HAL open science

Aménités environnementales et choix résidentiels

C. Tartui

To cite this version:

C. Tartui. Aménités environnementales et choix résidentiels. Sciences de l'environnement. 2014. hal-02605765

HAL Id: hal-02605765

https://hal.inrae.fr/hal-02605765v1

Submitted on 16 May 2020

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.

Aménités environnementales et choix résidentiels

Mémoire

Cosmina TARTIU

Encadré par : Claire SALMON *M2CEES IAE Savoie Mont-Blanc*

Yves SCHAEFFER *IRSTEA*

SOMMAIRE

INTRODUCTION

Partie 1 Analyse statistique

1. Analyse cartographique

2. Courbe de Lorenz spatialisée

Partie 2 Analyse économétrique

1. Méthodologie

2. Les données

3. Estimations et interprétation des résultats

CONCLUSION

INTRODUCTION

Problématique

L'objectif général de ce travail est d'étudier les disparités socioéconomiques spatiales d'accessibilité aux aménités naturelles.

L'aménité est une qualité agréable ou utile associée à un lieu. Les aménités naturelles font référence aux « aspects hédoniques et esthétiques associés aux caractéristiques naturelles et de paysages, telles que : les arbres, les espaces ouverts, l'eau (lacs, rivières, et côte) et la topographie (collines, montagnes, canyons) » (Chi, 2013).

L'accès aux aménités est socialement inégal. La notion d'accès fait ici référence à sa dimension géographique, plus précisément à la distance que les personnes ont à parcourir pour profiter des aménités naturelles. En effet, il semble logique de supposer que, toutes choses égales par ailleurs, résider à une plus grande proximité des aménités naturelles constitue un avantage pour en bénéficier. A l'instar de De Palma (2007), on fait l'hypothèse que l'inégalité dans la distribution des aménités a un impact différencié sur les ménages et influence leur choix de localisation.

Dans ce contexte, la **question de recherche est de savoir quel rôle jouent les aménités naturelles dans la différenciation sociale des choix résidentiels des ménages ?** Est-ce que les individus les plus défavorisés sur le plan socioéconomique le sont également en matière d'accessibilité aux aménités naturelles ? Une réponse affirmative impliquerait qu'une répartition inégale des aménités naturelles constitue un déterminant de la ségrégation sociale résidentielle, et contribue en cela à la production des inégalités socioéconomiques. L'hypothèse économique sous-jacente est celle que les aménités naturelles sont des biens normaux, dont la consommation augmente avec le revenu. En conséquence, toutes choses égales par ailleurs, les ménages les plus riches ont une plus forte disposition à payer pour résider à proximité des aménités environnementales. Ces préférences résidentielles se traduisent par une hausse des prix des logements qui évince les ménages les plus défavorisés des zones les plus proches des aménités naturelles.

Pour répondre à cette problématique, nous proposons un travail en deux temps, mobilisant respectivement des méthodes : statistiques, cartographiques et économétriques.

1. Le travail en amont c'est concrétisé par la constitution d'une base de données d'indicateurs d'aménités naturelles et d'indicateurs socio-économiques à l'échelle communale. Les indicateurs ainsi construits, seront utilisés pour vérifier que les disparités spatiales en matière d'accessibilité géographique aux aménités naturelles sont corrélées avec les disparités socioéconomiques spatiales.
2. Dans un deuxième temps, nous allons étudier le rôle des aménités naturelles dans la différenciation sociale des choix résidentiels et d'évaluer ainsi leur contribution aux processus de ségrégation résidentielle.

Méthodologie

Deux types d'analyses seront conjugués pour apporter des éléments de réponse à la question de l'impact des aménités naturelles sur le choix résidentiels des ménages, à savoir l'analyse statistique - cartographique et économétrique. L'analyse statistique et cartographique consiste dans l'étude des relations statistiques (linéaires ou non) entre les indicateurs socioéconomiques et les indicateurs des aménités naturelles. L'analyse économétrique utilise le cadre théorique de l'analyse des choix discrets, l'objectif étant de modéliser les choix résidentiels des individus (variable expliquée) parmi un set fini d'alternatives.

Données : nous mobiliserons un échantillon représentatif d'individus ayant effectué au moins une mobilité résidentielle durant la période 1999-2008, à l'intérieur de la zone d'étude. Les données proviendront du RP de l'Insee, Fichier Migrations résidentielles MIGCOM.

Le modèle économétrique retenu sera de type logit conditionnel (McFadden, 1973,1981). Par construction, ce type de modèle admet une hétérogénéité des caractéristiques entre les alternatives. De ce fait, le logit conditionnel est privilégié pour l'estimation des modèles où le choix est principalement déterminé par les caractéristiques des alternatives plutôt que par des caractéristiques individuelles. Les résultats d'estimation serviront de base pour calculer les probabilités de choix résidentiels de chaque individu, en fonction de ses caractéristiques sociodémographiques. L'agrégation de ces choix permettra d'obtenir une prédiction de la distribution spatiale des migrants et du niveau de ségrégation sociale associé.

Cadrage

A l'échelle régionale, la littérature empirique sur les migrations et le développement économique suggère que les aménités naturelles sont un déterminant important des choix résidentiels et des disparités économiques spatiales (Waltert et Schläpfer, 2010 ; Waltert *et al.*, 2011). Cependant, à une échelle urbaine plus fine, l'effet des aménités naturelles sur les migrations et les processus de production des disparités spatiales demeure mal connu. Brueckner *et al.* (1999) ont montré théoriquement le rôle que les aménités naturelles exogènes et les aménités sociales endogènes sont susceptibles de jouer dans la production de la ségrégation urbaine. Notre intention étant ici de quantifier empiriquement ce phénomène.

D'un point de vue méthodologique, l'analyse constituera un prolongement de Cremer-Schulte, D. (2013), qui étudie le rôle des aménités environnementales liées à la montagne (autour de Grenoble) dans la dynamique des processus de ségrégation socio-économique, et de Goffette-Nagot et Schaeffer (2013), elle-même inspirée de Schmidheiny (2006), qui examine l'effet de la distance au centre-ville et de la composition sociale du voisinage sur les choix résidentiels et la ségrégation au sein des 25 plus grandes aires urbaines françaises.

Elle mettra également à profit sur les résultats de Bertrand *et al.* (2010), qui s'intéresse aux déterminants des choix résidentiels selon la position du ménage dans le cycle de vie au sein de la région urbaine montpelliéraine, et de De Palma (2007), qui identifie le rôle de nombreuses aménités locales dans les choix résidentiels des ménages de la région urbaine parisienne. La valeur ajoutée portera sur la richesse des caractéristiques des localisations résidentielles prises en compte (en particulier les aménités naturelles), la modélisation explicite du réseau de transport et des accessibilités.

Partie 1 Analyse statistique

Cette partie consiste dans une étude des relations statistiques (linéaires ou non) entre les indicateurs socioéconomiques et les indicateurs des aménités naturelles. Dans un premier temps on propose une analyse cartographique de la zone d'étude. L'objectif est de montrer que les aménités naturelles sont distribuées de manière inégalitaire dans l'espace. Puis, à l'aide des indices de ségrégation calculés sur les caractéristiques socio-économiques des communes, nous concluons à l'existence d'une répartition sociospatiale inégalitaire des ménages dans la zone d'étude. Finalement, on utilisera des courbes de Lorentz spatialisées pour examiner la relation entre les inégalités dans les aménités et les inégalités socio-économiques. Ce travail d'analyse nous permettra de mettre en évidence l'existence des inégalités socio-économiques d'accès aux aménités naturelles.

Plusieurs indicateurs seront mobilisés dans chaque partie.

- **Caractéristiques socioéconomiques** des espaces ***résidentiels*** des ménages: densité de la population, composition socio-professionnelle, revenu médian, structuration par âge et par la taille du ménage

  Sources : Recensement de la population (RP), Revenus fiscaux locaux des ménages

- **Caractéristiques naturelles** : typologie d'occupation du sol pour les grandes catégories d'aménités spatiales; accessibilité (en termes de distance et d'exposition directe) aux aménités remarquables de la zone d'étude.

  Sources : La base de données d'occupation des sols Corine Land Cover; IGN BD_TOPO ORONYMES pour les aménités remarquables, relief Alpes, BD CARTO pour les centres des mairies, BD ALTI.

La zone d'étude

Le travail d'analyse se concentrera sur le bassin d'emploi autour de l'aire urbaine d'Aix-en Provence - Marseille- Toulon, tel qu'il est défini par l'INSEE dans le zonage en aire urbaine de 2010 à partir de flux domicile-travail des actifs occupés. Selon le recensement de la population 2009, parmi les 920 000 actifs en emploi qui résident au sein du réseau, 98%

travaillent à l'intérieur du réseau. Cette zone présente une forte dotation en aménités naturelles (exemples : la Côte Méditerranéenne, le Parc national des Calanques) et une forte densité de la population (540.4hab./km^2 contre 101.6 hab./km^2 en moyenne pour la France). L'échelle géographique privilégiée sera celle de la commune, soit 144 communes dans la zone d'étude.

1. Analyse cartographique[1]

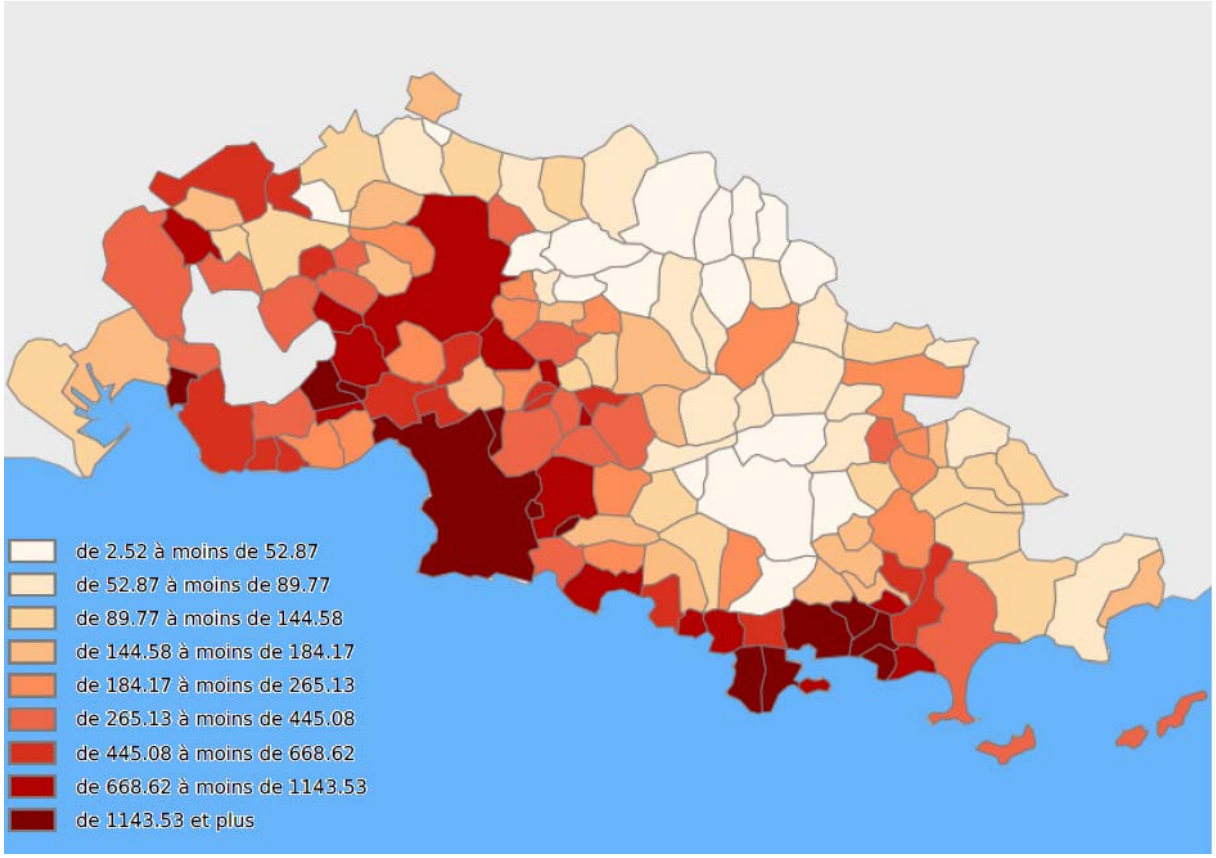
1.1. Caractéristiques socioéconomiques

1. La population des ménages

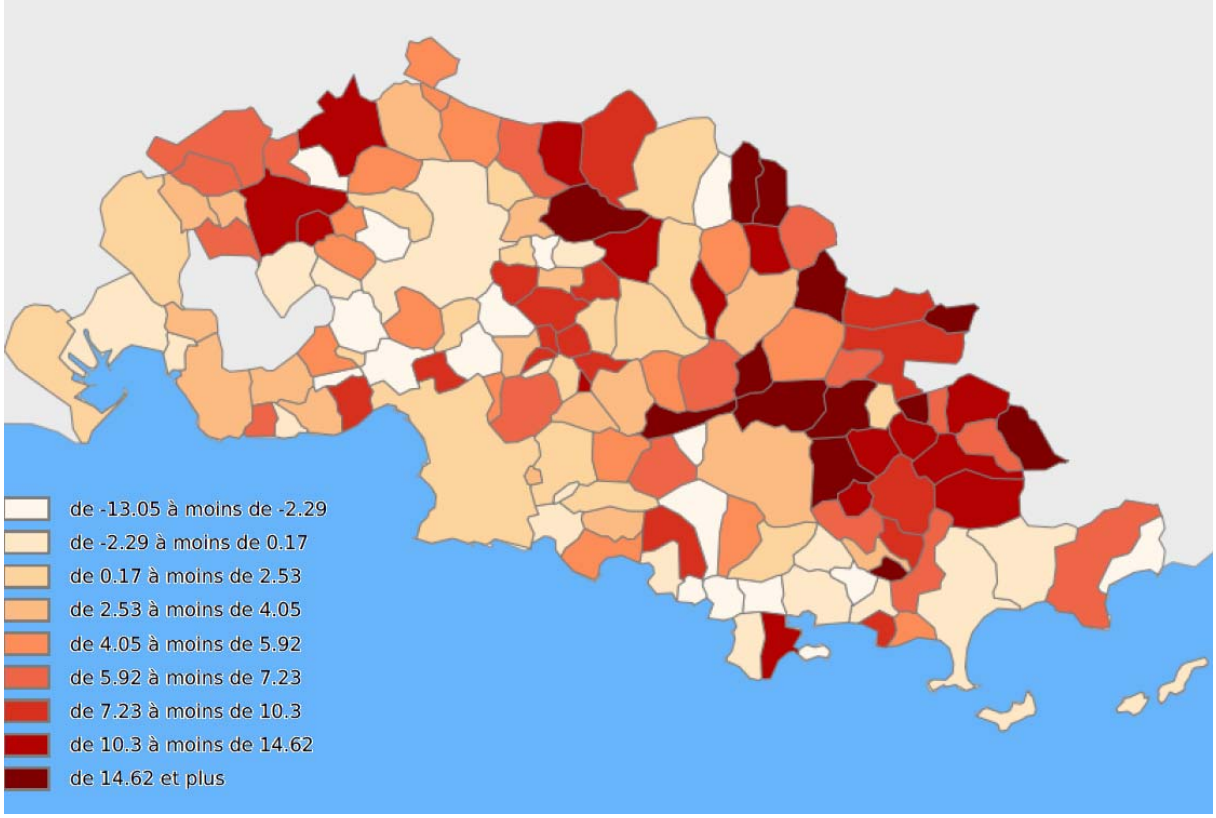
La population totale de la zone d'étude s'élevé à presque deux millions et demi d'habitants en 2011, soit un nombre de ménages d'un peu plus d'un million. C'est un territoire très dense avec 540.4hab./km^2 en moyenne contre 101.6 hab./km^2 pour la France. Comme on peut le voir su la carte, la population est surtout concentrée dans les trois grands pôles urbains : Toulon (3827 hab./km^2), Marseille (3535 hab./km^2) et Aix en Provence (757 hab./km^2) et le long de la Côte Méditerranéenne. Sur les derniers 5 années, on observe des taux de croissance de la population très hétérogènes. Ainsi, les zones les plus denses connaissent une plus faible croissance de la population. La population des grandes villes est en croissance atone, voire en baisse Toulon (-2.29%), Marseille (1.38%) et Aix en Provence (-1.3%), tandis que les communes situées dans le nord-ouest connaissent une très importante croissance de la population, allant jusqu'à 52.17%.

[1]Cartes réalisées par l'auteur à l'aide de la plateforme web OASIS (Outil d'Analyse de la Ségrégation et des Inégalités Spatiales), du portail SIDDT (Système d'Information Dédié aux Territoires) Irstea et de ArcGis ArcMap 10.2.

Carte1 : Densité de la population (2011)



Carte 2 : Evolution de la population (%) 2006-2011

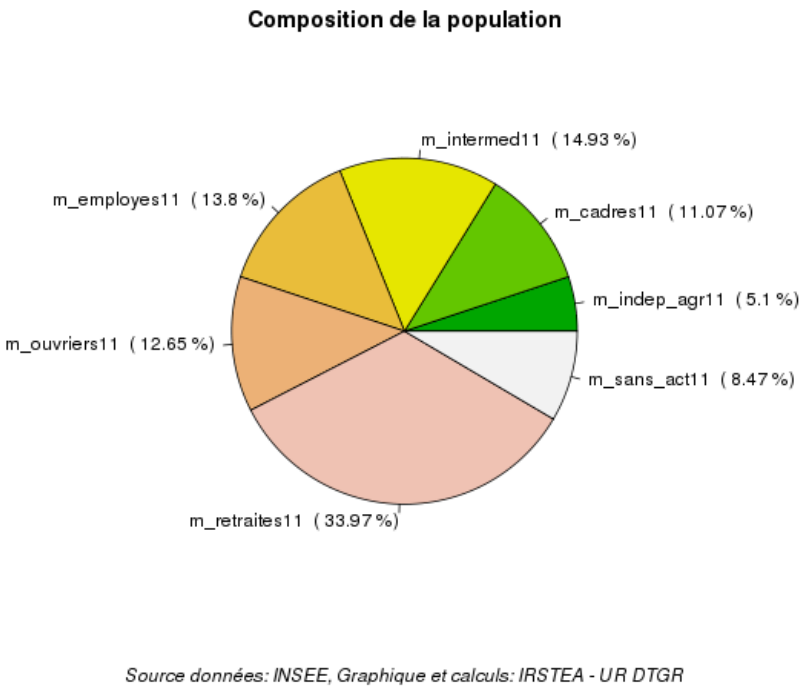


2. Catégories socio professionnelles

Etat des lieux

Les retraités correspondent à la catégorie socio - professionnelle la mieux représentée, avec 34 % des ménages. Ils sont, suivis dans l'ordre, par les professions intermédiaires, les employés, les ouvriers et les cadres, dans des proportions assez proches, allant de 14.93% jusqu'à 11.07%.

Figure 1 :



Par la suite on propose une illustration de la localisation géographique de ces catégories (en 2011), ainsi que leur évolution sur la période 1990 et 2011.

Pour mieux rendre compte des disparités socio spatiale des différents groupes, nous analyserons la cartographie des catégories socio-professionnelles à l'aide des quotients de localisations. Les quotients de localisation sont plus adaptés que les proportions simples lorsqu'il s'agit de mesurer la ségrégation. Un quotient de localisation (Isard 1960) est un indice permettant de mesurer la concentration relative d'un groupe au sein d'un ensemble. Il est calculé pour chaque groupe séparément et répond aux questions liées à la surreprésentation ou la sous-représentation d'un groupe dans le territoire. Il représente rapport entre la proportion du groupe dans la population de chaque unité spatiale et la proportion du groupe dans la zone d'étude.

$$QL_i^k = \frac{x_i^k / t^i}{X^k / T}$$

où k = catégorie de population ; i = unité spatiale ; x_i^k= population du groupe k dans l'unité spatiale i ; X^k = population totale de la catégorie k dans la zone d'étude ; t^i = population totale dans l'unité spatiale i; T = population totale de la zone d'étude.

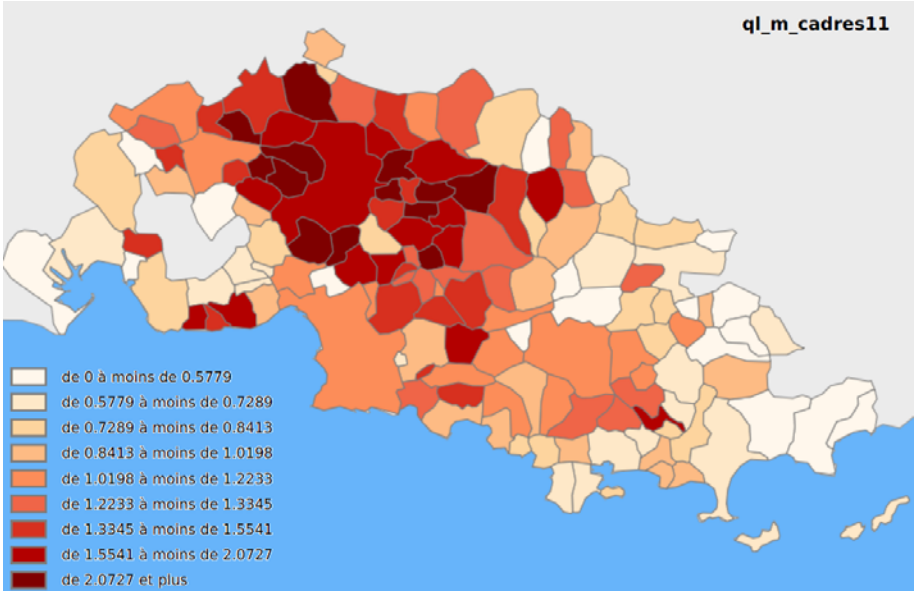
Interprétation :

- $Q_L < 1$ $Groupe\ cible\ sous - représenté$
- $Q_L > 1$ $Groupe\ cible\ surreprésenté$
- $Q_L = 1$ $Groupe\ représenté\ de\ manière\ égale$ à $la\ structure\ de\ la\ population$

Par la suite, nous focaliserons l'analyse que sur les catégories de population les plus représentatives des disparités spatiales dans la zone d'étude. Ces groupes sont : les cadres, les ouvriers, les retraités et les inactifs. Ces catégories se distinguent par le revenu, le statut, le degré de mobilité résidentielle et par la disposition à payer pour des aménités naturelles.

L'opposition entre les cadres et les ouvriers est une comparaison classique dans l'analyse de la ségrégation. Par rapport aux ouvriers, les cadres sont considères comme la catégorie à plus haut revenu, très dynamiques et mobiles, ayant une disposition à payer plus importante pour être exposés aux aménités. L'hypothèse économique sous-jacente, est ici celle considérant les aménités naturelles comme des biens normaux dont la consommation augmente avec le revenu.

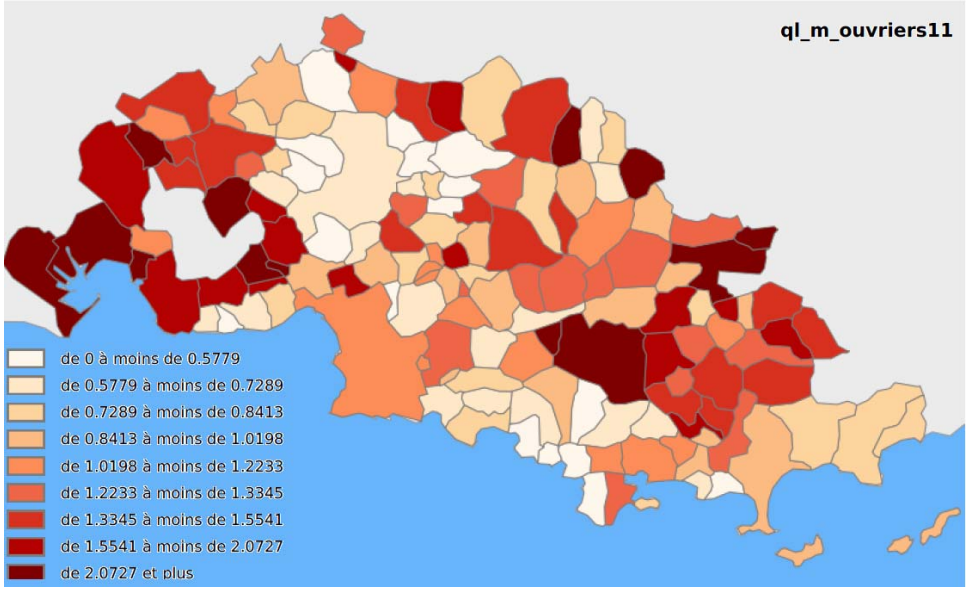
Carte 3 : Quotient de localisation des ménages cadres (2011)



Maximum: 3.1772 (SAINT-MARC-JAUMEGARDE)

Interprétation : Dans cette commune il y 3.17 fois plus de cadres que la moyenne des cadres dans la population de la zone.

Carte 4 : Quotient de localisation des ménages ouvriers (2011)

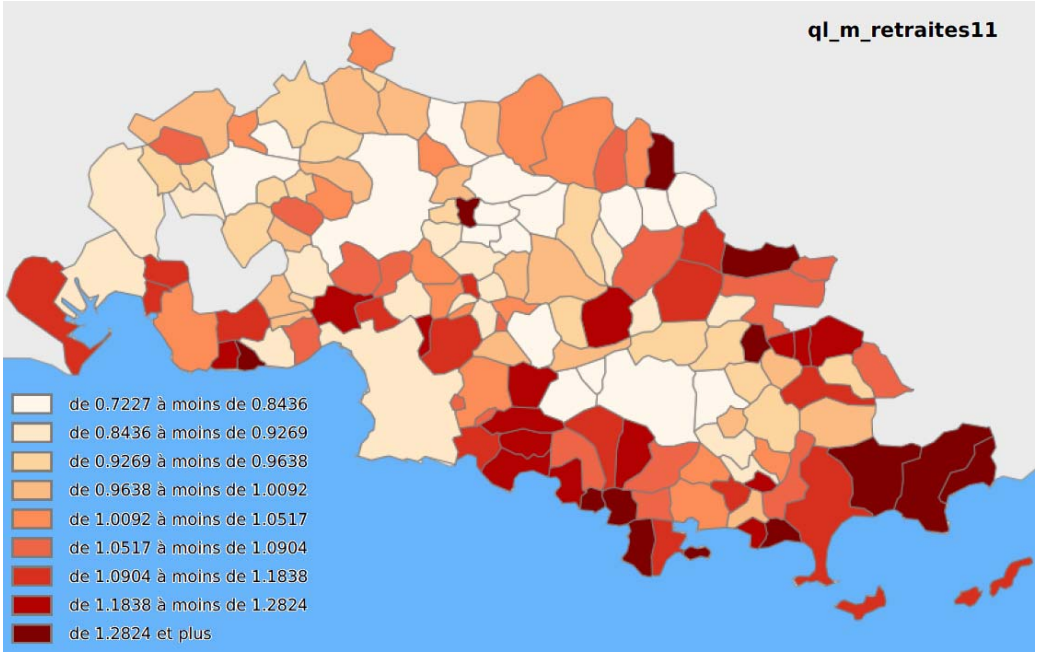


Maximum: 2.4962

Les résultats cartographiques illustrent une ségrégation évidente entre cadres et ouvriers. Les cadres sont surreprésentés à l'intérieur de la zone d'étude, notamment autour d'Aix en Provence et Marseille. Les ouvriers se situent en surnombre du côté ouest, autour de l'Etang de Berre et du côté est, autour de Brignoles.

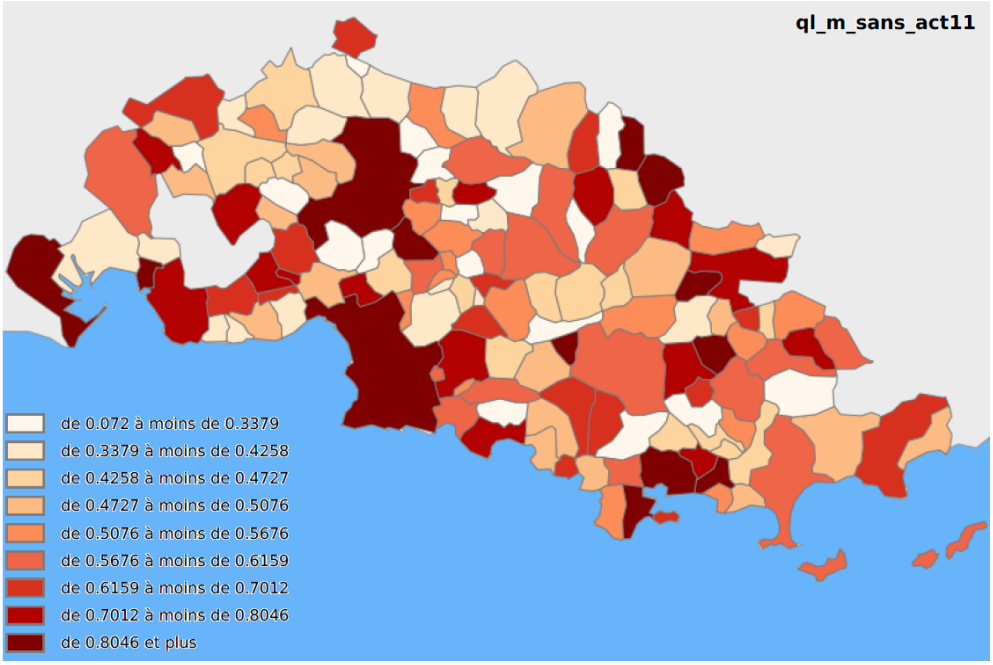
Les disparités entre les localisations des retraités et est celles des sans activité constituent une deuxième opposition marquante dans cette cartographie de la ségrégation. Les retraités correspondent, à priori, à une population à budget stable, à mobilité faible, dont les choix de localisation ne sont plus contraints par le travail mais par la recherche des aménités naturelles et des services publics. A l'opposée, les inactifs ont généralement un revenu variable, ils sont très mobiles en suivant l'emploi et recherchent les aménités urbaines (Bertrand 2010). Ces affirmations sont confirmées par l'analyse cartographique de la localisation des retraités et des inactifs.

Carte 5 : Quotient de localisation des ménages retraités (2011)



Les retraités sont concentrés dans les communes côtières (sauf Marseille), à la recherche des aménités littorales, mais pas très loin des centres urbains procurant des services publics. La distribution des inactifs est plus dispersée sur le territoire, avec une surreprésentation dans les grandes villes à recherche des aménités urbaines, autour de l'Etang de Berre à recherche d'emploi. En plus, on observe une sous-représentation ailleurs sur la Côte, ce qui témoigne d'un moindre intérêt pour les aménités naturelles.

Carte 6 : Quotient de localisation des ménages sans activité (2011)

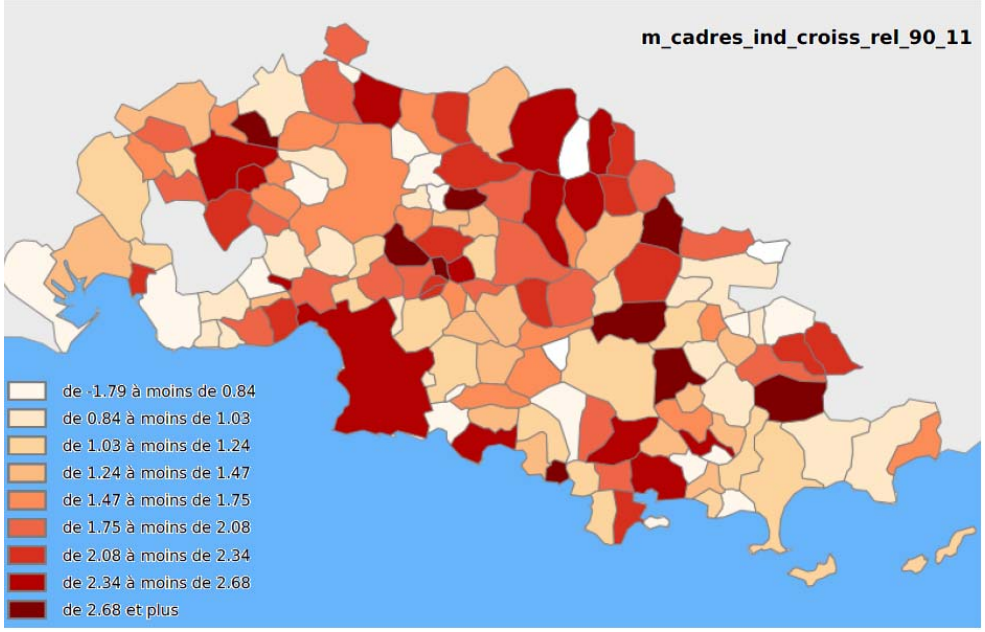


Maximum: 2.9503

Evolution

L'évolution des catégories socio-professionnelles, sur les 20 dernières années, montre des schémas de localisation différentes par rapport aux stocks existants.

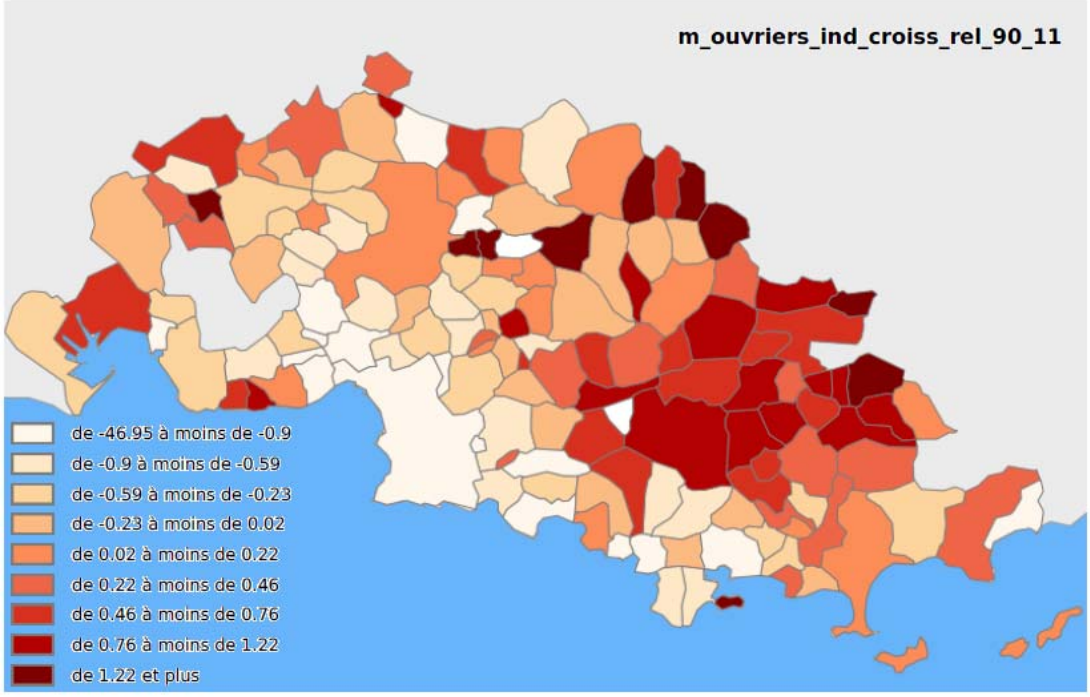
Carte 7 : Taux de croissance des ménages cadres (1990 -2011)



Maximum: 4.75 %

On constate que la croissance relative des cadres est très faible surtout dans les communes avec une forte présence de cette catégorie autour d'Aix en Provence. Cela pourrait être l'effet de saturation du marché foncier ou des contraintes de politique locale. Les « nouveaux » cadres se localisent plus vers la périphérie nord du territoire ou vers les deux autres pôles urbains : Marseille et Toulon.

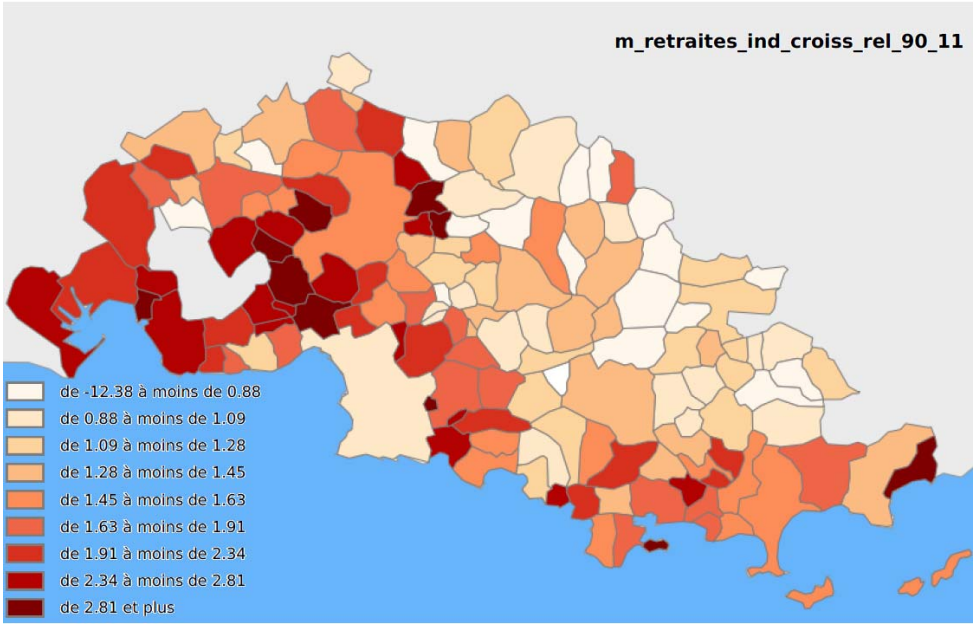
Carte 8 : Taux de croissance des ménages ouvriers (1990 -2011)



Maximum: 2.88 %

Les ouvriers semblent être à la base des mouvements de masse, car le territoire où leur croissance est la plus prononcée coïncide avec celui où l'augmentation de la population dans son ensemble la plus forte (voir la carte 2).

Carte 10 : Taux de croissance des ménages retraités (1990 -2011)



Maximum: 7.52 %

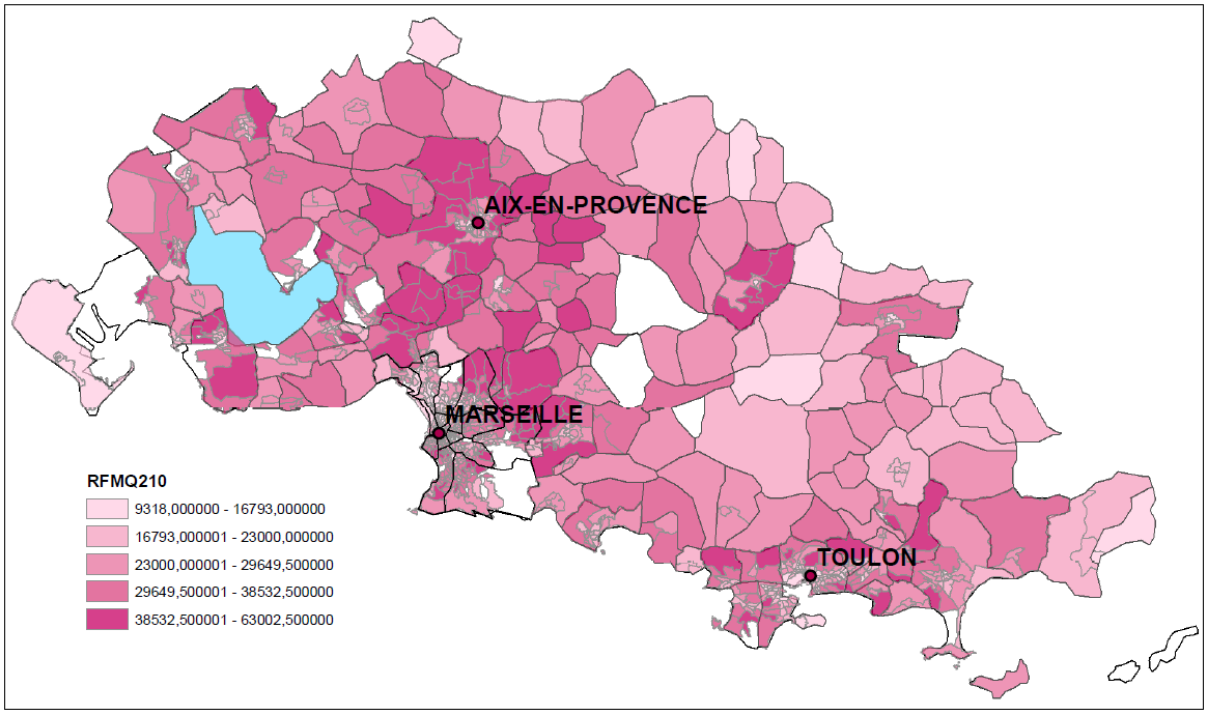
Le pôle de forte croissance de la population retraitée est constitué autour de l'Etang de Berre, jusqu'à Aix en Provence. Le reste de la suit la Côte Méditerranéenne suit la tendance de croissance moyenne de cette catégorie (autour de 1.5%).

3. Revenu médian disponible des ménages

Le revenu médian disponible des ménages est le niveau de revenu tel que 50% des ménages gagnent plus et 50% des ménages détiennent moins. Nous avons analysé cette variable au niveau de l'IRIS. L'IRIS est un découpage infracommunal proposé par l'Insee, des communes de plus de 10.000 habitants et de la plupart des communes de 5.000 à 10.000 habitants. Ce découpage constitue une partition du territoire en quartiers de 2.000 habitants. A ce niveau très fin d'analyse, la variable revenu médian montre d'importantes disparités entre les unités géographiques, l'étendue de sa distribution est comprise entre 9000 €/ ménage et 63 000 €/ ménage, pour une moyenne de 28 225 €/ménage. Ce type d'indicateur rend compte du milieu de la distribution des revenus dans la population. Un niveau élevé du revenu médian, ne témoigne pas forcement d'une localisation très riche, mais plutôt de l'existence d'une

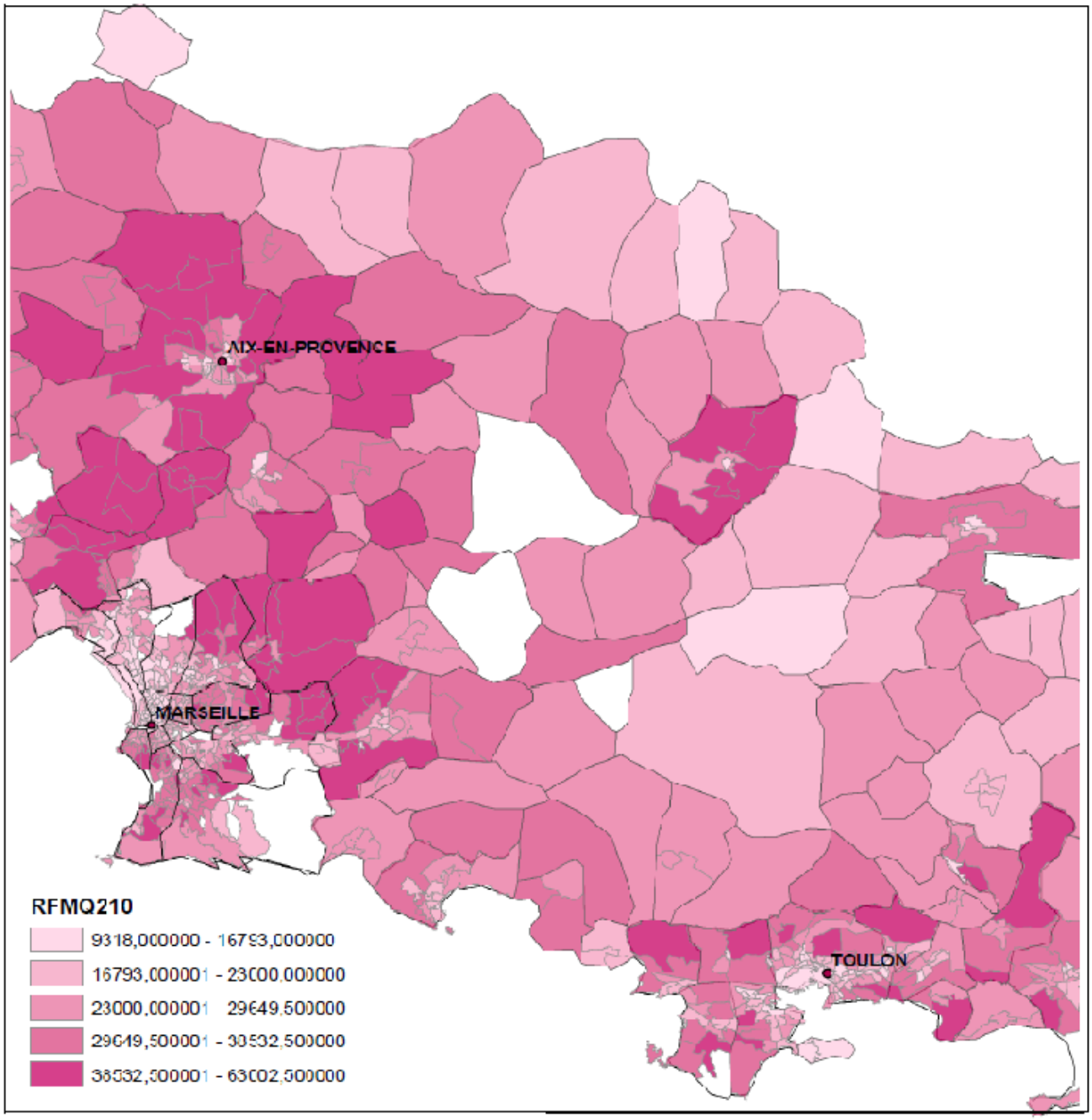
importante classe moyenne. Au contraire, un niveau bas de la variable témoigne de la précarité d'au moins 50% de la population concernée.

Carte 11a : Revenu médian disponible des ménages

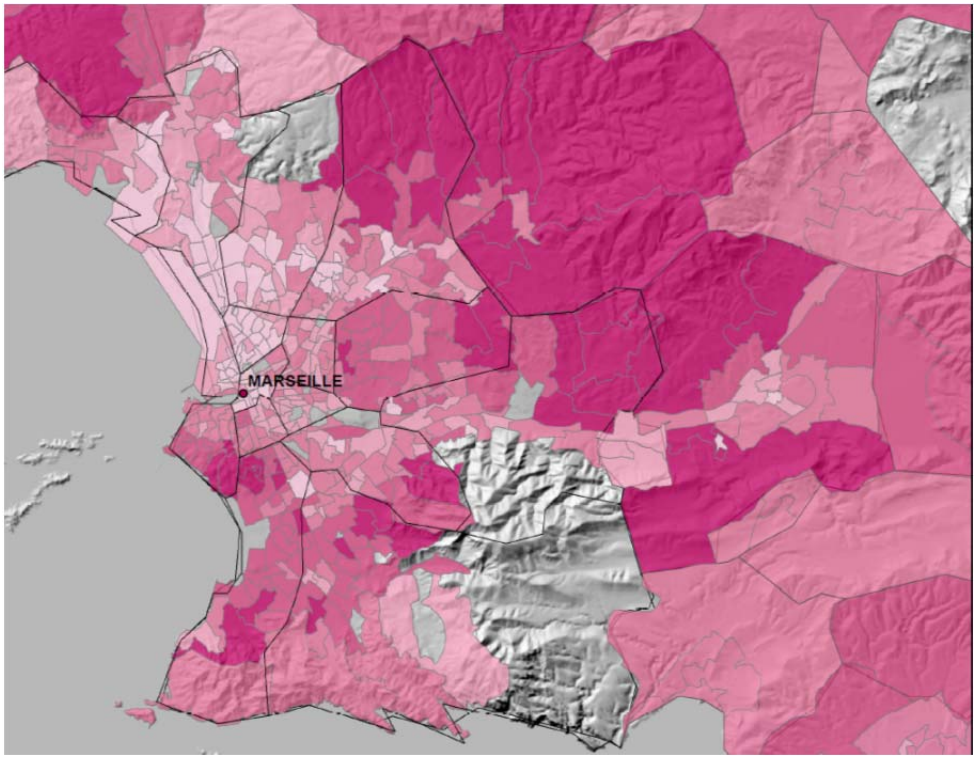


Comme on peut le constater sur les cartes, les unités spatiales les plus aisées se trouvent autour des trois grands pôles urbains de la zone : principalement Aix-Marseille et dans une moindre mesure Toulon. Lorsqu'on regarde de plus près la distribution des revenus à l'intérieur des pôles, on observe que les localisations périurbaines sont caractérisées par des revenus plus importants que les localisations centrales.

Carte 11b : Revenu médian disponible des ménages



Carte 11c : Revenu médian disponible des ménages : Marseille



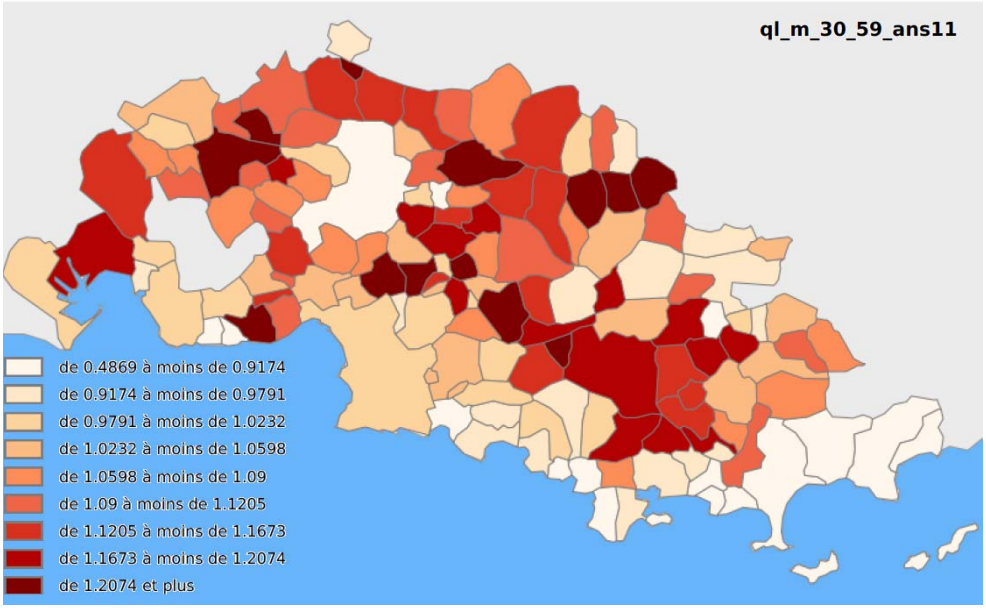
4. Age des ménages

La population de la zone d'étude peut être classifiée selon trois catégories d'âge : les jeunes (moins de 30 ans), les adultes (entre 30-59 ans) et les seniors (plus de 60 ans).

L'illustration cartographique de la localisation des trois catégories permet d'identifier des distributions spatiales inégales. La majorité (52.37%) correspond aux adultes de 30 à 59 ans. C'est une population surreprésentée à l'intérieur du territoire, en périphéries des grandes villes et à l'écart de la Côte.

La part des seniors dans la population est non négligeable, avec un poids largement supérieur la moyenne nationale (37.27 % contre 23.1%). La localisation de cette population suit fidèlement celle des retraités tout le long de la Cote (sauf à Marseille). Pour l'illustration, voir Carte 5 : Quotient de localisation des ménages retraités.

Carte 12 : Quotient de localisation des ménages âgés de 30 à 59 ans (2011)



Maximum: 1.4328

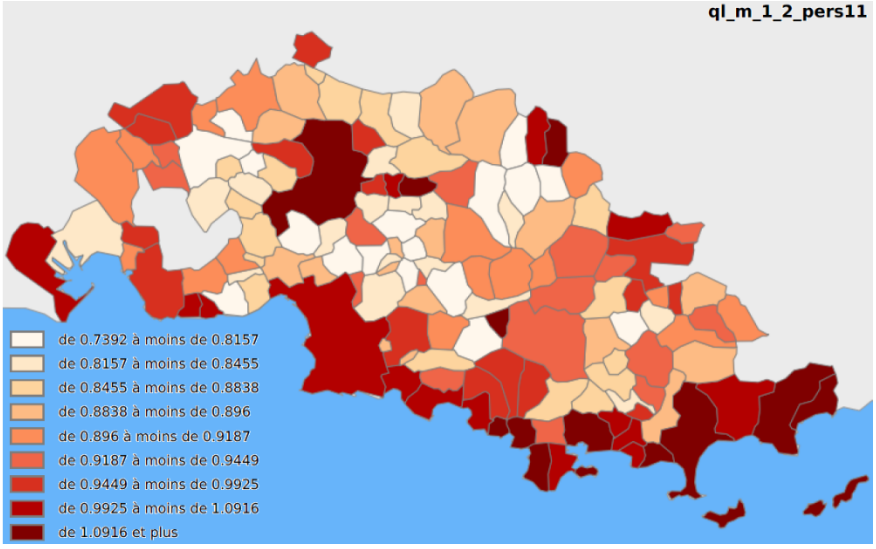
Les jeunes de 15 à 29 ans, correspondent à la catégorie la moins représentée avec 10.39% des ménages. L'analyse cartographique montre que leur répartition spatiale est très semblables aux inactifs : ils sont notamment concentrés dans les grandes villes. Ceci n'est pas surprenant, lorsque l'on sait que les étudiants et les individus n'ayant jamais travaillé (principalement des jeunes) comptabilisent la majeure partie de la catégorie des sans activité. Pour l'illustration, voir la Carte 6 : « Quotient de localisation des ménages sans activité ».

5. Taille des ménages

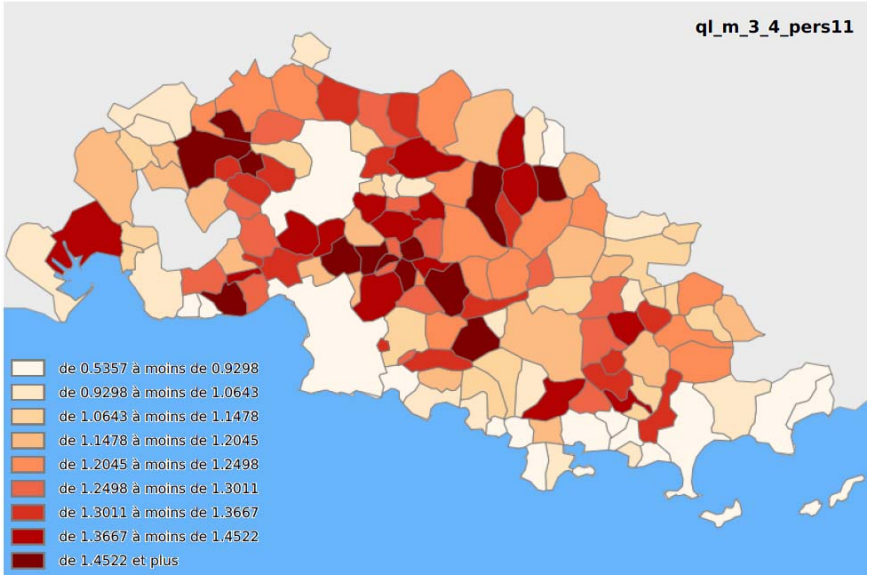
Le nombre des personnes dans le ménage est une autre caractéristique susceptible de structurer le comportement de localisation résidentielle des individus. Ceci peut se justifier par le fait que cette variable est étroitement liée à l'âge et au positionnement de l'individu dans le cycle de vie. Pour ce qui est des migrations non professionnelles, Baccaini (1994) montre que, avant 30 ans les raisons liées à la constitution de la famille dominent les migrations résidentielles. Ce sont des « migrations nécessaires » et de ce fait la mobilité des ménages est très importante. L'auteur affirme que les chances de changer de logement augmentent avec la naissance du premier enfant, alors que les naissances suivantes n'ont

aucun effet sur la mobilité des individus. Avec l'âge, une fois que le ménage est consolidé, on migre plutôt dans le but de trouver une localisation correspondant mieux à ses besoins. Il s'agit des « migrations de confort » et la mobilité des individus diminue. Les résultats cartographiques de notre zone d'étude mettent en évidence des comportements de localisation différents selon la taille du ménage. Les ménages de 1 à 2 personnes préfèrent résider à proximité de la Mer et dans les grandes villes. La tendance est inversée et stable pour les ménages à plusieurs personnes : on se décentralise en se localisant en périphérie des grandes villes, à l'intérieur des territoires et à l'écart de la Côte.

Carte 13 : Quotient de localisation des ménages âgés de 1 à 2 (2011)



Carte 13 : Quotient de localisation des ménages âgés de 3 à 4 (2011)



1.2. Caractéristiques naturelles

Dans cette partie on propose une illustration cartographique de la distribution inégalitaire des aménités naturelles dans la zone d'étude.

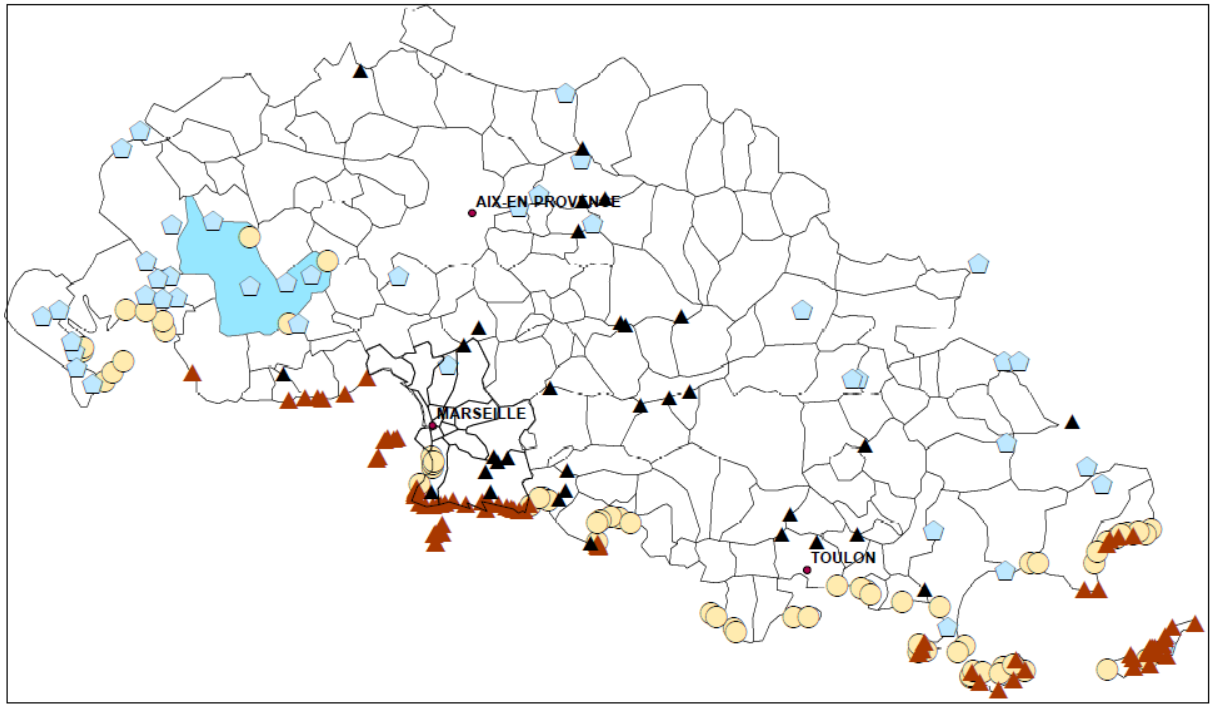
La première carte montre le positionnement des aménités naturelles spécifiques : les lacs (en bleu), les plages (marron sable) ; les montagnes (triangle noir) et les calanques (triangle marron). On observe que leur distribution dans l'espace est clairement hétérogène : la Côte concentre les aménités littorales (plages et calanques), la partie centrale est plus concernée par les aménités de montagne.

D'une manière générale les ressources en eau correspondent aux eaux permanentes, maritimes et continentales, et constituent des aménités naturelles appréciés par les ménages. Les rivières (Carte 15) sont reparties de façon assez homogène, tandis que les lacs sont surreprésentés soit dans l'ouest, soit dans le sud du territoire.

Les forêts (tous types confondus) semblent la ressource naturelle la plus égalitairement repartie sur la zone d'étude. Apriori, toutes les catégories de population devraient avoir un accès semblable à cette aménité.

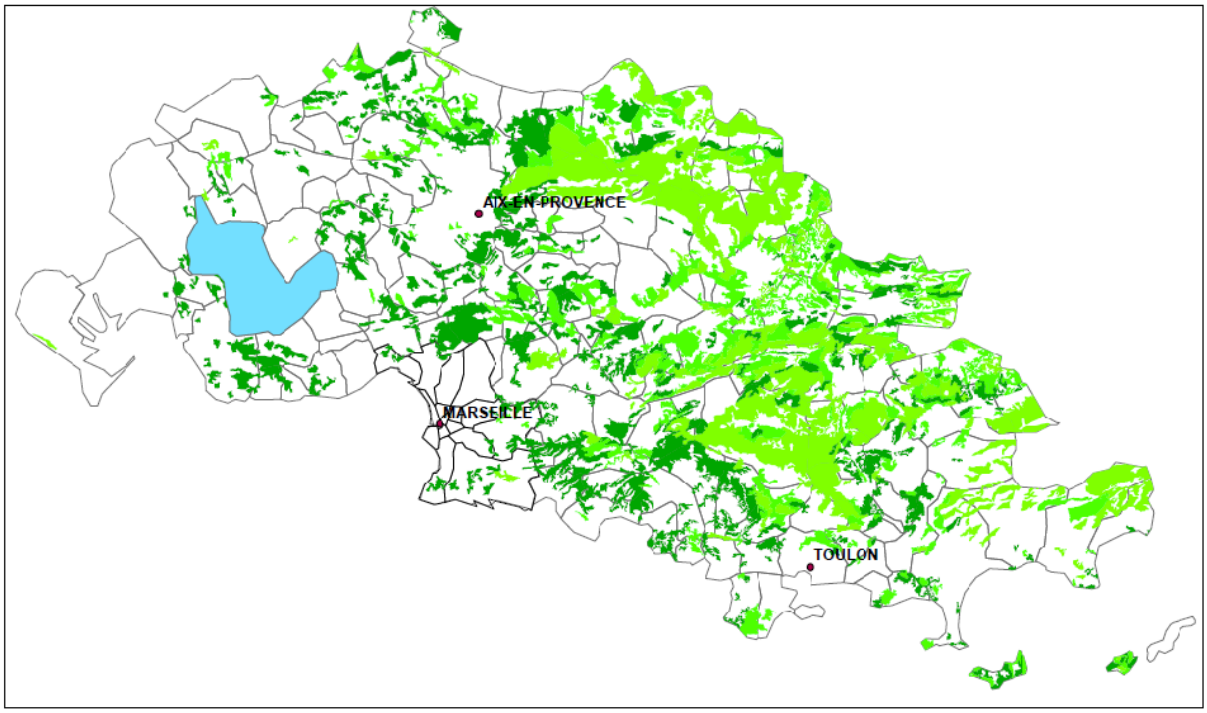
La végétation arbustive semble plus concentrée dans l'espace que les forêts. Lors de la superposition d'une couche cartographique de relief (carte 18), on observe que les surfaces couvertes par de la végétation arbustive suivent majoritairement le contour des montagnes. Ainsi, le taux de végétation arbustive dans la commune contribue à capter des aménités montagnardes de manière plus large. L'accès à cette aménité reste inégalitaire, mais il est plus important que l'accès aux montagnes.

Carte 14 : Localisation des aménités remarquables

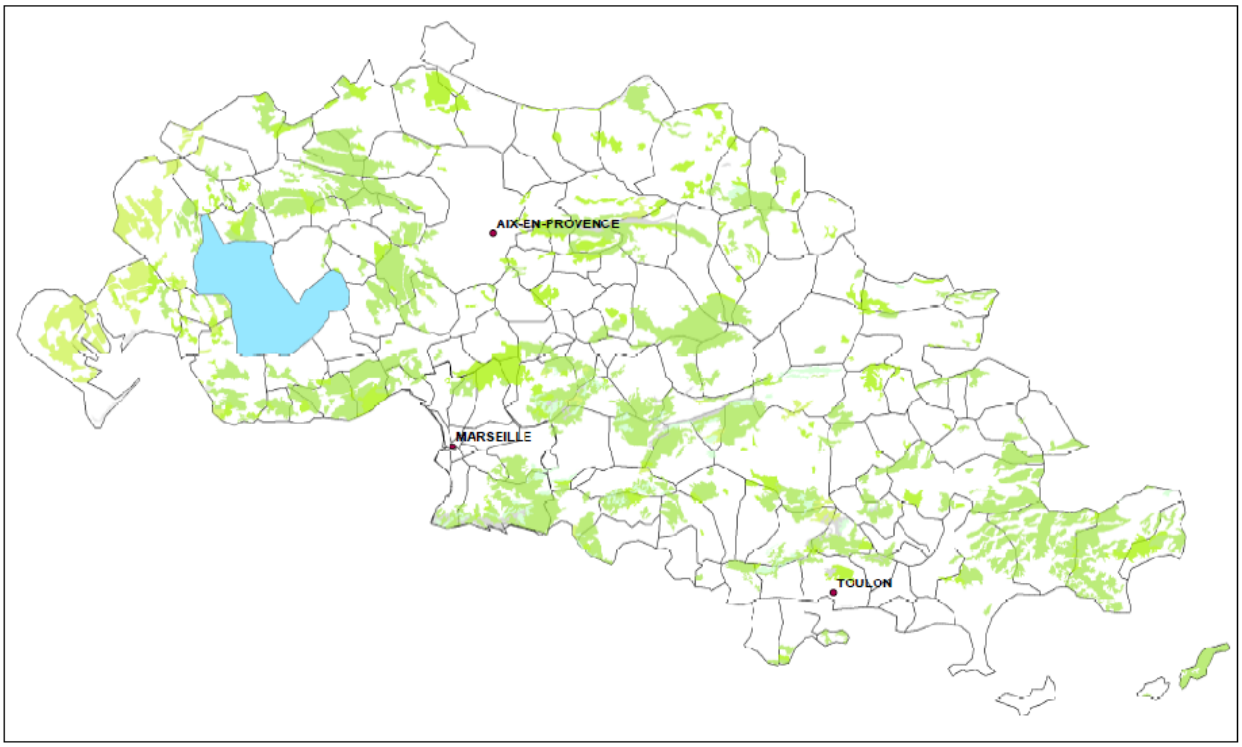


Carte 15 : Répartition des surfaces en eau

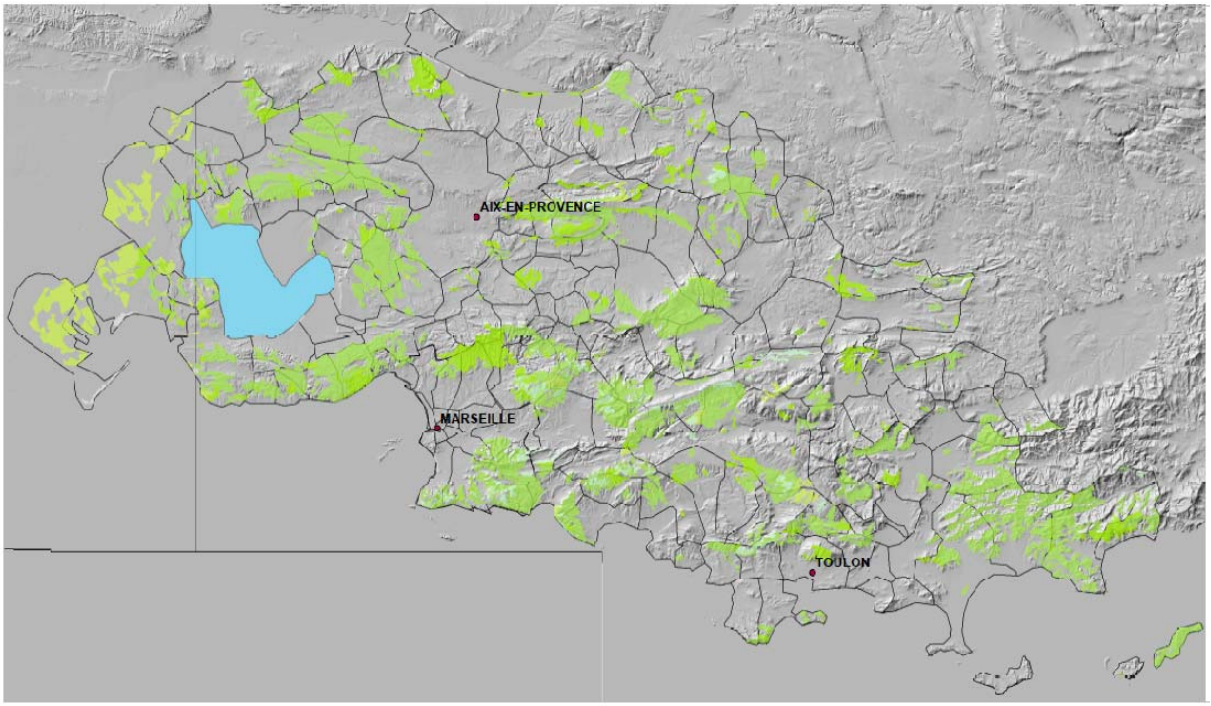
Carte 16 : Répartition des forets



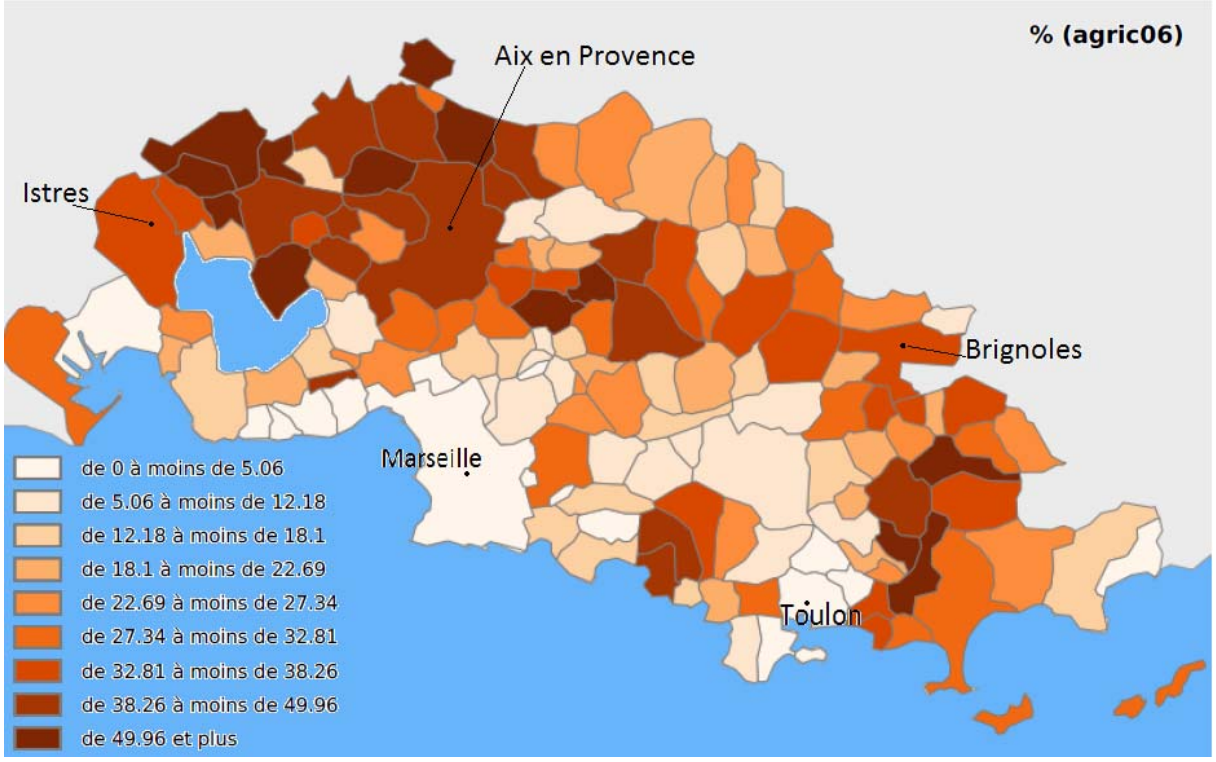
Carte 17 : Répartition de la végétation arbustive



Carte 18 : Corrélation végétation arbustive et relief



Carte 19 : Répartition des territoires agricoles



Le taux de terres agricoles de la commune est la dernière catégorie d'occupation du sol illustrée. Selon les résultats cartographiques, cette ressource est plus intensivement représentée dans les communes du nord-ouest et dans celles du sud-est de la zone d'étude. Les disparités de dotations en territoires agricoles sont importantes, car le pourcentage du sol occupé par cette ressource varie entre 0% et 50%.

2. Courbe de Lorenz spatialisée

Les résultats cartographiques précédents ont permis d'illustrer l'inégalité de la répartition des ressources humaines, naturelles et économiques dans l'espace.

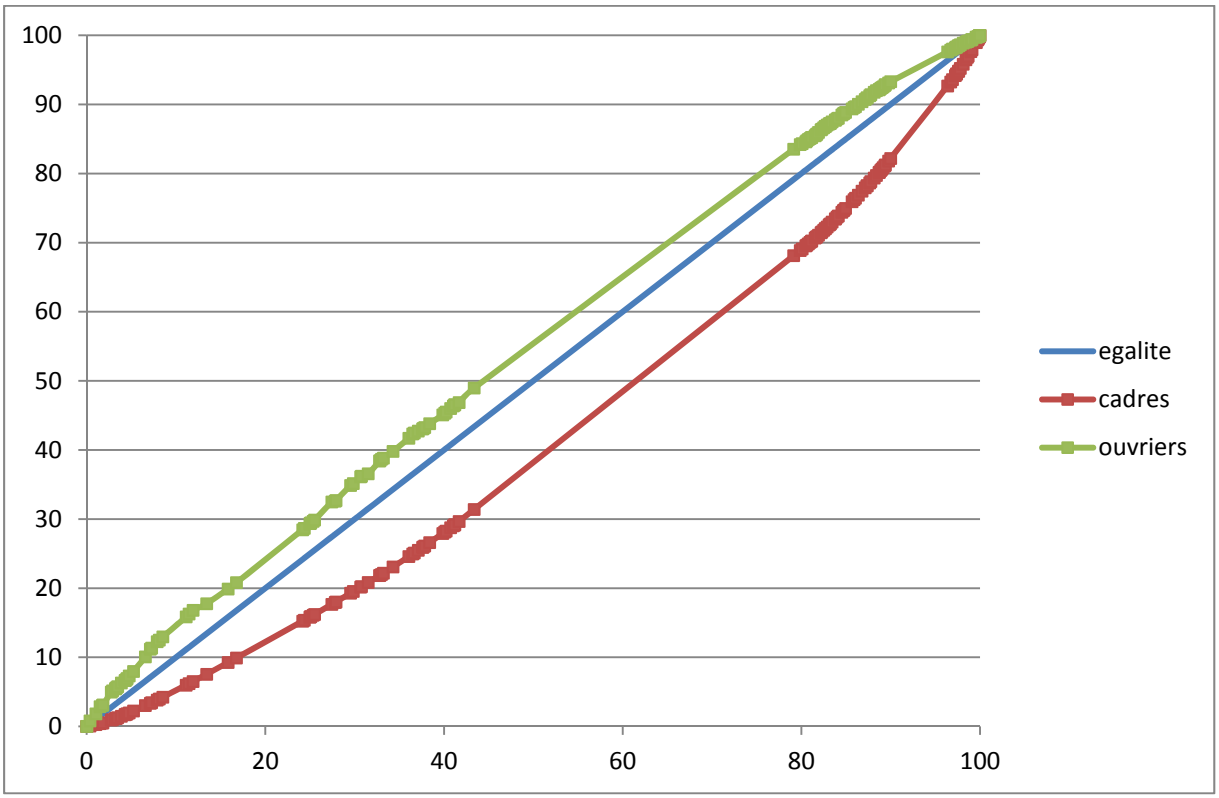
2.1. Cadres versus ouvriers

Un instrument classique d'analyse des inégalités dans une population (individus ou unités spatiales) et qui permet d'avoir une vision d'ensemble du phénomène, c'est la courbe de Lorentz. Dans sa version spatialisée, elle offre une mesure des inégalités plus complexe que la courbe de Lorentz simple. Construite avec des données spatialisées, elle prend en compte les différences de poids entre les communes. L'importance d'une unité spatiale est donnée par son poids relatif dans le nombre total des ménages de la commune « i » : $\left({}^{z_i}/_{z}\right)$. Comme la commune représente une part $\left({}^{x_i}/_{x}\right)$ dans les ménages « cadres », la pente de la courbe Lorenz au niveau de la commune « i » est égale au rapport entre $\left({}^{x_i}/_{x}\right)$ et : $\left({}^{z_i}/_{z}\right)$, qu'on appelle indicateur de spécificité « s_r ». Pour construire la courbe spatialisée, on dispose les communes, rangés par ordre croissant de la spécificité sr, dans un plan de coordonnées égal aux valeurs cumulées de $(\left({}^{z_i}/_{z}\right)$; $\left({}^{x_i}/_{x}\right)$;). Autrement dit, les communes sont classées par ordre croissant de la spécialisation en cadres.

Ce type d'indicateur permet de mesurer si la répartition d'une catégorie socio - professionnelle est égalitaire. Plus la courbe s'éloigne de la diagonale principale, plus la répartition de la variable est inégalitaire.

Sur le graphique suivant, nous représentons la courbe Lorentz spatialisée de la distribution des cadres (la courbe rouge). Les points qui apparaissent sur la courbe représentent les 144 communes de la zone d'étude. L'axe horizontal désigne le pourcentage cumulé de la part des communes dans la population totale alors que sur l'axe vertical on a le pourcentage cumulé de la part des communes dans la catégorie socioprofessionnelle. La courbe des ouvriers représente une courbe de Lorentz inversée. On observe que les deux distributions sont inégalitaires. De plus, en effectuant les représentations de toutes les autres catégories socio-professionnelles (non illustrés ici), on a pu constater que la distribution des ouvriers est la plus éloignée de celle des cadres. Les ouvriers sont donc la catégorie la plus ségrégée par rapport aux cadres.

Figure 2 : Courbe de Lorentz spatialisée des cadres

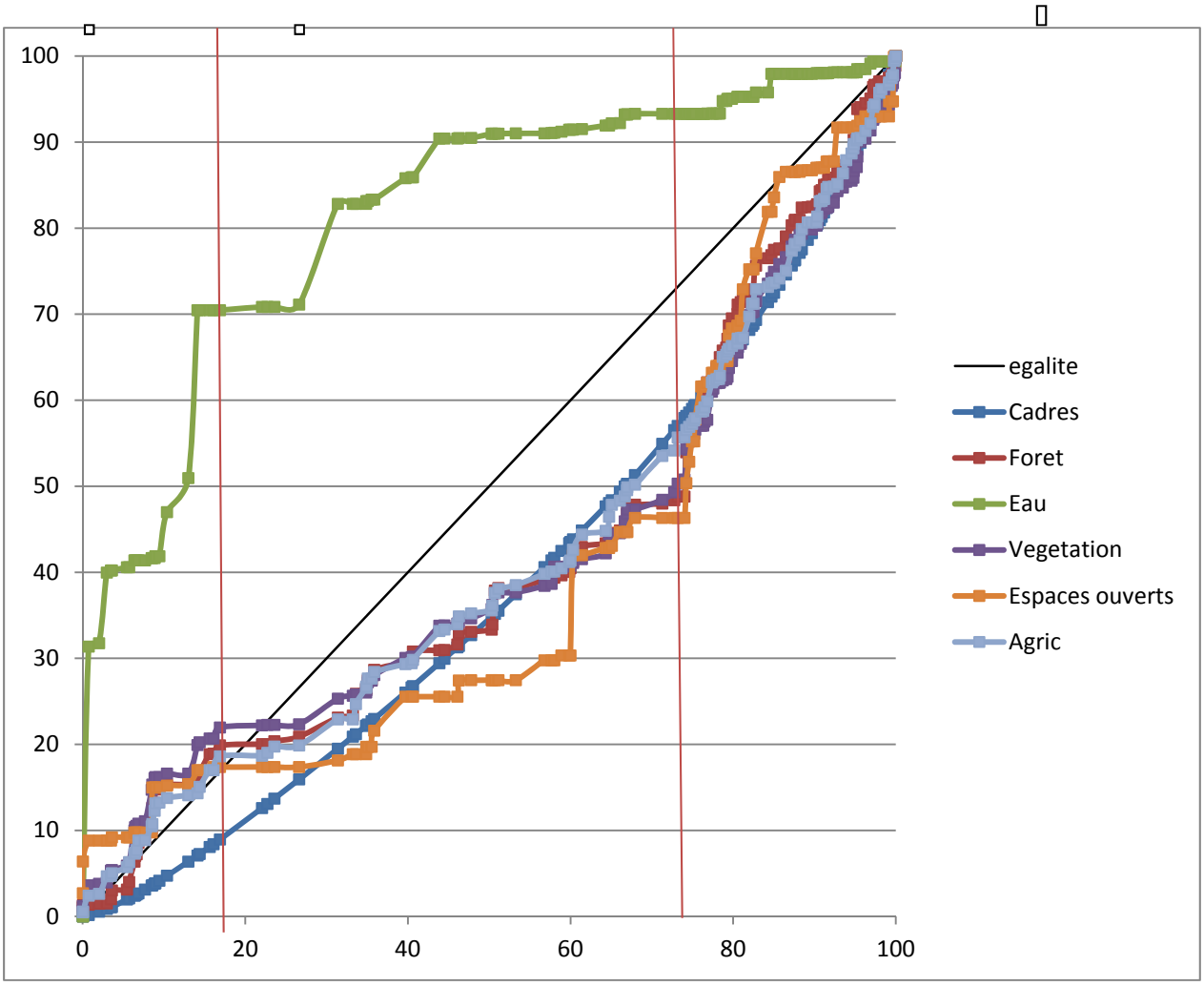


Exemple d'interprétation : Les communes les moins spécialisées en cadres, représentant 20% de la population totale cumulent 10% de la population des cadres et 20% de la population des ouvriers.

2.2. Courbe de Lorentz et aménités naturelles

Pour aller plus loin dans l'analyse, nous proposons une approche intégratrice, visant à capter simultanément de la répartition spatiale des catégories socio-professionnelles et celle des aménités environnementales. On garde le même cadre qu'avant, la courbe de Lorentz spatialisée des cadres, en ordonnant les communes selon leur spécialisation croissante en cadres. Sur le même graphique on représente les poids cumulés des communes dans la quantité totale de chaque ressource. Pour une meilleure illustration des résultats, les trois grandes villes de la zone d'étude ont été omises de l'analyse graphique. Ce choix est pertinent dans la mesure où l'on s'intéresse aux dotations en aménités naturelles et que ces dernières sont peu intensives dans les grands pôles urbains.

Selon la figure obtenue, on confirme que toutes les ressources sont distribuées de manière inégalitaire, et les dotations en eau apparaissent les plus éloignées d'une répartition égalitaire.



Partie 2 Analyse économétrique

Dans cette partie nous allons étudier le rôle des aménités naturelles dans la différenciation sociale des choix résidentiels et d'évaluer ainsi leur contribution aux processus de ségrégation résidentielle. Une première partie méthodologique expliquera les fondements théoriques des modèles économétriques appliqués, suivie par la présentation des données utilisées pour réaliser les estimations pour enchainer sur l'interprétation des résultats obtenus. Pour conclure on estimera si l'étude qu'on a menée est en mesure de répondre à la problématique posée.

1. Méthodologie

Le modèle utilisé pour estimer le rôle des aménités naturelles sur le choix résidentiel dans la zone d'étude est le modèle logit conditionnel, de la famille des modèles logit multinomial (McFadden, 1978, 2001; Train, 2009). Le modèle logit conditionnel est un des modèles empiriques de choix de localisation les plus populaires (Schmidheiny & Brülhart, 2011)

Les modèles de choix discrets tirent leurs fondements théoriques des modèles de fonctions d'utilité aléatoires (McFadden, 1973). On suppose qu'un ménage *n* doit choisir entre les alternatives discrètes mutuellement exclusives *i*=1..*I* afin de maximiser sa fonction d'utilité. Cette fonction est composée d'une partie déterministe (connue comme utilité systématique et souvent définie sous forme linéaire) et d'un terme aléatoire :

$$U_n^i = V_n^i\left(z_n^i, s_n, \beta\right) + \varepsilon_n^i$$

où z_n^i est un vecteur d'attributs de l'alternative *i* comme ils sont perçus par un ménage de catégorie *n*, s_n un vecteur des caractéristiques démographiques ou socio-économiques de l'agent *n*, β un vecteur des paramètres inconnus et ε_n^i est la composante non-observée de l'utilité qui capte la dispersion des choix observés faits par les ménages *n*.

Le modèle est complet si on spécifie la distribution de la variable ε_n^i, pour i=1..I. On peut alors calculer la probabilité du choix de l'alternative i:

$$P_n^i = \text{Prob}\left(U_n^i > U_n^j, \forall j \neq i\right)$$

$$= \text{Prob}\left(\varepsilon_n^j - \varepsilon_n^i < V_n^i - V_n^j, \forall j \neq i\right)$$

En fonction de la forme de la distribution des termes aléatoires on peut distinguer plusieurs modèles de choix discret. Pour notre étude de cas, nous avons choisi le modèle logit conditionnel où les termes aléatoires suivent une distribution de Gumbel. McFadden (1973) montre que les probabilités résultantes suivent un modèle logit :

$$P_n^i = \frac{\exp\left(V_n^i\right)}{\sum_{j=1}^{I} \exp\left(V_n^j\right)}$$

Une hypothèse implicite du modèle logit multinomial est que les termes aléatoires sont indépendants. L'avantage de cette hypothèse est la simplicité du modèle qui en résulte, par rapport à d'autres modèles, tel le probit, où il est nécessaire d'évaluer des intégrales multiples. Deuxièmement, la fonction de vraisemblance utilisée pour l'estimation des paramètres est concave, ce qui garanti la convergence de la maximisation de la vraisemblance (Börsch-Supan & Pitkin, 1988; Train, 2009). Egalement, cette méthodologie est très appréciée, par le fait qu'elle explique le comportement individuel en appelant aux principes de base de la microéconomie, la maximisation de l'utilité. Ces avantages techniques sont souvent accompagnés par un succès relatif de prédictions des modèles et par une agilité à traiter des quantités importantes de données (Frenkel et al., 2013).

Ces avantages ont en contrepartie une hypothèse forte : le choix entre deux alternatives est indépendant des attributs et de la disponibilité des autres alternatives. Cette propriété est intitulée « indépendance des alternative non pertinentes » (IIA - Independence of Irrelevant Alternatives). Ainsi, la probabilité de choisir une alternative i par rapport à une alternative j, ne dépend pas du nombre d'alternatives ou des attributs des alternatives autres que i et j. La condition d'indépendance des alternatives non pertinentes est facile à vérifier, parce que les rapports P_n^i / P_n^j dépendent seulement de V_n^i et V_n^j et non pas des utilités des autres alternatives.

Pour un modèle donné, les informations sur les choix réellement réalisés, en fonction des attributs z_n^i et s_n permettent d'estimer les valeurs des paramètres inconnus β et de réaliser des tests statistiques sur la validité des formes fonctionnelles. Les paramètres sont habituellement estimés en maximisant le logarithme de la fonction de vraisemblance (le log-likelyhood) :

$$L(\beta) = \sum_{n=1}^{N} \sum_{i=1}^{I} d_n^i \log P_n^i \left(\beta\right)$$

où N est la taille de l'échantillon. Les variables d_n^i sont des variables binaires de choix (dummy) qui sont égales à 1 si l'agent n choisit l'alternative i et égales à 0 sinon.

Dans l'estimation des paramètres de la fonction d'utilité, le plus souvent le vecteur des paramètres contient des constantes spécifiques des alternatives. Puisque l'utilité est ordinale, une de ces constantes est normalisée (habituellement à zéro). Ainsi l'alternative correspondante sert d'alternative de base pour des comparaisons. Ces constantes peuvent être interprétées comme l'utilité moyenne des caractéristiques non-observées de chaque alternative relativement à l'alternative de base.

2. Les données

Le jeu de données qu'on emploie pour l'estimation du modèle est constitué des caractéristiques spécifiques aux ménages (la table des ménages migrant) et des caractéristiques spécifiques aux localisations résidentielles constituant l'ensemble des choix.

A. Les ménages migrants

A. 1. Méthodologie

La base de données source est le fichier MIGCOM (Fichier Migrations résidentielles) de l'Insee tiré du Recensement de la population de 2008. Les données sont collectées entre 2004 et 2008. Chaque enregistrement du fichier correspond à un individu (d'au moins 5 ans) décrit selon sa résidence actuelle et sa résidence antérieure (5 ans auparavant), ses principales caractéristiques sociodémographiques, ainsi que celles du ménage auquel il appartient.

Après plusieurs manipulations de la base originale, nous avons extrait uniquement les données correspondant aux migrations internes à la zone d'étude. On s'intéresse qu'aux individus ayant changé de résidence à l'intérieur de la zone pendant la période 1999-2008. Notre population d'étude est constituée que par des migrants internes, car on fait l'hypothèse que les migrants externes, par méconnaissance du territoire se localisent la lors de leur arrivée dans les grands pôles urbains pour migrer ensuite.

Par rapport aux données d'origine, nous ajoutons quelques variables a fin de capter les faits : d'avoir des enfants, et d'être migrant, ou de migrer à l'intérieur de la même commune. Les différentes classes correspondant aux catégories socio professionnelles ont été dichotomisées. De plus, nous avons regroupé les indépendants et les agriculteurs, ces derniers ne pesant que 0.32% des ménages.

A.2. Statistiques descriptives de la table des ménages

Comme on peut le voir dans le tableau, la population étudiée est composée de 264 459 ménages, française à 94,4%, d'âge moyen 39 ans et dont 44% ont des enfants. La catégorie socioprofessionnelles la plus représentée correspond aux professions intermédiaires avec (20.9%), la part des retraités est la plus faible (11.7%), alors que dans la population des ménages stables[2] cette catégorie était la plus importante.

Les caractéristiques des ménages utilisés par la suite dans le modèle économétrique sont : l'age (« agerevq », les catégories socioprofessionnelles, le fait d'avoir des enfants, la taille de la famille et le fait de ne pas changer de commune de résidence lors de la migration.

Tableau 1 : Statistiques descriptives de la table des ménages

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
commune	102744	264459.745	32329.22	31287.24	13001	84026
arm	0	0				
dcran	102744	264459.745	32116.48	31115.53	13001	84026
achlr	0	0				
agemen8	0	0				
agerevq	102744	264459.745	39.26819	14.86385	10	105
csm_agri_i~p	102744	264459.745	.0613254	.2399267	0	1
csm_cadre	102744	264459.745	.1254689	.3312505	0	1
csm_prof_int	102744	264459.745	.2090672	.4066433	0	1
csm_employe	102744	264459.745	.2043874	.403254	0	1
csm_ouvriers	102744	264459.745	.1891521	.3916302	0	1
csm_retraite	102744	264459.745	.1170459	.3214755	0	1
csm_ss_act	102744	264459.745	.0935531	.2912065	0	1
nationalite	102744	264459.745	.9443762	.2291943	0	1
kids	102744	264459.745	.445129	.496981	0	1
migrant	102744	264459.745	1	0	1	1

Lorsqu'on compare la population des migrants internes à la population des non –migrants, on constante que les migrants sont plus jeunes (39 ans par rapport à 57 ans en moyennes) et que les retraités sont faiblement représentés (11% par rapport à 45%).

[2] Table représentant les ménages (migrants et non migrants) n'ayant pas quitté la zone entre 1999-2008.

B. Les caractéristiques des alternatives de choix

B. 1. Les statistiques descriptives

Dans le modèle économétrique, les ménages ont la possibilité de choisir une localisation résidentielle parmi les 144 communes de la zone d'étude.

Tableau 2 : Statistiques descriptives et sources de la table des communes :

Variable	Mean	Std. Dev.	Min	Max		Source
Revenu médian	144	15489.26	3205.092	0	25093	*Insee RFL 2001*
Altitude de la mairie	144	189.9931	141.0281	0	670	IGN BD ALTI
Distance à la plage	144	17021.17	11974.89	294.5114	48321.43	IGN BD TOPO
Distance aux calanques	144	22303.84	12069.15	831.658	51895.77	IGN BD TOPO
Distance aux lacs	144	10022.12	6861.465	351.647	30050.9	IGN BD TOPO
Distance au centre- ville	144	19814.16	9689.356	0	47513.06	IGN BD CARTO
Distance à la cote	144	17817.3	12687.34	51.4085	48239.48	IGN BD TOPO
Distance à la montagne	144	8748.488	5212.839	1257.371	28517.08	IGN BD TOPO
Distance aux rivières	144	4200.062	4115.785	6.574556	18741.1	IGN BD TOPO
% territoires artificialisés	144	18.29764	18.25195	0	74.63	CLC 2000
% territoires agricoles	144	25.09292	16.59359	0	73.44	CLC 2000
% végétation arbustive	144	20.59965	16.57317	0	89.8	CLC 2000
% espaces ouverts	144	1.979583	3.963919	0	25.77	CLC 2000
% espaces en eau	144	1.092431	3.875146	0	32.7	CLC 2000
% forets	144	32.93701	21.286	0	81.76	CLC 2000

CLC : Corine Land Cover, RFL : Revenus fiscalisés des ménages

B.2. Les variables

a. Les variables classiques

Goffette – Nagot et Schaeffer (2013) identifient essentiellement deux sources de ségrégation résidentielle au sein des aires urbaines françaises : le niveau de revenu et la distance au centre.

Le tri entre les communes selon le revenu de leurs résidents, à la Tiebout-Benabou, s'explique par les externalités de voisinage et l'hétérogénéité des offres de biens publics locaux.

Le tri selon la distance au centre, à la Alonso, s'explique par l'arbitrage entre un emploi plus accessible (au centre) et un logement moins cher (en périphérie).

b. Les aménités naturelles

L'altitude est une variable topographique souvent introduite dans les modèles économétriques en tant qu'aménité naturelle recherchée par les ménages (Wu and Cho 2003, Cremer-Schulte 2014). Cette variable à finalement été omise de l'estimation du à une forte corrélation avec la distance la plage (0.74).

La zone d'étude est caractérisée par des aménités spécifiques qu'on a considéré nécessaire d'identifier et pour pouvoir leur associer un indicateur d'accessibilité. Il s'agit des montagnes (Montagne Sainte Victoire…), la Côte Méditerranéenne, les plages, les calanques, les rivières et les lacs. A l'aide du logiciel ArcGiS, nous avons localisé précisément ces aménités et calculé ensuite des distances entre chaque commune et l'aménité spécifique la plus proche, par catégorie. Les variables : distance à la cote, distance à la plage et distance aux calanques sont fortement corrélés. Raison pour laquelle, nous avons retiré la variable le plus générale, la distance à la cote et estimé des modèles séparés pour les deux autres variables.

Les catégories d'occupation du sol sont les variables d'aménités les plus utilisées dans la littérature sur les migrations résidentielles (Chi 2013, Mc Granahan 2008). Elles ne jouent pas systématiquement de manière positive sur l'utilité des individus.

On note que les données pour les variables d'estimation correspondent aux périodes 1999 ou 2000, c'est au début de la période de migration (1999-2008). Ce choix minimise la probabilité (et le biais) de mesurer une variable explicative après que la migration résidentielle à eu lieu.

C. La table d’estimation

La table finale d’estimation est construite comme une combinaison de la table des ménages avec celle des alternatives. Chaque ménage est multiplié 144 fois, on s’assure ainsi qu’il a le choix parmi l’ensemble complet des communes. La variable expliquée choix est ajoutée, elle prend la valeur 1 si l’alternative correspond au choix du ménage et 0 sinon.

3. Estimations et interprétation des résultats

Nous avons estimée quatre types de modèles : deux spécifications logit conditionnel avec effets fixes pour les caractéristiques des alternatives et deux autres modèles logit conditionnel sans effets fixes. Comme présentée dans la partie « Données », pour chaque type de modèle nous avons testé séparément l'effet des plages et les calanques, aménités naturelles remarquables dans la zone d'étude mais trop corrélées pour etre intégrées simultanément dans un même modèle.

3.1 Estimations et qualité des résultats

A. Résultats

Le tableau suivant présenté les coefficients estimés par les modèles ainsi que leur niveau de significativité.

Tableau 3 : Estimations pour Logit Conditionnel avec et sans effets fixes (échantillon 10% des ménages)

Variable expliquée : choix de la localisation résidentielle	**EFFETS FIXES**				**SANS EFFETS FIXES**			
	Modèle 1		**Modèle 2**		**Modèle 3**		**Modèle 4**	
Variables explicatives :	Estimate		Estimate		Estimate		Estimate	
log(revenu_median)	-		-		-2.004e-01		-2.961e-01	
log(revenu_median) *csm_prof_int	*Ref*		*Ref*		-		-	
log(revenu_median) * csm_cadre	2.216e-01		1.943e-01		4.571e-01	.	4.981e-01	*
log(revenu_median) * csm_indep_agri	-6.741e-01	*	-6.988e-01	*	-3.679e-01		-3.722e-01	
log(revenu_median) * csm_employe	-1.881e+00	***	-1.886e+00	***	-1.543e+00	***	-1.498e+00	***
log(revenu_median) * csm_ouvrier	-2.116e+00	***	-2.136e+00	***	-1.854e+00	***	-1.864e+00	***
log(revenu_median) * csm_retraite	-2.793e+00	***	-2.799e+00	***	-2.243e+00	***	-2.247e+00	***
log(revenu_median) * csm_ss_act	-2.566e+00	***	-2.565e+00	***	-2.526e+00	***	-2.322e+00	***
log(revenu_median) * enfants	-5.375e-01	***	-5.132e-01	**	-1.172e-01		-1.161e-01	
log(revenu_median) * age	1.469e-01	.	1.609e-01	*	9.216e-02		1.241e-01	.
log(revenu_median) * age2	-7.486e-02	*	-7.229e-02	*	8.986e-03		1.217e-02	
log(revenu_median) * adultes	1.107e-01		9.299e-02		4.307e-01	**	3.527e-01	*
dist	-		-		1.643e-03		1.800e-03	
csm_prof_int	*Ref*		*Ref*		-		-	
dist * csm_cadre	-2.566e-02	***	-2.384e-02	***	-2.391e-02	***	-2.301e-02	***

dist * csm_indep_agri	-6.824e-03		-5.024e-03		-4.928e-03		-5.016e-03	
dist * csm_ouvrier	-1.861e-03		-6.251e-04		-3.657e-03		-4.613e-03	
dist * csm_employe	-8.426e-03	*	-6.072e-03	.	-7.245e-03	*	-5.991e-03	.
dist * csm_retraite	-6.637e-03		-2.553e-03		-5.248e-03		-4.905e-03	
dist * csm_ss_act	-4.226e-02	***	-3.750e-02	***	-4.070e-02	***	-3.621e-02	***
dist * adultes	2.055e-03		7.905e-04		6.142e-03	*	3.979e-03	
dist * enfants	1.723e-03		4.116e-03		5.272e-03	*	5.847e-03	*
dist * age	2.987e-03	*	3.382e-03	**	1.133e-03		1.898e-03	
dist * age2	-2.632e-03	***	-2.355e-03	***	-9.643e-04	.	-8.462e-04	
dist2	-		-		-3.269e-04		5.870e-05	
csm_prof_int	*Ref*		*Ref*		-		-	
dist2 * csm_cadre	-1.348e-03	***	-1.104e-03	***	-5.889e-04	*	-4.854e-04	.
dist2 * csm_indep_agri	-1.756e-03	***	-1.578e-03	***	-6.204e-04	.	-6.614e-04	.
dist2 * csm_ouvrier	-1.202e-03	***	-1.088e-03	***	-9.822e-05		-1.455e-04	
dist2 * csm_employe	-8.366e-04	**	-7.224e-04	**	2.516e-04		2.094e-04	
dist2 * csm_retraite	-9.941e-04	*	-8.393e-04	.	3.482e-04		2.482e-04	
dist2 * csm_ss_act	1.397e-03	***	1.763e-03	***	1.915e-03	***	2.125e-03	***
dist2 * adultes	-6.787e-04	***	-6.878e-04	***	-3.323e-04	.	-3.715e-04	*
dist2 * enfants	-1.433e-03	***	-1.336e-03	***	-8.985e-04	***	-8.464e-04	***
dist2 * age	-1.873e-04	*	-2.102e-04	*	-2.226e-04	**	-2.262e-04	**
dist2 * age2	-9.970e-05	*	-7.604e-05	*	-6.641e-05	.	-5.630e-05	
same_municp	-		-		4.498e+00	***	4.502e+00	***
csm_prof_int	*Ref*		*Ref*		-		-	
same_municp * csm_cadre	2.366e+00	***	2.363e+00	***	-3.373e-01	***	-3.330e-01	***
same_municp * csm_indep_agri	3.650e+00	***	3.648e+00	***	1.053e+00	***	1.047e+00	***
same_municp * csm_ouvrier	3.313e+00	***	3.311e+00	***	8.300e-01	***	8.250e-01	***
same_municp * csm_employe	3.619e+00	***	3.620e+00	***	7.072e-01	***	7.124e-01	***
same_municp * csm_retraite	4.132e+00	***	4.151e+00	***	5.528e-01	***	5.592e-01	***
same_municp * csm_ss_act	3.049e+00	***	3.052e+00	***	1.596e-01	*	1.718e-01	*
same_municp * enfants	1.631e+00	***	1.640e+00	***	4.424e-01	***	4.408e-01	***
same_municp * age	-1.913e-01	***	-1.911e-01	***	1.425e-01	***	1.418e-01	***
same_municp * age2	1.510e-01	***	1.500e-01	***	-3.811e-02	***	-3.880e-02	***
same_municp * adultes	7.160e-01	***	7.086e-01	***	-4.446e-01	***	-4.474e-01	***
Aménités naturelles :								
montagne_dist	-		-		-1.664e-03		-2.179e-04	
csm_prof_int	*Ref*		*Ref*		-		-	
montagne_dist * csm_cadre	2.145e-02	**	2.086e-02	**	1.297e-02	.	1.298e-02	.
montagne_dist * csm_indep_agri	7.087e-05		-1.975e-04		-1.138e-02		-1.107e-02	
montagne_dist * csm_employe	-2.526e-03		-1.970e-03		-1.561e-02	*	-1.418e-02	.
montagne_dist * csm_ouvrier	1.302e-02	.	1.267e-02	.	1.452e-03		6.029e-04	
montagne_dist * csm_retraite	1.008e-02		1.121e-02		-1.069e-02		-1.112e-02	
montagne_dist * csm_ss_act	1.109e-02		1.255e-02		6.548e-03		9.985e-03	
montagne_dist * enfants	2.548e-02	***	2.667e-02	***	1.214e-02	*	1.292e-02	*
montagne_dist * age	-3.837e-03	.	-3.384e-03		2.633e-04		1.104e-03	
montagne_dist * age2	5.960e-03	***	6.192e-03	***	4.081e-03	***	4.363e-03	***
montagne_dist * adultes	1.819e-02	***	1.747e-02	***	1.131e-02	*	9.621e-03	.

forets	-		-		4.944e-03	*	5.670e-03	**
csm_prof_int	*Ref*		*Ref*		-		-	
forets * csm_cadre	6.967e-03	*	1.145e-02	***	1.541e-03		3.140e-03	
forets * csm_indep_agri	1.809e-02	***	2.093e-02	***	1.524e-02	***	1.457e-02	***
forets * csm_employe	1.253e-02	***	1.364e-02	***	9.443e-03	***	7.591e-03	***
forets * csm_ouvrier	1.004e-02	***	1.144e-02	***	7.520e-03	**	7.074e-03	**
forets * csm_retraite	2.865e-02	***	2.977e-02	***	2.432e-02	***	2.216e-02	***
forets * csm_ss_act	1.692e-02	***	2.158e-02	***	1.379e-02	**	1.530e-02	***
forets * enfants	6.064e-03	**	7.010e-03	***	1.904e-03		2.996e-03	.
forets * age	-2.696e-03	**	-3.655e-03	***	-1.629e-03	.	-2.515e-03	***
forets * age2	-1.063e-03	*	-7.415e-04	*	-1.508e-03	***	-1.420e-03	***
forets * adultes	5.293e-03	**	5.927e-03	***	2.336e-03		2.917e-03	.
lac_dist	-		-		2.924e-03		2.654e-03	
csm_prof_int	*Ref*		*Ref*		-		-	
lac_dist * csm_cadre	1.406e-03		3.475e-04		9.469e-03	.	8.788e-03	.
lac_dist * csm_indep_agri	-1.492e-02	*	-1.510e-02	*	-3.966e-03		-4.078e-03	
lac_dist * csm_employe	-1.402e-02	**	-1.317e-02	**	-4.154e-03		-2.686e-03	
lac_dist * csm_ouvrier	-1.759e-02	***	-1.734e-02	***	-6.813e-03		-7.083e-03	
lac_dist * csm_retraite	-1.007e-02		-9.217e-03		5.529e-04		9.185e-04	
lac_dist * csm_ss_act	-3.555e-03		-1.260e-03		8.999e-04		4.074e-03	
lac_dist * enfants	-1.823e-02	***	-1.753e-02	***	-1.338e-02	***	-1.351e-02	***
lac_dist * age	2.885e-03	*	3.364e-03	*	2.901e-03	.	3.255e-03	*
lac_dist * age2	-1.231e-03	.	-1.241e-03	.	-5.489e-04		-4.835e-04	
lac_dist * adultes	-4.818e-03		-5.672e-03	.	2.133e-03		1.112e-03	
riviere_dist	-		-		-4.047e-04		-8.553e-03	
riviere_dist*csm_prof_int	*Ref*		*Ref*		-		-	
riviere_dist * csm_cadre	-1.141e-02		-1.192e-02		-1.099e-02		-1.138e-02	
csm_indep_agri	6.473e-03		4.128e-03		1.122e-02		1.103e-02	
riviere_dist * csm_employe	-1.435e-02		-2.129e-02	*	-1.375e-02		-1.979e-02	*
riviere_dist * csm_ouvrier	-6.856e-03		-9.415e-03		2.840e-03		4.432e-03	
riviere_dist * csm_retraite	-2.574e-03		-1.638e-02		-3.683e-03		-8.471e-03	
riviere_dist * csm_ss_act	6.671e-02	***	5.686e-02	***	6.486e-02	***	5.058e-02	***
riviere_dist * enfants	3.982e-04		-5.817e-03		-5.289e-03		-6.215e-03	
riviere_dist * age	-5.923e-03	.	-8.130e-03	**	-3.487e-03		-6.452e-03	*
riviere_dist * age2	-2.653e-03	.	-3.032e-03	*	-4.060e-03	**	-4.292e-03	**
riviere_dist * adultes	2.045e-02	**	2.549e-02	***	1.547e-02	*	2.253e-02	***
plage_dist	-		-		-8.228e-04			
csm_prof_int	*Ref*				-		-	
plage_dist * csm_cadre	1.016e-02	*			4.066e-03			
plage_dist * csm_indep_agri	5.053e-03				-2.579e-03			
plage_dist * csm_employe	-2.828e-03				-8.564e-03	.		
plage_dist * csm_ouvrier	9.249e-04				1.196e-04			
plage_dist * csm_retraite	-8.985e-03				-1.089e-02			
plage_dist * csm_ss_act	2.801e-03				-2.539e-03			
plage_dist * enfants	-2.555e-03				1.322e-03			
plage_dist * age	-3.585e-03	*			-3.616e-03	*		

plage_dist * age2	4.981e-04				1.391e-04			
plage_dist * adultes	5.356e-03	.			5.596e-03	.		
calanques_dist	-		-				-1.507e-02	***
csm_prof_int			*Ref*		-		-	
calanques_dist * csm_cadre			-3.734e-03				-2.014e-03	
calanques_dist * csm_indep_agri			-4.943e-03				-1.181e-03	
calanques_dist * csm_employe			-1.020e-02	**			-7.508e-03	*
calanques_dist * csm_ouvrier			-4.776e-03				2.955e-03	
calanques_dist * csm_retraite			-2.073e-02	***			-6.584e-03	
calanques_dist * csm_ss_act			-1.650e-02	**			-1.471e-02	**
calanques_dist * enfants			-9.755e-03	***			-1.946e-03	
calanques_dist * age			-2.677e-03	*			-3.906e-03	***
calanques_dist * age2			-9.976e-04	.			-6.614e-04	
calanques_dist * adultes			6.733e-03	**			1.026e-02	***
Vegetation.arbustive	-		-		2.766e-03		1.609e-03	
csm_prof_int	*Ref*		*Ref*		-		-	
Vegetation.arbustive * csm_cadre	1.546e-02	***	1.604e-02	***	1.105e-02	***	1.135e-02	***
Vegetation.arbustive * csm_indep_agri	1.236e-02	***	1.240e-02	***	8.065e-03	*	7.722e-03	*
Vegetation.arbustive * csm_employe	1.517e-02	***	1.454e-02	***	1.314e-02	***	1.186e-02	***
Vegetation.arbustive * csm_ouvrier	9.761e-03	***	9.763e-03	***	7.686e-03	**	7.804e-03	**
Vegetation.arbustive * csm_retraite	2.935e-02	***	2.760e-02	***	2.461e-02	***	2.300e-02	***
Vegetation.arbustive * csm_ss_act	8.548e-03	.	7.372e-03		7.333e-03		5.085e-03	
Vegetation.arbustive * enfants	-1.666e-03		-2.599e-03		-6.220e-03	**	-6.425e-03	**
Vegetation.arbustive * age	-3.775e-03	***	-4.268e-03	***	-2.681e-03	**	-3.381e-03	***
Vegetation.arbustive * age2	-7.135e-05		-1.864e-04		-7.633e-04	.	-8.892e-04	*
Vegetation.arbustive * adultes	8.503e-04		1.668e-03		-1.581e-03		4.999e-05	
espaces.ouverts	-		-		7.061e-02	***	6.521e-02	***
csm_prof_int	*Ref*		*Ref*		-		-	
espaces.ouverts * csm_cadre	1.002e-01	***	9.512e-02	***	1.435e-02		1.370e-02	
espaces.ouverts * csm_indep_agri	1.188e-01	***	1.132e-01	***	3.520e-02	**	3.527e-02	**
espaces.ouverts * csm_employe	8.267e-02	***	7.453e-02	***	3.564e-03		2.916e-04	
espaces.ouverts * csm_ouvrier	8.578e-02	***	8.057e-02	***	1.778e-02	*	1.880e-02	*
espaces.ouverts * csm_retraite	6.046e-02	***	4.474e-02	*	-3.399e-02	*	-3.636e-02	*
espaces.ouverts * csm_ss_act	1.069e-01	***	9.166e-02	***	4.871e-02	***	3.693e-02	**
espaces.ouverts * enfants	5.682e-02	***	5.089e-02	***	1.081e-02	.	1.087e-02	.
espaces.ouverts * age	-3.386e-03		-5.120e-03		4.797e-03	.	2.584e-03	
espaces.ouverts * age2	1.913e-03		1.242e-03		-2.349e-03	.	-2.770e-03	*
espaces.ouverts * adultes	4.167e-02	***	4.579e-02	***	3.536e-04		5.360e-03	
espaces.en.eau	-		-		3.674e-03		-2.309e-03	
csm_prof_int	*Ref*		*Ref*		-		-	
espaces.en.eau * csm_cadre	6.055e-03		2.445e-04		1.969e-03		-1.716e-04	
espaces.en.eau * csm_indep_agri	-2.408e-03		-6.544e-03		3.070e-04		7.280e-04	
espaces.en.eau * csm_employe	-9.607e-03		-1.240e-02		-1.271e-02		-1.207e-02	
espaces.en.eau * csm_ouvrier	3.971e-03		1.770e-03		9.564e-03		9.966e-03	
espaces.en.eau * csm_retraite	-3.039e-02	*	-3.586e-02	*	-1.790e-02		-1.750e-02	
espaces.en.eau * csm_ss_act	8.678e-03		1.723e-03		3.158e-03		-3.784e-04	

espaces.en.eau * enfants	-2.761e-02	***	-3.027e-02	***	-2.216e-02	***	-2.210e-02	***
espaces.en.eau * age	-1.439e-03		-1.035e-03		-2.851e-03		-2.601e-03	
espaces.en.eau * age2	-1.892e-04		-7.403e-04		-4.953e-04		-7.954e-04	
espaces.en.eau * adultes	-1.113e-02	*	-1.068e-02	*	-8.601e-03		-6.818e-03	
terres.agricoles	-		-		1.932e-03		4.070e-03	.
csm_prof_int	*Ref*		*Ref*		-		-	
terres.agricoles * csm_cadre	1.698e-03		5.123e-03	*	2.004e-03		3.489e-03	
terres.agricoles * csm_indep_agri	6.176e-04		2.860e-03		3.699e-03		3.359e-03	
terres.agricoles * csm_employe	-2.541e-03		-1.356e-03		1.975e-03		9.892e-04	
terres.agricoles * csm_ouvrier	-4.778e-03	.	-3.529e-03		-1.150e-03		-1.652e-03	
terres.agricoles * csm_retraite	5.300e-03		6.925e-03	.	9.696e-03	*	8.598e-03	*
terres.agricoles * csm_ss_act	2.099e-02	***	2.441e-02	***	2.380e-02	***	2.414e-02	***
terres.agricoles * enfants	-5.306e-03	**	-4.177e-03	*	-6.836e-03	***	-6.035e-03	***
terres.agricoles * age	-1.407e-03	.	-1.934e-03	**	-1.026e-03		-1.502e-03	*
terres.agricoles * age2	1.127e-04		4.100e-04		-2.528e-04		-1.327e-04	
terres.agricoles * adultes	-1.182e-03		-1.021e-03		-3.711e-04		-4.446e-04	
log(population)	-		-		1.076e+00	***	1.024e+00	***
fix_13001 : Aix en Provence	*Ref*		*Ref*					
fix_13002 : Allauch	-3.433e+00	***	-3.331e+00	***				
…								
fix_13055 : Marseille	1.858e+00	***	1.843e+00	***				
…								
fix_83137 : Toulon	1.425e-01	.	1.896e-01	*				
…								
fix_84026 : Cadenet	-3.054e+00	***	-2.885e+00	***				

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Echantillon des ménages	1 448 634 ménages			
Alternatives	144			
Ensemble de choix	141			
Qualité de l'estimation				
McFaddenR² (ajusté)	0.526	0.526	0.544	0.544
Estrella(ajusté)	0.888	0.888	0.899	0.9
LogLikelihood	2.380045e+04	2.379173e+04	2.303826e+04	2.301596e+04
Dissimilarité	0.03558736	0.03555795	0.1030259	0.1013607

B. Qualité globale des estimations

Deux types de spécifications ont été utilisés : une avec effets fixes (modèles 1 et 2) et une sans effets fixes (modèles 3 et 4). Pour chaque type de spécification, on isole l’impact des plages et des calanques dans deux modèles différents. Il s’agit de retracer des comportements particuliers de choix entrainés par deux aménités naturelles spécifiques à la zone d’étude,

toutes choses égales par ailleurs. Le fort taux de corrélation entre ces deux variables (0.95) ne permet pas leur estimation simultanée dans un même modèle de choix (outre l'estimation séparée, la construction des indicateurs d'aménités plus complexes à l'aide d'une analyse en composantes principales est une solution qui pourra être envisagée ultérieurement).

Les modèles à effets fixes ont une capacité prédictive nettement meilleure en ce qui concerne les flux totaux de ménage. Ils arrivent à correctement estimer les effectifs communaux de ménages dans 96.5% des cas, alors que sans la prise en compte des effets fixes, ce taux diminue à 89%. Néanmoins, les indicateurs de la qualité globale de l'estimation attestent que les deux types de modèle sont pertinents : le McFadden-R^2 (ajusté) s'établit à 0.52 et 0.54[3] et l'Estrella (ajusté) à 0.88 et 0.89. Les modèles sans effets fixes présentent même un ajustement global meilleur, ce qui signifie qu'en considérant la distribution spatiale de l'ensemble des caractéristiques des ménages (et non plus seulement de leurs effectifs totaux), leurs prédictions sont –très légèrement– meilleures.

C. Echantillon des ménages et ensemble de choix

Le jeu de données complet, comprenant 100% des migrants et toutes les alternatives de choix, représente un peu plus de 14 millions de lignes et son estimation est techniquement impossible avec les ressources informatiques disponibles. Pour dépasser cette limite et obtenir des résultats pertinents, les modèles sont estimés sur un échantillon représentatif de 10% les ménages migrants, tiré sans remise[4], soit 1 448 634 ménages. Le nombre total des alternatives de localisation résidentielle correspond aux communes de la zone d'étude et s'élève à 144. Parmi ces alternatives, nous définissons l'ensemble de choix des ménages comme les communes effectivement choisies par les 10% des ménages de l'échantillon. Cet ensemble comprend 141 communes.

On note que des tests de sensibilité des modèles ont été réalisés par des estimations sur un échantillon de 5% des ménages (représentant 693 495 ménages et 135 alternatives de choix). Malgré la taille différente du jeu de données, le sens des coefficients des variables significatives restent stables, attestant d'une bonne robustesse des modèles.

[3] Dans la littérature cette valeur dépasse rarement 0.3

[4] Nous avons aussi testé les modèles sur des échantillons tirés avec remise et il en résultait une précision de l'ajustement plus faible.

D. Variables

Les modèles emploient un mix de variables explicatives classiques, habituelles dans littérature des choix résidentiels et inspirée des modèles théoriques d'économie urbaine (Goffette-Nagot et Schaeffer, 2013), auxquelles on rajoute des variables d'aménités plus spécifiques.

Pour les spécifications à effets fixes (modèles 1 et 2), nous avons retenu 14 grands groupes de variables, dont 10 correspondant à des aménités naturelles. Il s'agit uniquement de variables d'interaction entre les caractéristiques des ménages et celles des communes. La spécification du modèle est complétée par des effets fixes spécifiques à chaque alternative. Ces effets captent toutes les particularités de chaque commune qui n'ont pas été contrôlées par les autres variables. Le nombre des effets fixes est nécessairement égal au nombre des choix possibles, soit 141 dans le cas présent.

Lorsque les effets fixes ne sont pas introduits dans l'estimation (modèle 3 et 4), on les remplace par des « variables pures » (sans interaction) censées capter les spécificités de chaque localisation. Ainsi, dans ces deux modèles, on retient les mêmes 14 grands groupes de variables (variables seules et en interaction). En suivant de Plama (2007), on ajoute également la variable population de la commune pour prendre en compte sa capacité d'accueil des nouveaux résidents.

E. Mesure des aménités naturelles

Les estimations utilisent des mesures simples d'accès aux aménités naturelles, à savoir les distances et les proportions. La distance euclidienne entre la mairie de la commune de résidence et l'aménité est une mesure simple et intuitive qui permet d'obtenir une indication sur le sens et l'ampleur de l'effet de l'accessibilité géographique à une aménité sur les probabilités de choix résidentiels. Leur intérêt est de capter l'impact de la plage, des calanques, des montagnes, des lacs, des rivières et du centre-ville (qui sont aussi les principaux centres d'emplois). Les mesures de proportion représentent la part de la commune correspondant aux différentes occupations du sol (les catégories selon Corine Land Cover : les forêts, la végétation arbustive, les espaces ouverts, les zones humides et les espace en eau, les

terres agricoles). Elles permettent de capter l'effet de la proximité immédiate de l'aménité sur les probabilités de choix en fonction de l'importance quantitative de cette aménité.

F. Pourquoi deux types de modèles ?

Dans leur forme actuelle, les modèles avec et sans effets fixes offrent des résultats complémentaires. Les deux spécifications permettent d'identifier une différenciation sociale des choix résidentiels. Les modèles sans effets fixes sont généralement préférables pour les modèles de choix, car ils dégagent l'effet absolu des variables. On pourra, par exemple, identifier pour quelles catégories socio-professionnelles les aménités sont attractives dans l'absolu. Dans les modèles à effets fixes nous avons une vue relative sur ces différences, l'interprétation se faisant toujours par rapport à une catégorie de référence dont on ne connait pas le comportement. Mais l'inconvénient des modèles sans effets fixes est que leur capacité prédictive risque d'être plus limitée, car seuls les effets des variables observées sont captés, tandis que les effets fixes permettent de capter également les effets des variables communales inobservées. Etant donné que, dans notre cas, les indicateurs d'ajustement global signalent des capacités prédictives semblables pour les deux types de modèles, nous mobilisons les deux lors de l'interprétation des résultats. Néanmoins, nous privilégierons ceux à effets fixes car nous les jugeons plus robustes. Les spécifications sans effets fixes offrent des informations complémentaires, mais elles sont à interpréter avec précaution, car certaines limites de la spécification (colinéarités, variables omises) pourraient biaiser le sens et l'ampleur des effets.

G. Limites économétriques

Plusieurs facteurs pourraient expliquer les plus faibles capacités des modèles 3 et 4 à prédire les flux d'ensemble de ménages.

Une limite importante provient des variables manquantes. Le nombre des variables observables est probablement insuffisant pour pallier à l'absence d'effets fixes. De plus, certaines variables explicatives sont caractérisées par une forte multicolinéarité, signifiant qu'elles sont approximativement déterminées par une combinaison linéaire des autres

variables indépendantes du modèle. La multicolinéarité modérée est habituelle car elle est entrainée par toute corrélation entre variables indépendantes. Mais lorsque la multicolinéarité dépasse un certain seuil, l'erreur standard tend à être très importante et cela peut engendrer une imprécision des coefficients. Dans le cas présent, les erreurs standard des modèles sans effets fixes restent bases, ce qui conforte l'idée d'un faible biais provoqué par colinéarité. Néanmoins, au regard des indicateurs de multicolinéarité, on ne peut exclure la présence de ce biais.

Les catégories d'occupation de sol fournies par la base de données Corine Land Cover sont une source de biais économétrique. Afin d'identifier les effets des différentes catégories d'aménités naturelles, nous introduisons l'ensemble des occupations de sol "naturelles" disponibles (forêts, végétation arbustive, espaces ouverts, zones humides et espace en eau, terres agricoles), en gardant comme référence la catégorie des territoires artificialisés. Or les communes faiblement artificialisées et à forte proportion d'aménités naturelles seront moins attractives pour les ménages du fait d'une insuffisante capacité d'accueil. Sans information complémentaire, le modèle conclura à une baisse d'attractivité de la commune due à une forte présence des aménités naturelles. Le rôle de ces dernières risque ainsi d'être sous-estimé. Pour réduire ce problème, le logarithme de la population est introduit en tant qu'indicateur de la capacité d'accueil de la commune. Il prend en compte le nombre de logements mais aussi le taille des ménages et est directement liée aux flux de population qui peuvent être accueillis. L'introduction de cette variable risque néanmoins d'augmenter le degré de multicolinéarité déjà élevée par ailleurs dans le modèle.

Par la suite, l'introduction de nouvelles variables et la prise en compte de manière plus fine des variables déjà spécifiées (type de mesure, maille géographique) pourra permettre de dépasser ces limites économétriques et d'arriver à une spécification sans effets fixes aussi robuste et précise que celle des modèles à effets fixes. Ceci nous rapprochera d'une explicitation complète des spécificités des communes.

3.2. Interprétation

La variable expliquée correspond au choix d'une localisation résidentielle. Les coefficients présentés dans le tableau d'estimation montrent l'impact de chaque variable sur ce choix.

Les coefficients du 1er et du 2eme modèle traduisent des différences relatives par rapport à une catégorie de référence. Toute choses égales par ailleurs, une valeur positive augmente relativement la probabilité de choix, tandis qu'une valeur négative la diminue. La non-significativité d'un coefficient s'interprète comme une absence de différence entre la variable et la catégorie de référence. L'effet fixe de chaque alternative indique l'attrait de chaque localisation relativement à la référence, qui est ici la commune d'Aix-en-Provence.

Dans les modèles sans effets fixes, les coefficients de la catégorie de référence correspondent aux variables sans interaction. Toutes choses égales par ailleurs, l'impact absolu des variables d'interaction est obtenu en sommant leurs coefficients avec ceux de la catégorie de référence.

Par la suite, nous discutons ces effets en insistant plus particulièrement sur ceux qui se sont avérés statistiquement significatifs.

A. Variables classiques

Les variables d'interaction entre le revenu médian des communes et les catégories socio-professionnelles sont toutes statistiquement significatives, sauf celles faisant intervenir les cadres. Toutes choses égales par ailleurs, les indépendants, les employés, les ouvriers, les retraités et les autres inactifs préfèrent des localisations à revenu médian plus bas que les professions intermédiaires (catégorie de référence). Les cadres semblent avoir les mêmes préférences que la catégorie de référence, mais le signe positif du coefficient pourrait traduire une préférence pour les communes plus aisés. Les ménages avec enfants ont tendance à préférer les locations plus modestes, par rapport aux ménages sans enfants. L'avancement en âge s'accompagne d'une recherche des localisations à revenu médian plus élevé, tandis qu'à

partir d’un certain âge (environ 55 ans) cette tendance s'inverse et le choix se porte sur des communes moins riches (la présence d'une variable d'âge élevée au carré permet de révéler cet effet non linéaire).

Les modèles sans effets fixes confirment l’intuition d'une préférence des cadres pour des localisations plus prospères (la valeur du coefficient de la variable d’interaction entre le revenu et l'indicatrice des cadres est positive et significative, et l'effet absolu reste positif lorsque cette valeur est additionnée à celle du coefficient de la catégorie de référence : -0.2.961+0.4981=+0.2567). Toutes les autres classes socio-professionnelles préfèrent des localisations plus modestes. Par ailleurs, les localisations aisées deviennent plus attractives avec l’âge et lorsque l’on vit en couple.

Les variables d’interaction faisant intervenir la distance au centre-ville le plus proche (parmi Aix en Provence, Marseille et Toulon) montrent que les cadres, les indépendants et les inactifs préfèrent des localisations plus centrales que les professions intermédiaires. Toutefois, au-delà d’une certaine distance (35 km), les ménages sans activités préfèrent des localisations plus périphériques que les autres ménages. Les modèles 3 et 4 donnent une indication sur le comportement de localisation des professions intermédiaires. Cette catégorie semble manifester la préférence la plus forte (non significative cependant) pour l'éloignement du centre-ville. Les variables d’interaction entre la distance et le fait d’avoir des enfants sont significatives et suggèrent que les ménages avec enfants ont une préférence pour s’installer en périphérie des grandes villes, mais sans trop s’éloigner (pas plus de 30 km). Les ménages ont tendance à se décentraliser avec l’âge, mais uniquement jusqu’à 28 km environ, tandis qu'au-delà la probabilité de localisation diminue avec la distance.

Les coefficients associés à l'indicatrice "same_municp" renseignent sur la préférence pour le fait choisir un logement dans la même commune que sa résidence antérieure (). Ils sont significatifs quelle que soit la variable d’interaction ou le modèle appliqué. Les modèles 3 et 4 confirment que, toute catégorie socio-professionnelle confondue, les ménages migrants auront toujours une préférence plus prononcée pour leur localisation antérieure. On observe, toutefois, une certaine hiérarchie dans les préférences entre les différentes catégories, mais qui n'est pas la même dans nos différents modèles. Ainsi, selon les modèles à effet fixes, les professions intermédiaires seraient les plus mobiles, alors que les modèles sans effets fixes identifient les cadres comme la catégorie ayant la probabilité la plus élevée de changer de commune de résidence. Cette différence de classement entre le deux types de modèles peut

être expliquée, au moins en partie, par les limites économétriques spécifiées auparavant. En revanche, toutes les spécifications indiquent qu'avoir des enfants augmente la probabilité de rester dans la même commune. L'effet de l'âge joue en deux temps : avec l'âge on préfère changer de commune de résidence, mais ce phénomène n'est pas linéaire. A partir de 50 ans, la probabilité de rester dans la même commune augmente. Enfin, les ménages à plusieurs adultes changent plus difficilement de commune que ceux à un seul adulte.

B. Variables d'aménités

Dans le but d'identifier l'existence des inégalités sociales d'accès aux aménités naturelles, nos modèles intégrent deux grands types d'indicateurs permettant de quantifier l'exposition à ces biens : la distance entre la mairie de la commune de destination et une aménité remarquable la plus proche d'une part, et la part d'une aménité dans l'occupation du sol de la commune d'autre part. En termes de distance, on s'intéresse au positionnement des ménages migrants par rapport aux montagnes, lacs, rivières, plages et calanques. En termes de surface, l'estimation intègre l'exposition aux forêts, à la végétation arbustive, aux espaces ouverts, aux espaces en eau et aux terres agricoles. Parmi les catégories d'occupation du sol, la proportion des territoires artificialisées est omise, car elle serait parfaitement colinéaire avec les autres types d'occupation du sol.

Selon les résultats d'estimation des quatre modèles, les cadres préfèrent – relativement aux autres catégories socioprofessionnelles et dans l'absolu – des localisations résidentielles éloignées des montagnes. La montagne est aussi moins attractive pour les ménages avec plusieurs adultes, avec enfants, et pour les ménages âgés. Autrement dit, des estimations robustes (confirmées par tous les modèles) indiquent que la montagne attire plus les jeunes célibataires sans enfants n'étant pas cadres. Des informations complémentaires sont fournies, sans contradiction, par un des deux types de modèles. Les deux premières spécifications identifient les ouvriers comme étant moins attirés par la montagne que les autres catégories socioprofessionnelles (à l'exception des cadres), tandis que les employés sont les plus attirés par les montagnes d'après les deux derniers modèles.

Contrairement aux montagnes, les lacs constituent des aménités naturelles attractives pour les ménages avec enfants et pour les ménages à plusieurs adultes. L'âge est plutôt un facteur d'éloignement par rapport aux lacs, mais dans les modèles à effets fixes, on observe que l'âge

au carré est significatif et négatif : ainsi, à partir de 57 ans, les ménages cherchent à résider à proximité des lacs. D'après les modèles 3 et 4, les cadres ont tendance à plus s'éloigner des lacs que les autres catégories socioprofessionnelles. Les modèles 1 et 2 ne distinguent pas le comportement des cadres de celui des professions intermédiaires, mais ils montrent néanmoins que les indépendants, les employés et les ouvriers ont des probabilités plus grandes de se localiser près des lacs que ces deux catégories. Le fort positionnement des ouvriers autours des lacs peut être en partie expliqué par l'introduction de l'Etang de Berre dans la catégorie lacs. Ce lac est particulier car il occupe une surface plus importante que les autres et, par ailleurs, sa spécialisation est plus industrielle et moins orientée vers le loisir. De ce fait, on retrouve une forte concentration d'ouvriers et d'employé dans les communes adjacentes[5].

Les localisations à proximité des rivières sont privilégiées par les employés, tandis qu'elles sont évitées par les inactifs et les ménages à plusieurs adultes. L'âge est un facteur favorisant le rapprochement par rapport à cette aménité naturelle. Toutes les autres variables ne sont pas significatives.

L'impact des plages sur le choix d'une localisation résidentielle est estimé à l'aide des modèles 1 (avec effets fixes) et 3 (sans effets fixes). Les variables significatives dans les deux spécifications sont l'interaction avec le nombre d'adultes et l'interaction avec l'âge. Le coefficient étant positif, pour la première, on peut conclure qu'un ménage à un seul adulte tirera plus d'utilité d'une localisation proche des plages qu'un ménage à plusieurs adultes, toutes choses égales par ailleurs. Avec l'âge, les individus ont tendance à plus préférer les plages. Pour ce qui est de la différenciation socio-professionnelle des choix, le modèle sans effets fixes indique que les employés sont plus attirés par les plages que toutes les autres catégories socio-professionnelles, alors que lorsqu'avec effets fixes, ce sont les cadres ont tendance à s'en éloigner plus que tous les autres ménages.

Symétriquement aux plages, l'impact des calanques sur le choix d'une localisation résidentielle est estimé à l'aide des modèles 2 et 4. Les interactions avec l'indicatrice « adultes » et avec les variables d'« âge » ont le même impact que lors de l'introduction des plages : quel que soit le modèle, la probabilité de résider à proximité des calanques est plus élevée pour les ménages composés d'un seul adulte et augmente avec l'âge. Le fait d'avoir des enfants augmente aussi cette probabilité. Les catégories socio-professionnelles les plus

[5] Voir les statistiques descriptives de la zone d'étude.

sensibles à la distance aux calanques sont les professions intermédiaires (modèle 4), les employés et les retraités et les sans activités (modèles 3 et 4). Tous leurs coefficients montrent un attrait des calanques.

Presque toutes les variables d'interaction avec les forêts (mesurées comme proportion dans l'occupation du sol de la commune de résidence) sont significatives. Les modèles 3 et 4 indiquent qu'à l'âge de référence (45 ans), ce sont dans l'absolu des aménités naturelles appréciées de toutes les catégories socio-professionnelles. Néanmoins, les interactions avec l'âge montrent que ce n'est pas nécessairement le cas pour les ménages les plus jeunes et surtout pour les plus âgés. En termes relatif, on peut observer un classement des préférences entre les différentes catégories. Dans tous les modèles, ce sont les retraités qui ont la préférence la plus forte pour les forêts, tandis que les moins sensibles sont les professions intermédiaires. Par ailleurs, les ménages à plusieurs adultes ou avec des enfants ont une préférence plus forte pour les communes les mieux dotées en forêts.

La variable végétation arbustive est significative et positive pour toutes les catégories socio-professionnelles hormis la référence. Les professions intermédiaires constituent donc le groupe le moins attiré par ce type d'aménité. Comme précédemment, le fait d'être retraité augmente la probabilité de se localiser à proximité de ces aménités, mais, paradoxalement, l'âge diminue la préférence pour la végétation arbustive. Selon les modèles 3 et 4, les ménages avec enfants ont une probabilité moindre que ceux sans enfants de se localiser dans des zones abondantes en végétation arbustive.

La proportion des espaces ouverts dans la commune est une variable largement significative et correspond le plus souvent à un effet absolu positif sur les probabilités de choix résidentiel. Ce résultat peut paraître surprenant, sachant que ce type d'aménité occupe une faible proportion de la zone d'étude. Sa significativité pourrait s'expliquer par la proximité géographique avec les montagnes. En intégrant de la végétation clairsemée et du relief rocheux, les espaces ouverts captent des aménités paysagères hétérogènes et des espaces de loisir attractives.

La variable zone humides et espaces en eau représente une mesure assez brute des dotations en eau des communes. Elle agrège des étendues en eau différentes avec un potentiel d'attraction très hétérogène (marais, tourbières, zones intertidales, cours et plans d'eau, eaux maritimes). Les coefficients estimés sont donc à interpréter avec précaution car on ne connait pas le poids de chaque type d'eau intégrée dans la variable. Les modèles montrent seulement

trois effets significatifs et ils sont négatifs. Ainsi, toutes choses égales par ailleurs, les ménages avec enfants, les retraités et les ménages à plusieurs adultes seraient moins attirés (que leurs références respectives) par les communes présentant une forte proportion en eau. L'utilisation des données à une échelle géographique plus petite et une construction plus fine de cette variable permettra certainement de nuancer ces estimations.

L'interaction entre la proportion des terres agricoles et les catégories socio-professionnelles des ménages montrent que les ménages sans activités, les cadres et les retraités sont plus attirés que les autres catégories de ménages par ce type d'occupation du sol. A l'opposé, les ouvriers préfèrent s'éloigner des terres agricoles. Le fait d'avoir des enfants et l'âge jouent aussi négativement sur la recherche de proximité avec ce type d'aménité.

Le logarithme de la population (modèles 3 et 4) est un indicateur de la capacité d'accueil de la commune. Il prend en compte le nombre de logements mais aussi le taille des ménages et est directement liée aux flux de population qui peuvent être accueillis. Comme attendu, le coefficient de cette variable est positif et très significatif, ce qui signifie que plus il y a de capacité d'accueil des nouveaux résidents dans la commune, plus la probabilité de s'y installer est importante.

Les effets fixes des modèles 1 et 2 captent toutes les spécificités des communes. Leur valeur indique l'attrait de la localisation relativement à notre référence : Aix en Provence. Les effets fixes sont tous significatifs et leur valeur est négative pour toutes les alternatives, sauf Marseille et Toulon. Ceci signifie que, toutes choses égales par ailleurs, la probabilité de se localiser à Marseille ou à Toulon est plus forte que celle de migrer à Aix en Provence, tandis qu'elle est plus faible pour toutes les autres communes.

Suite à cette analyse des résultats obtenus à l'aide des différents indicateurs de mesure d'accès aux aménités, on peut conclure que les indicateurs simples de distance semblent moins pertinents que les indicateurs de proportion. Par exemple, les variables d'occupation du sol « végétation arbustive et « espaces ouverts » présentent plus d'interactions significatives que la variable « distance à la montagne », alors que la localisation de ces aménités est très proche géographiquement. Plusieurs critiques peuvent être apportées à l'indicateur de distance. Il a été, effectivement, employé dans sa forme la plus simple : la distance linéaire, et ne prend pas

en compte le fait que l'importance de la distance s'estompe à partir d'un certain niveau d'éloignement (effet non linéaire de la distance). De plus, ce type d'indicateur ne prend pas en compte la taille de la ressource. Ainsi, on intègre les montagnes sans distinction, alors qu'il sera plus fin de distinguer les plus importantes (comme la Montagne Sainte Victoire, par exemple). Il semble donc nécessaire de développer des nouveaux indicateurs d'accès aux aménités, plus complexes, qui puissent rendre compte à la fois de la proximité à la ressource et de son importance.

CONCLUSION

Notre hypothèse de départ est que les ménages les plus favorisés bénéficient d'une meilleure accessibilité aux aménités naturelles, et que cet état est la conséquence de migrations résidentielles socialement sélectives, où les ménages les plus favorisés expriment les plus fortes préférences pour les aménités naturelles. On s'attend donc à ce que nos estimations révèlent une préférence marquée des cadres, relativement aux autres catégories socioprofessionnelles moins aisées, pour se localiser à proximité des aménités naturelles, toutes choses égales par ailleurs.

Dans quelles mesures les résultats obtenus à ce stade étayent-ils cette hypothèse ? Nous pouvons rappeler en quelques phrases les éléments de différenciation sociale – plus précisément socioprofessionnelle – des choix résidentiels.

Tout d'abord, un résultat robuste est que les cadres se différencient des autres ménages en préférant des localisations plus éloignées des montagnes. De plus, nos modèles à effets fixes indiquent que de toutes les catégories socioprofessionnelles, les ouvriers présentent les préférences les plus semblables à celles des cadres. Ce premier résultat va clairement à l'encontre de notre hypothèse. De même, les cadres se différencient des autres ménages (à l'exception des professions intermédiaires) en préférant des localisations plus éloignées des lacs. Par ailleurs, leur comportement ne se différencie pas de celui des autres catégories socioprofessionnelles (hormis les employés qui sont plus attirés) pour ce qui concerne la distance aux rivières, et de même pour la distance aux plages dans le modèle sans effets fixes. Ils préfèrent aussi, plus que les autres, s'éloigner des plages dans le modèle à effets fixes, et s'éloigner des calanques, plus que de nombreux autres groupes, dans les modèles avec ou sans effets fixes. Tous ces résultats concernant les distances aux aménités nous conduisent à la même conclusion : notre hypothèse de travail n'est pas vérifiée.

Les résultats portant sur les variables d'occupation du sol ne disent pas autre chose. Les cadres n'ont pas une préférence sensiblement plus marquée que les autres pour la végétation arbustive, ni pour les espaces ouverts, ni pour les espaces en eau, ni pour les terres agricoles. Et ils sont moins attirés par la forêt que la plupart des autres ménages.

Bien que nos résultats soient perfectibles – des problèmes de variables omises et de colinéarité sont présents dans le modèle sans effets fixes et les variables d'aménités demeurent très simples et imparfaites dans les deux modèles – nous pouvons présumer que notre hypothèse de départ n'est pas vérifiée sur notre zone d'étude.

Comment expliquer ce résultat, alors que l'analyse descriptive semble aller dans le sens de l'existence des inégalités socio-économiques d'accès aux aménités naturelles ?

Une explication possible à cette dualité des résultats part du fait que les deux types d'analyses ne se font pas sur les mêmes populations. L'analyse descriptive porte sur les stocks des ménages, tandis que le modèle économétrique traite les flux migratoires, sans la prise en compte de leur évolution historique.

Les résultats cartographiques montrent que les zones les plus denses et notamment celles caractérisées par une forte présence des cadres (par exemple les communes autour d'Aix en Provence) ont eu une très faible évolution de la population ces 20 dernières années. Au contraire, les territoires moins denses, présentent des taux de croissance de la population importants. Ces phénomènes démographiques suggèrent fortement l'existence des effets de saturation foncière dans certaines communes. Ainsi, les nouveaux ménages, de certaines catégories sociales sont confrontés avec une contrainte d'offre immobilière et donc ils se retrouvent dans l'impossibilité de se localiser selon les mêmes schémas que leurs prédécesseurs. Si cette hypothèse est confirmée, les inégalités sociales d'accès aux aménités naturelles entre les catégories socioprofessionnelles existent toujours, mais des nouvelles formes d'inégalités émergent entre les membres d'une même catégorie.

Bibliographie

Alonso W. (1964) “Location and land use: toward a general theory of land rent”, Cambridge, *Harvard University Press*.

Baccaïni B. (1994) “Comportement migratoire et cycle de vie ”, *Espaces, populations, sociétés*, vol. 12, n. 1, pp 61-74.

Baccaïni B., Levy D. (2009) “Recensement de la population de 2006: Les migrations entre départements : le Sud et l’Ouest toujours très attractifs”, *Insee Première*, n. 1248.

Boarnet M. G. (1994) “An empirical model of intrametropolitan population and employment growth”, *Papers in Regional Science*, vol. 73, n. 2, p. 135-152.

Börsch-Supan, A., & Pitkin, J. (1988) “Choice Models of Housing” *Journal of Urban Economics*, 24, 153–172.

Bertrand N., Rambonilaza T., Tivadar M., Lyser S. (2010) “Response functions for housing demand”, *PLUREL project report.*

Brueckner J.K., Thisse J., Zenou Y. (1999) “Why is central Paris rich and downtown Detroit poor?: An amenity-based theory”, *European Economic Review*, vol. 43, p. 91-107.

Caruso G., Peeters D., Cavailhes J., & Rounsevell M. (2007) “Spatial configurations in a periurban city. A cellular automata-based microeconomic model”, *Regional Science and Urban Economics*, *37*(5), p. 542–567.

Carlino G. A., Mills E. S. (1987) “The determinants of county growth”, *Journal of Regional Science*, vol. 27, n. 1, pp. 39-54.

Chi, G. Q. and D. W. Marcouiller (2013) "In-migration to remote rural regions: The relative impacts of natural amenities and land developability" *Landscape and Urban Planning* 117, p. 22-31.

Clark D. E., Murphy C. A. (1996) “Countywide employment and population growth: An analysis of the 1980s” *Journal of Regional Science*, vol. 36, n. 2, p. 235-256.

Cremer-Schulte, D. (2013) “Residential choices, natural amenities & segregation – the case of an urban mountain context” Paper presented at the 53rd *ERSA congress. Regional*

Integration: Europe, The Mediterranean and the World Economy, 27th to 31st August 2013, Palermo, Italy.

De Palma A., Motamedi K.,Picard N., Waddell, P. (2007) "Accessibility an environmental quality: inequality in the Paris housing market", *European transport* n 36, p. 47-74.

Deller S. C., Tsai T. S. H., Marcouiller D. W., English D. B. K. (2001) "The role of amenities and quality of life in rural economic growth", *American Journal of Agricultural Economics*, vol. 83, n. 2, p. 352-365

Duffy-Deno K. T. (1997) "The effect of state parks on the country economies of the West", *Journal of Leisure Research*, vol. 29, n. 2, p. 201-224.

Fusco G., Scarella F (2011) "Métropolisation et ségrégation sociospatiale. Le flux de mobilités résidentielles en Provence-Alpes-Côte d'Azur, *L'espace géographique*, 2011/4 Tome40, p.319-336.

Frenkel, A., Bendit, E., & Kaplan, S. (2013). Residential location choice of knowledge-workers: The role of amenities, workplace and lifestyle. Cities, 35, 33–41.

Goffette-Nagot F., Schaeffer Y. (2013) "Accessibilité ou voisinage? Une analyse des sources de la ségrégation résidentielle au sein des aires urbaines françaises", *Revue économique*, Vol. 64, n. 5, p. 857-882 ;

Graves P. E. (1983) "Migration with a composite amenity: The role of rents", *Journal of Regional Science*, vol. 23, n. 4, p. 541-546

Graves P.E. (1980) "Migration and climate", *Journal of Regional Science*, vol. 20, n. 2, p. 227-237

Graves P.E. (1979) "A life-cycle empirical analysis of migration and climate, by race", *Journal of Urban Economics*, vol. 6, n. 2, p. 135-147;

Isard W. (1960) "Methods of regional analysis: an introduction to regional science" *The MIT Press*, Cambridge.

Jim C. Y., Chen W. Y., (2006) "Impacts of urban environmental elements on residential housing prices in Guangzhou (China)", *Landscape and Urban Planning* 78, p. 422-434.

Johnson K. M., Fuguitt G. V. (2000) “Continuity and change in rural migration patterns, 1950-1995”, *Rural Sociology*, vol. 65, n. 1, p. 27-49.

Kong F, Yin H, Nakagoshi N, 2007, “Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China”, *Landscape and Urban Planning* 79 240-252.

McFadden, D. (1973) “Conditional logit analysis of qualitative choice behavior”, in P. Zarembka (ed.), *Frontiers in Econometrics New York: Wiley*, p. 105-135.

McFadden, D. (1978) “Modelling the choice of residential location” In A. Karlquist (Ed.), *Spatial Interaction Theory and Residential Location* (pp. 75–96) Amsterdam, NL: North Holland.

McFadden, D. (1981) “Econometric models of probabilistic choice”, in F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data With Econometric Applications*, Cambridge, Mass. MIT Press, p. 198-272.

McFadden, D. (2001). Economic choices. The American Economic Review, 91(3), 351–378.

McGranahan D. A. (1999) “Natural amenities drive rural population change”, *Agricultural Economic Report* n. 781, Washington, DC: U.S. Dept. of Agriculture, Economic Research Service.

McGranahan D. A. (2008) “Landscape influence on recent rural migration in the U.S.”, *Landscape and Urban Planning*, vol. 85, n. 3-4, p. 228-240.

Mueser P. R., Graves P. E. (1995) “Examining the role of economic opportunity and amenities in explaining population redistribution” *Journal of Urban Economics*, vol. 37, n. 2, p. 176-200.

Schmidheiny K. (2006) “Income segregation and local progressive taxation: Empirical evidence from Switzerland”, Journal of Public Economics, 90(3), p. 429-458;

Schmidheiny, K., & Brülhart, M. (2011). On the equivalence of location choice models: Conditional logit, nested logit and Poisson. Journal of Urban Economics, 69(2), 214–222.

Tiebour C. (1956) “A pure theory of local expeditures”, *Journal of Political Economy* 64 (5), p. 416- 424.

Train, K. E. (2009). Discrete Choice Methods with Simulation (2nd ed.). New York, USA: Cambridge University Press.

Waltert F., Schläpfer F. (2010) “Landscape amenities and local development: A review of migration, regional economic and hedonic pricing studies” *Ecological Economics*, vol. 70, n. 2, p. 141-152.

Wu J, 2006, “Environmental amenities, urban sprawl, and community characteristics” *Journal of Environmental Economics and Management* **52** 527-547

Wu J, Plantinga A J, (2003) “The influence of public open space on urban spatial structure” *Journal of Environmental Economics and Management* **46** 288-309

Wu, J. and S. H. Cho (2003) “Estimating households' preferences for environmental amenities using equilibrium models of local jurisdictions”, *Scottish Journal of Political Economy* 50(2): p.189-206.