – *rural* – over-represents lacemakers (11%), maids (6%) and servants (15% of women, 7% of men), as well as cordwainers, carpenters, male workers and farm servants (13% of men are farm servants) and female cultivators. It under-represents literate spouses (58% of women, 73% of men), day-laborers (4% of women, 9% of men), carters and joiners. It is also distinctly Flemish-speaking, over-represents small villages and includes few spouses born outside of the sample.