

---

**DETERMINANTS DES SALAIRES HORAIRES**  
**DES ASSISTANTES MATERNELLES :**  
**ESTIMATION D'UN MODELE ECONOMETRIQUE A EQUATIONS**  
**SIMULTANEEES ET AUTOCORRELATION SPATIALE**

Nicolas PALIOD (\*)

(\*) Insee, Direction de la méthodologie et de la coordination statistique et internationale

[nicolas.paliiod@insee.fr](mailto:nicolas.paliiod@insee.fr)

**Mots-clés** : assistantes maternelles, économétrie spatiale, autocorrélation spatiale, équations simultanées, mode de garde

---

## Résumé

La garde des jeunes enfants (enfants de moins de 6 ans) représente socialement et économiquement un enjeu de taille pour les pouvoirs publics puisque le nombre d'enfants de moins de 6 ans en France hors Mayotte est estimé à 4,8 millions par l'Insee. L'objectif de cette étude réalisée à la Drees (Direction de la recherche, des études, de l'évaluation et des statistiques du Ministère des Solidarités et de la Santé) est de comprendre pourquoi certaines assistantes maternelles ont un salaire horaire plus élevé que d'autres. S'il apparaît que la localisation de l'assistante maternelle ou que son âge ont une influence sur le salaire horaire de l'assistante maternelle, l'étude montre également que le salaire horaire moyen des assistantes maternelles par bassin de vie présente une forte corrélation spatiale. En l'occurrence, cela signifie que lorsque les salaires horaires des assistantes maternelles sont élevés dans un bassin de vie, ils ont également une probabilité importante d'être élevés dans les bassins de vie voisins.

Le marché de la garde d'enfants fait intervenir les assistantes maternelles d'un côté et les parents qui sont également employeurs de l'autre. Lors de la signature du contrat, les parties se mettent d'accord sur le salaire horaire de l'assistante maternelle, mais également sur le nombre d'heures de garde et sur le nombre d'enfants gardés, dans un cadre fixé par la loi. Un modèle économique reposant sur un modèle de concurrence de Cournot a été établi afin d'expliquer comment sont déterminés le nombre d'enfants gardés, le salaire horaire et le nombre d'heures de garde. Afin de tenir compte des contraintes de déplacement des parents pour faire garder leurs enfants et compte-tenu des effets de corrélation spatiale mis en évidence sur le salaire horaire par bassin de vie, ce modèle de concurrence à salaire constant a été étudié au niveau du bassin de vie.

La résolution économique débouche sur l'analyse économétrique d'un modèle à trois équations simultanées expliquant respectivement le nombre d'enfants gardés par les assistantes maternelles dans un bassin de vie, le salaire horaire moyen des assistantes maternelles dans le bassin de vie et le nombre d'heures de garde dans le bassin de vie. Les méthodes « classiques » de résolution ne permettent pas de corriger de la corrélation spatiale et les résidus ne sont donc pas indépendants de la structure territoriale. Des techniques d'économétrie spatiale, spécialement développées pour la résolution de systèmes à multiples équations, sont donc utilisées pour obtenir des résultats plus robustes. Néanmoins, il apparaît dans la résolution numérique que l'estimateur de variance de la méthode utilisée converge trop lentement. En outre, les méthodes de bootstrap classiques ne sont pas exploitables lorsque les données sont dépendantes d'une structure géographique. Une méthode de bootstrap des résidus a ainsi été utilisée pour estimer la variance des estimateurs afin de ne pas déformer la structure géographique sous-jacente des observations, avec des résultats contrastés.

Les conclusions de l'étude mettent en évidence des résultats différents quantitativement voire qualitativement de ceux obtenus sans tenir compte de la corrélation spatiale entre les observations. Surtout, tenir compte de l'autocorrélation spatiale des salaires horaires des assistantes maternelles relève d'une logique différente de l'économétrie plus classique et permet de mieux rendre compte de la concurrence du marché de la garde d'enfants. L'étude caractérise quelques déterminants causaux secondaires du salaire horaire des assistantes maternelles. Mais l'effet principal mis en évidence est celui de l'autocorrélation spatiale des salaires horaires.

L'objectif de cette présentation est double :

- Montrer combien certains phénomènes économiques nécessitent une prise en compte particulière des effets de dépendance spatiale. Sans cela, les résultats de cette étude auraient été biaisés et différents de ceux finalement obtenus ;
- Expliciter la démarche d'estimation de modèles économétriques à équations simultanées et autocorrélation spatiale.

## Abstract

This study aims at improving the comprehension of the profession of childminders. Understanding which factors influence this profession is important for public authorities since there are 4,8 million of children in France excluding Mayotte. A theoretical model explaining the interaction between childminders and employers is analysed with econometric methods in order to explain the hourly wage of childminders. The chosen econometric method gathers spatial autocorrelation and simultaneous equations since the hourly wage of the childminder is determined in the contract at the same time as the duration of work and the number of minded children. The main effect on the hourly wage of childminders is a diffusion effect: it is very dependent on the hourly wage of childminders in the areas surrounding the area where the childminder lives. Nevertheless, some demand and offer effects also are also at stake.

## Introduction

La garde des jeunes enfants (enfants de moins de 6 ans) représente socialement et économiquement un enjeu de taille pour les pouvoirs publics. L'Institut national de la statistique et des études économiques (Insee) chiffre à 4,8 millions le nombre d'enfants de moins de 6 ans en France hors Mayotte en 2014 (Beaumel, Bellamy, 2016). Au-delà de l'école préélémentaire qui présente un taux de scolarisation des enfants de 3 à 5 ans proche de 100 %, la garde des enfants de moins de 3 ans et la garde des enfants de 3 à 5 ans en dehors des heures d'école représentent un enjeu public. En effet, lorsqu'ils le souhaitent, les parents doivent être en mesure de trouver une solution pour faire garder leurs enfants et continuer une activité professionnelle. À l'inverse, l'État permet aux parents souhaitant garder leurs enfants d'avoir une compensation financière à la cessation ou à la réduction de leur activité professionnelle avec certaines allocations telles que la prestation partagée d'éducation de l'enfant (PréParE) ou le complément de libre choix d'activité (CLCA). L'Observatoire national de la petite enfance (ONPE) chiffre ainsi en 2015 à 31,4 milliards d'euros les dépenses publiques engendrées par les différents modes de garde du jeune enfant, qu'il s'agisse d'allocations versées aux familles ou de dépenses de fonctionnement d'établissements comprenant les écoles préélémentaires ou encore d'autres dépenses liées aux jeunes enfants (Observatoire national de la petite enfance, 2015).

Parmi ces dépenses importantes, à hauteur de plus d'1 % du Produit Intérieur Brut français, une part élevée est consacrée au complément de libre choix du mode de garde (CMG) du jeune enfant par des assistantes maternelles. En 2014, 5,6 milliards d'euros d'allocations ont été versées aux familles

grâce au CMG pour une assistante maternelle<sup>1</sup> d'après l'ONPE. Outre ces dépenses importantes, les assistantes maternelles représentent une part importante de l'offre de travail en France avec près de « 320 000 assistantes maternelles employées par des particuliers en France » au deuxième trimestre 2014 (Vroylandt, 2016). Comprendre le marché du travail des assistantes maternelles est donc crucial pour les pouvoirs publics du fait de l'importance que cela représente en termes d'emplois et de dépenses publiques.

Cette étude a pour objet de mieux appréhender le marché du travail des assistantes maternelles et tout particulièrement de comprendre les déterminants de leur salaire horaire. Ces salaires horaires sont-ils plutôt dus à des effets de demande des employeurs ou à des effets d'offre des assistantes maternelles ? Ou s'expliquent-ils plutôt par une harmonisation des salaires entre concurrentes ? Pour répondre à ces questions, la Drees (Direction de la recherche, des études, de l'évaluation et des statistiques du Ministère des Affaires Sociales et de la Santé) a disposé de la base PAJEMPLOI, contenant les déclarations de salaires des particuliers employeurs et les caractéristiques des employées.

D'abord, la bibliographie existante concernant les assistantes maternelles est présentée afin d'explicitier un modèle microfondé du marché du travail des assistantes maternelles (section 1). Une fois le cadre théorique posé, le cadre économétrique le plus adapté aux prévisions du modèle microéconomique est établi dans le but d'analyser des résultats économétriques pertinents et fiables et de les confronter aux prévisions théoriques (section 2).

La méthode retenue dont les résultats seront développés dans cette étude est un modèle économétrique à équations simultanées avec autocorrélation spatiale. Dans ce cadre, les déterminants du nombre d'enfants gardés par les assistantes maternelles, du salaire horaire des assistantes maternelles, ainsi que du nombre d'heures de garde effectuées par les assistantes maternelles sont expliqués simultanément.

La principale conclusion est que les salaires horaires des assistantes maternelles suivent un phénomène de diffusion spatiale, c'est-à-dire que les assistantes maternelles ayant des salaires élevés vont influencer positivement les salaires horaires de leurs concurrentes. Cependant, l'étude n'explique pas les mécanismes par lesquels cette influence s'opère. Des déterminants élevés de la demande, tels que le salaire médian des ménages, semblent également favoriser des salaires horaires élevés pour les assistantes maternelles. Néanmoins, le modèle retenu ne permet pas d'avoir des estimateurs convergents des écarts-types des coefficients estimés. Il n'est donc pas possible de conclure sur les déterminants significatifs des salaires horaires des assistantes maternelles.

## **1. Modélisation microéconomique de l'interaction des assistantes maternelles et des employeurs**

Un des objectifs initiaux de cette étude était d'utiliser la théorie de la microéconomie pour prévoir les déterminants théoriques des salaires horaires des assistantes maternelles, afin de les tester empiriquement et de vérifier si les hypothèses formulées semblent ou non être cohérentes avec les résultats obtenus. Cela a nécessité une étude bibliographique des études quantitatives et qualitatives disponibles sur la garde des assistantes maternelles et sur le choix des employeurs de faire ou non garder leurs enfants. Dès lors, il est devenu possible de modéliser dans un premier temps les différents agents de ce marché du travail. Dans un second temps se sont posées les questions du choix du modèle microéconomique le plus adapté au marché du travail des assistantes maternelles et de sa résolution.

---

<sup>1</sup> La profession d'assistante maternelle est exercée à plus de 99 % par des femmes (Vroylandt, 2016). Le féminin sera donc systématiquement utilisé pour la décrire dans ce document.

## 1.1. Modélisation des différents agents

### 1.1.1. Les différents agents du marché du travail des assistantes maternelles

Le marché du travail des assistantes maternelles est avant tout la rencontre entre un besoin de travailler de l'assistante maternelle et un besoin de garde de ses enfants par l'employeur. Ce dernier peut en effet souhaiter disposer de temps durant lequel il n'est pas avec son enfant pour des raisons professionnelles ou extraprofessionnelles. C'est dans ce cadre précis avec deux types d'agents, les assistantes maternelles et les employeurs, que cette étude a été menée.

Toutefois, il s'agit d'un cadre d'étude simplifié qui correspond plutôt à la profession révolue de nourrice qui est l'ancêtre du métier d'assistante maternelle. En réalité, les assistantes maternelles doivent depuis la loi du 17 mai 1977 être agréées pour exercer leur profession. Cet agrément est délivré par le conseil départemental de résidence de l'assistante maternelle. En outre, les pouvoirs publics subventionnent la garde des enfants de moins de 6 ans par des assistantes maternelles. Une partie des charges liées à l'emploi de l'assistante maternelle n'est ainsi pas payée par les employeurs et ces derniers peuvent également bénéficier du complément mode de garde de la Caisse d'Allocations Familiales (CAF) qui prend en charge une partie du salaire des assistantes maternelles.

La délivrance de l'agrément ne se fait pas de manière aléatoire. D'après l'article R. 421-3 et l'article R. 421-5 du Code de l'action sociale et des familles (CASF), l'obtention d'un agrément nécessite un logement de taille suffisante et de confort suffisant pour accueillir les enfants de l'employeur. Ces critères ne sont pas indépendants des déterminants au centre de cette étude puisqu'ils intègrent le lieu de vie des assistantes maternelles. Or, savoir si les salaires horaires peuvent être expliqués par des variables telles que la densité ou le salaire médian d'une localité ou encore par les loyers de cette localité est un enjeu de l'étude. La délivrance d'un agrément est donc endogène au modèle. Concernant les allocations, le constat est identique. Les montants maxima versés à l'employeur à travers l'allocation complément mode de garde dépendent de ses revenus. Cette allocation n'est donc pas exogène au modèle qui sera ultérieurement établi puisque les revenus des ménages et plus précisément les revenus médians d'unités géographiques seront des variables explicatives. Néanmoins, pour permettre d'établir un modèle théorique, la délivrance d'agrément est supposée exogène dans toute la suite du document et n'est plus évoquée.

Ainsi, la modélisation se limite à deux types d'acteurs (les assistantes maternelles et les employeurs) pour simplifier l'étude théorique. Il convient de garder à l'esprit qu'il s'agit d'une modélisation simplifiée d'une réalité plus complexe.

### 1.1.2. Modélisation des acteurs

Lors de la négociation entre assistantes maternelles et employeurs, trois points sont abordés implicitement ou explicitement :

- L'assistante maternelle accepte-t-elle de garder un ou plusieurs enfants supplémentaires ?
- Quel salaire horaire sera fixé dans le contrat ?
- Combien d'heures de garde par mois seront effectuées par l'assistante maternelle ?

Les employeurs vont ainsi suggérer un nombre d'heures de garde proche de leurs besoins. Néanmoins, selon le résultat de la négociation, rien n'empêche l'assistante maternelle de préférer ne pas garder cet enfant si les conditions négociées pour le contrat ne lui permettent pas de dégager un bénéfice suffisant de son travail. L'assistante maternelle peut ainsi préférer ne pas prendre en charge l'enfant compte-tenu des termes du marché. De même, les potentiels employeurs sont libres de ne pas choisir ce mode d'accueil s'ils le trouvent trop coûteux en considérant leurs revenus par exemple.

Afin de modéliser cette interaction, il faut donc définir les besoins et les contraintes des assistantes maternelles, ainsi que ceux des employeurs. En reprenant la terminologie de la théorie économique, le bénéfice global des assistantes maternelles est représenté par une fonction de profit puisqu'elles offrent un service, de même qu'une entreprise offre un bien. Dans ce profit, il faut séparer d'une part les gains de l'assistante maternelle qui seraient équivalents au chiffre d'affaires d'une entreprise, et d'autre part les coûts physiques et psychologiques  $C_{i,s,t}$  engendrés par l'exercice du métier.

Les employeurs ont, tels des consommateurs, été modélisés par une fonction d'utilité. Celle-ci représente l'intérêt de faire garder leur enfant dont sont déduits les coûts financiers qui sont égaux aux gains de l'assistante maternelle<sup>2</sup>. Outre ces coûts et cette fonction d'utilité, les employeurs choisissent ou non de participer au marché du travail des assistantes maternelles en fonction d'un seuil de participation. Si leur utilité dont sont déduits les coûts est supérieure au seuil de participation, ils feront garder leur enfant par une assistante maternelle. Si ce n'est pas le cas, ils préféreront garder leur enfant par eux-mêmes ou par un autre moyen.

Formellement, les fonctions définies ci-dessus se traduisent par :

- un profit  $\Pi_{i,s,t} = w_{i,s,t} q_{i,s,t} h_{i,s,t} - C_{i,s,t}(\cdot)$  de l'assistante maternelle  $i$  à la période  $t$  d'observation et vivant dans l'unité géographique  $s$  ;
- un profit  $u_{j,s,t} - \int q_{i,j,s,t} h_{i,j,s,t} w_{i,j,s,t} di$  de l'employeur  $j$  vivant dans l'unité géographique  $s$ , employant les assistantes maternelles  $i$  à la période  $t$  d'observation et  $u_{j,s,t}$  est l'utilité que les parents retirent de la garde ;
- un seuil de participation  $k_{j,s,t}$  au marché de la garde par assistante maternelle de l'employeur  $j$  à la période  $t$  d'observation et vivant dans l'unité géographique  $s$ .

L'explication de ce que représente l'unité géographique  $s$  sera donnée dans le paragraphe 1.1.3.  $q_{i,s,t}$  est le nombre d'enfants gardés par l'assistante maternelle  $i$  à la période  $t$  et vivant dans l'unité géographique  $s$ .  $h_{i,s,t}$  est le nombre moyen d'heures de garde par enfant gardé par cette assistante maternelle  $i$ .  $w_{i,s,t}$  est le salaire horaire net moyen de l'assistante maternelle  $i$  par enfant gardé.  $q_{i,j,s,t}$  est le nombre d'enfants de l'employeur  $j$  gardés par l'assistante maternelle  $i$ .  $w_{i,j,s,t}$  est le salaire horaire issu du contrat de l'assistante maternelle  $i$  et de l'employeur  $j$  par enfant sur le contrat.  $h_{i,j,s,t}$  est le nombre d'heures de garde indiqués pour ce contrat par enfant. Ceci implique que :

$$q_{i,s,t} = \int q_{i,j,s,t} dj \quad h_{i,s,t} q_{i,s,t} = \int q_{i,j,s,t} h_{i,j,s,t} dj \quad h_{i,s,t} q_{i,s,t} w_{i,s,t} = \int q_{i,j,s,t} h_{i,j,s,t} w_{i,j,s,t} dj$$

### 1.1.3.Choix d'une unité géographique de référence

En pratique, toutes les assistantes maternelles ne seront pas en contact avec tous les employeurs potentiels. En utilisant la base brute de Pajemploi qui est la principale base de données pour cette étude et qui est présentée en annexe A.a il est possible de trouver une unité géographique cohérente au sein de laquelle les employeurs vivant dans cette unité restent pour faire garder leurs enfants. L'idée sous-jacente de la détermination de cette unité géographique est de segmenter

<sup>2</sup> Ceci représente une approximation. Les familles faisant garder un enfant de moins de 6 ans par une assistante maternelle reçoivent généralement une allocation, le complément mode de garde, versée par la Caisse d'Allocations Familiales. Cependant, le montant de cette allocation pour les employeurs ne pouvant être calculé, les allocations versées par la CAF n'ont pas été prises en compte dans la modélisation. De même, les prélèvements sur le salaire brut versé par l'employeur ne sont pas pris en compte puisqu'ils sont payés par la Caisse d'Allocations Familiales dans la majorité des cas. Enfin, les indemnités versées par les employeurs aux assistantes maternelles sont négligées puisqu'en théorie, il ne s'agit pas d'un bénéfice pour l'assistante maternelle.

géographiquement le marché du travail des assistantes maternelles pour répondre à la réalité de la contrainte de déplacement de l'employeur. Il est en effet hautement improbable qu'un employeur en Provence-Alpes-Côte d'Azur fasse garder un enfant par une assistante maternelle en Normandie par exemple, à part éventuellement durant ses vacances.

Cette intuition est confirmée empiriquement car la grande majorité des contrats (97 %) sont signés par une assistante maternelle et un employeur au sein d'une même région. De plus, des contrats peuvent être signés entre deux parties habitant des communes proches mais dans deux régions différentes. Peu de contrats sont donc signés pour faire garder des enfants dans une destination de vacances.

Dès lors, deux cas de figure sont envisageables : soit l'employeur préfère faire garder son enfant à proximité de chez lui et ainsi le déposer et le récupérer au moment de partir et de rentrer à son domicile, soit il préfère employer une assistante maternelle à proximité de son lieu de travail ou sur le chemin de son domicile à son lieu de travail.

La détermination de l'unité géographique au sein de laquelle à la fois les assistantes maternelles et les employeurs d'un même contrat habitent est donc une étape importante pour l'étude puisque ce choix conditionnera les résultats. Différentes solutions ont été envisagées : la commune, le canton, l'unité urbaine<sup>3</sup>, l'EPCI<sup>4</sup>, le bassin de vie<sup>5</sup>, l'aire urbaine<sup>6</sup>, la zone d'emploi<sup>7</sup>, le département, l'ancienne région et la nouvelle région.

Afin de segmenter géographiquement le marché du travail des assistantes maternelles, il faut une unité géographique dont la majorité des flux de gardes sont internes. La commune et le canton ont donc été éliminés parmi les possibilités d'unité géographique de référence puisque 59 % et 74 % seulement des contrats sont signés par des assistantes maternelles et des employeurs habitant respectivement la même commune et le même canton.

D'autre part, il n'est pas souhaitable de choisir une unité géographique trop étendue, au risque de générer une erreur écologique<sup>8</sup> : les déterminants du salaire moyen des assistantes maternelles de l'unité géographique ne sont pas forcément ceux qui influent localement sur leur salaire horaire individuel. Ainsi, la région, le département et la zone d'emploi ne peuvent pas être considérés comme des unités géographiques de référence, car ils existent en trop petits nombre (respectivement 17, 100 et 321 dans lesquels se trouve au moins une assistante maternelle active).

Les EPCI ne conviennent pas non plus car certaines communes ne font pas partie d'un EPCI et se retrouveraient donc dans une unité géographique similaire alors qu'elles ne sont pas proches géographiquement. C'est le cas par exemple de certaines îles de Bretagne et d'une partie de la région parisienne en 2014. De même, l'aire urbaine ne peut être choisie comme l'unité géographique de

---

<sup>3</sup> « On appelle unité urbaine une commune ou un ensemble de communes présentant une zone de bâti continu (pas de coupure de plus de 200 mètres entre deux constructions) qui compte au moins 2 000 habitants » (Insee)

<sup>4</sup> Établissements publics de coopération intercommunale : « regroupements de communes ayant pour objet l'élaboration de projets communs de développement au sein de périmètres de solidarité » (Insee)

<sup>5</sup> « Le bassin de vie est le plus petit territoire sur lequel les habitants ont accès aux équipements et services les plus courants » (Insee)

<sup>6</sup> « Une aire urbaine est un ensemble de communes, d'un seul tenant et sans enclave, constitué par un pôle urbain (unité urbaine) de plus de 10 000 emplois, et par des communes rurales ou unités urbaines (couronne périurbaine) dont au moins 40 % de la population résidente ayant un emploi travaille dans le pôle ou dans des communes attirées par celui-ci » (Insee)

<sup>7</sup> « Une zone d'emploi est un espace géographique à l'intérieur duquel la plupart des actifs résident et travaillent, et dans lequel les établissements peuvent trouver l'essentiel de la main d'œuvre nécessaire pour occuper les emplois offerts. » (Insee)

<sup>8</sup> Une erreur écologique intervient lorsque des conclusions sont déduites d'une étude sur les agrégats mais lorsque ces conclusions ne sont pas vraies sur les unités de départ qui ont été agrégées.

référence car il existe la catégorie des communes multipolarisées dans laquelle se situent des communes éloignées géographiquement.

Ainsi, le choix d'unité géographique de référence se limite au bassin de vie et à l'unité urbaine. 85 % des gardes font intervenir un employeur et une assistante maternelle du même bassin de vie contre 80 % pour l'unité urbaine. En outre, la définition du bassin de vie par rapport à l'accès aux équipements et services correspond plus à l'accessibilité qu'ont les employeurs vis-à-vis du service offert par les assistantes maternelles qu'une définition par rapport à la continuité du bâti, bien que ce service précis ne soit pas pris en compte dans la définition des bassins de vie. Enfin, la continuité du bâti ne présage en rien des flux domicile-travail des employeurs qui peuvent influencer sur les flux de garde. Cela est moins vrai concernant l'accessibilité aux équipements et services qui semble déjà plus corrélée aux flux d'emploi.

Pour toutes ces raisons, résumées pour certaines dans le tableau 1, on pose comme hypothèse que les marchés de garde sont segmentés géographiquement par les bassins de vie.

Unité géographique	Pourcentage de gardes intra-unité	Nombre d'unités avec au moins une assistante maternelle active
Commune	59%	27 182
Canton	74%	3 750
Unité urbaine	80%	2 370
EPCI	82%	2 133
Bassin de vie	85%	1 666
Aire urbaine	88%	794
Zone d'emploi	92%	321
Département	95%	100
Nouvelle région	97%	17

Champ : France entière, hors Mayotte

Sources : Acoess - Centre Pajemploi, Exploitation Drees

Tableau 1 : Pourcentages de contrats de garde faisant intervenir une assistante maternelle de la même unité géographique que l'employeur

Deux améliorations seraient toutefois possibles pour l'unité géographique de référence. La première consisterait à définir des bassins de garde à partir de l'accessibilité des différents modes de garde pour les employeurs potentiels. Il faudrait ainsi connaître les caractéristiques de la commune des parents et de la commune du mode de garde choisi par les parents (établissement d'accueil du jeune enfant, garde à domicile, assistante maternelle, ...). Dès lors, « le plus petit territoire sur lequel les habitants ont accès à un mode de garde » pourrait être déterminé. À défaut de disposer de ces données pour les structures telles que les établissements d'accueil du jeune enfant, la base de données Pajemploi les fournit pour les gardes à domicile et pour les assistantes maternelles. Cette solution serait probablement la meilleure pour cette étude mais s'est avérée trop coûteuse pour être implémentée.

Une autre alternative, compte-tenu des modèles développés par la suite faisant intervenir de l'autocorrélation spatiale, aurait été d'explorer les potentialités permises par l'indice de Moran<sup>9</sup> qui aurait permis d'identifier les niveaux géographiques où il existe une corrélation spatiale entre les salaires des assistantes maternelles.

<sup>9</sup> L'indice de Moran est un indicateur dont une déclinaison permet de calculer localement l'autocorrélation d'une variable.

Ces méthodes n'ont pas été approfondies. C'est pourquoi le choix de l'unité géographique de référence s'est limité aux bassins de vie.

#### 1.1.4. Détermination des fonctions modélisant les acteurs

Après avoir modélisé par des fonctions au paragraphe 1.1.2 les différents agents du marché de la garde par assistante maternelle, il convient d'affiner et de déterminer précisément la forme de ces fonctions.

##### Coûts $C_{i,s,t}(\cdot)$ des assistantes maternelles

La bibliographie décrivant le métier d'assistante maternelle, et principalement les études qualitatives, permet de mieux appréhender la réalité de ce métier.

Plusieurs composantes du travail d'assistante maternelle ont ainsi pu être mis en valeur. Il s'agit d'un travail dont la pénibilité<sup>10</sup> ne peut être négligée. En effet, les assistantes maternelles n'ayant pas toujours un pouvoir de négociation sur les enfants qu'elles acceptent, elles sont contraintes de s'adapter aux horaires des parents, à leur avance éventuelle pour déposer un enfant ou à leur retard éventuel pour le récupérer. La structure de l'emploi local peut rendre difficile la gestion de ces horaires pour les assistantes maternelles. En outre, si s'occuper de plusieurs enfants peut présenter certaines synergies comme la préparation du repas qui n'est pas forcément plus pénible à faire pour deux enfants que pour un seul enfant, s'occuper de plusieurs enfants nécessite de répondre à des demandes individuelles, généralement seule. La pénibilité de ce travail augmente ainsi avec le nombre d'enfants gardés. L'âge intervient également dans la pénibilité du travail d'assistante maternelle : une jeune assistante maternelle ressentira moins le coût de certains sacrifices qu'une assistante maternelle plus âgée.

De plus, l'expérience, le réseau, le lieu de vie des assistantes maternelles sont autant de variables qui peuvent influencer sur la facilité des assistantes maternelles à trouver des enfants à garder et donc aux coûts de temps de recherche d'enfants à garder. Les coûts passés peuvent également être ressentis dans les coûts présents à travers une accumulation de fatigue qui peut amener l'assistante maternelle à réguler son temps de garde et le nombre d'enfants qu'elle garde.

Si des explications plus fournies sont présentes en annexe C, la fonction de coûts suivante dont les notations qui n'auraient pas été mentionnées jusqu'à présent sont précisées en annexe B sera celle utilisée pour l'étude :

$$C_{i,s,t} = \beta_{1+} \mathbb{1} \left\{ \frac{\int_S h_{i,j,s,t} q_{i,j,s,t} dj}{\int_S \text{jours}_{i,j,s,t} q_{i,j,s,t} dj} \geq 8 \right\} \left( \frac{\int_S h_{i,j,s,t} q_{i,j,s,t} dj}{\int_S \text{jours}_{i,j,s,t} q_{i,j,s,t} dj} - 8 \right)^{p_1} + \beta_{2+} q_{i,s,t}^{p_2} \\ + \beta_{3-} \Pi_{i,s,t-1} + \beta_{4+} C_{i,s,t-1} h_{i,s,t} + \beta_{5+} C_{i,s,t-1} q_{i,s,t} + \beta_{6+} \text{conc}_{s,t} q_{i,s,t} \\ + \beta_{7+} \text{age}_{i,s,t}^{p_3} q_{i,s,t} + \beta_{8+} \text{age}_{i,s,t}^{p_4} h_{i,s,t} + \beta_{9-} \text{dens}_{i,s,t} q_{i,s,t} + \beta_{10-} \text{exp}_{i,s,t} q_{i,s,t} \\ + \beta_{11+} \text{loyer}_{i,s,t} q_{i,s,t} + F_{i,s,t}$$

Le premier terme est un terme de pénibilité horaire. Sa complexité apparente cache une réalité bien plus intuitive : au-delà d'un certain nombre d'heures de garde par enfant et par jour, les désagréments liés au métier d'assistante maternelle augmentent.  $q_{i,s,t}$  est le nombre d'enfants gardés par l'assistante maternelle  $i$  vivant dans le bassin de vie  $s$  à l'instant  $t$ .  $h_{i,s,t}$  est le nombre d'heures de garde qu'elle effectue par enfant qu'elle garde.  $w_{i,s,t}$  est le salaire horaire de l'assistante maternelle  $i$  par enfant qu'elle garde.  $q_{i,j,s,t}$ ,  $h_{i,j,s,t}$  et  $\text{jours}_{i,j,s,t}$  sont respectivement le nombre d'enfants mentionnés dans le contrat entre l'assistante maternelle  $i$  et l'employeur  $j$ , le nombre d'heures mentionnées dans le contrat par enfant  $q_{i,j,s,t}$  et le nombre de jours de garde mentionnés dans le contrat par enfant  $q_{i,j,s,t}$ .

<sup>10</sup> Qui se distingue de la notion de souffrance au travail.

Néanmoins, conserver des aspects dynamiques dans les coûts impliquerait que les assistantes maternelles effectuent leurs choix en début de carrière en fonction des coûts à venir, même plusieurs années plus tard. Cela n'est pas réaliste. En outre, les données de recensement de l'Insee utilisées par la suite ne se prêtent pas à une utilisation en panel puisqu'uniquement un cinquième de la base des communes de moins de 10 000 habitants est renouvelé chaque année (Insee, 2014). Concernant les communes de 10 000 habitants ou plus, l'Insee avertit que le recensement par sondage engendre des écarts types souvent plus élevés que les variations des variables de recensement de cette commune. Ainsi, intégrer une dynamique temporelle dans les coûts ne semble pas être judicieux. La fonction de coûts suivante a donc été retenue :

$$C_{i,s,t} = \beta_1 \mathbb{1}_{\left\{ \frac{\int_S h_{i,j,s,t} q_{i,j,s,t} dj}{\int_S \text{jours}_{i,j,s,t} q_{i,j,s,t} dj} \geq 8 \right\}} \left( \frac{\int_S h_{i,j,s,t} q_{i,j,s,t} dj}{\int_S \text{jours}_{i,j,s,t} q_{i,j,s,t} dj} - 8 \right)^{p_1} + \beta_2 q_{i,s,t}^{p_2} + \beta_6 \text{conc}_{s,t} q_{i,s,t} + \beta_7 \text{age}_{i,s,t}^{p_3} q_{i,s,t} + \beta_8 \text{age}_{i,s,t}^{p_4} h_{i,s,t} + \beta_9 \text{dens}_{i,s,t} q_{i,s,t} + \beta_{10} \text{exp}_{i,s,t} q_{i,s,t} + \beta_{11} \text{loyer}_{i,s,t} q_{i,s,t} + F_{i,s,t} \quad (1)$$

Sans perte de généralité puisqu'aucun phénomène dynamique ne sera étudié, le symbole  $t$  sera ôté de l'écriture des variables. Il sera toujours sous-entendu.

#### Utilité $u_{j,s,t}(\cdot)$ des employeurs

Le complément du libre choix d'activité (CLCA) et la prestation partagée d'éducation pour l'enfant (PréParE) permettent au ménage de choisir s'il souhaite ou non réduire ou cesser l'activité d'un parent pour garder son enfant. Il s'agit d'une alternative au mode de garde par assistante maternelle. L'hypothèse que le recours d'un ménage à une assistante maternelle est motivé par le choix de travailler des deux parents du ménage est donc plausible. Si, en réalité, ce n'est pas le cas de tous les employeurs, il sera néanmoins supposé que cela en concerne une majorité.

Une des hypothèses de cette étude est que tous les employeurs déposent leur enfant avant de partir au travail et qu'ils le récupèrent après le travail. On suppose donc que les parents embauchent une assistante maternelle à proximité de leur domicile plutôt qu'à proximité du travail de l'un des deux parents. Cela permet aux deux parents de pouvoir chercher l'enfant et est donc plus plausible. Cette hypothèse suppose également que l'enfant est gardé durant le temps de travail cumulé au temps de trajet des parents. Une alternative est que l'enfant soit déposé à l'école préélémentaire par l'employeur et cherché par l'assistante maternelle. Étant donné, d'après la base permanente des équipements de l'Insee, que plus d'un tiers des communes disposent de classes préélémentaires, l'enfant va probablement à l'école à proximité de son domicile. Afin que l'assistante maternelle puisse chercher l'enfant après l'école préélémentaire, il est probable que les employeurs choisissent des assistantes maternelles à proximité de leur domicile plutôt qu'à proximité de leur travail, dans ce cas également. Cette hypothèse d'embauche de l'assistante maternelle à proximité du domicile des parents sera testée par la significativité de la variable de temps de trajet à la fin de l'étude. Si elle est significative, cette hypothèse sera plutôt cohérente. Sinon, elle sera rejetée.

Au regard des hypothèses préalablement formulées, la journée de l'employeur peut être décomposée en deux phases : le trajet domicile-travail aller-retour et la journée de travail d'un côté, la partie de la journée en famille de l'autre. Durant la journée de travail et son trajet, l'employeur a besoin de faire garder son enfant. Le faire garder une heure ne lui servirait à presque rien alors que le faire garder durant la durée de son temps de travail cumulé à son temps de trajet lui est très utile. Ainsi, durant cette partie de la journée, chaque heure de garde supplémentaire est plus utile à l'employeur que la précédente. En revanche, une fois que l'employeur est rentré de sa journée de

travail, faire garder son enfant une heure supplémentaire lui est de moins en moins profitable. Ces hypothèses sont réductrices car elles ne tiennent pas compte d'activités extra-professionnelles de l'employeur. Mais elles permettent de modéliser de manière cohérente les besoins de garde des employeurs. La figure 1 rend compte de ces hypothèses.



Figure 1 - Forme attendue de la fonction d'utilité de l'employeur

Les heures de garde affichées en abscisse tiennent compte de tous les modes de garde confondus. Dans cette étude, il s'agit des modes de garde par assistante maternelle et par l'école préélémentaire.

Deux cas doivent désormais être distingués : l'enfant a moins de 3 ans et n'est pas scolarisé d'un côté, l'enfant a 3 ans ou plus et est scolarisé de l'autre<sup>11</sup>. Deux fonctions d'utilité  $u_{j,s,t}$  différentes sont donc définies selon que l'enfant ait moins de 3 ans ou entre 3 et 5 ans. Les autres cas sont négligés puisque les enfants de 6 ans et plus ne représentent en 2014 que 6 % des enfants gardés par les assistantes maternelles (Vroylandt, 2016). En notant  $te_{j,s,t}$  le temps d'école de l'enfant de l'employeur  $j$  et  $h_{j,s,t}$  le temps de garde par enfant, deux fonctions d'utilité sont obtenues pour ces deux catégories d'âge :

$$u_{j,s,t}^{<3\text{ ans}} = \phi_1 \mathbb{1}\{h_{j,s,t} \leq tt_{j,s,t} + traj_{j,s,t}\} h_{j,s,t}^{r_1} + \mathbb{1}\{h_{j,s,t} > tt_{j,s,t} + traj_{j,s,t}\} [\phi_1 (tt_{j,s,t} + traj_{j,s,t})^{r_1} - \phi_2 (tt_{j,s,t} + traj_{j,s,t})^{r_2}] + \phi_2 \mathbb{1}\{h_{j,s,t} > tt_{j,s,t} + traj_{j,s,t}\} h_{j,s,t}^{r_2}$$

$$u_{j,s,t}^{3\text{ à }5\text{ ans}} = \phi_1 \mathbb{1}\{h_{j,s,t} + te_{j,s,t} \leq tt_{j,s,t} + traj_{j,s,t}\} (h_{j,s,t} + te_{j,s,t})^{r_1} + \mathbb{1}\{h_{j,s,t} + te_{j,s,t} > tt_{j,s,t} + traj_{j,s,t}\} [\phi_1 (tt_{j,s,t} + traj_{j,s,t})^{r_1} - \phi_2 (tt_{j,s,t} + traj_{j,s,t})^{r_2}] + \phi_2 \mathbb{1}\{h_{j,s,t} + te_{j,s,t} > tt_{j,s,t} + traj_{j,s,t}\} (h_{j,s,t} + te_{j,s,t})^{r_2}$$

Pour respecter les hypothèses ayant abouti à la figure 1, les contraintes suivantes s'appliquent sur les paramètres :  $r_1 > 1$  et  $0 < r_2 < 1$ .  $h_{j,s,t}$  est le nombre d'heures de garde de l'enfant de l'employeur  $j$  par l'ensemble des assistantes maternelles qui le gardent. Les autres notations sont détaillées en annexe B. Pour simplifier l'étude, il sera supposé que chaque employeur n'a qu'un enfant. Cela ne changera pas les résultats globaux de l'étude mais simplifiera les calculs.

<sup>11</sup> D'après les statistiques de la Depp (direction de l'évaluation, de la prospective et de la performance) du Ministère de l'Education nationale, de l'Enseignement supérieur et de la Recherche, 98,5 % des enfants de 3 ans étaient scolarisés et 100 % des enfants de 4 à 5 ans étaient scolarisés en 2013 contre seulement 11,9 % des enfants de 2 ans. Cela justifie la distinction des deux cas.

Par ailleurs, une autre hypothèse simplificatrice qui sera faite consiste à considérer qu'il n'existe qu'une seule tranche d'âge d'enfants. Autrement dit, le modèle sera résolu comme si les enfants de 3 à 5 ans n'allaient pas à l'école maternelle. Le cadre théorique plus complexe dans lequel sont distingués les enfants de moins de 3 ans et les enfants de 3 à 5 ans est présenté en annexe F.a.

Dans le cadre de cette hypothèse où il n'existe qu'une catégorie unique d'enfants, l'utilité de l'employeur s'écrit :

$$u_{j,s,t} = \phi_1 \mathbb{1}\{h_{j,s,t} \leq tt_{j,s,t} + traj_{j,s,t}\} h_{j,s,t}^{r_1} + \mathbb{1}\{h_{j,s,t} > tt_{j,s,t} + traj_{j,s,t}\} [\phi_1 (tt_{j,s,t} + traj_{j,s,t})^{r_1} - \phi_2 (tt_{j,s,t} + traj_{j,s,t})^{r_2}] + \phi_2 \mathbb{1}\{h_{j,s,t} > tt_{j,s,t} + traj_{j,s,t}\} h_{j,s,t}^{r_2}$$

La première partie de la fonction d'utilité  $\phi_1 \mathbb{1}\{h_{j,s,t} \leq tt_{j,s,t} + traj_{j,s,t}\} h_{j,s,t}^{r_1}$  correspond à la période de journée de travail et de déplacement de l'employeur. La deuxième partie,  $\mathbb{1}\{h_{j,s,t} > tt_{j,s,t} + traj_{j,s,t}\} [\phi_1 (tt_{j,s,t} + traj_{j,s,t})^{r_1} - \phi_2 (tt_{j,s,t} + traj_{j,s,t})^{r_2}]$  correspond simplement à une constante servant à s'assurer de la continuité de la fonction d'utilité en  $tt_{j,s,t} + traj_{j,s,t}$ . La troisième partie de la fonction d'utilité  $\phi_2 \mathbb{1}\{h_{j,s,t} > tt_{j,s,t} + traj_{j,s,t}\} h_{j,s,t}^{r_2}$  correspond à la période où l'employeur est rentré de son travail.

#### *Seuil de participation $k_{j,s,t}$ des employeurs*

Modéliser l'utilité et les coûts des parents à faire garder leurs enfants n'est pas suffisant pour déterminer si les parents souhaitent ou non faire garder leurs enfants par une assistante maternelle. En effet, les parents peuvent préférer un autre mode d'accueil car ce dernier serait plus avantageux qu'un mode de garde par assistante maternelle. Mais ils peuvent aussi souhaiter qu'un des deux parents arrête totalement ou partiellement son activité pour s'occuper de leurs enfants.

Dans le cadre de cette étude, il a été choisi par simplification de considérer comme exogène la possibilité de recourir à d'autres modes de garde. L'hypothèse que le choix de travailler à temps partiel pour un parent est exogène est également faite. Ainsi, seule la possibilité pour un des deux parents de cesser complètement toute activité professionnelle est une alternative endogène à la garde par assistante maternelle.

Enfin, une dernière hypothèse concernant la participation des parents au marché de garde par les assistantes maternelles est émise. Dans le cas où les parents préfèrent que l'un des deux arrête de travailler pour s'occuper de leurs enfants plutôt que de faire garder leur enfant par une assistante maternelle, ils préfèrent également cette possibilité à tous les modes de garde possibles.

Afin d'endogénéiser la cessation totale d'activité, un seuil de participation  $k_{j,s,t}$  des employeurs au marché de la garde par assistante maternelle est modélisé. Quelles peuvent-être les raisons pour les parents à préférer la cessation totale d'activité de l'un des deux parents plutôt qu'à participer à tout mode de garde payant ? Une étude portant en partie sur ce sujet a été réalisée en 2004 par Robert-Bobée, Le Blanc et Choné (Robert-Bobée, Le Blanc, Choné, 2004). Leurs hypothèses et leurs conclusions sont utilisées pour modéliser le seuil de participation des parents au marché de garde des enfants par les assistantes maternelles.

Dans cette étude, les auteurs montrent que « le recours [à une garde payante hors domicile] est fortement corrélé à l'activité de la femme : 46 % des ménages avec enfants de moins de 7 ans font appel à un service d'aide rémunéré hors domicile lorsque la femme est salariée, 4 % si elle est au chômage ou inactive ». Donc lorsqu'un des deux parents est au chômage ou inactif, le ménage a moins tendance à faire garder ses enfants par des assistantes maternelles. Les auteurs supposent par ailleurs « qu'un fort taux de chômage local peut dissuader des femmes de se présenter sur le marché du travail ». Cette hypothèse est confirmée par les résultats de leurs régressions économétriques. Ainsi, un fort taux de chômage ou un faible taux d'activité peuvent dissuader un des parents de chercher un travail car la situation économique du bassin de vie n'est pas propice à l'emploi. Dans le cas où l'un des

deux parents ne travaille pas, il peut ainsi, s'il le souhaite, garder ses enfants. Ainsi, le seuil de participation des parents au marché de garde par les assistantes maternelles dépend positivement du taux de chômage et négativement du taux d'activité. Les parents participent d'autant plus au marché de la garde des assistantes maternelles que leur seuil de participation est faible.

D'autre part, toujours selon ces auteurs, « toutes les études trouvent un impact significativement positif du salaire de la mère sur l'emploi et les heures travaillées ». Cela s'explique pour eux par le fait que « les dépenses occasionnées par le recours à un service de garde payant peuvent demeurer élevées relativement aux revenus d'activité potentiels de la mère ». Il y a donc un arbitrage entre le choix de travailler et le choix de recourir à un mode d'accueil payant pour les parents. Ainsi, si le salaire médian du bassin de vie est élevé, ne pas travailler présente un coût d'opportunité important pour le ménage. Il intervient donc négativement dans le seuil de participation des parents au marché de la garde par les assistantes maternelles.

Le seuil de participation des ménages  $k_{j,s,t}$  pour faire garder leur enfant est donc de la forme :

$k_{j,s,t} = \eta_0 + \eta_1 \lambda_{s,t} + \eta_2 \mu_{s,t} + \eta_3 med_{s,t} + \varepsilon_{j,s,t} \quad (2)$
---

$\lambda_{s,t}$ ,  $\mu_{s,t}$  et  $med_{s,t}$  sont respectivement le taux d'activité, le taux de chômage et le salaire médian du bassin de vie où vit le ménage  $j$ .  $\varepsilon_{j,s,t}$  correspond à la situation particulière du ménage.

## 1.2. Choix du cadre d'interaction entre assistantes maternelles et employeurs

### 1.2.1. Un modèle de concurrence de Cournot

Modéliser les acteurs n'est pas suffisant pour construire un modèle afin d'en déduire le salaire théorique des assistantes maternelles. La manière dont ils interagissent entre eux doit aussi être définie. Plusieurs options sont alors envisageables comme un modèle de négociation salariale en économie du travail ou un modèle de concurrence issu de la théorie de l'organisation industrielle.

Bien qu'étudiant des individus négociant un salaire, la deuxième solution a été préférée. En effet, un modèle de concurrence de Cournot sera développé pour modéliser l'interaction entre assistantes maternelles. Ce choix est motivé par plusieurs raisons.

Premièrement, l'offre de garde des assistantes maternelles est un service offert par un agent  $i$  ayant des coûts de garde  $C_{i,s,t}$ . De ce point de vue, dans un cadre théorique, les choix de l'assistante maternelle  $i$  sont assimilables à ceux d'une entreprise vendant un bien dont la production engendre des coûts. L'assistante maternelle propose en effet ses services à des employeurs, qui ont une utilité et des coûts à profiter de ce service. L'employeur est donc proche d'un consommateur. Cela justifie l'intérêt de la théorie de l'organisation industrielle dans le cadre de cette étude.

Deuxièmement, les assistantes maternelles ont une capacité de garde limitée légalement par l'article L421-4 du Code de l'action sociale et des familles. Cette capacité de garde limitée justifie l'utilisation d'un modèle de concurrence de Cournot dans lequel les entreprises ont des capacités de production limitées.

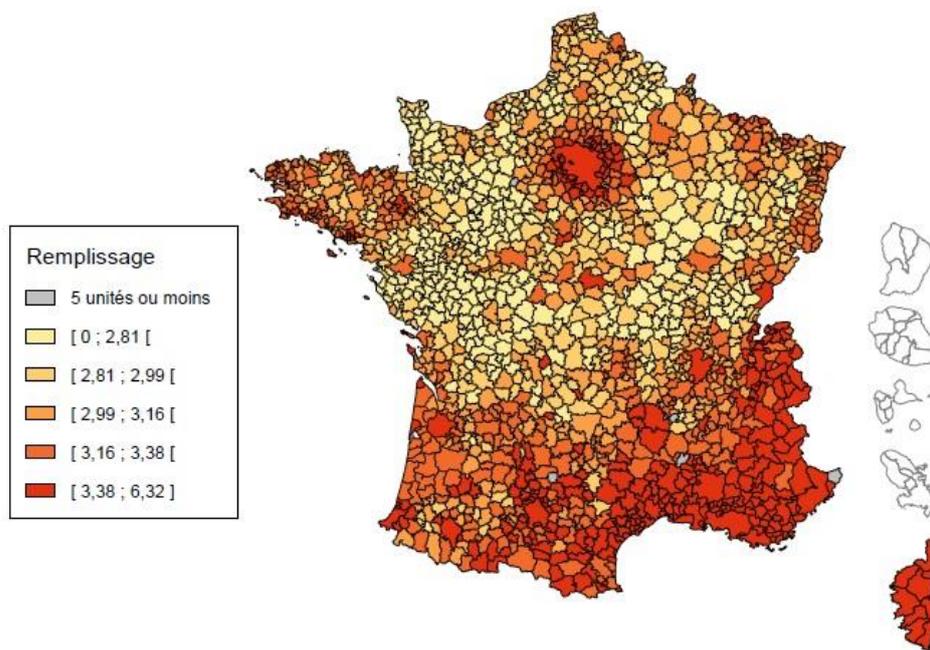
Troisièmement, le modèle de concurrence de Cournot suppose une interaction où la concurrence entre entreprises se limite à un nombre fixé d'entreprises. Ainsi, les choix des unes se ressentent pour les autres. C'est exactement le cas du marché de garde par les assistantes maternelles. Les assistantes maternelles ne sont en concurrence qu'avec les assistantes maternelles vivant à proximité<sup>12</sup>. C'est en tout cas ce qui a été supposé dans la section 1.1.3 sur le choix d'une unité géographique de référence. Le but de cette section était de trouver une unité géographique au sein

<sup>12</sup> Elles sont aussi en concurrence avec les établissements d'accueil de jeunes enfants, les gardes à domicile et, dans une moindre mesure, les petites sections de maternelle du même lieu géographique.

de laquelle toutes les gardes s'effectuent. Les bassins de vie ont été retenus car 85 % des gardes impliquent une assistante maternelle et un employeur vivant dans un même bassin de vie.

Quatrièmement, un tel modèle de concurrence suppose une concurrence en quantité et des prix identiques pour toutes les firmes. Ici, cela suppose que les assistantes maternelles d'un bassin de vie donné se font concurrence par le nombre d'enfants qu'elles gardent ou par le nombre d'heures de garde qu'elles effectuent mais que leur salaire horaire par enfant gardé est identique. Les salaires des assistantes maternelles semblent être en effet relativement proches (Figure 2).

### Salaire horaire moyen des assistantes maternelles par bassin de vie Juin 2012



Carte réalisée avec Cartes & Données - © Artique

Figure 2 - Salaire horaire des assistantes maternelles par enfant gardé

Cette segmentation géographique de la concurrence par bassin de vie se justifie également quantitativement par des variations locales des salaires horaires plus faibles qu'au niveau national. En effet, l'écart-type des salaires horaires des assistantes maternelles est inférieur à 20 % du salaire horaire moyen des assistantes maternelles sur le bassin de vie pour plus de 50 % des bassins de vie, alors que l'écart-type en France métropolitaine des salaires horaires des assistantes maternelles atteint 30 % du salaire horaire moyen des assistantes maternelles. Ceci confirme qu'étudier une concurrence locale en quantité et à salaire local identique pour les assistantes maternelles a du sens. Néanmoins, le fait que les variations des salaires entre assistantes maternelles restent importantes peut remettre en cause le choix du bassin de vie comme unité géographique de référence. Il faut cependant faire un compromis entre choisir une unité suffisamment petite pour avoir des salaires proches entre assistantes maternelles et choisir une unité suffisamment grande pour concentrer les flux de garde en son intérieur. C'est pourquoi le bassin de vie a été conservé comme unité géographique de référence.

Ainsi, pour les raisons précédemment exposées, ce document étudie un modèle de concurrence entre assistantes maternelles dérivé de celui de Cournot. Cela suppose principalement

que le salaire horaire  $w_{i,s,t}$  individuel de l'assistante maternelle  $i$  sera égal dans le modèle au salaire horaire moyen  $w_{s,t}$  des assistantes maternelles du bassin de vie  $s$ .

### 1.2.2. Jeu entre les acteurs

3 questions précédemment énoncées concernant l'établissement d'un contrat important dans cette étude :

- L'assistante maternelle accepte-t-elle de garder un ou plusieurs enfants supplémentaires ?
- Quel salaire horaire sera fixé dans le contrat ?
- Combien d'heures de garde par mois seront effectuées par l'assistante maternelle ?

Pour répondre à ces questions, il est nécessaire d'explicitier le jeu représentant l'interaction entre assistantes maternelles et employeurs. Les hypothèses suivantes ont été réalisées : les assistantes maternelles choisissent en amont le nombre d'enfants qu'elles gardent mais les parents choisissent en aval le nombre d'heures de garde. Ceci rejoint l'étude de Cartier, d'Halluin, Rousseau et Lechien (Cartier, d'Halluin, Rousseau, Lechien, 2012) qui estiment que certaines assistantes maternelles peuvent plus facilement choisir les employeurs : « ce sont les parents appartenant aux classes moyennes du secteur public qui vont être les employeurs privilégiés de ces assistantes maternelles ». Ce choix des employeurs par les assistantes maternelles montre que, connaissant la demande des parents concernant les horaires de garde, les assistantes maternelles acceptent de garder leur enfant ou refusent car les contraintes horaires sont trop importantes. Elles peuvent alors préférer garder l'enfant d'un autre employeur avec des horaires de garde allégés. Les assistantes maternelles choisissent ainsi leurs horaires de garde en choisissant les enfants qu'elles gardent.

Un jeu en plusieurs étapes est donc envisagé.

- I. Choix de l'assistante maternelle de participer ou non au marché : détermination du nombre d'assistantes maternelles actives
- II. Choix du nombre d'enfants gardés par les assistantes maternelles
- III. Choix des parents de participer ou non
- IV. Choix par les parents du nombre d'heures de garde pour leur enfant

Une étape de négociation salariale est inutile : le salaire horaire  $w_{s,t}$  se fixera de lui-même d'après le marché.

### 1.3. Résolution du modèle dans une situation à tranche d'âge unique

#### 1.3.1. Étapes de la résolution

Compte-tenu des étapes précédemment définies, le jeu doit être résolu en partant de la dernière étape puis il s'agira de remonter progressivement jusqu'à la première. Désormais, sans perte de généralité, l'indice temporel  $t$  sera systématiquement omis des notations.

#### *Résolution de l'étape du choix d'heures côté employeurs*

Chaque employeur a une fonction d'utilité qu'il maximise selon le nombre d'heures de garde  $h_{j,s}$ . Il en ressortira donc que  $h_{j,s} = F(w_s)$  dans un premier temps puisque chaque employeur fait garder un unique enfant par hypothèse développée précédemment.

#### *Résolution de l'étape du choix de participation côté employeurs*

Un employeur ayant des enfants en âge d'être gardés va prendre en considération le nombre d'heures de garde  $h_{j,s} = F(w_s)$  optimal pour lui. Il va alors comparer d'une part son utilité à faire garder son enfant tout en en déduisant ses coûts avec d'autre part son seuil de participation  $k_{j,s}$ . La probabilité de participation du ménage s'exprime alors en fonction de  $w_s$ . En agrégeant, la demande

de garde  $q_s$  qui s'exprime en fonction de  $w_s$  est obtenue. Cela permet de déduire la fonction de demande inverse  $w_s = G(q_s)$ .

### Résolution de l'étape du choix du nombre d'enfants côté assistantes maternelles

Par symétrie des ménages, il sera supposé que  $h_{j,s}$  est indépendant de l'employeur. Donc l'assistante maternelle se voit imposer un horaire de garde  $h_{i,s} = h_{j,s}$  par l'employeur identique pour tous ses employeurs. L'assistante maternelle en tient compte et sait également que  $w_s = G(q_s)$ . Elle maximise donc sa fonction de profit qui peut se réécrire comme  $H(q_{i,s}, q_s)$  et donc comme  $J(q_{i,s}, q_{-i,s})$ , ce qui lui permet de choisir son nombre d'enfants gardés  $q_{i,s}$  en supposant  $q_{-i,s}$  connu où  $q_{-i,s}$  est la demande de garde auprès des autres assistantes maternelles. Cela permet de déduire  $q_{i,s}$  et donc  $q_s$  puis  $w_s$ .

Enfin, la première étape du jeu défini dans la section 1.2.2 sera considérée exogène. Cela signifie que le nombre d'assistantes maternelles  $N_{s,assmat}$  participant au marché du travail dans un bassin de vie sera supposé donné. L'implantation des assistantes maternelles sur certains bassins de vie plutôt que d'autre ne sera donc pas expliquée par cette étude. Dans le cas contraire, il aurait fallu définir un seuil de participation des assistantes maternelles au marché de la garde.

### 1.3.2. Résolution du modèle théorique

#### Choix du nombre d'heures de garde par les parents

Par hypothèse posée dans la section 1.1.4, le modèle se limite à une tranche d'âge pour les enfants. L'utilité des employeurs est ainsi celle définie comme en 1.1.4 :

$$u_{j,s} = \phi_1 \mathbb{1}\{h_{j,s} \leq tt_{j,s} + traj_{j,s}\} h_{j,s}^{r_1} + \mathbb{1}\{h_{j,s} > tt_{j,s} + traj_{j,s}\} [\phi_1 (tt_{j,s} + traj_{j,s})^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{r_2}] + \phi_2 \mathbb{1}\{h_{j,s,t} > tt_{j,s} + traj_{j,s}\} h_{j,s}^{r_2}$$

D'autre part, les employeurs font face à un coût  $\int q_{i,j,s} h_{i,j,s} w_{i,j,s} di$ . Or, 98 % des familles n'ont recours qu'à une assistante maternelle (Vroylandt, 2016). Comme il a été également supposé que l'employeur ne faisait garder qu'un enfant pour simplifier le modèle, cela signifie que  $q_{i,j,s} = 0$  ou 1. De plus, comme le salaire horaire est identique sur tout le bassin de vie par hypothèse du modèle de Cournot,  $w_{i,j,s} = w_s \forall i, j$ . Enfin, comme l'employeur ne fait garder son enfant que par une assistante maternelle, le nombre d'heures de garde total  $h_{j,s}$  que paye l'employeur est aussi le nombre d'heures de garde  $h_{i,j,s}$  de cet enfant par l'unique assistante maternelle employée. Donc les coûts de l'employeur sont  $h_{j,s} w_s$ .

La fonction que les employeurs maximisent est donc :

$u_{j,s} - h_{j,s} w_s = \phi_1 \mathbb{1}\{h_{j,s} \leq tt_{j,s} + traj_{j,s}\} h_{j,s}^{r_1} + \mathbb{1}\{h_{j,s} > tt_{j,s} + traj_{j,s}\} [\phi_1 (tt_{j,s} + traj_{j,s})^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{r_2}] + \phi_2 \mathbb{1}\{h_{j,s,t} > tt_{j,s} + traj_{j,s}\} h_{j,s}^{r_2} - h_{j,s} w_s$	(3)
--	-----

Les employeurs ont un coût linéaire en  $h_{j,s}$ . Étant donné que l'utilité est une fonction continue par morceaux, il n'est pas possible de la maximiser aisément. Les différents cas possibles sont donc analysés graphiquement.

Il convient d'analyser les graphiques d'utilité et de coûts des figures 3 et 4 en se référant à la situation particulière de l'étude. Pour les employeurs, faire garder un enfant durant leur temps de travail et leur temps de trajet est très utile. Par contre, quand ils sont rentrés de leur travail, l'utilité marginale qu'ils retirent d'une heure de garde supplémentaire de leur enfant diminue brusquement.

La dérivée de l'utilité des employeurs est donc discontinue en un point : celui où le nombre d'heures de garde mensuelle est égal à leur temps de travail cumulé à leur temps de trajet. Une situation comme celle du cas 2 de la figure 3, où l'utilité marginale augmente en ce point de discontinuité est donc irréaliste. Donc le cas 2 ne correspond pas à une situation concrète dans le cadre de l'étude. Par contre, le cas 1 de la figure 3 peut se produire et signifie simplement que l'employeur n'a pas intérêt à faire garder ses enfants par une assistante maternelle, du fait de coûts trop élevés.

Dans le cas 3 de la figure 4, le maximum est atteint au niveau de la discontinuité de la fonction d'utilité. Les parents ont donc intérêt à faire garder leur enfant durant la somme de leur temps de trajet et de leur temps de travail. Dans le cas 4, l'employeur a intérêt à faire garder son enfant un peu plus longtemps que le cumul de son temps de travail et de son temps de trajet. Concrètement, cette situation intervient lorsque les parents souhaitent conserver du temps de « loisir »<sup>13</sup>. Le cas 1 est le même que celui de la figure 3. Il n'y a donc que 3 maxima possibles : le parent ne fait pas garder son enfant, le parent fait garder son enfant sur son temps de travail cumulé à son temps de trajet, le parent fait garder son enfant sur le maximum de la deuxième partie de sa fonction d'utilité dont on déduit la fonction de coûts.

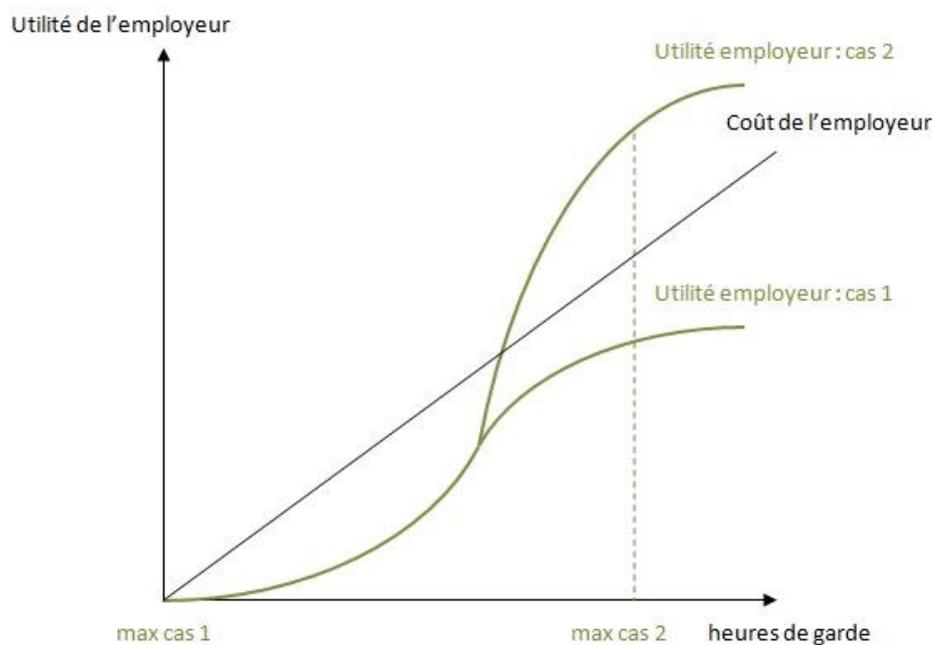


Figure 3 - Situation possible et situation impossible pour la garde d'un enfant

<sup>13</sup> Ce qui suppose par ailleurs que si le coût était nul, le parent ferait garder son enfant 24h/24. Cette situation, digne d'un conte de Perrault en un peu plus cynique, met en exergue les limites du modèle économique.



$$\begin{aligned}
& P(u_{j,s} - h_{j,s} w_s \geq k_{j,s}) \\
& = P(\phi_1 (tt_s + traj_s)^{r_1} - (tt_s + traj_s) w_s \geq \eta_0 + \eta_1 \lambda_s + \eta_2 \mu_s + \eta_3 med_s + \varepsilon_{j,s}) \\
& = P(\varepsilon_{j,s} \leq \phi_1 (tt_s + traj_s)^{r_1} - (tt_s + traj_s) w_s - \eta_0 - \eta_1 \lambda_s - \eta_2 \mu_s - \eta_3 med_s) \\
& = \phi_1 (tt_s + traj_s)^{r_1} - (tt_s + traj_s) w_s - \eta_0 - \eta_1 \lambda_s - \eta_2 \mu_s - \eta_3 med_s
\end{aligned}$$

La probabilité que les parents préfèrent la garde par assistante maternelle à la solution où l'un d'eux ne travaille pas est donc calculée. Or, il a été supposé que préférer une garde avec cessation d'activité professionnelle à un mode de garde payant revenait à préférer cet arrêt d'activité à tous les modes de garde payants. La probabilité calculée est donc la probabilité de préférer un mode d'accueil payant à une inactivité. Il faut donc maintenant introduire la préférence  $\alpha_s$  des employeurs potentiels du bassin de vie pour le mode de garde par assistante maternelle aux autres modes de garde payants.  $\alpha_s$  avait été supposé exogène dans le paragraphe « Seuil de participation  $k_{j,s,t}$  des employeurs » de la section 1.1.4. En sommant sur le bassin de vie, la demande de garde suivante est obtenue :

$$q_s = N_{s,enfants} \alpha_s (\phi_1 (tt_s + traj_s)^{r_1} - (tt_s + traj_s) w_s - \eta_0 - \eta_1 \lambda_s - \eta_2 \mu_s - \eta_3 med_s) \quad (6)$$

car  $q_s = N_{s,enfants} \alpha_s P(u_{j,s} - h_{j,s} w_s \geq k_{j,s})$ . Il s'agit ici d'une approximation puisque tous les bassins de vie ne se prêtent pas à la loi des grands nombres. Donc à salaire horaire fixé de l'assistante maternelle, comme  $\eta_1 < 0$ ,  $\eta_2 > 0$  et  $\eta_3 < 0$ , quand les conditions économiques sont favorables à l'emploi dans un bassin de vie (taux d'activité élevé, taux de chômage faible et salaire médian élevé), la demande de garde est plus forte.

#### Détermination du nombre d'enfants gardés par les assistantes maternelles

Les assistantes maternelles maximisent leur profit par le nombre d'enfants gardés. La fonction de profit des assistantes maternelles est (cf. annexe C pour les détails) :

$$\begin{aligned}
\Pi_{i,s} = q_{i,s} h_{i,s} w_{i,s} - \beta_1 \mathbb{1} \left\{ \frac{\int_s h_{i,j,s} q_{i,j,s} dj}{\int_s jours_{i,j,s} q_{i,j,s} dj} \geq 8 \right\} & \left( \frac{\int_s h_{i,j,s} q_{i,j,s} dj}{\int_s jours_{i,j,s} q_{i,j,s} dj} - 8 \right)^{p_1} - \beta_2 q_{i,s}^{p_2} \\
& - \beta_6 conc_s q_{i,s} - \beta_7 age_{i,s}^{p_3} q_{i,s} - \beta_8 age_{i,s}^{p_4} h_{i,s} - \beta_9 dens_{i,s} q_{i,s} - \beta_{10} exp_{i,s} q_{i,s} \\
& - \beta_{11} loyer_{i,s} q_{i,s} - F_{i,s}
\end{aligned}$$

Afin d'explicitier une solution théorique, le coût de pénibilité horaire faisant intervenir une indicatrice et une intégrale doit être simplifié. Il est donc approximé que ce coût de pénibilité horaire intervient au niveau du nombre d'heures  $h_{i,s}$  effectuées par mois et par enfant. La fonction de profit se réécrit alors :

$$\begin{aligned}
\Pi_{i,s} = q_{i,s} h_{i,s} w_{i,s} - \beta_1 h_{i,s}^{p_1} - \beta_2 q_{i,s}^{p_2} - \beta_6 conc_s q_{i,s} - \beta_7 age_{i,s}^{p_3} q_{i,s} - \beta_8 age_{i,s}^{p_4} h_{i,s} \\
- \beta_9 dens_{i,s} q_{i,s} - \beta_{10} exp_{i,s} q_{i,s} - \beta_{11} loyer_{i,s} q_{i,s} - F_{i,s}
\end{aligned}$$

Or, par hypothèse, les salaires horaires des assistantes maternelles sont identiques sur un bassin de vie. Donc  $w_{i,s} = w_s \forall i$ . Par ailleurs, les assistantes maternelles se voient imposer leur temps de garde par enfant puisque chaque parent souhaite que son enfant soit gardé durant  $h_{j,s} = tt_s + traj_s$  heures. Comme chaque enfant est gardé par une seule assistante maternelle par hypothèse et comme tous les parents sont supposés faire garder leur enfant pendant la même durée, une assistante maternelle garde en moyenne chaque enfant durant  $h_{i,s} = tt_s + traj_s$ . La relation précédente se réécrit donc :

$$\begin{aligned}
\Pi_{i,s} = q_{i,s} (tt_s + traj_s) w_s - \beta_1 (tt_s + traj_s)^{p_1} - \beta_2 q_{i,s}^{p_2} - \beta_6 conc_s q_{i,s} - \beta_7 age_{i,s}^{p_3} q_{i,s} \\
- \beta_8 age_{i,s}^{p_4} (tt_s + traj_s) - \beta_9 dens_{i,s} q_{i,s} - \beta_{10} exp_{i,s} q_{i,s} - \beta_{11} loyer_{i,s} q_{i,s} \\
- F_{i,s}
\end{aligned}$$

La suite de la maximisation du profit de l'assistante maternelle est développée en annexe D. Afin de simplifier la lecture du document, les notations suivantes sont adoptées.  $g_{i,s} = \beta_6 conc_s + \beta_7 age_{i,s}^{p_3} + \beta_9 dens_{i,s} + \beta_{10} exp_{i,s} + \beta_{11} loyer_{i,s}$  sont les coûts des assistantes maternelles et  $D_s = \eta_0 + \eta_1 \lambda_s + \eta_2 \mu_s + \eta_3 med_s$  sont les effets de demande dus à l'arbitrage des employeurs entre travailler et garder eux-mêmes leurs enfants. D'après l'équation (21) en annexe D obtenue en combinant l'équation (6) de demande de garde avec le profit  $\Pi_{i,s}$  des assistantes maternelles maximisé ensuite selon  $q_{i,s}$ , le nombre d'enfants  $q_s$  gardés par bassin de vie est :

$$q_s = \frac{1}{\frac{N_{s,assmat} + 1}{N_{s,enfants} \alpha_s} + 2 \beta_2} \left( \phi_1 (tt_s + traj_s)^{r_1} N_{s,assmat} - D_s N_{s,assmat} - \sum_i g_{i,s} \right) \quad (7)$$

#### 1.4. Analyse des prédictions du modèle

L'objectif de cette section est d'interpréter les équations (7), (6) et (4) permettant respectivement de calculer  $q_s$ ,  $w_s$  et  $h_s$  :

$$\left\{ \begin{array}{l} \frac{q_s}{N_{s,assmat}} = \frac{1}{\frac{N_{s,assmat} + 1}{N_{s,enfants} \alpha_s} + 2 \beta_2} \left( \phi_1 (tt_s + traj_s)^{r_1} - D_s - \sum_i \frac{g_{i,s}}{N_{s,assmat}} \right) \\ (tt_s + traj_s) w_s = \phi_1 (tt_s + traj_s)^{r_1} - \frac{q_s}{N_{s,enfants} \alpha_s} - D_s \\ \frac{h_s}{N_{s,assmat}} = (tt_s + traj_s) \frac{q_s}{N_{s,assmat}} \end{array} \right.$$

##### 1.4.1. Les effets de demande

Les effets de demande interviennent principalement à travers le seuil de participation  $\eta_0 + \eta_1 \lambda_s + \eta_2 \mu_s + \eta_3 med_s + \varepsilon_{j,s}$  défini dans l'équation (2) en section 1.1.4. Ces variables de demande sont contenues dans  $D_s = \eta_0 + \eta_1 \lambda_s + \eta_2 \mu_s + \eta_3 med_s$ . Si  $D_s$  diminue, le nombre d'enfants gardés par assistante maternelle et le salaire horaire des assistantes maternelles augmentent. Or  $D_s$  diminue si la conjoncture économique est favorable dans le bassin de vie (taux d'activité élevé, taux de chômage faible, salaire médian élevé). Donc le nombre d'enfants gardés dans un bassin de vie et le salaire horaire moyen des assistantes maternelles du bassin de vie dépendent positivement du taux d'activité et du salaire médian du bassin de vie et négativement du taux de chômage. C'est rassurant car les bassins de vie où l'activité économique est florissante voient à la fois plus de ménages participer au monde du travail ( $D_s$  faible) et donc aussi plus d'enfants gardés par assistante maternelle  $\frac{q_s}{N_{s,assmat}}$  puisque la demande de garde est plus forte. De plus, la demande étant alors élevée, les assistantes maternelles bénéficient de conditions favorables de négociation et ont des salaires horaires  $w_s$  plus élevés.

La demande intervient également à travers le temps de travail  $tt_s$  et le temps de trajet  $traj_s$ . Plus ils sont élevés, plus la demande de garde augmente. Cela semble plutôt naturel car les besoins de garde sont alors plus importants. L'effet de ces deux variables sur le salaire horaire est plus ambigu. Par contre, naturellement, plus le temps de trajet moyen et le temps de travail moyen sont élevés dans le bassin de vie, plus les besoins horaires de garde des employeurs sont élevés et plus les assistantes maternelles travaillent d'heures par enfant gardé  $\frac{h_s}{N_{s,assmat}}$ .

Dans l'équation du nombre d'enfants gardés  $\frac{q_s}{N_{s,assmat}}$  par assistante maternelle, les deux derniers effets demande interviennent au dénominateur à travers le nombre d'enfants total en âge d'être gardés  $N_{s,enfants}$  et la préférence  $\alpha_s$  pour le mode de garde par assistante maternelle. Tout naturellement, quand ces deux paramètres augmentent à nombre d'assistantes maternelles fixé, la demande de garde adressée aux assistantes maternelles augmente et le nombre d'enfants gardés par assistante maternelle également.

#### 1.4.2. Les coûts des assistantes maternelles

Les coûts moyens des assistantes maternelles  $g_{i,s}$  interviennent dans le nombre d'enfants gardés par assistante maternelle  $\frac{q_s}{N_{s,assmat}}$ . S'ils sont élevés (si la concurrence est forte, si les assistantes maternelles sont inexpérimentées, si les assistantes maternelles sont isolées, si les loyers sont élevés) alors le nombre d'enfants gardés par assistante maternelle diminue. Cet effet est un effet d'offre. Ce résultat suppose que le salaire horaire soit suffisamment élevé pour dépasser les coûts horaires des assistantes maternelles et pour qu'elles participent donc au marché de garde des enfants. Il sera supposé que cela est vérifié car le nombre d'assistantes maternelles a été considéré exogène et la participation des assistantes maternelles à ce marché n'a pas été étudiée.

#### 1.4.3. Les effets endogènes

Outre les effets de demande et d'offre précédemment mis en valeur, le nombre d'heures travaillées par assistante maternelle  $\frac{h_s}{N_{s,assmat}}$  dépend positivement du nombre d'enfants gardé par assistante maternelle  $\frac{q_s}{N_{s,assmat}}$ . C'est logique car plus l'assistante maternelle garde d'enfants, plus elle travaille d'heures.

D'autre part, l'équation (5) de participation des employeurs au marché de la garde implique que le salaire horaire moyen  $w_s$  des assistantes maternelles du bassin de vie dépend négativement du nombre d'enfants gardés par assistante maternelle.

## 2. Modélisation économétrique adaptée à l'étude

La partie précédente a permis de modéliser théoriquement un fonctionnement possible du marché de la garde d'enfants par les assistantes maternelles. De cette modélisation ont été déduites des expressions pour le nombre d'enfants  $\frac{q_s}{N_{s,assmat}}$  gardés par assistante maternelle du bassin de vie, pour le salaire horaire moyen  $w_s$  des assistantes maternelles sur un bassin de vie et pour le nombre moyen d'heures de garde  $\frac{h_s}{N_{s,assmat}}$  par assistante maternelle du bassin de vie.

L'objectif initial de cette étude est de comprendre les déterminants du salaire des assistantes maternelles. Le modèle théorique développé a ainsi mis en avant des effets attendus de certaines variables sur le salaire des assistantes maternelles. Ces effets ont été détaillés dans la section 1.4. L'analyse économétrique qui suit permettra de tester ces effets attendus empiriquement.

### 2.1. Modélisation économétrique à équations simultanées à variables endogènes

Comme précédemment montré dans l'équation (4), le choix du nombre d'heures de garde par assistante maternelle n'est pas indépendant du nombre d'enfants qu'elle garde. De même, déduit de l'équation (6), le salaire des assistantes maternelles dépend du nombre d'enfants gardés.

Or le nombre d'enfants gardés par assistante maternelle est déterminé par le modèle dans l'équation (7). Dès lors, cette variable ne peut être considérée exogène. Si cette variable était traitée comme exogène dans un simple modèle OLS (Ordinary Least Square), les coefficients de régression seraient biaisés du fait de l'endogénéité de  $\frac{q_s}{N_{s,assmat}}$ .

Dans une telle situation, le modèle économétrique adapté est un modèle à équations simultanées. Ce modèle prend en compte dans les équations de régression l'endogénéité des variables qui sont déterminées par une autre équation de régression. C'est exactement le cas ici.

### 2.1.1. Forme structurelle

Il existe plusieurs modélisations possibles dans un modèle économétrique à équations simultanées. Dans la section 15.2.3 de son livre (Greene, 2003), Greene distingue la forme structurelle du modèle et la forme réduite. La forme structurelle du modèle regroupe les équations déduites de la théorie. Comme l'économétrie qui sera développée dans ce document est de l'économétrie linéaire, les équations (7), (6) et (4) donnent respectivement la forme structurelle suivante :

$$\begin{aligned} \frac{q_s}{N_{s,assmat}} = & \gamma_0^q + \gamma_1^q \frac{N_{s,assmat}}{N_{s,enfants} \alpha_s} + \gamma_2^q tt_s + \gamma_3^q traj_s + \gamma_4^q \lambda_s + \gamma_5^q \mu_s + \gamma_6^q med_s \\ & + \gamma_7^q \frac{\sum_i conc_s}{N_{s,assmat}} + \gamma_8^q \frac{\sum_i dens_{i,s}}{N_{s,assmat}} + \gamma_9^q \frac{\sum_i exp_{i,s}}{N_{s,assmat}} + \gamma_{10}^q \frac{\sum_i loyer_{i,s}}{N_{s,assmat}} \\ & + \gamma_{11}^q \frac{\sum_i age_{i,s}}{N_{s,assmat}} + \gamma_{12}^q \frac{\sum_i age_{i,s}^2}{N_{s,assmat}} + \varepsilon_{s,q} \end{aligned} \quad (8)$$

$$\begin{aligned} w_s = & \gamma_0^w + \gamma_1^w \frac{N_{s,assmat}}{N_{s,enfants} \alpha_s} + \gamma_2^w tt_s + \gamma_3^w traj_s + \gamma_4^w \lambda_s + \gamma_5^w \mu_s + \gamma_6^w med_s \\ & + \gamma_7^w \frac{q_s}{N_{s,assmat}} + \varepsilon_{s,w} \end{aligned} \quad (9)$$

$$\frac{h_s}{N_{s,assmat}} = \gamma_0^h + \gamma_1^h \frac{q_s}{N_{s,assmat}} + \gamma_2^h tt_s + \gamma_3^h traj_s + \varepsilon_{s,h} \quad (10)$$

$\varepsilon_{s,q}$ ,  $\varepsilon_{s,w}$  et  $\varepsilon_{s,h}$  sont les résidus structurels du modèle. Greene distingue de cette forme structurelle la forme réduite du modèle. Dans la forme structurelle, les variables endogènes interviennent directement parmi les variables explicatives des autres variables endogènes. C'est le cas ici de  $\frac{q_s}{N_{s,assmat}}$  qui est une variable explicative des équations déterminant  $w_s$  et  $\frac{h_s}{N_{s,assmat}}$ . Au contraire, dans la forme réduite du modèle, le système d'équations est inversé afin que seules les variables exogènes interviennent parmi les variables explicatives des variables endogènes. Concrètement, cela reviendrait à remplacer  $\frac{q_s}{N_{s,assmat}}$  dans les équations déterminant  $w_s$  et  $\frac{h_s}{N_{s,assmat}}$  par son expression définie dans l'équation (8).

Ainsi, dans la forme réduite,  $\frac{q_s}{N_{s,assmat}}$  n'intervient pas dans l'expression de  $w_s$  mais l'inversion du système d'équations fait intervenir dans  $w_s$  les variables  $\frac{\sum_i conc_s}{N_{s,assmat}}$ ,  $\frac{\sum_i dens_{i,s}}{N_{s,assmat}}$ ,  $\frac{\sum_i exp_{i,s}}{N_{s,assmat}}$ ,  $\frac{\sum_i loyer_{i,s}}{N_{s,assmat}}$ ,  $\frac{\sum_i age_{i,s}}{N_{s,assmat}}$ ,  $\frac{\sum_i age_{i,s}^2}{N_{s,assmat}}$  et  $\varepsilon_{s,q}$ .

La forme structurelle traduit les effets directs, même endogènes, qui influent sur les variables d'intérêt. La forme réduite ne fait, au contraire, intervenir que les effets exogènes même quand ceux-ci sont indirects et n'ont en réalité d'influence qu'à travers un effet endogène. L'étude s'intéressant uniquement aux effets directs influençant les salaires des assistantes maternelles, seule la forme structurelle présentée dans le système d'équations (8), (9) et (10) sera utilisée par la suite.

### 2.1.2. Identification du modèle

Un modèle à équations endogènes n'est pas nécessairement identifiable, c'est-à-dire que le calcul des coefficients des régressions n'est pas toujours possible ou alors il peut exister plusieurs solutions.

Dans sa section 15.3.1, Greene (Greene, 2003) donne les conditions nécessaires et suffisantes à l'identification du modèle. Dans ce contexte précis, comme il n'y a qu'une variable endogène  $\frac{q_s}{N_{s,assmat}}$  intervenant dans les équations des autres variables endogènes, il faut et il suffit que des variables exogènes expliquant  $\frac{q_s}{N_{s,assmat}}$  dans l'équation (8) n'interviennent pas dans l'équation (9) et dans l'équation (10). Or, les variables  $\frac{\sum_i conc_s}{N_{s,assmat}}$ ,  $\frac{\sum_i dens_{i,s}}{N_{s,assmat}}$ ,  $\frac{\sum_i exp_{i,s}}{N_{s,assmat}}$ ,  $\frac{\sum_i loyer_{i,s}}{N_{s,assmat}}$ ,  $\frac{\sum_i age_{i,s}}{N_{s,assmat}}$ ,  $\frac{\sum_i age_{i,s}^2}{N_{s,assmat}}$  expliquent  $\frac{q_s}{N_{s,assmat}}$  dans l'équation (8) mais ne sont présentes ni dans l'équation (9), ni dans l'équation (10). Donc ces variables exogènes jouent le rôle d'instruments dans les équations (9) et (10) et permettent d'identifier le modèle.

Comme les conditions d'identification nécessaires et suffisantes au calcul des coefficients du système d'équations (8), (9) et (10) sont vérifiées, les régressions économétriques peuvent être présentées.

### 2.1.3. Premiers résultats économétriques

La construction des variables est explicitée dans l'annexe E.

#### *La colinéarité des variables*

Les résultats de la régression du système d'équations (8), (9) et (10) sont notés ci-dessous dans le tableau 2.

Les coefficients des variables de concurrence et de taux d'offre sont très élevés par rapport aux autres et le  $R^2$  ajusté est très proche de 1. Il est possible que cela traduise un problème de colinéarité. En effet :

$$\begin{aligned} \log\left(\frac{q_s}{N_{s,assmat}}\right) &= \log(q_s) - \log(N_{s,assmat}) \\ &\approx \log(\text{Capacité de garde des assistantes maternelles}) - \log(N_{s,assmat}) \\ &\approx \log(\text{Capacité de garde sur le bassin de vie}) - \log(N_{s,assmat}) \\ &\approx \log\left(\frac{\text{Capacité de garde sur le bassin de vie}}{N_{s,enfants}}\right) + \log(N_{s,enfants}) \\ &\quad - \log\left(\frac{N_{s,assmat}}{\alpha_s N_{s,enfants}}\right) - \log(\alpha_s N_{s,enfants}) \\ &\approx \log(conc_s) - \log(\text{taux d'offre}) - \log(\alpha_s) \end{aligned}$$

Certes ces égalités sont très approximées mais elles donnent une bonne idée de la relation quasiment colinéaire entre les variables de nombre d'enfants gardés par assistante maternelle  $\frac{q_s}{N_{s,assmat}}$ , de taux d'offre  $\frac{N_{s,assmat}}{\alpha_s N_{s,enfants}}$  et de concurrence  $conc_s$ . Les deux dernières variables ne peuvent donc toutes deux être gardées parmi les variables explicatives. Il n'est possible d'en conserver qu'une seule. Le taux d'offre  $\frac{N_{s,assmat}}{\alpha_s N_{s,enfants}}$  intervient également dans l'équation de salaire  $w_s$  alors que la variable de concurrence n'intervient pas dans d'autres équations. En outre, la construction de la variable de concurrence repose sur plusieurs approximations dans l'annexe E.d. Cette variable sera donc supprimée dans la suite de cette étude puisque sa construction en fait une variable endogène. Elle n'a pas été remplacée par une autre variable de concurrence, par exemple le nombre de places en EAJE (établissement d'accueil du jeune enfant), car il s'agit d'après le modèle d'une variable de concurrence globale et non uniquement par rapport à un autre mode de garde.

Variabes explicatives	Coefficient
Constante	-0,3718***
Taux d'offre $\log \left( \frac{N_{s, assmat}}{\alpha_s N_{s, enfants}} \right)$	-0,9719***
$tt_s$	0,0007
$traj_s$	0,0119
$\lambda_s$	-0,0508*
$\mu_s$	0,0128
$\log (med_s)$	0,0317***
$\log (conc_s)$	0,9648***
$\log \left( \frac{\sum_i dens_{i,s}}{N_{s, assmat}} \right)$	0,0016***
$\frac{\sum_i exp_{i,s}}{N_{s, assmat}}$	0,0009
$\frac{\sum_i age_{i,s}}{N_{s, assmat}}$	0,0043*
$\frac{\sum_i age_{i,s}^2}{N_{s, assmat}}$	-0,0001**
R <sup>2</sup> ajusté	0,98

Note : \*\*\*p<0.01, \*\*p<0.05, \*p<0.1

Tableau 2 - Résultats de la régression de la variable du nombre d'enfants gardés par assistante maternelle  $\log \left( \frac{q_s}{N_{s, assmat}} \right)$

#### 2.1.4. Rejet du modèle simple à équations simultanées

En supprimant la variable de concurrence des variables explicatives, le problème de colinéarité a été résolu. Cela permet d'analyser les nouveaux résultats économétriques.

Pour que les résultats ne soient pas biaisés, il est nécessaire que  $E(\varepsilon_{s,w} | \text{variables exogènes}) = 0$  où  $\widehat{\varepsilon}_{s,w}$  sont les résidus estimés de l'équation de salaire et  $\varepsilon_{s,w}$  les résidus théoriques. La figure 5 montre cependant que  $E(\widehat{\varepsilon}_{s,w} | \text{localisation du bassin de vie}) \neq 0$ . En effet, les résidus sont quasiment tous positifs en Provence-Alpes-Côte d'Azur et très majoritairement négatifs dans le Grand Est. Bien que n'ayant pas été montré quantitativement par le calcul des espérances empiriques des résidus par région, il est évident graphiquement que les résidus estimés ne sont pas répartis aléatoirement dans l'espace.

Pour comprendre pourquoi les résidus ne sont pas aléatoires, il faut se référer à la figure 2 de la partie 1.2.1. Cette carte montre qu'une des variables endogènes, le salaire horaire, est autocorrélée positivement spatialement<sup>14</sup>. Cela peut provenir de plusieurs raisons différentes analysées ci-dessous.

<sup>14</sup> Cela signifie ici que des bassins de vie voisins ont des salaires horaires similaires.

Premièrement, il se peut que l'autocorrélation spatiale des salaires horaires provienne de l'autocorrélation spatiale d'une variable exogène. Concrètement, cela signifie que le fait que le salaire horaire d'un bassin de vie soit élevé n'est pas lié au fait que ceux des bassins de vie voisins soient élevés. Mais comme une des variables explicatives du salaire horaire aurait une distribution spatialement autocorrélée, elle induirait exactement la même distribution spatiale pour le salaire horaire. Cette hypothèse est à exclure. En effet, si tel était le cas, les résidus ne seraient pas autocorrélés spatialement puisque la prise en compte de cette variable explicative dans le modèle économétrique expliquerait toute l'autocorrélation spatiale du salaire horaire. Les résidus seraient donc en moyenne nuls dans chaque région. Ce n'est pas ce que montre la figure 5.

### Résidus de l'équation de salaire - modèle simultané

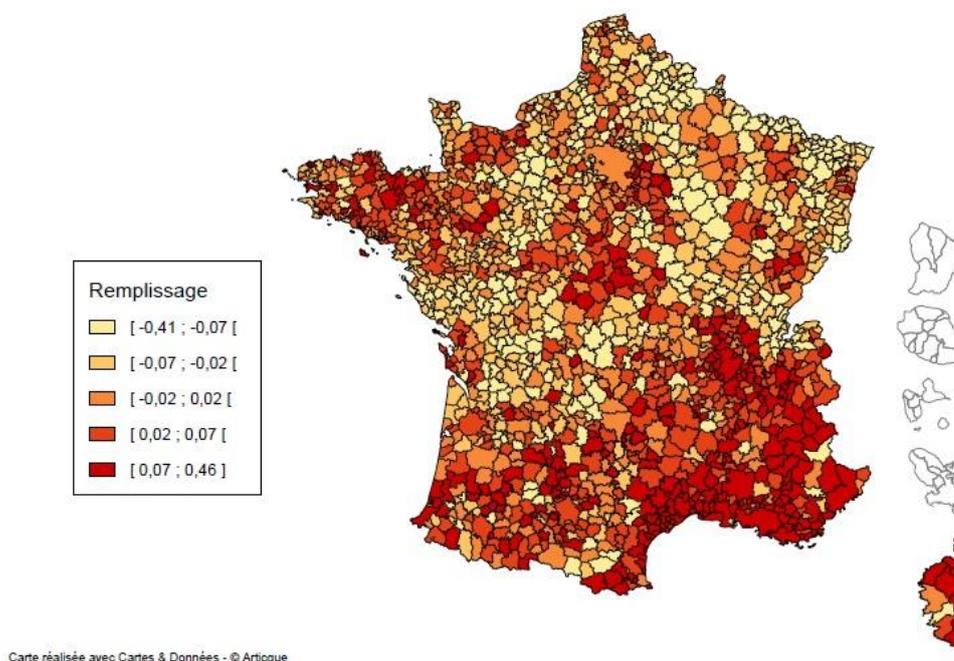


Figure 5 - Résidus de l'équation de salaire dans le modèle à équations simultanées

Deuxièmement, une variable qui serait autocorrélée spatialement a pu être omise. Son omission conduirait à avoir sur les résidus la même corrélation spatiale que celle de la variable omise. L'autocorrélation spatiale du salaire horaire et *a fortiori* des résidus pourrait ainsi être due à la variable de concurrence ou à une variable qui n'a pas été identifiée.

Troisièmement, contrairement à ce qui avait été supposé dans la section 1.1.3, les 15 % de contrats faisant intervenir un employeur et une assistante maternelle de deux bassins de vie différents ne sont pas forcément négligeables. Dès lors, trois options peuvent être envisagées. D'abord, cela peut signifier que l'unité géographique de référence choisie n'est pas la bonne. Si tel est le cas, une régression avec un modèle à équations simultanées en utilisant la bonne unité géographique permettra d'avoir des résidus qui ne sont pas autocorrélés spatialement. Cette option ne sera pas étudiée car les résidus semblent être corrélés spatialement sur de trop grands espaces et car les deux options suivantes sont privilégiées.

Ces deux options exploitent les 15 % de contrats sus-mentionnés. La première consiste à affirmer que cette perméabilité des marchés, pourtant segmentés géographiquement, conduit à

propager les salaires horaires d'un bassin de vie au bassin de vie voisin. Ainsi, les employeurs faisant garder leur enfant dans un bassin de vie voisin rendraient concurrentes les assistantes maternelles de leur bassin de vie et du bassin de vie voisin. Cela conduirait les assistantes maternelles de ces deux bassins de vie à harmoniser leur salaire horaire pour être en concurrence en nombre d'enfants gardés.

Dans la seconde option, le salaire horaire d'un bassin de vie varie en fonction de la demande et de l'offre (variables exogènes) des bassins de vie voisins. En effet, les marchés n'étant pas réellement segmentés géographiquement, une variation de la demande ou de l'offre dans les bassins de vie voisins se répercutera également sur le bassin de vie observé.

Les deux dernières options envisagées impliquent que les résidus de la figure 5 du modèle à équations simultanées sont autocorrélés car l'hypothèse supposant que les marchés de garde des assistantes maternelles sont segmentés géographiquement était trop forte. Ce sont ces options qui sont développées dans la suite de l'étude.

## 2.2. Modélisation économétrique intégrant la problématique d'autocorrélation spatiale

Afin de corriger le problème de répartition géographique non aléatoire des résidus de la figure 5, les solutions existant en économétrie spatiale sont présentées ci-dessous.

### 2.2.1. Modèles d'économétrie spatiale

Il existe différents types de modèles économétriques prenant en compte l'autocorrélation spatiale (Floch, Le Saout, 2016). Les modèles à équations simultanées et les modèles d'économétrie spatiale ne sont pas encore combinés dans cette sous-partie. Pour l'instant, les modèles d'économétrie spatiale présentés ne comportent qu'une équation.

Il existe 3 méthodes courantes pour tenir compte de l'autocorrélation spatiale en économétrie : le modèle SAR (Spatial AutoRegressive), le modèle SEM (Spatial Error Model) et le modèle SLX (Spatial Lag of X).

Le modèle SAR est un modèle d'autorégression spatiale dans lequel la variable endogène est expliquée par la variable endogène des bassins de vie voisins. Mathématiquement, cela donnerait pour le salaire horaire  $w$ , en notant  $W$  la matrice de poids spatiaux (permettant de savoir si deux bassins de vie sont proches ou voisins) et en considérant  $\frac{q_s}{N_{s,assmat}}$  exogène :

$$w = \rho^w W w + \gamma_0^w + \gamma_1^w \frac{N_{assmat}}{N_{enfants}} \alpha + \gamma_2^w tt + \gamma_3^w traj + \gamma_4^w \lambda + \gamma_5^w \mu + \gamma_6^w med + \gamma_7^w \frac{q}{N_{assmat}} + \varepsilon_w \quad (11)$$

L'indice  $s$  n'est plus présent en indice des variables car la notation doit être vectorielle en économétrie spatiale, ce qui est le cas pour l'équation (11). Dans l'étude, ce modèle correspond à l'option évoquée dans la troisième raison de la section 2.1.4 selon laquelle les salaires horaires d'un bassin de vie se propagent au bassin de vie voisin du fait des 15 % de contrats faisant intervenir une assistante maternelle et un employeur de bassins de vie différents. Les assistantes maternelles de deux bassins de vie voisins deviennent alors concurrentes.

Le modèle SEM est un modèle visant à expliquer l'autocorrélation spatiale par l'autocorrélation du terme d'erreur. Il s'écrit de la manière suivante :

(12)

avec  $\varepsilon_w = \lambda^w W \varepsilon_w + u_w$   
 où  $u_w$  est un bruit blanc ne présentant pas d'autocorrélation spatiale

Concrètement, cela permet de prendre en compte les situations où il y a des variables omises autocorrélées spatialement. Cela correspond à la deuxième raison de la section 2.1.4. Cependant, cette solution a été écartée car elle traduit un oubli dans la modélisation du marché des assistantes maternelles et n'apporte de ce point de vue aucun caractère explicatif. Elle ne servirait ici qu'à obtenir des coefficients plus fiables alors que l'objectif de l'étude est de comprendre les phénomènes sous-jacents à l'établissement des salaires des assistantes maternelles. Du point de vue de l'économétrie, cette solution reste néanmoins envisageable. Par ailleurs, lorsque les modèles SAR et SEM ont été couplés avec un modèle à équations simultanées, les coefficients  $\rho$  d'autocorrélation de la variable endogène de l'équation (11) obtenus étaient très élevés par rapport aux coefficients  $\lambda$  d'autocorrélation spatiale des erreurs de l'équation (12) qui étaient proches de 0,1. Cela légitimise la décision de privilégier le modèle SAR au modèle SEM.

Le modèle SLX est un modèle expliquant l'autocorrélation spatiale par l'influence des variables exogènes des bassins de vie voisin sur la variable endogène du bassin de vie concerné. En considérant  $\frac{q}{N_{assmat}}$  exogène :

$$\begin{aligned}
 w = & \gamma_0^w + \gamma_1^w \frac{N_{assmat}}{N_{enfants} \alpha} + \gamma_2^w tt + \gamma_3^w traj + \gamma_4^w \lambda + \gamma_5^w \mu + \gamma_6^w med \\
 & + \gamma_7^w \frac{q}{N_{assmat}} + \gamma_8^w W \frac{N_{assmat}}{N_{enfants} \alpha} + \gamma_9^w W tt + \gamma_{10}^w W traj + \gamma_{11}^w W \lambda + \gamma_{12}^w W \mu \\
 & + \gamma_{13}^w W med + \gamma_{14}^w W \frac{q}{N_{assmat}} + \varepsilon_w
 \end{aligned} \quad (13)$$

Ce modèle correspond à l'option de la troisième raison développée dans la section 2.1.4 selon laquelle le salaire horaire d'un bassin de vie varie en fonction de la demande et de l'offre (variables exogènes) du bassin de vie voisin, en plus de varier en fonction de la demande et de l'offre en son sein. Concrètement, cela signifie qu'une variation locale du taux d'activité dans un bassin de vie peut avoir une influence sur le salaire horaire des assistantes maternelles des bassins de vie voisins. Mais cette influence ne s'étendra pas au-delà des voisins directs, contrairement à ce qui sera le cas du modèle SAR et à ce qui semble être le cas sur la figure 5.

Par la suite, les modèles SAR et SLX ((11) et (13)) seront étudiés dans le cadre de l'économétrie à équations simultanées.

### 2.2.2. Choix de la matrice de poids spatiaux $W$

Les équations précédentes de la partie 2.2.1 font intervenir une matrice de poids spatiaux qui n'a pas été définie plus précisément jusqu'à présent. Elle traduit la proximité de deux bassins de vie. C'est en quelque sorte une matrice de voisins. Le choix de cette matrice est controversé dans la littérature (Floch, Le Saout, 2016). Étant donné que dans ce cas précis, les bassins de vie ne choisissent pas leurs bassins de vie voisins, il s'agit d'une situation où la matrice de poids spatiaux est définie de manière exogène.

Il existe cependant diverses possibilités comme le souligne de Bellefon (de Bellefon, 2016) : doit-on choisir une matrice de poids 0 ou 1 prenant pour valeur 1 si deux bassins de vie sont voisins et 0 sinon ? Ou cette même matrice normalisée selon les lignes ? Peut-être faudrait-il choisir une matrice de distance géographique ? Concernant cette dernière option, elle a été rejetée car deux bassins de vie éloignés ne s'influencent nullement en termes de garde par assistante maternelle. La matrice des voisins normalisée a été choisie car un bassin de vie comme le bassin de vie parisien compte une trentaine de voisins. Il serait osé d'estimer que chacun de ces voisins influence le bassin de vie parisien au même degré qu'un bassin de vie qui n'a qu'un voisin est influencé par son voisin. Compte-tenu de

cette hétérogénéité du nombre de voisins, il est donc nécessaire de normaliser la matrice des voisins. Le poids de l'influence d'un voisin  $s'$  pour un bassin de vie  $s$  ayant  $n_s$  voisins est donc  $W_{s,s'} = \frac{1}{n_s}$ . Chaque bassin de vie a donc un nombre de voisins normalisé à 1.

Ce choix d'une matrice de poids normalisés a également un intérêt dans l'interprétation des coefficients. Dans la régression d'un modèle SLX (équation (13)), l'augmentation d'une variable exogène de 1 % ou de 1 point dans tous les bassins de vie voisins du bassin de vie  $s$  se répercutera sur le salaire horaire des assistantes maternelles du bassin de vie  $s$  à hauteur du coefficient  $\gamma$  correspondant calculé dans la régression.

Cependant, aucune matrice de voisinage des bassins de vie préconstruite n'était à disposition pour cette étude. Le logiciel Metric développé par l'Insee a été utilisé pour construire cette matrice. Le temps de trajet entre tous les couples de communes du même département, de deux départements adjacents et de deux départements distants de moins d'une vingtaine de kilomètres a été calculé. Le critère de voisinage suivant a ensuite été défini : deux bassins de vie sont voisins s'ils contiennent des communes dont le temps de trajet de l'une à l'autre est inférieur ou égal à 15 minutes en heures creuses. En effet, ayant fait l'hypothèse que les parents emploient une assistante maternelle à proximité de leur domicile, un détour de plus de 15 minutes pour aller travailler est peu envisageable.

Cette définition ne repose cependant pas sur les véritables frontières entre bassins de vie et pose le problème suivant : 4 bassins de vie n'ont pas de voisins. Il s'agit de deux îles françaises, d'un bassin de vie en Corse et d'un bassin de vie dans les calanques marseillaises. Comme ces bassins de vie ont 0 voisin dans la matrice d'adjacence  $W$  quand les autres en ont 1 du fait de la normalisation, leur présence dans les observations sous-estime très largement l'influence des variables spatiales dans le modèle. Ces bassins de vie ont donc par la suite été supprimés des observations pour ne pas fausser les régressions.

### 2.2.3. Modèles SESAR et SESLX

Les modèles vus en partie 2.2.1 ne suffisent pas à décrire la situation du marché de garde des assistantes maternelles. En effet, dans les équations (11) et (13), le phénomène d'autocorrélation spatiale est pris en compte mais le caractère endogène du modèle est omis. Dans ces équations,  $\frac{q}{N_{assmat}}$  est considérée exogène puisqu'aucune équation ne détermine cette variable, contrairement à ce qui était observé dans le système d'équations (8), (9) et (10) de la section 2.1.1.

Ce système d'équations simultanées (8), (9) et (10) doit donc être combiné aux équations à autocorrélation spatiale (11) et (13). Pour ce faire, le cadre théorique développé ci-dessous repose sur le papier fondateur de Kelejian et Prucha (Kelejian, Prucha, 2004), ainsi que la présentation de ce papier à une conférence (Gebremariam, 2007) qui distingue des sous-cas tout particulièrement intéressants pour cette étude.

Dans le cas d'un modèle où l'autocorrélation spatiale est expliquée par une autocorrélation des variables endogènes, le système SESAR (Simultaneous Equations Spatial AutoRegressive) suivant<sup>15</sup>, résultant de la combinaison du système d'équations simultanées (8), (9) et (10) avec le modèle SAR de l'équation (11), est étudié :

$$\frac{q}{N_{assmat}} = \rho_q^q W \frac{q}{N_{assmat}} + \gamma_0^q + \gamma_1^q \frac{N_{assmat}}{N_{enfants}} \alpha + \gamma_2^q tt + \gamma_3^q traj + \gamma_4^q \lambda + \gamma_5^q \mu + \gamma_6^q med + \gamma_8^q \frac{\sum_i dens_i}{N_{assmat}} + \gamma_9^q \frac{\sum_i exp_i}{N_{assmat}} + \gamma_{11}^q \frac{\sum_i age_i}{N_{assmat}} + \varepsilon_q \quad (14)$$

<sup>15</sup> Du fait de la corrélation très forte entre la variable d'âge et d'âge au carré, seule la variable d'âge sera gardée dans la suite de l'étude. De même, le loyer sera omis des variables explicatives comme justifié en annexe E.b.

$$w = \rho_w^w W w + \gamma_0^w + \gamma_1^w \frac{N_{assmat}}{N_{enfants} \alpha} + \gamma_2^w tt + \gamma_3^w traj + \gamma_4^w \lambda + \gamma_5^w \mu + \gamma_6^w med + \gamma_7^w \frac{q}{N_{assmat}} + \rho_q^w W \frac{q}{N_{assmat}} + \varepsilon_w \quad (15)$$

$$\frac{h}{N_{assmat}} = \rho_h^h W \frac{h}{N_{assmat}} + \gamma_0^h + \gamma_1^h \frac{q}{N_{assmat}} + \gamma_2^h tt + \gamma_3^h traj + \rho_q^h W \frac{q}{N_{assmat}} + \varepsilon_h \quad (16)$$

où interviennent dans une équation donnée à la fois le retard spatial de la variable endogène concernée comme dans un modèle SAR, les autres variables endogènes comme dans un modèle à équations simultanées et les retards spatiaux de ces autres variables endogènes dus à la combinaison de ces deux modèles. Il s'agit du sous-cas décrit dans la section 2.1 par Gebremariam (Gebremariam, 2007). Les coefficients  $\rho$  sont les coefficients d'autocorrélation spatiale.

Dans le cas où l'autocorrélation spatiale est expliquée par les variables exogènes, le modèle SESLX (Simultaneous Equations Spatial Lag of X) suivant, résultant de la combinaison du système d'équations simultanées (8), (9) et (10) avec le modèle SLX de l'équation (13), est étudié :

$$\begin{aligned} \frac{q}{N_{assmat}} = & \gamma_0^q + \gamma_1^q \frac{N_{assmat}}{N_{enfants} \alpha} + \gamma_2^q tt + \gamma_3^q traj + \gamma_4^q \lambda + \gamma_5^q \mu + \gamma_6^q med \\ & + \gamma_8^q \frac{\sum_i dens_i}{N_{assmat}} + \gamma_9^q \frac{\sum_i exp_i}{N_{assmat}} + \gamma_{11}^q \frac{\sum_i age_i}{N_{assmat}} + \gamma_{13}^q W \frac{N_{assmat}}{N_{enfants} \alpha} + \gamma_{14}^q W tt \\ & + \gamma_{15}^q W traj + \gamma_{16}^q W \lambda + \gamma_{17}^q W \mu + \gamma_{18}^q W med + \gamma_{20}^q W \frac{\sum_i dens_i}{N_{assmat}} \\ & + \gamma_{21}^q W \frac{\sum_i exp_i}{N_{assmat}} + \gamma_{23}^q W \frac{\sum_i age_i}{N_{assmat}} + \varepsilon_q \end{aligned} \quad (17)$$

$$\begin{aligned} w = & \gamma_0^w + \gamma_1^w \frac{N_{assmat}}{N_{enfants} \alpha} + \gamma_2^w tt + \gamma_3^w traj + \gamma_4^w \lambda + \gamma_5^w \mu + \gamma_6^w med \\ & + \gamma_7^w \frac{q}{N_{assmat}} + \gamma_8^w W \frac{N_{assmat}}{N_{enfants} \alpha} + \gamma_9^w W tt + \gamma_{10}^w W traj + \gamma_{11}^w W \lambda + \gamma_{12}^w W \mu \\ & + \gamma_{13}^w W med + \varepsilon_w \end{aligned} \quad (18)$$

$$\frac{h}{N_{assmat}} = \gamma_0^h + \gamma_1^h \frac{q}{N_{assmat}} + \gamma_2^h tt + \gamma_3^h traj + \gamma_4^h W tt + \gamma_5^h W traj + \varepsilon_h \quad (19)$$

Si l'estimation du modèle SESLX (équations (17), (18) et (19)) peut se faire par les fonctions classiques utilisées pour la résolution d'un système à équations simultanées (en l'occurrence sous le logiciel SAS), il n'en est pas de même de l'estimation du modèle SESAR (équations (14), (15) et (16)). Ce modèle a été estimé grâce au package *splm* implémenté sous R et qui utilise l'estimateur proposée par Kelejian et Prucha (Kelejian, Prucha, 2004).

#### 2.2.4. Le SESAR, seul modèle dont les résidus ne sont pas autocorrélés spatialement

Les modèles SESAR (équations (14), (15) et (16)) et SESLX (équations (17), (18) et (19)) ont été implémentés pour prendre en compte l'autocorrélation spatiale mise en avant dans le modèle à équations simultanées et visible sur la figure 5 de la partie 2.1.4. Le critère principal pour choisir entre les modèles SESAR et SESLX concerne donc l'autocorrélation spatiale des résidus. Pour ne pas avoir des coefficients biaisés, il faut que  $E(\varepsilon | \text{géographie}) = 0$ . N'ayant pas utilisé d'indice d'autocorrélation spatiale, cette appréciation est purement graphique.

Les résidus obtenus par un modèle à équations simultanées, par un modèle SESLX et par un modèle SESAR doivent donc être comparés. Les figures 6 et 7 montrent que les résidus d'un modèle à équations simultanées ne prenant nullement en compte l'autocorrélation spatiale sont à peu près autant corrélés spatialement que les résidus du modèle SESLX. Combiner le modèle SLX vu dans l'équation (13) et le modèle à équations simultanées (8), (9) et (10) ne résout donc nullement le problème d'autocorrélation spatiale vu en partie 2.1.4.

Au contraire, les résidus du salaire horaire dans le modèle SESAR (équations (14), (15) et (16)) de la figure 15 en annexe G.b semblent ne pas être autocorrélés spatialement. Il en est de même pour les résidus de nombre d'enfants gardés par assistante maternelle et du nombre d'heures de garde par assistante maternelle (figures 14 et 16 en annexe G.b). Bien qu'un indicateur d'autocorrélation spatiale devrait être utilisé pour s'en assurer, le modèle SESAR semble ainsi être bien meilleur que le modèle SESLX. C'est donc le modèle SESAR qui est retenu pour la suite de l'analyse.

### Résidus de l'équation de salaire - modèle simultané

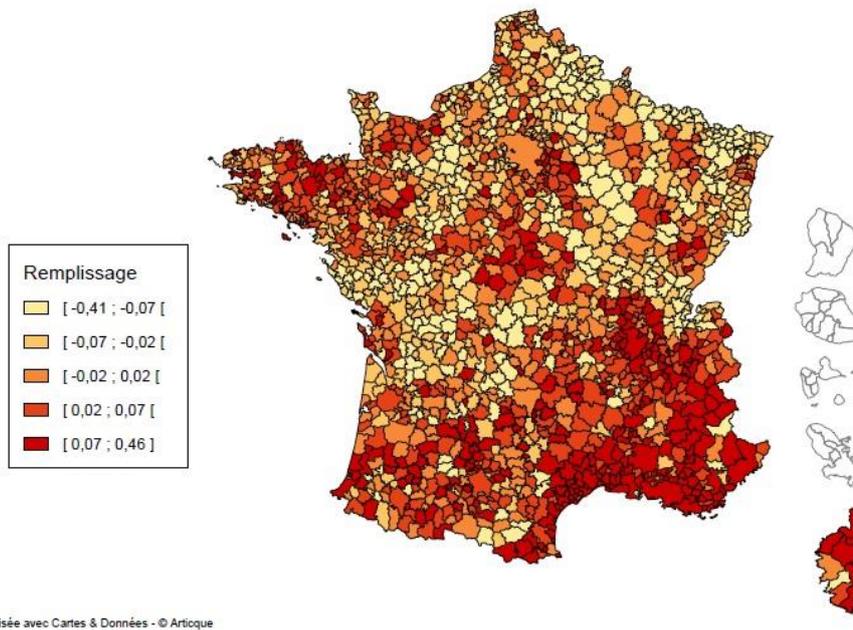


Figure 6 - Résidus de l'équation 9 de salaire du modèle à équations simultanées

### Résidus de l'équation de salaire – Modèle SESLX

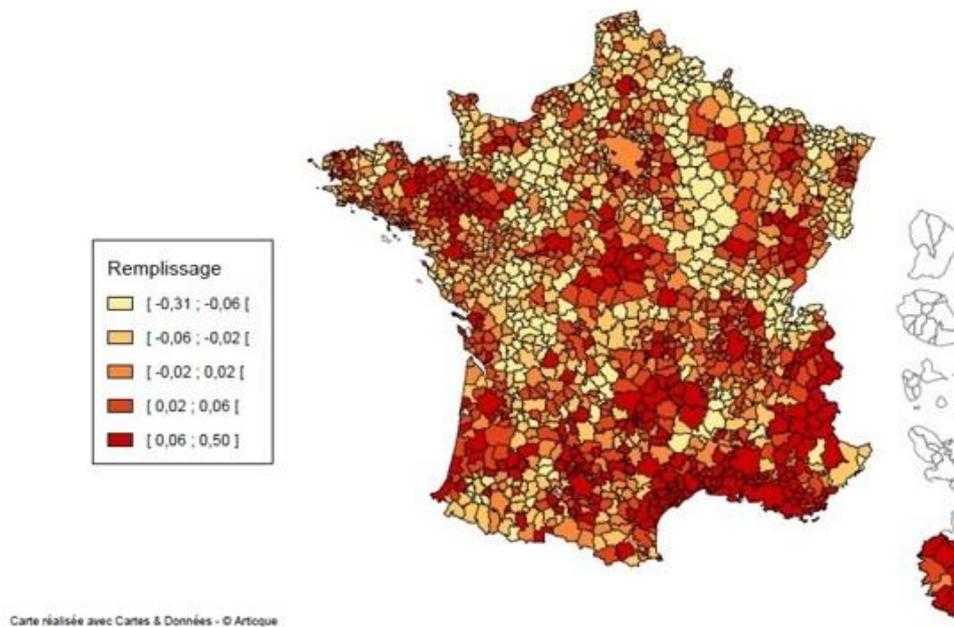


Figure 7 - Résidus de l'équation 18 de salaire du modèle SESLX

### 2.2.5. Évaluer les écarts types du modèle SESAR : un enjeu non-résolu de l'étude

Bien que le modèle SESAR soit retenu, l'utilisation du package `splm` de R ne fait pas ressortir de coefficient significatif. Plusieurs éléments peuvent l'expliquer. Dans l'estimation du SESAR, la matrice  $I_{3n} - \rho_n W_3$  est inversée. Il est par exemple possible que la matrice soit presque singulière, auquel cas cela peut expliquer des écarts types élevés. Une simulation de Monte-Carlo<sup>16</sup> a permis de montrer que, si les coefficients calculés ne sont pas biaisés, les écarts-types sont largement surestimés par la fonction R utilisée. Les résidus ont donc plutôt été calculés par une méthode de bootstrap.

$$\text{avec } \rho_n = \begin{pmatrix} \rho_q^a I_n & 0_n & 0_n \\ \rho_q^w I_n & \rho_w^w I_n & 0_n \\ \rho_q^h I_n & 0_n & \rho_h^h I_n \end{pmatrix} \text{ et } W_3 = \begin{pmatrix} W & 0_n & 0_n \\ 0_n & W & 0_n \\ 0_n & 0_n & W \end{pmatrix}$$

Une méthode de bootstrap classique en tirant avec remise 1640 bassins de vie (correspondant aux 1640 observations de régression) ne pouvait être utilisée. Procéder ainsi signifie tirer à chaque fois des bassins de vie qui n'auront pas de voisin dans les échantillons ainsi constitués. Certains n'ont effectivement que 2 voisins. Les tirer sans tirer leurs voisins revient donc à considérer qu'ils n'ont plus de voisins. Dès lors, leur présence dans l'échantillon fausse l'estimation des coefficients, au même titre que les 4 bassins de vie sans voisin faussaient l'estimation du modèle SESAR comme cela a été détaillé dans la section 2.2.2.

Ainsi, une méthode de bootstrap des résidus décrite par Efron et Tibshirani (Efron, Tibshirani, 1998) dans leur section 9 a été étudiée. Cette méthode résout le problème de réseau et de voisinage puisque tous les bassins de vie sont gardés dans les échantillons de bootstrap par cette méthode. Ainsi, tous auront des voisins. Le bootstrap consiste ici à conserver la valeur prédite par le modèle pour chaque bassin de vie et à lui associer un résidu tiré aléatoirement avec remise parmi les résidus calculés lors de la régression du modèle SESAR. Cette méthode de bootstrap des résidus permet ainsi de calculer des intervalles de confiance à 95 % largement restreints en comparaison à ceux calculés par le package `splm` de R. Néanmoins, des simulations de Monte-Carlo ont été réalisées pour vérifier si les intervalles de confiance obtenus par la méthode de bootstrap des résidus étaient proches de ceux obtenus par simulation de Monte-Carlo sur des données avec une structure d'autocorrélation spatiale similaire à celle du modèle SESAR. En conclusion de ces simulations, la méthode de bootstrap des résidus ne permet pas d'obtenir des intervalles de confiance convergents des coefficients estimés par le modèle SESAR.

Ainsi, cette étude ne propose pas de solution pour estimer la variance ou les intervalles de confiance des coefficients estimés par le modèle SESAR (équations (14), (15) et (16)). Ceci limite grandement l'analyse de la partie 2.3 puisqu'il n'est pas possible de déterminer quels coefficients sont significatifs dans l'estimation du modèle SESAR. Cela explique que les tableaux de résultats soient présentés sans écart-type et sans intervalle de confiance pour le modèle SESAR.

## 2.3. Analyse des résultats du modèle à équations simultanées avec autocorrélation spatiale des variables endogènes

### 2.3.1. Confrontation des résultats aux prévisions du modèle

Les sections 2.1.4 puis 2.2.4 ont abouti au choix du modèle SESAR d'équations simultanées avec autocorrélation spatiale des variables endogènes comme modèle le plus adapté au marché de garde des assistantes maternelles. Les résultats de l'estimation de ce modèle par la méthode explicitée par Kelejian et Prucha (Kelejian, Prucha, 2004) sont présentés dans le tableau 3.

Ces résultats sont à comparer avec les prévisions du modèle théorique développé précédemment et qui a abouti à la forme structurelle (8), (9) et (10) développée en section 2.1.1.

---

<sup>16</sup> Les simulations de Monte-Carlo évoquées dans cette sous-partie n'ont pas été intégrées à ce document d'étude.

L'étude s'intéressant principalement au salaire horaire des assistantes maternelles, les commentaires se focalisent sur l'équation (15), en ce qui concerne l'interprétation qualitative des résultats du tableau 3. Néanmoins, les résultats commentés ici sont qualitatifs étant donné qu'il n'a pas été possible d'estimer la significativité des coefficients.

Les coefficients d'autocorrélation spatiale sont très élevés, notamment en ce qui concerne le salaire horaire des assistantes maternelles dont le retard spatial a un coefficient de 0,85. Cela signifie que l'intuition formulée en section 2.1.4 selon laquelle les 15 % des employeurs faisant garder leur enfant dans un bassin de vie voisin rendraient concurrentes les assistantes maternelles de leur bassin de vie et du bassin de vie voisin est vérifiée. Cela peut signifier que le choix du bassin de vie comme unité géographique de référence en section 1.1.3 n'est pas optimal. En effet, un coefficient d'autocorrélation aussi élevé peut venir d'une trop petite unité géographique de référence (de Bellefon, 2016). Les effets d'autocorrélation spatiale sont donc potentiellement surestimés dans les résultats présentés dans le tableau 3.

Variables explicatives	Équation 14 $\log \left( \frac{q}{N_{assmat}} \right)$	Équation 15 $\log (w)$	Équation 16 $\log \left( \frac{h}{N_{assmat}} \right)$
Constante	0,0051	-2,3827	0,5834
Taux d'offre $\log \left( \frac{N_{assmat}}{\alpha N_{enfants}} \right)$	-0,1310	-0,0957	
Taux d'actifs occupés à temps partiel $tt$	0,1246	0,2502	-0,4660
Taux d'actifs à faible temps de trajet $traj$	-0,0053	0,0827	-0,3335
Taux d'actifs de 25 à 54 ans $\lambda$	0,7698	0,4041	
Taux de chômage de 25 à 54 ans $\mu$	-0,4612	0,0350	
Salaire médian $\log (med)$	-0,0564	0,2055	
Densité $\log \left( \frac{\sum_i dens_i}{N_{assmat}} \right)$	-0,0093		
Expérience $\frac{\sum_i exp_i}{N_{assmat}}$	0,0467		
Âge $\frac{\sum_i age_i}{N_{assmat}}$	-0,0024		
$\log \left( \frac{q}{N_{assmat}} \right)$		-0,5573	0,8246
$W \log \left( \frac{q}{N_{assmat}} \right)$	0,5668	0,4536	-0,6116
$W \log (w)$		0,8511	
$W \log \left( \frac{h}{N_{assmat}} \right)$			0,8757

**Note :** l'écart-type et la significativité des coefficients ne sont pas précisés car aucun estimateur de variance convergent n'a été implémenté pour le modèle SESAR.

Tableau 3 : Résultats de la régression SESAR du système d'équations 14, 15 et 16

L'importance de la demande des employeurs à travers le coefficient du salaire médian du bassin de vie (0,21) est également notable : si les employeurs potentiels ont accès à des salaires plus élevés, ils vont a priori plutôt participer au marché de la garde des assistantes maternelles car ne pas

travailler présente un coût d'opportunité. Ceci engendre alors des conditions de négociation favorables pour les assistantes maternelles qui peuvent augmenter leur salaire horaire. Ce raisonnement ne semble pas totalement fonctionner puisque le coefficient du salaire médian dans la détermination du nombre d'enfants gardés par assistante maternelle est négatif (-0,06). Cela peut s'interpréter par l'omission du loyer payé par les assistantes maternelles, justifiée en annexe E.b, parmi les variables explicatives. En effet, les loyers et le salaire du bassin de vie sont corrélés positivement et l'équation (8) de la section 2.1.1 prévoit qu'une augmentation de loyer fasse diminuer le nombre d'enfants gardés par assistante maternelle. Il y aurait donc un biais négatif dans l'estimation du coefficient du salaire médian du bassin de vie dans l'équation (14). Une autre possibilité serait que le salaire médian capte un phénomène d'offre : face à des employeurs plus riches et un coût de la vie plus élevé, l'assistante maternelle peut élever ses tarifs. Les effets de demande ont aussi une influence sur le salaire horaire des assistantes maternelles à travers le taux d'activité des 25-54 ans du bassin de vie (0,40). C'était ce qui était escompté dans la section 1.4.1. Si une assistante maternelle vit dans un bassin de vie à fort taux d'activité et à faible taux de chômage, elle fera face à une demande plus élevée (coefficients du signe attendu dans l'équation (14)), ce qui conduit indirectement à une hausse du salaire horaire des assistantes maternelles due à l'effet endogène du nombre d'enfants gardés par assistante maternelle sur le salaire horaire dans l'équation (15). Elle bénéficiera en outre directement de conditions favorables de négociation pour son salaire du fait du coefficient du taux d'activité dans l'équation de salaire horaire. Cette interprétation est toutefois conditionnée à la significativité des coefficients qui n'a pas pu être évaluée.

Il existe également un effet d'offre : à nombre d'enfants gardés par assistante maternelle donné, plus le taux d'offre de nombre d'assistantes maternelles comparé au nombre d'enfants potentiels à garder est élevé, plus le salaire horaire des assistantes maternelles diminue (-0,10 dans l'équation (15)). Ainsi, si la concurrence entre assistantes maternelles augmente, ces dernières ont moins de pouvoir de négociation et leur salaire horaire est plus faible. De même, si le nombre d'assistantes maternelles augmente ou si le nombre d'enfants pouvant être gardés par les assistantes maternelles diminue, le nombre d'enfants gardés par assistante maternelle diminue (coefficient -0,13 dans l'équation (14)).

Des effets d'offre et de demande interviennent aussi indirectement sur le salaire horaire des assistantes maternelles par l'équation (14) déterminant le nombre d'enfants gardés par assistante maternelle  $\frac{q}{N_{assmat}}$ . En effet, les effets d'offre et de demande influent sur ce nombre d'enfants gardés par les assistantes maternelles, qui lui-même a un effet sur le salaire horaire des assistantes maternelles. L'expérience des assistantes maternelles n'a par exemple pas un impact direct sur leur salaire horaire. Mais si cette expérience varie, le nombre d'enfants qu'elles gardent variera, ce qui influera sur leur salaire horaire.

Concernant l'équation (16) du nombre d'heures de garde par assistante maternelle, toutes les variables jouent dans le sens attendu et explicité dans le système d'équations (8), (9) et (10) de la section 2.1.1. Les sens qui ne correspondent pas proviennent simplement de la manière dont ont été construites les différentes variables dans l'annexe E. Les variables  $tt$  et  $traj$  sont en fait décroissantes en le temps de travail et le temps de trajet des employeurs, contrairement à ce qui était le cas dans la spécification du modèle de la partie 1. Si le coefficient de la variable de temps de trajet a le sens escompté, l'impossibilité de montrer sa significativité dans l'équation du nombre d'heures ne permet pas de valider l'hypothèse selon laquelle les employeurs font garder leur enfant à proximité de chez eux plutôt qu'à proximité de leur lieu de travail. De plus, du fait de la construction de cette variable comme le taux d'actifs occupés se rendant au travail à vélo ou à pied, il est aussi possible que cette variable capte un effet géographique comme le fait qu'un bassin de vie est plutôt rural ou urbain.

### 2.3.2. L'importance de la propagation spatiale

Dans le tableau 3, les coefficients d'autocorrélation spatiale des endogènes sur elles-mêmes sont élevés voire proches de 1 : 0,57 pour le nombre d'enfants gardés par assistante maternelle, 0,85

pour le salaire horaire des assistantes maternelles et 0,88 pour le nombre d'heures de garde par assistante maternelle. Ces coefficients sont les plus élevés dans leur équation respective.

Sous forme matricielle, le système d'équations (14), (15) et (16) de la section 2.2.3 se réécrit  $(I_{3n} - \rho_n W_3) Y = \gamma X + \varepsilon$  en notant  $X$  les variables exogènes et  $Y$  les variables endogènes. En inversant le système,  $Y = (I_{3n} - \rho_n W_3)^{-1} (\gamma X + \varepsilon)$ . Cela signifie que, si les hypothèses permettant cette écriture sont respectées<sup>17</sup> :

$$Y = (I_{3n} + \rho_n W_3 + \rho_n^2 W_3^2 + \dots + \rho_n^N W_3^N)(\gamma X + \varepsilon) \quad (20)$$

Ainsi, le fait que les coefficients d'autocorrélation spatiale soient très élevés est crucial : cela signifie que les variations des variables exogènes à plusieurs bassins de vie d'écart se ressentent pleinement sur le bassin de vie observé. Par exemple,  $0,85^4 = 0,52$ . Donc des variations exogènes au marché de la garde à 4 bassins de vie d'écart du bassin de vie observé aura une incidence non négligeable sur le marché de la garde de ce bassin de vie. Cela peut signifier que le choix du bassin de vie comme unité géographique de référence n'est pas forcément optimal puisque les bassins de vie ont une influence sur d'autres bassins de vie éloignés.

Par ailleurs, un inconvénient de ce modèle SESAR est de ne pas expliquer par quel vecteur s'opère la propagation spatiale des effets observés. Est-ce un lié à une concurrence qui s'étend sur une distance plus grande qu'attendue ou à une raison qui n'a pas été envisagée ?

### 2.3.3. Comparaison qualitative des différents modèles

Au-delà de l'analyse du modèle SESAR, il est important de comparer les différents modèles évoqués au cours de cette présentation, afin d'évaluer l'apport de cette méthodologie prenant en compte l'autocorrélation spatiale. L'annexe G.a regroupe les résultats des modèles OLS, à équations simultanées, SESLX et SESAR. D'un point de vue qualitatif, pour une variable donnée, lorsqu'un coefficient est significatif, son signe ne dépend pas du modèle. Ce résultat est rassurant car même si un modèle OLS aboutit à des coefficients biaisés, les résultats sont qualitativement comparables.

Néanmoins, l'analyse qualitative est limitée par le fait qu'il n'est pas possible de déterminer les coefficients significatifs dans les équations du modèle SESAR. Il n'est donc pas possible de comparer la significativité des coefficients biaisés des modèles OLS, à équations simultanées et SESLX à celle des coefficients a priori sans biais du modèle SESAR.

Ainsi, l'intérêt du modèle SESAR réside dans le cas de cette étude plus en l'estimation sans biais des coefficients de l'équation de salaire horaire des assistantes maternelles que dans l'interprétation des variables influant significativement sur ce salaire horaire. C'est une limite de la méthode utilisée : elle ne permet pas de conclure aux déterminants du salaire horaire des assistantes maternelles. Elle permet simplement d'estimer l'ampleur de l'effet de ces déterminants, dans l'hypothèse selon laquelle ils sont significatifs.

### 2.3.4. Une comparaison quantitative difficile entre le modèle SESAR et les autres modèles

Il est aussi intéressant de comparer les coefficients quantitativement. Ceci n'est pas chose aisée. Dans le modèle OLS simple, l'effet d'une variation locale d'une variable exogène sur les différentes variables endogènes est la valeur du coefficient pourvu que les variables exogènes soient dans la même échelle (logarithmique ou non).

Dans le modèle SESAR, déterminer l'effet de la variation d'une variable sur le salaire horaire des assistantes maternelles est bien plus complexe. En effet, celle-ci va se répercuter sur le salaire

<sup>17</sup> Attention,  $W^k$  n'est pas la k-ième puissance de  $W$  mais la matrice des voisins d'ordre  $k$ , c'est-à-dire la matrice dont les coefficients valent 1 pour les bassins de vie séparés géographiquement par exactement  $k-1$  bassins de vie.

horaire des assistantes maternelles dans le bassin de vie concerné. Mais cet effet va, du fait de la présence d'autocorrélation spatiale, avoir une influence sur le salaire horaire des assistantes maternelles dans les bassins de vie voisins. L'autocorrélation spatiale va également être à l'origine de la rétroaction de ce dernier effet puisque les salaires horaires des bassins de vie voisins ayant changé, cela influe sur le salaire horaire des assistantes maternelles sur le bassin de vie observé.

Ainsi, pour une interprétation quantitative du modèle SESAR, il n'est possible de déterminer l'effet d'une variation locale d'une variable exogène sur les variables endogènes qu'en effectuant une simulation numérique avec un système informatique. L'interprétation quantitative de ce modèle n'est pas immédiate. Cela n'est pas gênant pour prévoir les effets d'une politique publique ou d'un changement conjoncturel de l'économie sur le marché de garde des assistantes maternelles. Mais pour évaluer l'importance des déterminants du salaire horaire des assistantes maternelles, il peut paraître décevant de ne pas voir immédiatement quelle sera l'influence d'une variation d'un de ces déterminants sur le salaire horaire des assistantes maternelles.

### **Conclusion**

La question initiale était de savoir si les salaires horaires des assistantes maternelles sont influencés par des effets de demande des employeurs, par des effets d'offre des assistantes maternelles ou par une harmonisation des salaires entre concurrentes. D'après les résultats, l'effet prédominant sur le salaire horaire des assistantes maternelles réside en l'harmonisation de leur salaire. Les régressions économétriques mettent en valeur un coefficient d'autocorrélation spatiale très élevé des salaires horaires. L'étude montre également l'importance de la demande de garde notamment à travers le salaire médian des différents bassins de vie et l'importance de l'offre à travers la concurrence entre assistantes maternelles. Ces effets d'offre et de demande ont une influence à la fois directe et indirecte sur les salaires des assistantes maternelles puisqu'ils influent également sur le nombre d'enfants gardés par assistante maternelle qui lui-même a une influence sur le salaire horaire des assistantes maternelles.

Pour aboutir à ces résultats, cette étude sur les assistantes maternelles a nécessité l'utilisation d'un modèle économétrique peu courant. Le salaire horaire d'une assistante maternelle n'étant pas déterminé indépendamment du nombre d'enfants qu'elle garde et du nombre d'heures de garde qu'elle effectue, un modèle à équations simultanées a été utilisé. Cependant, les résultats obtenus par ce modèle sont peu satisfaisants car ils ne prennent pas en compte l'autocorrélation spatiale du salaire horaire des assistantes maternelles : des assistantes maternelles vivant à proximité l'une de l'autre ont tendance à avoir des salaires horaires proches. Un modèle à équations simultanées a donc été combiné avec des modèles d'économétrie spatiale pour tenir compte de tous les aspects particuliers (salaires horaires autocorrélés spatialement et déterminés simultanément avec le nombre d'enfants gardés et le nombre d'heures de garde) du marché de la garde des assistantes maternelles. Le modèle ainsi utilisé permet de ne pas se tromper sur la valeur des coefficients des variables influant sur le salaire horaire des assistantes maternelles. Cependant, il présente une vraie limite d'interprétation puisqu'il n'a pas été possible de déterminer la significativité des coefficients du modèle. Il n'a donc pas été possible de déterminer les déterminants des salaires horaires des assistantes maternelles. Par contre, conditionnellement à la significativité d'un coefficient, le modèle utilisé permet d'évaluer l'ampleur d'un déterminant du salaire horaire des assistantes maternelles. La principale piste d'amélioration de ce document réside donc dans le calcul d'écarts-types convergents des coefficients du modèle à équations simultanées avec autocorrélation spatiale. Cela permettrait une interprétation bien plus riche des résultats présentés dans cette étude.

Les résultats doivent également être nuancés du point de vue de leur interprétation spatiale. Compte-tenu de la forte autocorrélation spatiale des salaires horaires des assistantes maternelles, il est légitime de se demander si l'unité géographique de référence a été correctement choisie dans cette étude. Peut-être aurait-il fallu choisir une unité géographique plus grande, réduisant ainsi le nombre d'observations mais rendant plus fidèlement compte de l'autocorrélation spatiale des salaires horaires

des assistantes maternelles. Ce choix n'est pas anodin : la modification de l'unité géographique de référence influe quantitativement (voire même qualitativement) sur les coefficients de régression calculés dans les modèles. Une des améliorations possibles de cette étude réside donc en l'utilisation d'outils adaptés à l'analyse spatiale tels que l'indice de Moran. Cela permettrait également de mieux comparer les résultats des modèles économétriques entre eux, en mesurant par exemple quantitativement l'autocorrélation spatiale des résidus de chaque modèle.

D'autres pistes de poursuite de cette étude sont envisageables. Dans cette étude, tous les enfants ont été considérés identiques pour les assistantes maternelles et les employeurs, qu'ils aient moins de 3 ans ou entre 3 et 5 ans. La réalité est cependant différente. Le taux de scolarisation des enfants de 3 à 5 ans s'approchant de 100 %, les enfants de 3 à 5 ans sont gardés bien moins d'heures par les assistantes maternelles que les enfants de moins de 3 ans. Une suite possible à cette étude consisterait donc à étudier l'influence de certaines variables telles que l'âge ou l'expérience des assistantes maternelles sur la proportion d'enfants de moins de 3 ans et sur la proportion d'enfants de 3 à 5 ans qu'elles gardent.

## Bibliographie

Beaumont C., Bellamy V., « La situation démographique en 2014 », *Insee Résultats* n°182 Société, juin 2016.

Cartier M., d'Halluin E., Rousseau J., Lechien M.-H., 2012, « Temps partiel ou irrégularité de l'activité des assistantes maternelles ? Une enquête exploratoire », *Dossier d'études*, CAF, n°150.

de Bellefon M.-P., 2016, « Introduction à l'analyse spatiale », *Séminaire de Méthodologie Statistique du département des méthodes statistiques de l'Insee*.

DEPP, 2016, « Repères & Références statistiques 2016 sur les enseignements, la formation et la recherche ».

Direction générale de l'action sociale, « Référentiel de l'agrément des assistants maternels à l'usage des services de protection maternelle et infantile », Ministère du Travail, des Relations sociales, de la Famille, de la Solidarité et de la Ville.

Efron B., Tibshirani R. J., 1998, « An Introduction to the Bootstrap », *Chapman & Hall/CRC*.

Floch J.-M. and Le Saout R., 2016, « Économétrie spatiale : une introduction pratique », *Séminaire de Méthodologie Statistique du département des méthodes statistiques de l'Insee*.

Gebremariam G. H., 2007, « Modeling and Estimation Issues in Spatial Simultaneous Equations Models », *54th Annual North American Meetings of the Regional Science Association International in Savannah*.

Greene W.-H., 2003, « Econometric Analysis (International Edition), Fifth Edition », *Prentice Hall*.

Insee, 2014, « Recensement de la population. Évolutions : pourquoi privilégier les évolutions quinquennales ou celles par rapport à 1999 ? ».

Insee, 2016, « Tableaux de l'économie française », *Insee Références*.

Kelejian H.-H. and Prucha I.-R., 2004, « Estimation of simultaneous systems of spatially interrelated cross-sectional equations », *Journal of Econometrics*, n°118, p. 27-50.

Observatoire national de la petite enfance, 2015, « L'accueil du jeune enfant en 2014 », Caisse Nationale d'Allocations Familiales.

Robert-Bobée I., Le Blanc D., Choné P., 2004, « Offre de travail féminine et garde des jeunes enfants », *Economie et prévision*, n°162, p. 23-50.

Ulmann A.-L., Betton E., Jobert G., 2011, « L'activité des professionnelles de la petite enfance », *Dossier d'études*, CAF, n°145.

Vroylandt T., 2016, « Les assistantes maternelles gardent 8 % d'enfants de plus en 2014 qu'en 2010 », *Études et Résultats*, DREES, n°978, octobre.

Vroylandt T., Paliot N., 2017, « Les assistant·e·s maternel·le·s ont gagné en moyenne 1 108 euros en juin 2014 », *Études et Résultats*, DREES, n°1020, août.

## Annexes

### A. Traitement de la base de données

#### A.a. Données Pajemploi

Pajemploi est un service des Urssaf visant notamment à centraliser le recueil des déclarations salariales des employeurs d'assistantes maternelles. La base de données utilisée dans cette étude est fournie par l'ACOSS qui est la Caisse nationale du réseau des Urssaf. Il s'agit donc de la base fiscale des déclarations des employeurs d'assistantes maternelles.

Cette base de données est composée de quatre tables dont trois sont intéressantes pour l'étude : une table recensant les assistantes maternelles, une table recensant les employeurs, une table recensant les contrats entre employeurs et assistantes maternelles. Ces tables contiennent des informations telles que le salaire net perçu par une assistante maternelle  $i$  et versé par l'employeur  $j$ , ainsi que le nombre d'heures effectuées ou les indemnités versées à l'employée. Mais elles fournissent également la commune de résidence de l'employeur et celle de l'assistante maternelle, ainsi que son année de naissance, par exemple.



Tableau 8 - Présentation de la base de données Pajemploi

#### A.b. Nettoyage de la base de données

La base Pajemploi est une base fiscale et est donc sujette à des erreurs de remplissage de la part des employeurs. Cette étude ne visant pas à redresser des données, il a fallu décider de la manière adéquate de traiter les données aberrantes. D'un côté, les salaires des assistantes maternelles sont fixés en fonction du nombre d'enfants gardés par assistante maternelle (équation (6)). Cela nécessite donc d'avoir un nombre d'enfants gardés fiable par assistante maternelle. De l'autre côté, garder des contrats aberrants mène à des erreurs car l'erreur peut provenir du nombre d'enfants gardés, du nombre d'heures de garde effectué par l'assistante maternelle par enfant sur le contrat ou du salaire net déclaré par l'employeur.

Il a été choisi de supprimer les contrats aberrants de la base de données. En effet, les erreurs présentes dans un contrat de la base de données peuvent provenir de plusieurs sources. Par exemple, certains contrats indiquent plus de 600 heures de garde mensuelles par enfant pour un seul mois. Cela peut provenir d'une erreur de remplissage concernant le nombre d'enfants gardés ou concernant le nombre d'heures de garde effectuées. Conserver ce type de contrats aberrants aurait donc conduit à ajouter du bruit au nombre d'enfants gardés par bassin de vie, au nombre d'heures de garde par enfant et par bassin de vie et au salaire horaire des assistantes maternelles. Cela paraît plus préjudiciable que de supprimer ces contrats de l'étude. L'inconvénient est que le nombre d'enfants gardés par assistante maternelle est ainsi sous-estimé.

Un certain nombre de critères ont été utilisés pour estimer si un contrat était aberrant ou non. Par exemple, un contrat présentant un salaire horaire strictement supérieur à 30 euros de l'heure a été jugé aberrant. En effet, plus de 90 % de ces contrats font état de strictement moins d'une heure de garde par jour et par enfant sur le contrat. Ces contrats présentent donc vraisemblablement des erreurs de remplissage. Un autre critère concerne les salaires horaires nets trop faibles. Un seuil légal existe en effet. En 2014, celui-ci était de 2,07€ pour la France métropolitaine hors Alsace et Moselle et de 2,03€ en Alsace et en Moselle. Il est défini proportionnellement au SMIC (salaire minimum interprofessionnel de croissance) horaire. Tous les contrats établissant un salaire horaire trop faible ont donc systématiquement été filtrés en utilisant un seuil dépendant du SMIC horaire. Les contrats présentant un nombre d'heures de garde irréaliste ont également été supprimés. Pour cela, il a été estimé qu'une assistante maternelle travaillait au maximum 25 jours par mois et environ 11 heures par jour travaillé et par enfant sur le contrat. D'autres filtres ont été utilisés mais cela nécessiterait de plus amples explications sur les choix effectués, qui ne sont pas l'objet de cette étude.

## B. Notations

Tableau 4 : Notations employées dans le document

Notation	Définition
$i$	assistante maternelle $i$
$j$	employeur $j$
$s$	bassin de vie $s$ (unité géographique de référence)
$t$	période d'observation $t$
$\Pi_{i,s,t}$	profit de l'assistante maternelle $i$ à la date $t$ et vivant dans $s$
$C_{i,s,t}$	coûts de l'assistante maternelle $i$ à la date $t$ et vivant dans $s$
$u_{j,s,t}$	utilité de l'employeur $j$ à faire garder son enfant à la date $t$ et vivant dans $s$
$k_{j,s,t}$	seuil de participation de l'employeur $j$ vivant dans $s$ au marché de garde des assistantes maternelles $s$ dans à la date $t$
$q_{i,j,s,t}$	nombre d'enfants de l'employeur $j$ gardés par l'assistante maternelle $i$
$h_{i,j,s,t}$	nombre d'heures de garde effectuées par l'assistante maternelle $i$ par enfant $q_{i,j,s,t}$ de l'employeur $j$
$jours_{i,j,s,t}$	nombre de jours de garde effectués par l'assistante maternelle $i$ par enfant $q_{i,j,s,t}$ de l'employeur $j$
$w_{i,j,s,t}$	saire horaire du contrat de l'assistante maternelle $i$ avec l'employeur $j$ par enfant $q_{i,j,s,t}$
$q_{i,s,t}$	nombre d'enfants gardés par l'assistante maternelle $i$
$h_{i,s,t}$	nombre moyen d'heures de garde effectuées par l'assistante maternelle $i$ par enfant gardé
$w_{i,s,t}$	saire horaire moyen par enfant gardé de l'assistante maternelle $i$ tous contrats confondus
$q_{s,t}$	nombre d'enfants gardés par des assistantes maternelles dans le bassin de vie $s$
$h_{s,t}$	nombre d'heures total de garde effectuées par des assistantes maternelles dans $s$
$w_{s,t}$	saire horaire moyen parmi toutes les assistantes maternelles exerçant dans $s$ (poids identique pour toutes les assistantes maternelles)
$conc_{s,t}$	concurrence tous modes de gardes confondus rencontrée par les assistantes maternelles de $s$
$age_{i,s,t}$	âge de l'assistante maternelle $i$
$dens_{i,s,t}$	densité de la commune de l'assistante maternelle $i$
$exp_{i,s,t}$	expérience de l'assistante maternelle $i$ dans la profession d'assistante maternelle
$loyer_{i,s,t}$	loyer payé par l'assistante maternelle $i$
$F_{i,s,t}$	autres coûts liés au travail rencontrés par l'assistante maternelle $i$
$tt_{j,s,t}$	temps de travail de l'employeur $j$
$traj_{j,s,t}$	temps de trajet de l'employeur $j$ pour se rendre à son travail
$te_{j,s,t}$	temps passé à l'école maternelle pour l'enfant de l'employeur $j$
$tt_{s,t}$	temps de travail moyen des employeurs du bassin de vie $s$
$traj_{s,t}$	temps de trajet moyen des employeurs du bassin de vie $s$ pour se rendre à leur travail
$h_{j,s,t}$	nombre d'heures de garde de l'enfant de l'employeur $j$ cumulées parmi toutes les assistantes maternelles

$\lambda_{s,t}$	taux d'activité du bassin de vie $s$
$\mu_{s,t}$	taux de chômage du bassin de vie $s$
$med_{s,t}$	salaire médian du bassin de vie $s$
$\varepsilon_{j,s,t}$	situation particulière de l'employeur $j$
$N_{s,t,assmat}$	nombre d'assistantes maternelles actives dans le bassin de vie $s$
$N_{s,t,enfants}$	nombre d'enfants en âge d'être gardés dans le bassin de vie $s$
$\alpha_{s,t}$	préférence des employeurs pour le mode de garde par assistante maternelle par rapport aux autres modes de garde

### C. Coûts des assistantes maternelles

Il s'agit ici d'expliciter en détail les équations de la partie 1.1.4 et notamment la fonction de coûts introduite dans l'équation (1). Dans l'ensemble de cette annexe, deux documents de travail publiés dans la collection des dossiers de la Caisse Nationale d'Allocations Familiales (Ulmann, Betton, Jobert, 2011 et Cartier, d'Halluin, Rousseau, Lechien, 2012) sont largement utilisés.

#### C.a. Coûts de pénibilité

##### C.a.i. Pénibilité horaire

Dans la quatrième étape de l'interaction entre assistantes maternelles et les employeurs définie en section 1.2.2, il a été supposé que ces derniers choisissent le nombre d'heures de garde pour leur enfant. En effet, si l'assistante maternelle peut refuser de garder un enfant car elle n'accepte pas les contraintes horaires liées à cette garde, ce sont bien les employeurs qui ont des contraintes d'emploi du temps dont ils peuvent difficilement s'abstraire. Généralement, l'emploi d'une assistante maternelle se justifie par l'impossibilité de l'employeur de garder son enfant durant ses heures de travail.

Ainsi, l'assistante maternelle se voit dans une certaine mesure imposer des heures de travail. Le dossier de la CAF sur le temps de travail des assistantes maternelles (Cartier, d'Halluin, Rousseau, Lechien, 2012) confirme ce constat : « Les employeurs aux horaires atypiques ou aux conditions d'emploi et de travail difficiles imposent aux assistantes maternelles de longues amplitudes de garde, parfois à la source d'un sentiment de "ras le bol", qui peut déboucher sur des interruptions ou des réductions d'activité ». Les amplitudes élevées d'horaire de garde peuvent ainsi être pesantes pour les assistantes maternelles, comme le confirme implicitement le témoignage d'une assistante maternelle dans ce dossier : « Et quand elle est venue, je lui ai bien demandé les horaires. Parce que moi, garder des enfants jusqu'à 20 heures, non ! ».

Sur la base de ce témoignage, la pénibilité horaire est ressentie par les assistantes maternelles comme un coût plutôt journalier. Il en va de même pour de nombreux métiers. Dès lors, lorsque les horaires de travail dans une journée deviennent trop élevés pour une assistante maternelle, elle ressent de plus en plus ces coûts. Il existe donc un seuil pouvant être défini arbitrairement (par exemple 8 heures par jour et par enfant), au-delà duquel les assistantes maternelles sont affectées par cette pénibilité horaire. Une heure supplémentaire travaillée par enfant et par jour est de plus en plus fatigante au-delà de ce seuil. Ce coût de pénibilité horaire peut donc s'écrire :

$$\beta_1^+ \mathbb{1} \left\{ \frac{\int_S h_{i,j,s,t} q_{i,j,s,t} dj}{\int_S \text{jours}_{i,j,s,t} q_{i,j,s,t} dj} \geq 8 \right\} \left( \frac{\int_S h_{i,j,s,t} q_{i,j,s,t} dj}{\int_S \text{jours}_{i,j,s,t} q_{i,j,s,t} dj} - 8 \right)^{p_1} \text{ avec } p_1 > 1$$

##### C.a.ii. Pénibilité du nombre d'enfants gardés

Les assistantes maternelles gardent généralement plusieurs enfants puisqu'en juin 2014, seules 15% des assistantes maternelles actives gardaient un unique enfant. La question qui se pose est alors de savoir s'il est aussi pénible de s'occuper d'un enfant que de deux enfants ou plus. Intuitivement, s'occuper de quatre enfants est plus prenant que de ne garder qu'un enfant car il convient de les surveiller correctement. Néanmoins, pour certaines tâches, il peut être pratique de garder plusieurs enfants plutôt qu'un. Par exemple, préparer un déjeuner sera presque aussi contraignant lorsque l'assistante maternelle s'occupe d'un enfant que de quatre enfants. Mais il est aussi possible d'objecter que chacun a des contraintes alimentaires, certains enfants ne pouvant pas encore mâcher certains aliments, d'autres n'appréciant pas certains ingrédients, etc. Ainsi, contrairement à l'intuition, s'occuper de plusieurs enfants ne présenterait pas forcément d'économie d'échelle. C'est une conclusion du dossier de la CAF sur les professionnelles de la petite enfance (Ulmann, Betton, Jobert, 2011) : « On peut noter ici que si les contraintes de devoir s'occuper de plusieurs enfants à la fois sont tout aussi présentes pour les assistantes maternelles, ces dernières ne

peuvent pas aussi facilement recourir à la dimension du collectif pour différer leurs réponses aux sollicitations ».

S'occuper d'un enfant supplémentaire est donc plus fatigant pour l'assistante maternelle que de s'occuper du précédent. Le coût de pénibilité dû au nombre d'enfants gardés semble donc être de la forme :

$$\beta_2 q_{i,s,t}^{p_2} \text{ avec } p_2 > 1$$

#### C.b. Propagation des coûts dans la durée

##### C.b.i. Coûts intertemporels

Les assistantes maternelles choisissent rarement ce métier comme première vocation. Devenir assistante maternelle peut être un choix pour pouvoir garder son enfant au quotidien après sa naissance. Cela peut aussi répondre à un besoin financier. Le graphique 9 traduit bien ces motivations : il n'y a que très peu de jeunes assistantes maternelles, mais, jusqu'à 42 ans, le nombre d'assistantes maternelles augmente. Alors que 8,8 % des femmes actives en 2014 avaient entre 15 et 24 ans (Insee, 2016) et alors que les assistantes maternelles sont très majoritairement des femmes (Vroylandt, 2016), seules 0,4 % des assistantes maternelles actives avaient entre 15 et 24 ans en juin 2014. Compte-tenu de la pénibilité de ce métier et des différentes motivations à devenir assistantes maternelles exposées précédemment en s'appuyant sur le document analysant le temps de travail des assistantes maternelles (Cartier, d'Halluin, Rousseau, Lechien, 2012), ce métier n'est pas toujours exercé durant plusieurs dizaines d'années par les assistantes maternelles. Cela se traduit dans le graphique 9 par une diminution du nombre d'assistantes maternelles actives au-delà de 50 ans. En outre, des modulations du temps de travail sont possibles selon les années pour compenser la pénibilité. Malgré tout, le métier d'assistante maternelle reste exercé par une population relativement âgée : 2,5 % des assistantes maternelles ayant gardé au moins un enfant en juin 2014 étaient âgées de 65 ans ou plus tandis que la part d'actives de 65 ans ou plus en France est inférieure à 0,1 %.

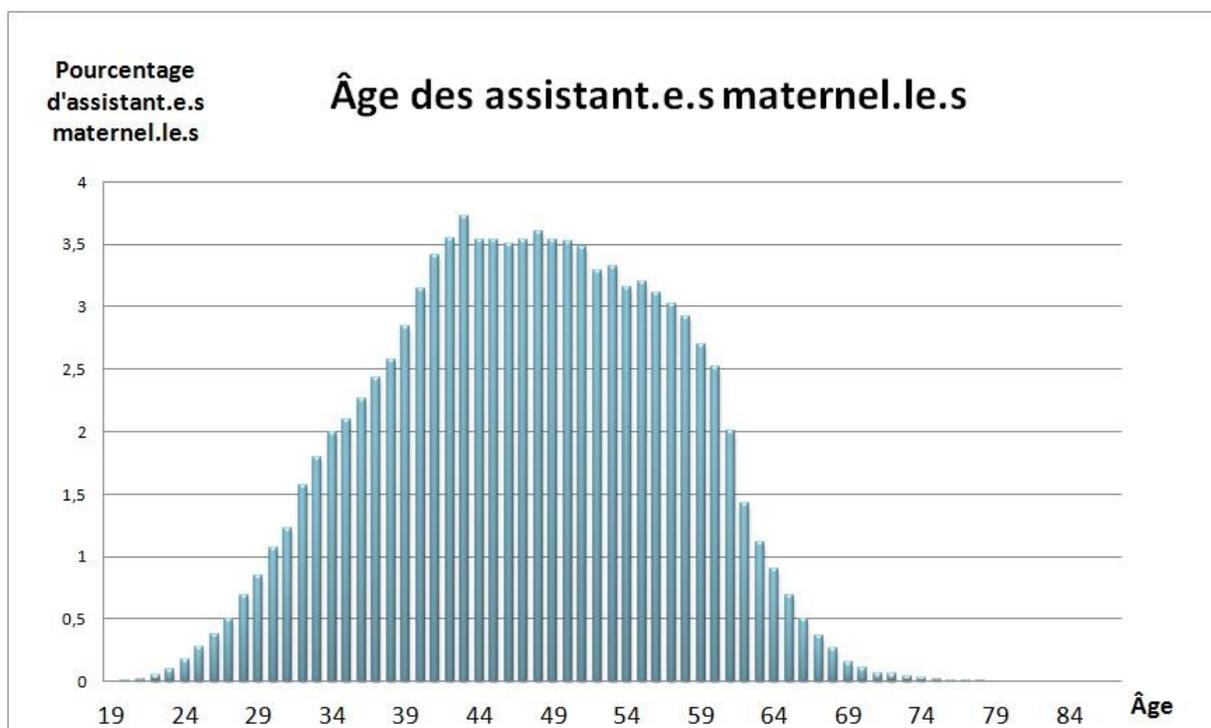


Figure 9 - Âge des assistantes maternelles actives en juin 2014

Ainsi, la pénibilité du travail d'assistante maternelle ne saurait être dissociée de la pénibilité passée. L'étude qualitative de Cartier, d'Halluin, Rousseau et Lechien (Cartier, d'Halluin, Rousseau, Lechien, 2012) souligne ainsi que « si les pressions du conjoint et des enfants, nous l'avons vu, peuvent contribuer à l'entrée dans le métier, elles peuvent aussi être à l'origine de la réduction de l'activité une fois les contraintes du métier éprouvées ». Les choix du nombre d'heures de garde et d'enfants gardés par les assistantes maternelles dépendent donc de leurs coûts passés. Une forte pénibilité passée peut ainsi entraîner une réduction du temps de travail de l'assistante maternelle puisqu'il peut en découler « une alternance entre des périodes de travail intense et des périodes de relâchement, tantôt voulu, tantôt subi » pour les assistantes maternelles. De tels coûts peuvent amener une assistante maternelle à travailler moins d'heures par enfant et à accepter la garde de moins d'enfants en sélectionnant les employeurs. Un tel coût peut s'écrire :

$$\beta_3 \underset{-}{\Pi}_{i,s,t-1} + \beta_4 \underset{+}{C}_{i,s,t-1} h_{i,s,t} + \beta_5 \underset{+}{C}_{i,s,t-1} q_{i,s,t}$$

Les coûts passés influenceront donc positivement les coûts présents et une période précédente de coûts élevés apportera probablement une modulation du temps de travail à travers les termes  $C_{i,s,t-1} h_{i,s,t}$  et  $C_{i,s,t-1} q_{i,s,t}$ . Le profit passé  $\Pi_{i,s,t-1}$  a plutôt un rôle de seuil et peut amener l'assistante maternelle à cesser son activité : un profit passé trop faible peut conduire l'assistante maternelle à cesser de faire des sacrifices coûteux pour elle et sa famille en travaillant à domicile.

#### C.b.ii. Coûts liés à l'âge

Cette pénibilité du travail mise en avant précédemment doit également tenir compte de l'âge de l'assistante maternelle. Jusqu'à présent, que ce soit en matière de pénibilité horaire ou de propagation des coûts dans le temps, l'influence de l'âge de l'assistante maternelle sur les coûts ressentis n'a pas été pris en considération. Pourtant, « dans le cas des assistantes maternelles approchant la soixantaine, les entretiens témoignent d'une réduction du temps de travail ou du nombre d'enfants gardés liée au vieillissement, à la fatigue mentale associée au travail, mais aussi à l'entrée en retraite des proches » (Cartier, d'Halluin, Rousseau, Lechien, 2012). Les coûts doivent donc tenir compte de la fatigue croissante engendrée par ce métier lorsque l'âge de l'assistante maternelle augmente. Un terme de la forme suivante aurait pu être ajouté en se basant sur les annexes C.a.i et C.a.ii :

$$\beta_1 \underset{+}{age}_{i,s,t}^{p_3} \mathbb{1} \left\{ \frac{\int_S h_{i,j,s,t} q_{i,j,s,t} dj}{\int_S \text{jours}_{i,j,s,t} q_{i,j,s,t} dj} \geq 8 \right\} \left( \frac{\int_S h_{i,j,s,t} q_{i,j,s,t} dj}{\int_S \text{jours}_{i,j,s,t} q_{i,j,s,t} dj} - 8 \right)^{p_1} + \beta_2 \underset{+}{age}_{i,s,t}^{p_4} q_{i,s,t}^{p_2}$$

Par souci de résolution du modèle, des coûts linéaires en le nombre d'heures et en le nombre d'enfants gardés mais convexes en l'âge ont été privilégiés. Cela signifie que chaque année supplémentaire pour l'employée rend le travail plus pénible.

$$\beta_7 \underset{+}{age}_{i,s,t}^{p_3} q_{i,s,t} + \beta_8 \underset{+}{age}_{i,s,t}^{p_4} h_{i,s,t} \text{ avec } p_3 > 1 \text{ et } p_4 > 1$$

#### C.c. Les coûts de recherche d'enfants à garder

##### C.c.i. Les coûts liés à la concurrence

La rencontre de l'offre et de la demande varie selon les marchés considérés. Du point de vue du marché d'accueil des assistantes maternelles, les bassins de vie sont loin d'être équivalents. Au sein même des bassins de vie existent des divergences puisque « les assistantes maternelles "urbaines" n'ont souvent pas de difficultés à trouver des enfants à garder et peuvent même se permettre de sélectionner leurs employeurs » d'après le dossier de la CAF portant sur l'analyse du temps de travail des assistantes maternelles (Cartier, d'Halluin, Rousseau, Lechien, 2012). Cette difficulté ou facilité à trouver un enfant à garder pour les assistantes maternelles en fonction de leur lieu de vie ou de la concurrence doit être prise en compte dans le modèle. Des dispositifs locaux comme celui indiqué ci-après ne peuvent, certes, pas être pris en compte : « Christine Dumont a quant à elle remarqué des

différences d'organisation entre le RAM de Vertou, où elle exerçait précédemment, et celui de Savenay, beaucoup plus récent et qui, selon elle, met en contact l'offre et la demande de garde de façon moins efficace » (Cartier, d'Halluin, Rousseau, Lechien, 2012). Mais il est néanmoins possible de considérer une certaine forme de concurrence.

Au vu de ces éléments, ces coûts de concurrence peuvent être ajoutés dans les coûts de l'assistante maternelle :

$$\beta_6 \underset{+}{conc}_{s,t} q_{i,s,t} + \beta_9 \underset{-}{dens}_{i,s,t} q_{i,s,t}$$

Plus la concurrence est élevée, avec par exemple beaucoup de places en établissement d'accueil du jeune enfant et beaucoup de gardes à domicile comparé au nombre de jeunes enfants, plus il sera coûteux de trouver des enfants à garder. Même si l'assistante maternelle garde déjà un enfant, en trouver un autre à garder sera tout aussi coûteux que cela l'a été pour trouver le premier enfant à garder. En effet, il y a tellement d'enfants dans le bassin de vie qu'un enfant de plus ou de moins à garder dans le bassin de vie ne change pas les coûts de recherche pour l'assistante maternelle. Pour ce coût de recherche  $conc_{s,t} q_{i,s,t}$ , une dépendance linéaire en  $q_{i,s,t}$  est introduite. De même, plus les assistantes maternelles sont dans un environnement urbain et dense, plus leurs coûts de recherche sont faibles au vu du constat sur les assistantes maternelles « urbaines ». Cela est représenté par le coût  $dens_{i,s,t} q_{i,s,t}$ .

#### C.c.ii. Les coûts dus à l'établissement d'un réseau professionnel

Comme dans toute profession, le métier d'assistante maternelle est influencé par le réseau professionnel. Pour une assistante maternelle qui s'établit, trouver un enfant à garder est relativement compliqué. Les employeurs potentiels ne savent pas forcément qu'une assistante maternelle vient de s'installer à proximité de leur domicile ou de leur emploi. Au contraire, une assistante maternelle expérimentée aura, si cela se passe bien avec ses employeurs, beaucoup de facilité à trouver un enfant supplémentaire à garder. En effet, les anciens employeurs pourront témoigner vis-à-vis de leurs relations de la compétence de cette assistante maternelle, facilitant ainsi son embauche. D'autre part, il est également facile de garder le même nombre d'enfants qu'à la période précédente puisque, si la relation entre l'assistante maternelle, l'enfant et l'employeur est bonne, l'enfant peut être gardé plusieurs années d'affilée par la même assistante maternelle. Ceci peut être traduit par un coût linéaire :

$$\beta_{10} \underset{-}{exp}_{i,s,t} q_{i,s,t}$$

La linéarité de ce coût en l'expérience est tout à fait discutable. Cependant, il n'y avait pas d'argument suffisamment pertinent pour choisir entre un coût concave ou convexe en l'expérience. La linéarité de ce coût en le nombre d'enfants gardés semble être cohérente à partir du moment où il n'inclut pas l'aspect intertemporel de la garde (un enfant gardé l'étant pour plusieurs périodes généralement). Les arguments utilisés pour la linéarité en le nombre d'enfants gardés  $q_{i,s,t}$  du coût de recherche de garde d'un enfant par l'assistante maternelle dû à la concurrence justifient alors aussi que ce coût de recherche de garde dû au manque d'expérience soit linéaire en  $q_{i,s,t}$ .

#### C.d. Autres coûts liés à la garde

##### C.d.i. Le coût de l'accueil du jeune enfant

La délivrance de l'agrément d'assistante maternelle est accordée par la présidence du conseil départemental. Parmi les critères d'obtention de l'agrément figure la taille du logement (Direction générale de l'action sociale). Celui-ci doit être suffisamment grand pour accueillir des enfants de tiers. En outre, la taille du logement peut être un critère pour délivrer un agrément d'accueil pour trois enfants plutôt que pour quatre enfants. À prix du mètre carré égal, accueillir beaucoup d'enfants

nécessite une grande surface de lieu de vie et donc un prix d'achat ou de loyer élevé pour l'assistante maternelle. Les coûts des assistantes maternelles incluent donc :

$$\beta_{11} \text{loyer}_{i,s,t} q_{i,s,t}$$

Ce coût n'est pas réellement ressenti par l'assistante maternelle. Mais, en faisant intervenir ce coût, la non-prise en compte de la délivrance d'agrément pour un nombre d'enfants inférieur à la demande de la personne souhaitant devenir assistante maternelle est compensée. Cela aura son importance dans l'interprétation des résultats.

#### C.d.ii. Autres coûts

Au-delà des différents coûts spécifiques au métier d'assistante maternelle qui ont été évoqués, travailler présente des contraintes qui peuvent être plus ou moins bien vécues par l'assistante maternelle. « Depuis 7 H 30, Mme A se prépare à cet accueil : faire déjeuner rapidement ses deux enfants pour les faire partir à l'école mais surtout pour ranger leur chambre, nettoyer la cuisine, sortir leurs affaires de la salle principale de l'appartement et faire les lits, car il ne serait pas possible d'accueillir les enfants dans un appartement en désordre » (Ulmann, Betton, Jobert, 2011). Ces contraintes sont ressenties différemment par chacune et un coût  $F_{i,s,t}$  spécifique à l'assistante maternelle  $i$  lié au fait de travailler peut être introduit.

#### D. Étape de résolution du modèle théorique principal : maximisation du nombre d'enfants gardés par assistante maternelle

Dans cette annexe, la résolution du modèle théorique à une tranche d'âge est détaillée. Cette section complète le paragraphe « Détermination du nombre d'enfants gardés par les assistantes maternelles » de la section 1.3.2. Il y a été démontré que les assistantes maternelles ont une fonction de profit de la forme :

$$\Pi_{i,s} = q_{i,s} (tt_s + traj_s) w_s - \beta_1 (tt_s + traj_s)^{p_1} - \beta_2 q_{i,s}^{p_2} - \beta_6 conc_s q_{i,s} - \beta_7 age_{i,s}^{p_3} q_{i,s} - \beta_8 age_{i,s}^{p_4} (tt_s + traj_s) - \beta_9 dens_{i,s} q_{i,s} - \beta_{10} exp_{i,s} q_{i,s} - \beta_{11} loyer_{i,s} q_{i,s} - F_{i,s}$$

Or,  $(tt_s + traj_s) w_s = \phi_1 (tt_s + traj_s)^{r_1} - \frac{q_s}{N_{s,enfants} \alpha_s} - \eta_0 - \eta_1 \lambda_s - \eta_2 \mu_s - \eta_3 med_s$ ,  
fonction de demande inverse issue de l'équation (6). Donc :

$$\Pi_{i,s} = q_{i,s} \left( \phi_1 (tt_s + traj_s)^{r_1} - \frac{q_{i,s} + q_{-i,s}}{N_{s,enfants} \alpha_s} - \eta_0 - \eta_1 \lambda_s - \eta_2 \mu_s - \eta_3 med_s \right) - \beta_1 (tt_s + traj_s)^{p_1} - \beta_2 q_{i,s}^{p_2} - \beta_6 conc_s q_{i,s} - \beta_7 age_{i,s}^{p_3} q_{i,s} - \beta_8 age_{i,s}^{p_4} (tt_s + traj_s) - \beta_9 dens_{i,s} q_{i,s} - \beta_{10} exp_{i,s} q_{i,s} - \beta_{11} loyer_{i,s} q_{i,s} - F_{i,s}$$

$p_2 > 1$  d'après l'annexe C.a.ii. Pour des fins de résolution, il est raisonnable de poser  $p_2 = 2$ . Les assistantes maternelles maximisent leur profit selon le nombre d'enfants gardés  $q_{i,s}$  en supposant le nombre d'enfants gardés par leurs concurrentes  $q_{-i,s}$  donné.  $q_s = q_{i,s} + q_{-i,s}$ .

$$\begin{aligned} \frac{\partial \Pi_{i,s}}{\partial q_{i,s}} &= 0 \\ \Leftrightarrow \left( \phi_1 (tt_s + traj_s)^{r_1} - \frac{q_s}{N_{s,enfants} \alpha_s} - \eta_0 - \eta_1 \lambda_s - \eta_2 \mu_s - \eta_3 med_s \right) - \frac{q_{i,s}}{N_{s,enfants} \alpha_s} - 2 \beta_2 q_{i,s} - \beta_6 conc_s - \beta_7 age_{i,s}^{p_3} - \beta_9 dens_{i,s} - \beta_{10} exp_{i,s} - \beta_{11} loyer_{i,s} &= 0 \\ \Leftrightarrow 2 \beta_2 q_{i,s} + \frac{q_{i,s} + q_s}{N_{s,enfants} \alpha_s} &= \phi_1 (tt_s + traj_s)^{r_1} - \eta_0 - \eta_1 \lambda_s - \eta_2 \mu_s - \eta_3 med_s - \beta_6 conc_s - \beta_7 age_{i,s}^{p_3} - \beta_9 dens_{i,s} - \beta_{10} exp_{i,s} - \beta_{11} loyer_{i,s} \end{aligned}$$

Les notations suivantes sont introduites :  $g_{i,s} = \beta_6 conc_s + \beta_7 age_{i,s}^{p_3} + \beta_9 dens_{i,s} + \beta_{10} exp_{i,s} + \beta_{11} loyer_{i,s}$  sont les coûts des assistantes maternelles et  $D_s = \eta_0 + \eta_1 \lambda_s + \eta_2 \mu_s + \eta_3 med_s$  sont les effets demande dûs aux coûts d'opportunité. En sommant sur les assistantes maternelles :

$$\begin{aligned} 2 \beta_2 q_s + (N_{s,assmat} + 1) \frac{q_s}{N_{s,enfants} \alpha_s} &= \phi_1 (tt_s + traj_s)^{r_1} N_{s,assmat} - D_s N_{s,assmat} - \sum_i g_{i,s} \\ \Leftrightarrow \left( \frac{N_{s,assmat} + 1}{N_{s,enfants} \alpha_s} + 2 \beta_2 \right) q_s &= \phi_1 (tt_s + traj_s)^{r_1} N_{s,assmat} - D_s N_{s,assmat} - \sum_i g_{i,s} \end{aligned}$$

$$\Leftrightarrow q_s = \frac{1}{\frac{N_{s,assmat} + 1}{N_{s,enfants} \alpha_s} + 2 \beta_2} \left( \phi_1 (tt_s + traj_s)^{r_1} N_{s,assmat} - D_s N_{s,assmat} - \sum_i g_{i,s} \right) \quad (21)$$

## E. Construction des variables

Dans cette annexe sont détaillées la manière dont a été construite chaque variable ainsi que la source des données.

Tout d'abord, des données de 2012 sont utilisées. En effet, de nombreuses données issues du recensement de l'Insee sont exploitées. Le recensement Insee d'une année N fait intervenir les échantillons des années N-2, N-1, N, N+1 et N+2 (Insee, 2014). Donc, en 2016, année de cette étude, ont été publiées les données de recensement de l'année 2013 puisque celles-ci intègrent l'échantillon de 2015. Ainsi, comme des variables de recensement sont utilisées, les données les plus récentes à disposition sont issues de 2013. De plus, l'une des variables de 2013 utilisée, faisant partie de la base Filosofi de l'Insee, a été diffusée après la fin de cette étude. Cela justifie le choix de travailler sur des données de 2012.

La construction de toutes les variables intervenant dans les équations (8), (9) et (10) est explicitée ci-dessous. Il s'agit de  $q_s$ ,  $N_{s,assmat}$ ,  $N_{s,enfants}$ ,  $\alpha_s$ ,  $tt_s$ ,  $traj_s$ ,  $\lambda_s$ ,  $\mu_s$ ,  $med_s$ ,  $conc_s$ ,  $\sum_i dens_{i,s}$ ,  $\sum_i exp_{i,s}$ ,  $\sum_i loyer_{i,s}$ ,  $\sum_i age_{i,s}$ ,  $\sum_i age_{i,s}^2$ ,  $w_s$  et  $h_s$ .

Étant donné que les variables ont été construites par bassin de vie, la table d'appartenance géographique des communes de l'Insee a été utilisée pour associer chaque assistante maternelle de la base Pajemploi à son bassin de vie.

### E.a. Construction des variables relatives à la demande

Les variables relatives à la demande sont  $tt_s$ ,  $traj_s$ ,  $\lambda_s$ ,  $\mu_s$ ,  $med_s$  et  $N_{s,enfants}$ .

$tt_s$

Le temps de travail moyen  $tt_s$  des employeurs du bassin de vie a été approximé par le taux de personnes de plus de 15 ans travaillant à temps partiel sur le bassin de vie parmi toutes les personnes actives occupées de plus de 15 ans. Ces variables sont issues de la table de caractéristiques de l'emploi de l'Insee. La variable ainsi calculée est inversement proportionnelle au temps de travail  $tt_s$  du bassin de vie. Les signes obtenus dans les régressions devront donc être interprétés en tenant compte de cette inversion de signe attendu.

$$tt_s = \frac{\text{Nombre d'actifs occupés de plus de 15 ans à temps partiel}}{\text{Nombre d'actifs occupés de plus de 15 ans}}$$

Cette variable est quelque peu endogène puisque le choix de travailler à temps partiel est une alternative à faire garder son enfant par une assistante maternelle. Mais travailler à temps partiel et faire garder son enfant par une assistante maternelle le reste du temps est aussi possible. L'interprétation du signe de cette variable dans l'équation (8) déterminant  $\frac{q_s}{N_{s,assmat}}$  sera donc difficile. L'intérêt de cette variable est qu'elle capte bien la demande horaire adressée aux assistantes maternelles : si le taux de personnes travaillant à temps partiel dans un bassin de vie augmente, le temps de travail des employeurs potentiels des assistantes maternelles diminue.

$traj_s$

Le temps de trajet moyen  $traj_s$  des employeurs du bassin de vie a été calculé comme étant le taux de personnes se rendant au travail à pied ou n'utilisant pas de transport pour aller travailler parmi les actifs occupés. Ces données sont issues de la table des caractéristiques de l'emploi de l'Insee. Les autres options sont d'aller travailler en deux-roues, en voiture ou en transports en commun. Autant les deux premières options (marche et aucun transport) utilisées pour construire  $traj_s$  ne sont possibles que sur de courtes distances, autant les trois autres ne présagent pas de la distance domicile-travail. En effet, prendre sa voiture en grande ville et en milieu rural ne se fait pas sur les mêmes distances ou durées.

La variable ainsi construite est inversement proportionnelle au temps de trajet. Les signes attendus lors des régressions sont donc opposés aux signes attendus dans le système d'équations (8), (9) et (10).

Par ailleurs, cette variable est une source potentielle d'endogénéité car un parent travaillant depuis son domicile garde peut-être son enfant lui-même. Ceci est surtout valable pour l'équation (8) déterminant  $\frac{q_s}{N_{s,assmat}}$ .

$$traj_s = \frac{\text{Actifs occupés de plus de 15 ans utilisant la marche ou ne se déplaçant pas pour aller travailler}}{\text{Nombre d'actifs occupés de plus de 15 ans}}$$

$\lambda_s$  et  $\mu_s$

Pour construire ces variables, la table emploi - population active de l'Insee a été exploitée. Plutôt que d'utiliser des taux d'activité et de chômage classiques, les variables construites sont les taux d'activité et de chômage des personnes de 25 à 54 ans qui représentent probablement mieux les taux d'activité et de chômage des employeurs potentiels d'assistantes maternelles.

$$\lambda_s = \frac{\text{Nombre d'actifs ayant entre 25 et 54 ans}}{\text{Population totale ayant entre 25 et 54 ans}}$$

$$\mu_s = \frac{\text{Nombre de chômeurs ayant entre 25 et 54 ans}}{\text{Nombre d'actifs ayant entre 25 et 54 ans}}$$

$med_s$

Le salaire médian du bassin de vie n'est pas calculable en tant que tel à partir des données Insee. Les revenus disponibles médians par unité de consommation des communes sont donc pondérés par le nombre de ménages fiscaux de la commune. Cette moyenne des revenus médians est sans doute proche du revenu médian du bassin de vie, même si cela n'est pas vérifiable. Ces variables sont obtenues à partir de la table revenus et pauvreté des ménages du dispositif Filosofi de l'Insee.

$$med_s = \frac{\sum_c \text{Nombre de ménages fiscaux} * \text{Revenu disponible médian par unité de consommation}}{\sum_c \text{Nombre de ménages fiscaux de la commune } c}$$

$N_{s,enfants}$

Pour cette variable, la table POP1A de la base de recensement de l'Insee a été utilisée afin de sommer le nombre d'enfants de moins de 3 ans et de 3 à 5 ans.

$$N_{s,enfants} = \text{Nombre d'enfants de moins de 3 ans} + \text{Nombre d'enfants de 3 à 5 ans}$$

#### E.b. Construction des variables relatives à l'offre

Les variables relatives à l'offre sont  $\sum_i dens_{i,s}$ ,  $\sum_i exp_{i,s}$ ,  $\sum_i loyer_{i,s}$ ,  $\sum_i age_{i,s}$ ,  $\sum_i age_{i,s}^2$ .

Aucune variable pour mesurer le loyer  $loyer_{i,s}$  représentatif d'une commune n'était à disposition pour cette étude. Plus généralement, hormis le salaire médian d'une commune, aucune variable de coût de la vie pouvant approcher le loyer représentatif de la commune n'a été trouvée. Le salaire médian intervenant déjà à travers la variable  $med_s$ , il n'a pas été possible d'utiliser le salaire pour approcher le loyer. Cette dernière variable a donc été abandonnée. Cependant, le salaire médian étant probablement corrélé au loyer, il est possible que  $med_s$  capte une partie des coûts des assistantes maternelles dans les loyers. Cela peut avoir une influence sur le signe du coefficient de  $med_s$  dans l'équation (8) déterminant  $\frac{q_s}{N_{s,assmat}}$  puisqu'à la fois les variables  $med_s$  et  $\sum_i loyer_{i,s}$  interviennent dans des sens opposés.

Pour les variables  $\frac{\sum_i exp_{i,s}}{N_{s,assmat}}$ ,  $\frac{\sum_i age_{i,s}}{N_{s,assmat}}$  et  $\frac{\sum_i age_{i,s}^2}{N_{s,assmat}}$ , le même procédé de calcul a été mis en œuvre. Les variables relatives aux assistantes maternelles et disponibles dans la base Pajemploi ont été moyennées parmi les assistantes maternelles en activité dans le bassin de vie.

De même, la variable  $\frac{\sum_i dens_{i,s}}{N_{s,assmat}}$  a été construite en moyennant parmi les assistantes maternelles la densité (nombre d'habitants par km<sup>2</sup>) de leur commune disponible dans la table Résumé statistique de l'Insee.

#### E.c. Construction des variables endogènes

Les variables endogènes sont  $q_s$ ,  $w_s$  et  $h_s$ .

La base Pajemploi a été utilisée pour construire ces variables. Les données nettoyées comme précisé en annexe A.b ont été exploitées.  $q_s$  et  $h_s$  sont respectivement la somme du nombre d'enfants et la somme du nombre d'heures de garde des assistantes maternelles du bassin de vie.  $w_s$  est calculé comme la moyenne non pondérée des salaires horaires des assistantes maternelles du bassin de vie.

#### E.d. Construction des variables relatives à la concurrence

Les variables relatives à la concurrence sont  $N_{s,assmat}$ ,  $\alpha_s$  et  $conc_s$ .

$N_{s,assmat}$  est le nombre d'assistantes maternelles actives sur le bassin de vie et est calculé à partir de la base Pajemploi.

$\alpha_s$  est la préférence des employeurs du bassin de vie pour le mode de garde par assistante maternelle par rapport aux autres modes de garde payants.

$\alpha_s$

$$= \frac{\text{Nombre d'enfants gardés par des assistantes maternelles}}{\text{Nombre d'enfants gardés par des assistantes maternelles ou par des gardes à domicile ou en EAJE}}$$

Le nombre d'enfants gardés en EAJE (établissement d'accueil du jeune enfant) n'est pas fourni en tant que tel : le site de la CAF communique uniquement la table des places en EAJE par commune. Il sera supposé que les places sont saturées et que le nombre de places en EAJE est une bonne approximation du nombre d'enfants gardés en EAJE.

De même, Pajemploi contient des informations sur les gardes à domicile mais la base ne fournit que le nombre d'employeurs par garde à domicile et non le nombre d'enfants gardés. Pour contourner cette difficulté, le rapport du nombre d'enfants gardés par employeur par un système de garde à domicile est supposé identique au rapport du nombre d'enfants gardés par employeur par un système de garde par assistante maternelle sur un bassin de vie donné. Cela permet de calculer une approximation du nombre d'enfants gardés par garde à domicile.

Enfin, il reste à construire la variable de concurrence  $conc_s$ . Elle est construite de la manière suivante :

$$conc_s = \frac{\text{Capacité de garde des assistantes maternelles, des gardes à domicile et des EAJE}}{\text{Nombre de jeunes enfants } N_{s,enfants}}$$

La capacité de garde des EAJE est disponible puisque la CAF fournit le nombre de places disponibles en EAJE. Pour les assistantes maternelles, la capacité de garde de chaque assistante maternelle en 2012 est estimée comme étant le maximum du nombre d'enfants qu'elle a gardé parmi les mois d'avril, mai et juin 2012. Concernant les capacités de garde des gardes à domicile, la même règle de trois est utilisée que pour calculer le nombre d'enfants gardés par des gardes à domicile, en remplaçant les nombres d'enfants par les capacités de garde.

## F. Modèles alternatifs

### F.a. Modèle à deux catégories d'enfants

#### F.a.i. Trois catégories différentes d'enfants, ramenées à deux

En réalité, comme introduit dans le paragraphe « Utilité  $u_{j,s,t}(\cdot)$  des employeurs » en 1.1.4, il y a plusieurs catégories d'enfants : les enfants de moins de 3 ans, les enfants de 3 à 5 ans et les enfants de 6 ans et plus. Pourquoi introduire différentes catégories d'enfants et pourquoi définir ces seuils en particulier ? Ces seuils sont motivés par l'universalité de l'école maternelle, par celle de l'école élémentaire et par la faible propension en France à scolariser son enfant à 2 ans. En effet, à la rentrée de l'année scolaire 2015-2016, seuls 11,5 % des enfants de 2 ans étaient scolarisés et aucun enfant de moins de 2 ans n'était scolarisé. Le taux de scolarisation des enfants de 3 à 10 ans était, pour chaque âge, supérieur à 98 % (DEPP, 2016). Distinguer les enfants de 3 ans ou plus des enfants de moins de 3 ans est donc nécessaire puisque l'école maternelle à partir de 3 ans puis l'école élémentaire impliquent presque systématiquement la scolarisation des enfants de 3 ans ou plus. Pour simplifier, la possibilité de l'école maternelle à 2 ans n'est pas considérée.

Seule la situation duale dans laquelle un employeur choisit de travailler ou de ne pas travailler est prise en compte. Il n'y a pas de possibilité intermédiaire. Les enfants de moins de 3 ans sont gardés toute la journée par les assistantes maternelles ou le mode de garde adéquat lorsque l'employeur travaille. Par contre, presque tous les enfants de 3 à 5 ans étant scolarisés à l'école maternelle, les assistantes maternelles ou le mode de garde complémentaire de l'école les garderont beaucoup moins longtemps. La troisième catégorie d'enfants concerne les enfants de 6 ans et plus qui sont en totalité scolarisés à l'école élémentaire. Ils peuvent, en outre, être gardés lors d'heures extrascolaires (les heures d'étude surveillée par exemple) au sein de l'école élémentaire quand les parents ne peuvent pas venir les chercher à la fin de la journée scolaire. Les assistantes maternelles sont donc moins souvent sollicitées pour les enfants de 6 ans et plus et sur des durées a priori inférieures aux enfants de 3 à 5 ans. Cette hypothèse est doublement vérifiée par les chiffres. D'une part, les enfants de 6 ans et plus ne représentent que 6 % des enfants gardés par les assistantes maternelles (Vroylandt, 2016). D'autre part, il apparaît dans la base Pajemploi que les heures de garde sont nettement inférieures pour les enfants de 6 ans et plus que pour les enfants de 3 à 5 ans. Distinguer les enfants de 3 à 5 ans des enfants de 6 ans et plus s'impose donc, du fait de la structure du système scolaire en France. En outre, cette distinction entre les trois tranches d'âge se justifie également par la structure du CMG financé par la CAF qui distingue les enfants de moins de 3 ans, des enfants de 3 à 5 ans pour lesquels la CAF prend en charge une partie moindre du salaire des assistantes maternelles employées par les parents, des enfants de 6 ans et plus pour lesquels la CAF ne finance pas le salaire de l'assistante maternelle et les cotisations et contributions et sociales des employeurs d'après l'article L. 531-5 du Code de la sécurité sociale.

Trois catégories distinctes d'enfants doivent donc être considérées. Toutefois, les enfants de 6 ans et plus seront négligés puisqu'ils ne représentent que 6 % des enfants gardés.

#### F.a.ii. Un marché unique ou deux marchés distincts ?

Les assistantes maternelles peuvent choisir de plutôt se spécialiser dans la garde des enfants de moins de 3 ans ou dans celle de ceux de 3 à 5 ans. La différence intervient principalement en termes de nombre d'heures travaillées, ce qui se ressent pour l'assistante maternelle à travers ses coûts.

Il s'agit alors de savoir si les deux catégories d'enfants forment un marché unique de garde ou deux marchés différents. Si les marchés sont différents alors les salaires horaires pour les deux catégories d'enfants sont distincts ou identiques. S'il n'y a qu'un marché, les salaires horaires pour les deux catégories d'enfants sont identiques. Réciproquement, si les salaires horaires sont distincts, il y a nécessairement deux marchés distincts. Mais si les salaires sont identiques, cela n'indique pas si les marchés sont distincts ou identiques.

En observant les salaires médians des assistantes maternelles regroupées par le nombre d'enfants qu'elles gardent et par la répartition de cette garde entre les enfants de moins de 3 ans et

les enfants de 3 à 5 ans dans la figure 10, il est très net que les salaires horaires sont distincts pour la garde des enfants de 3 à 5 ans et pour la garde des enfants de moins de 3 ans. Ceci est visible de deux manières différentes. En diagonale (colonnes de couleur identique) pour un nombre d'enfants gardés de moins de 3 ans donné, le salaire horaire par enfant des assistantes maternelles diminue quand le nombre d'enfants accueillis augmente. Le salaire horaire de l'assistante maternelle diminue donc quand la part d'enfants gardés de moins de 3 ans diminue. D'autre part, pour un nombre d'enfants gardés donné, le salaire horaire par enfant augmente nettement lorsque le nombre d'enfants de moins de 3 ans augmente (colonnes de couleur différentes). Ainsi, cette observation plaide sans aucun doute pour deux marchés différents.

Des salaires horaires distincts entre les assistantes maternelles gardant plutôt des enfants de moins de 3 ans et les assistantes maternelles gardant plutôt des enfants de 3 à 5 ans peuvent provenir de différences dues à l'offre. Par exemple, il est possible que des assistantes maternelles gardent des enfants plus jeunes dans des zones où le coût de la vie est plus élevé. Les salaires horaires plus élevés pour ces assistantes maternelles s'expliqueraient alors par le coût de la vie de l'assistante maternelle et non par une différence de rémunération liée à l'âge de l'enfant gardé. Cela ne semble cependant pas être le cas ici puisque les résultats de la figure 10 sont conservés même lorsque les salaires horaires sont contrôlés par la tranche d'âge des assistantes maternelles ou par la géographie en utilisant les tranches d'unité urbaine définies par l'Insee.

Néanmoins, cette observation qui ne saurait être complète doit être largement nuancée. Jusqu'ici seuls les revenus salariaux des assistantes maternelles ont été pris en considération en négligeant les indemnités versées pour les frais d'entretien, pour les repas et pour les transports.

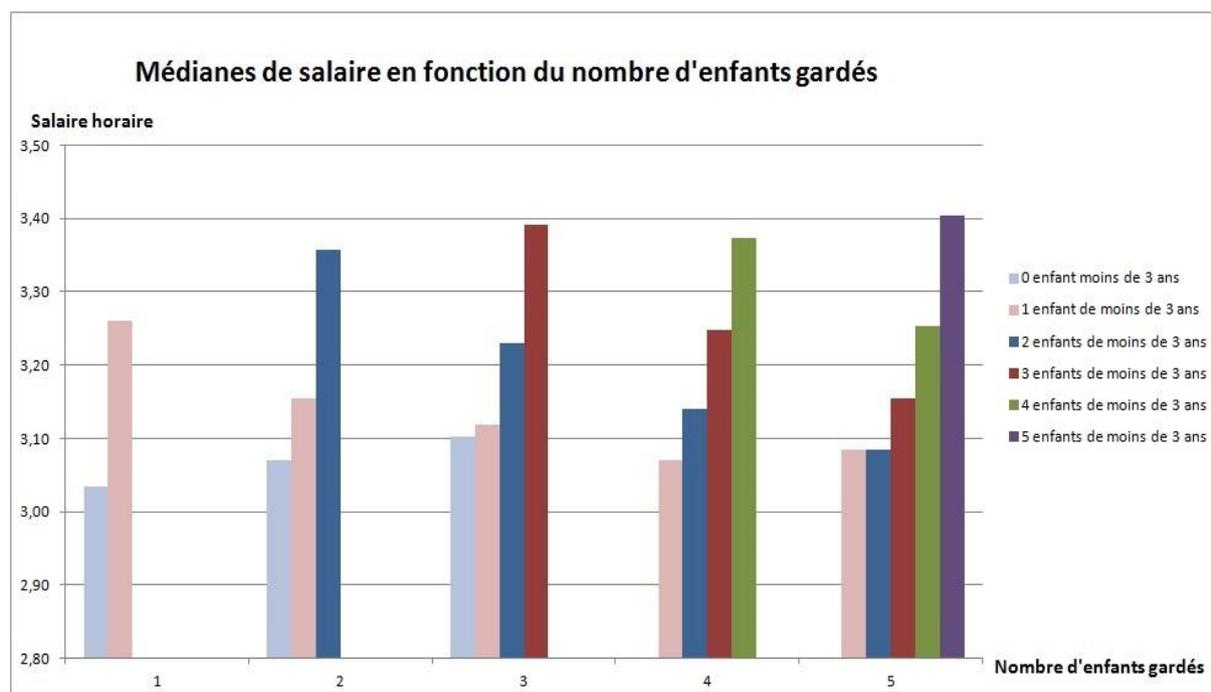


Figure 10 - Médiane des salaires parmi les assistantes maternelles, regroupées par le nombre d'enfants gardés et par la répartition de leur garde entre les différentes catégories d'enfants

La même analyse est donc refaite en sommant salaire horaire et indemnités horaires (figure 11). Cette fois, la tendance selon laquelle le salaire horaire pour un enfant de moins de 3 ans serait plus élevé que pour un enfant de 3 à 5 ans est moins nette. Lorsque ces statistiques sont contrôlées par les tranches d'unité urbaine (figure 12), l'âge des enfants n'influe plus sur le salaire horaire des assistantes maternelles. Ainsi, rien ne semble prouver que les assistantes maternelles tarifent plus cher leurs services pour des enfants de moins de 3 ans que pour des enfants de 3 à 5 ans. L'hypothèse d'un

marché unique est donc plausible. Cette hypothèse ne sera cependant pas suivie car la résolution du modèle n'a pas abouti à une solution interprétable avec celle-ci.

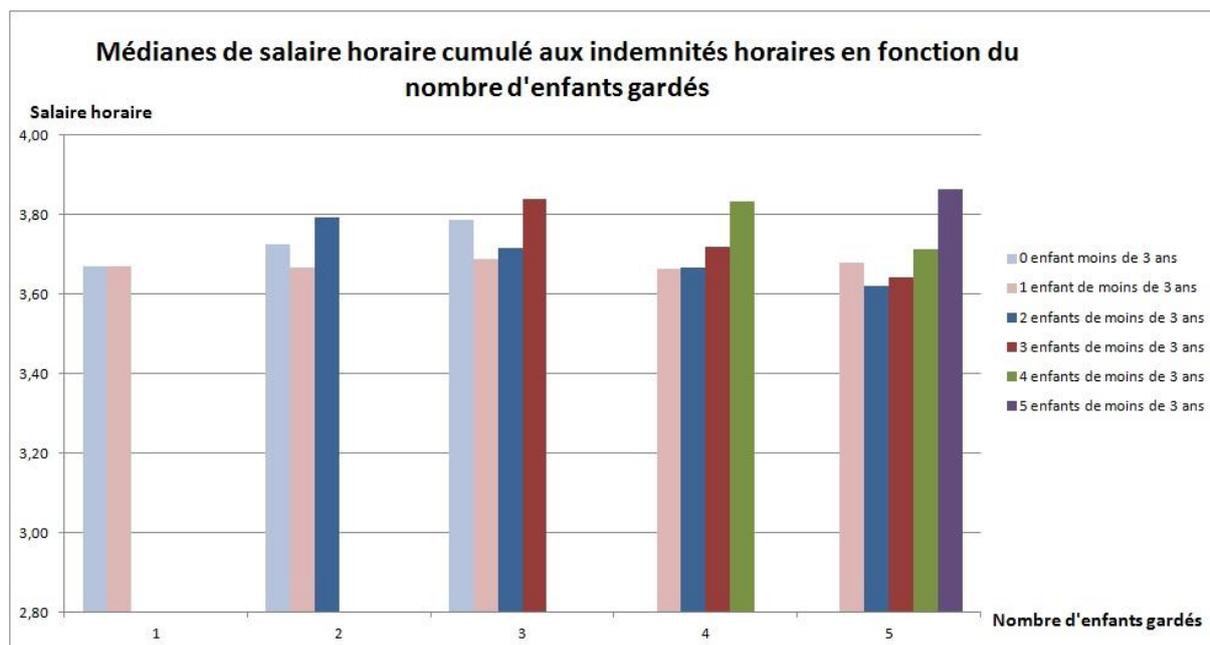


Figure 11 - Médiane des salaires cumulés aux indemnités parmi les assistantes maternelles, regroupées par le nombre d'enfants gardés et par la répartition de leur garde entre les différentes catégories d'enfants

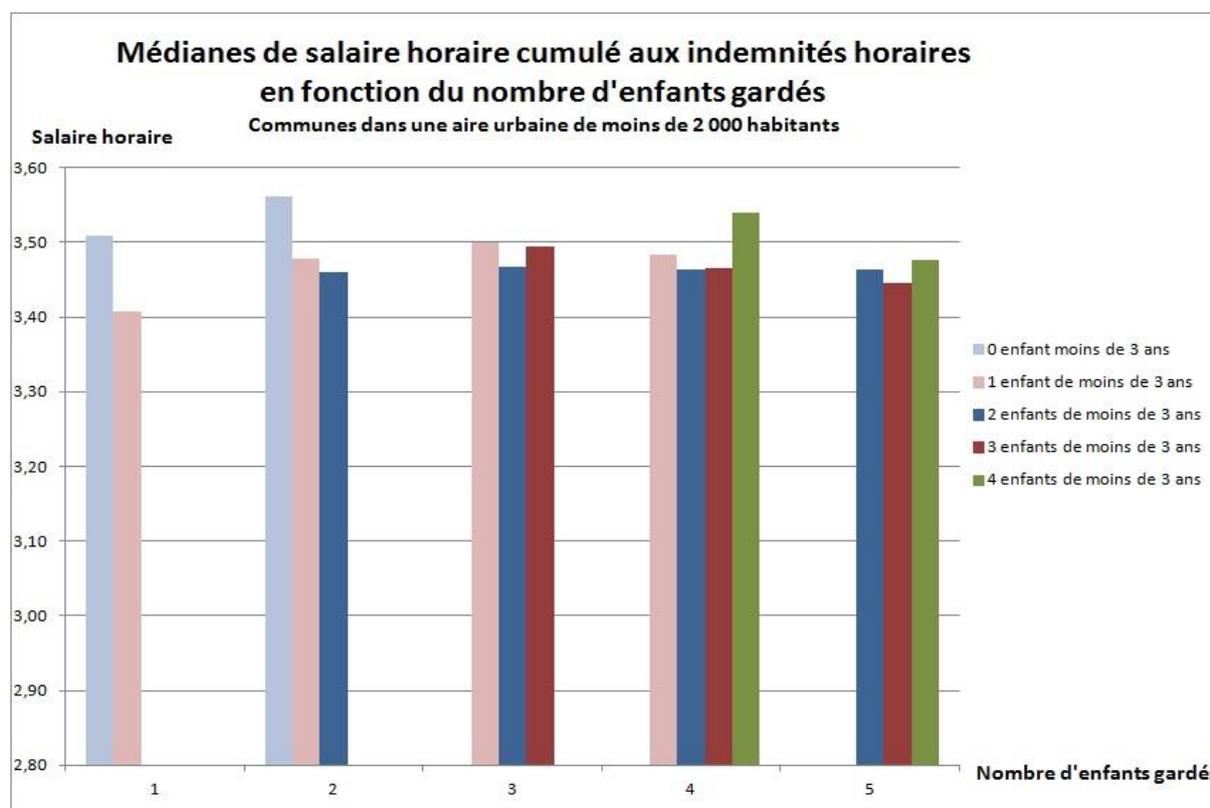


Figure 12 - Médiane des salaires cumulés aux indemnités parmi les assistantes maternelles, regroupées par le nombre d'enfants gardés et par la répartition de leur garde entre les différentes catégories d'enfants dans les aires urbaines de moins de 2000 habitants

Comment expliquer cependant ces différences de structure entre les salaires horaires et les indemnités des assistantes maternelles pour la garde des enfants de moins de 3 ans et enfants de 3 à 5 ans ? Tout d'abord, la part du salaire des salaires des assistantes maternelles prise en charge par la CAF dépend de l'âge des enfants à charge. Elle est deux fois plus importante pour les enfants de moins de 3 ans que pour les enfants de 3 à 5 ans du fait de l'article R. 421-3 du Code de la sécurité sociale. Dès lors, le fait que les assistantes maternelles aient un salaire plus élevé pour les enfants de moins de 3 ans que pour les enfants de 3 à 5 ans peut s'expliquer par une répartition différente entre les salaires et les indemnités. Toutefois, les indemnités à disposition dans la base Pajemploi ne semblent pas toujours parfaitement renseignées, ce qui peut fausser l'analyse précédente.

### F.a.iii. Détermination des fonctions modélisant les acteurs

#### Coûts $C_{i,s}$ des assistantes maternelles

La construction des coûts de l'assistante maternelle ne change pas. La pénibilité horaire reste la même et la pénibilité augmente toujours avec le nombre d'enfants par exemple. Les changements interviendront par l'intermédiaire des variables endogènes. Le nombre d'enfants de moins de 3 ans  $q_{i,s}^{0a2}$  gardés par l'assistante maternelle est désormais distingué du nombre d'enfants de 3 à 5 ans  $q_{i,s}^{3a5}$  qu'elle garde. Cela amène aussi à distinguer le nombre d'heures de garde par enfant gardé de moins de 3 ans  $h_{i,s}^{0a2}$  et le nombre d'heures par enfant gardé de 3 à 5 ans  $h_{i,s}^{3a5}$ . Enfin, dans le cas d'un marché unique, les salaires horaires sont identiques pour chaque catégorie d'enfants et égaux à  $w_s$ . Au contraire, dans le cas de deux marchés de garde distincts, le salaire horaire  $w_s^{0a2}$  pour les enfants de moins de 3 ans est différent du salaire horaire  $w_s^{3a5}$  pour les enfants de 3 à 5 ans.

Qu'il s'agisse d'un marché unique de garde ou de deux marchés distincts, le coût de l'assistante maternelle déterminé dans la section 1.1.4 par l'équation (1) qui a été simplifiée dans le paragraphe « Détermination du nombre d'enfants gardés par les assistantes maternelles » de la section 1.3.2 se réécrit :

$$\begin{aligned}
 C_{i,s} = & \beta_1 \left( h_{i,s}^{0a2} \frac{q_{i,s}^{0a2}}{q_{i,s}^{0a2} + q_{i,s}^{3a5}} + h_{i,s}^{3a5} \frac{q_{i,s}^{3a5}}{q_{i,s}^{0a2} + q_{i,s}^{3a5}} \right)^{p_1} + \beta_2 (q_{i,s}^{0a2} + q_{i,s}^{3a5})^{p_2} \\
 & + \beta_6 conc_s (q_{i,s}^{0a2} + q_{i,s}^{3a5}) + \beta_7 age_{i,s}^{p_3} (q_{i,s}^{0a2} + q_{i,s}^{3a5}) \\
 & + \beta_8 age_{i,s}^{p_4} \left( h_{i,s}^{0a2} \frac{q_{i,s}^{0a2}}{q_{i,s}^{0a2} + q_{i,s}^{3a5}} + h_{i,s}^{3a5} \frac{q_{i,s}^{3a5}}{q_{i,s}^{0a2} + q_{i,s}^{3a5}} \right) + \beta_9 dens_{i,s} (q_{i,s}^{0a2} + q_{i,s}^{3a5}) \\
 & + \beta_{10} exp_{i,s} (q_{i,s}^{0a2} + q_{i,s}^{3a5}) + \beta_{11} loyer_{i,s} (q_{i,s}^{0a2} + q_{i,s}^{3a5}) + F_{i,s}
 \end{aligned} \tag{22}$$

#### Utilité $u_{j,s}$ des employeurs pour les deux types d'enfants

Dans le paragraphe « Utilité  $u_{j,s,t}$  des employeurs » de la section 1.1.4, l'utilité de faire garder respectivement un enfant de moins de 3 ans et un enfant de 3 à 5 ans pour un parent était de la forme, en supposant qu'ils ne font garder qu'un enfant :

$$\begin{aligned}
 u_{j,s}^{<3\text{ ans}} = & \phi_1 \mathbb{1}\{h_{j,s} \leq tt_{j,s} + traj_{j,s}\} h_{j,s}^{r_1} \\
 & + \mathbb{1}\{h_{j,s} > tt_{j,s} + traj_{j,s}\} [\phi_1 (tt_{j,s} + traj_{j,s})^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{r_2}] \\
 & + \phi_2 \mathbb{1}\{h_{j,s} > tt_{j,s} + traj_{j,s}\} h_{j,s}^{r_2} \\
 u_{j,s}^{3\text{ à }5\text{ ans}} = & \phi_1 \mathbb{1}\{h_{j,s} + te_{j,s} \leq tt_{j,s} + traj_{j,s}\} (h_{j,s} + te_{j,s})^{r_1} \\
 & + \mathbb{1}\{h_{j,s} + te_{j,s} > tt_{j,s} + traj_{j,s}\} [\phi_1 (tt_{j,s} + traj_{j,s})^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{r_2}] \\
 & + \phi_2 \mathbb{1}\{h_{j,s} + te_{j,s} > tt_{j,s} + traj_{j,s}\} (h_{j,s} + te_{j,s})^{r_2}
 \end{aligned}$$

Les puissances  $r_1$  et  $r_2$  sont identiques pour les enfants de moins de 3 ans et les enfants de 3 à 5 ans car les besoins en garde (école et/ou mode de garde complémentaire) des parents sont les mêmes d'une catégorie d'enfants à l'autre.

#### Seuil de participation $k_{j,s}$ des employeurs

Dans le paragraphe « Seuil de participation  $k_{j,s,t}$  des employeurs » de la section 1.1.4 », l'équation (2) indiquait que le seuil de participation des employeurs pouvait s'écrire :

$$k_{j,s} = \eta_0 + \underset{-}{\eta_1} \lambda_s + \underset{+}{\eta_2} \mu_s + \underset{-}{\eta_3} med_s + \varepsilon_{j,s}$$

Cependant, rien ne garantit que ce seuil ait les mêmes paramètres pour des enfants de moins de 3 ans et pour des enfants de 3 à 5 ans.

Le seuil de participation des employeurs est probablement plus élevé pour les enfants de 3 à 5 ans que pour les enfants de moins de 3 ans. En effet, le coût d'opportunité de travailler à temps réduit ou d'aménagement du temps de travail pour pouvoir chercher son enfant à l'école est bien plus faible que celui de ne pas travailler. Le seuil de participation de l'employeur étant inversement proportionnel à ce coût d'opportunité, il est donc plus élevé pour les employeurs ayant des enfants de 3 à 5 ans.

Il faut donc définir une fonction de seuil de participation  $k_{j,s}^{3a5}$  pour les parents d'enfants de 3 à 5 ans. De la même manière qu'en déterminant l'équation (2) dans la partie 1.1.4 :

$$k_{j,s}^{0a2} = \eta_0^{0a2} + \underset{-}{\eta_1}^{0a2} \lambda_s + \underset{+}{\eta_2}^{0a2} \mu_s + \underset{-}{\eta_3}^{0a2} med_s + \varepsilon_{j,s}^{0a2}$$

$$k_{j,s}^{3a5} = \eta_0^{3a5} + \underset{-}{\eta_1}^{3a5} \lambda_s + \underset{+}{\eta_2}^{3a5} \mu_s + \underset{-}{\eta_3}^{3a5} med_s + \varepsilon_{j,s}^{3a5}$$

où le coefficient  $\eta_0^{3a5}$  est censé être supérieur à  $\eta_0^{0a2}$  puisque le coût d'opportunité de travailler à temps partiel est moindre que celui de ne pas travailler. Le coefficient  $\eta_1^{3a5}$  est supposé identique au coefficient  $\eta_1^{0a2}$ . En effet, les parents inactifs peuvent a priori tout autant chercher leurs enfants de 3 à 5 ans à l'école qu'ils peuvent s'occuper en journée de leurs enfants de moins de 3 ans, surtout lorsque le choix d'être inactif relève de la volonté d'être plus présent auprès de ses enfants. Au contraire, un chômeur en recherche active de travail pourra plus simplement aménager sa recherche de travail pour chercher ses enfants à l'école que pour les garder toute la journée. Il semble donc plus probable qu'un chômeur ayant des enfants participe au marché de la garde s'il a un enfant de moins de 3 ans qu'un enfant de 3 à 5 ans. Donc  $\eta_2^{3a5} > \eta_2^{0a2} > 0$ . D'autre part travailler à temps partiel est financièrement plus simple que de ne pas travailler du tout afin de garder ses enfants de moins de 3 ans toute la journée. Donc les employeurs seront plus sensibles à une évolution de leur salaire pour les enfants de 3 à 5 ans que de moins de 3 ans. Donc  $|\eta_3^{3a5}| > |\eta_3^{0a2}|$ .

#### F.a.iv. Résolution du modèle théorique dans le cas général

##### Choix du nombre d'heures de garde par les parents

D'après le paragraphe « Utilité  $u_{j,s}$  des employeurs pour les deux types d'enfants » de la section F.a.iii, l'utilité des parents a la forme suivante :

$$u_{j,s}^{<3\text{ ans}} = \phi_1 \mathbb{1}\{h_{j,s}^{0a2} \leq tt_{j,s} + traj_{j,s}\} h_{j,s}^{0a2 r_1} + \mathbb{1}\{h_{j,s}^{0a2} > tt_{j,s} + traj_{j,s}\} [\phi_1 (tt_{j,s} + traj_{j,s})^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{r_2}] + \phi_2 \mathbb{1}\{h_{j,s}^{0a2} > tt_{j,s} + traj_{j,s}\} h_{j,s}^{0a2 r_2}$$

$$u_{j,s}^{3\text{ à }5\text{ ans}} = \phi_1 \mathbb{1}\{h_{j,s}^{3a5} + te_{j,s} \leq tt_{j,s} + traj_{j,s}\} (h_{j,s}^{3a5} + te_{j,s})^{r_1} + \mathbb{1}\{h_{j,s}^{3a5} + te_{j,s} > tt_{j,s} + traj_{j,s}\} [\phi_1 (tt_{j,s} + traj_{j,s})^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{r_2}] + \phi_2 \mathbb{1}\{h_{j,s}^{3a5} + te_{j,s} > tt_{j,s} + traj_{j,s}\} (h_{j,s}^{3a5} + te_{j,s})^{r_2}$$

Sans perte de généralité,  $w_s^{0a2}$  et  $w_s^{3a5}$  sont supposés distincts, quitte à ajouter une contrainte d'égalité pour la suite. Toujours par souci de simplification, les parents ne font garder qu'un enfant et les parents sont supposés symétriques. Cette dernière condition implique que tous les employeurs d'un bassin de vie ont le même temps de travail  $tt_s$  et le même temps de trajet  $traj_s$ . Le coût de l'employeur ayant un enfant de moins de 3 ans est alors  $h_{j,s}^{0a2} w_s^{0a2}$ . Celui d'un employeur ayant un enfant de 3 à 5 ans est  $h_{j,s}^{3a5} w_s^{3a5}$ .

La figure 4 de la section 1.3.2 donne les différents choix possibles pour les employeurs ayant des enfants de moins de 3 ans. Comme dans la situation à une catégorie unique d'enfants, la solution où les parents font garder leur enfant de moins de 3 ans durant leur temps de trajet et leur temps de travail est retenue. Donc  $h_{j,s}^{0a2} = tt_s + traj_s \forall j$ . Et  $h_s^{0a2} = \sum_j h_{j,s}^{0a2} = (tt_s + traj_s) q_s^{0a2}$ .

Pour les enfants de 3 à 5 ans, la fonction d'utilité et la fonction de coûts ont une forme présentée dans la figure 13.

La figure 13 montre qu'en maximisant le nombre d'heures de garde, trois choix s'offrent aux parents des enfants de 3 à 5 ans, tout comme dans le cas des enfants de moins de 3 ans (figure 4). La première option est qu'ils ne font garder leurs enfants que pendant le temps d'école qui est gratuit et n'ont donc pas recours à une assistante maternelle. Il s'agit du cas 1 : les parents font garder leur enfant uniquement par l'école maternelle durant le temps d'école  $te_s = te_{j,s}$  par symétrie des parents. Les durées de journée de cours par école maternelle sont supposées identiques dans un bassin de vie puisque la semaine d'enseignement est fixée légalement à 24 heures par semaine. Mais les parents peuvent aussi décider de recourir à une assistante maternelle durant le temps qui sépare leur temps de travail et de trajet du temps d'école de l'enfant. Les parents font alors garder leur enfant par l'assistante maternelle durant  $tt_s + traj_s - te_s = \max$  cas 2 - max cas 1. Il s'agit du cas 2. Enfin, les parents peuvent souhaiter faire garder leur enfant plus longtemps car le salaire horaire des assistantes maternelles est suffisamment faible pour qu'il soit rentable pour eux de conserver du temps de loisir. Il s'agit du cas 3. Les employeurs font alors garder leur enfant durant  $\max$  cas 3 -  $te_s = \max$  cas 3 -  $\max$  cas 1.

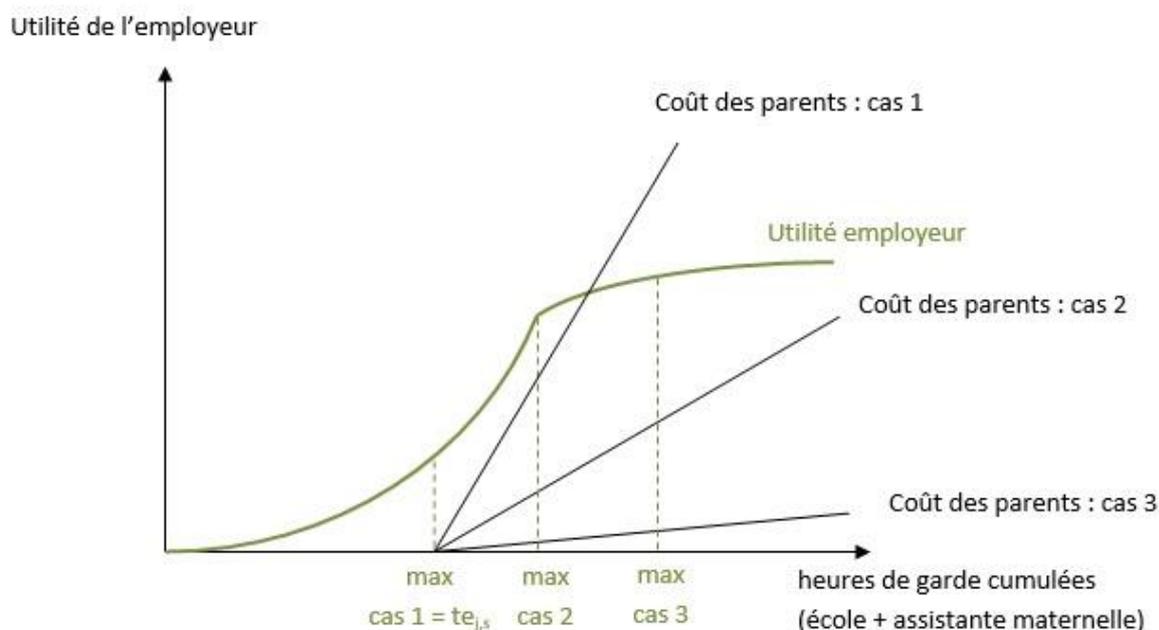


Figure 13 - Fonctions d'utilité et de coût de l'employeur pour les enfants de 3 à 5 ans

Par souci de cohérence et de comparaison avec la situation dans laquelle il n'y a qu'une catégorie d'enfants, seul le cas 2 sera étudié : les parents font garder leur enfant par l'assistante maternelle durant  $tt_s + traj_s - te_s$ . Ainsi,  $h_s^{3a5} = \sum_j h_{j,s}^{3a5} = (tt_s + traj_s - te_s) q_s^{3a5}$ .

#### Détermination de la demande locale

Une fois déterminées les heures de garde  $h_{j,s}^{0a2}$  et  $h_{j,s}^{3a5}$ , les mêmes arguments et la même démarche que dans le paragraphe « Détermination de la demande locale » de la section 1.3.2 sont conservés pour calculer les équations de demande de garde :

$$\frac{q_s^{0a2}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} = \phi_1 (tt_s + traj_s)^{r_1} - (tt_s + traj_s) w_s^{0a2} - \eta_0^{0a2} - \eta_1^{0a2} \lambda_s - \eta_2^{0a2} \mu_s - \eta_3^{0a2} med_s \quad (23)$$

$$\frac{q_s^{3a5}}{N_{s,enfants}^{3a5} \alpha_s^{3a5}} = \phi_1 (tt_s + traj_s)^{r_1} - (tt_s + traj_s - te_s) w_s^{3a5} - \eta_0^{3a5} - \eta_1^{3a5} \lambda_s - \eta_2^{3a5} \mu_s - \eta_3^{3a5} med_s \quad (24)$$

#### F.a.v. Résolution du modèle côté assistantes maternelles dans le cas de deux marchés distincts

Jusqu'à présent, le fait qu'il existe potentiellement un marché unique de garde ou deux marchés distincts n'a pas été utilisé. L'option du marché unique de garde n'ayant pas abouti à une solution explicite, le cas développé ci-dessous est la situation dans laquelle les deux marchés sont distincts.

Les assistantes maternelles maximisent leur profit :

$$\Pi_{i,s} = q_{i,s}^{0a2} h_{i,s}^{0a2} w_s^{0a2} + q_{i,s}^{3a5} h_{i,s}^{3a5} w_s^{3a5} - \beta_1 h_{i,s}^{p_1} - \beta_2 q_{i,s}^{p_2} - g_{i,s} q_{i,s} - \beta_8 age_{i,s}^{p_4} h_{i,s} - F_{i,s}$$

en notant  $g_{i,s} = \beta_6 conc_s + \beta_7 age_{i,s}^{p_3} + \beta_9 dens_{i,s} + \beta_{10} exp_{i,s} + \beta_{11} loyer_{i,s}$  afin de simplifier les notations, et en repartant de la fonction de coûts déterminée en équation (22) dans la section F.a.iii.

Toutefois, afin d'obtenir une solution explicite et interprétable, le coût de pénibilité horaire  $\beta_1 h_{i,s}^{p_1}$  a été négligé. Ce choix n'est pas idéal puisqu'il amène à négliger le fait qu'il est plus pénible pour une assistante maternelle qui garde un nombre donné d'enfants de garder ces enfants plus longtemps.

De même, dans le but d'obtenir une solution explicite, le terme de pénibilité horaire de l'assistante maternelle croisé à son âge  $\beta_8 age_{i,s}^{p_4} h_{i,s}$  a été remplacé par  $\beta_8 age_{i,s}^{p_4} h_{i,s} q_{i,s}$ . Enfin, dans le même but,  $p_2 = 2$  est supposé.

La fonction de profit se réécrit donc :

$$\Pi_{i,s} = q_{i,s}^{0a2} h_{i,s}^{0a2} w_s^{0a2} + q_{i,s}^{3a5} h_{i,s}^{3a5} w_s^{3a5} - \beta_2 (q_{i,s}^{0a2} + q_{i,s}^{3a5})^2 - g_{i,s} (q_{i,s}^{0a2} + q_{i,s}^{3a5}) - \beta_8 age_{i,s}^{p_4} (h_{i,s}^{0a2} q_{i,s}^{0a2} + h_{i,s}^{3a5} q_{i,s}^{3a5}) - F_{i,s}$$

En utilisant les relations  $h_{i,s}^{0a2} = (tt_s + traj_s)$  et  $h_{i,s}^{3a5} = (tt_s + traj_s - te_s)$  montrées dans le paragraphe « Choix du nombre d'heures de garde par les parents » de la section F.a.iv<sup>18</sup>, la fonction de profit s'écrit :

$$\begin{aligned} \Pi_{i,s} = & q_{i,s}^{0a2} (tt_s + traj_s) w_s^{0a2} + q_{i,s}^{3a5} (tt_s + traj_s - te_s) w_s^{3a5} - \beta_2 (q_{i,s}^{0a2} + q_{i,s}^{3a5})^2 \\ & - g_{i,s} (q_{i,s}^{0a2} + q_{i,s}^{3a5}) - \beta_8 age_{i,s}^{p_4} \left( (tt_s + traj_s) q_{i,s}^{0a2} + (tt_s + traj_s - te_s) q_{i,s}^{3a5} \right) \\ & - F_{i,s} \end{aligned}$$

En utilisant les équations (23) et (24) de la section F.a.iv et en notant  $D_s^{0a2} = \eta_0^{0a2} + \eta_1^{0a2} \lambda_s + \eta_2^{0a2} \mu_s + \eta_3^{0a2} med_s$  et  $D_s^{3a5} = \eta_0^{3a5} + \eta_1^{3a5} \lambda_s + \eta_2^{3a5} \mu_s + \eta_3^{3a5} med_s$  :

$$\begin{aligned} \Pi_{i,s} = & q_{i,s}^{0a2} \left( \phi_1 (tt_s + traj_s) r_1 - \frac{q_s^{0a2}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} - D_s^{0a2} \right) \\ & + q_{i,s}^{3a5} \left( \phi_1 (tt_s + traj_s) r_1 - \frac{q_s^{3a5}}{N_{s,enfants}^{3a5} \alpha_s^{3a5}} - D_s^{3a5} \right) - \beta_2 (q_{i,s}^{0a2} + q_{i,s}^{3a5})^2 \\ & - g_{i,s} (q_{i,s}^{0a2} + q_{i,s}^{3a5}) - \beta_8 age_{i,s}^{p_4} \left( (tt_s + traj_s) q_{i,s}^{0a2} + (tt_s + traj_s - te_s) q_{i,s}^{3a5} \right) \\ & - F_{i,s} \end{aligned}$$

Les assistantes maternelles maximisent alors leur profit selon le nombre d'enfants gardés de chaque tranche d'âge :

$$\frac{\partial \Pi_{i,s}}{\partial q_{i,s}^{0a2}} = 0$$

$$\begin{aligned} \Leftrightarrow & \phi_1 (tt_s + traj_s) r_1 - \frac{q_s^{0a2}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} - D_s^{0a2} - \frac{q_{i,s}^{0a2}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} \\ & = 2 \beta_2 (q_{i,s}^{0a2} + q_{i,s}^{3a5}) + g_{i,s} + \beta_8 age_{i,s}^{p_4} (tt_s + traj_s) \end{aligned}$$

$$\frac{\partial \Pi_{i,s}}{\partial q_{i,s}^{3a5}} = 0$$

$$\begin{aligned} \Leftrightarrow & \phi_1 (tt_s + traj_s) r_1 - \frac{q_s^{3a5}}{N_{s,enfants}^{3a5} \alpha_s^{3a5}} - D_s^{3a5} - \frac{q_{i,s}^{3a5}}{N_{s,enfants}^{3a5} \alpha_s^{3a5}} \\ & = 2 \beta_2 (q_{i,s}^{0a2} + q_{i,s}^{3a5}) + g_{i,s} + \beta_8 age_{i,s}^{p_4} (tt_s + traj_s - te_s) \end{aligned}$$

En sommant sur les assistantes maternelles :

$$\begin{aligned} & \phi_1 N_{s,assmat} (tt_s + traj_s) r_1 - \frac{(N_{s,assmat} + 1)}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} q_s^{0a2} - N_{s,assmat} D_s^{0a2} \\ & = 2 \beta_2 (q_s^{0a2} + q_s^{3a5}) + \sum_i g_{i,s} + \beta_8 \sum_i age_{i,s}^{p_4} (tt_s + traj_s) \end{aligned} \quad (25)$$

<sup>18</sup> En effet, les parents étant symétriques, le nombre d'heures de garde  $h_{i,s}^{0a2}$  par enfant de moins de 3 ans pour une assistante maternelle est égal au nombre d'heures de garde  $h_{j,s}^{0a2}$  de l'enfant de chaque employeur  $j$ . De même  $h_{i,s}^{3a5} = h_{j,s}^{3a5}$ .

Et

$$\begin{aligned} \phi_1 N_{s,assmat} (tt_s + traj_s)^{r_1} - \frac{(N_{s,assmat} + 1)}{N_{s,enfants}^{3a5} \alpha_s^{3a5}} q_s^{3a5} - N_{s,assmat} D_s^{3a5} \\ = 2 \beta_2 (q_s^{0a2} + q_s^{3a5}) + \sum_i g_{i,s} + \beta_8 \sum_i age_{i,s}^{p_4} (tt_s + traj_s - te_s) \end{aligned}$$

En faisant la différence entre ces deux relations :

$$\frac{q_s^{3a5}}{N_{s,enfants}^{3a5} \alpha_s^{3a5}} - \frac{q_s^{0a2}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} = \frac{N_{s,assmat}}{N_{s,assmat} + 1} (D_s^{0a2} - D_s^{3a5}) + \beta_8 te_s \frac{\sum_i age_{i,s}^{p_4}}{N_{s,assmat} + 1} \quad (26)$$

$\frac{q_s^{3a5}}{N_{s,enfants}^{3a5} \alpha_s^{3a5}}$  et  $\frac{q_s^{0a2}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}}$  sont des taux de garde. Il s'agit du nombre d'enfants effectivement gardés sur le nombre de parents qui, s'ils faisaient garder leur enfant solliciteraient une assistante maternelle. Il s'agit donc du nombre de parents qui participent à la première étape du jeu défini dans la section 1.2.2, avant donc de choisir s'ils participent ou non au marché de la garde des assistantes maternelles. Le rapport des enfants de 3 à 5 ans augmente par rapport à celui des moins de 3 ans quand  $D_s^{0a2} - D_s^{3a5}$  augmente, c'est-à-dire quand le seuil de participation des parents d'enfants de moins de 3 ans augmente par rapport à celui des parents d'enfants de 3 à 5 ans. Ils participent alors moins à la troisième étape du jeu par rapport aux parents d'enfants de 3 à 5 ans. Cela se produit quand le taux de chômage diminue et quand le salaire médian du bassin de vie augmente. En effet, les parents d'enfants de moins de 3 ans peuvent moins moduler leur temps de travail que les parents d'enfants de 3 à 5 ans. Ces derniers peuvent passer du temps partiel au temps complet plus facilement que les parents d'enfants de moins de 3 ans dont l'arbitrage est plutôt d'être inactif pour garder leur enfant ou travailler.

Par ailleurs, quand les assistantes maternelles du bassin de vie sont plus âgées, elles vont plutôt satisfaire la demande des parents des enfants de 3 à 5 ans que celle des parents des enfants de moins de 3 ans. Cela semble cohérent au niveau individuel mais discutable au niveau agrégé. En effet, il semble peu probable que l'offre de garde pour les enfants de moins de 3 ans dans un bassin de vie diminue lorsque les assistantes maternelles sont en moyenne un peu plus âgées dans ce bassin de vie par rapport à la moyenne d'âge des assistantes maternelles en France. Il est donc probable que le coefficient de la variable d'âge ne soit pas significatif empiriquement dans cette équation.

En remplaçant  $\frac{q_s^{3a5}}{N_{s,enfants}^{3a5} \alpha_s^{3a5}}$  dans l'équation (25) par sa formule dans l'équation (26) :

$$\begin{aligned}
& \phi_1 N_{s,assmat} (tt_s + traj_s)^{r_1} - \frac{(N_{s,assmat} + 1) q_s^{0a2}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} - N_{s,assmat} D_s^{0a2} \\
& = 2 \beta_2 \left( q_s^{0a2} + \frac{N_{s,enfants}^{3a5} \alpha_s^{3a5}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} q_s^{0a2} \right. \\
& \quad + N_{s,enfants}^{3a5} \alpha_s^{3a5} \frac{N_{s,assmat}}{N_{s,assmat} + 1} (D_s^{0a2} - D_s^{3a5}) \\
& \quad \left. + \beta_8 te_s N_{s,enfants}^{3a5} \alpha_s^{3a5} \frac{\sum_i age_{i,s}^{p_4}}{N_{s,assmat} + 1} \right) + \sum_i g_{i,s} + \beta_8 \sum_i age_{i,s}^{p_4} (tt_s + traj_s) \\
\Leftrightarrow & \phi_1 N_{s,assmat} (tt_s + traj_s)^{r_1} - N_{s,assmat} D_s^{0a2} - \sum_i g_{i,s} - \beta_8 \sum_i age_{i,s}^{p_4} (tt_s + traj_s) \\
& - 2 \beta_2 N_{s,enfants}^{3a5} \alpha_s^{3a5} \frac{N_{s,assmat}}{N_{s,assmat} + 1} (D_s^{0a2} - D_s^{3a5}) - 2 \beta_2 \beta_8 te_s N_{s,enfants}^{3a5} \alpha_s^{3a5} \frac{\sum_i age_{i,s}^{p_4}}{N_{s,assmat} + 1} \\
& = \left[ \frac{N_{s,assmat} + 1}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} + 2 \beta_2 + 2 \beta_2 \frac{N_{s,enfants}^{3a5} \alpha_s^{3a5}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} \right] q_s^{0a2} \\
\Leftrightarrow q_s^{0a2} = & \frac{\phi_1 N_{s,assmat} (tt_s + traj_s)^{r_1} - N_{s,assmat} D_s^{0a2} - \sum_i g_{i,s} - \beta_8 \sum_i age_{i,s}^{p_4} (tt_s + traj_s)}{\frac{N_{s,assmat} + 1}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} + 2 \beta_2 + 2 \beta_2 \frac{N_{s,enfants}^{3a5} \alpha_s^{3a5}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}}} \\
& - \frac{2 \beta_2 N_{s,enfants}^{3a5} \alpha_s^{3a5} \frac{N_{s,assmat}}{N_{s,assmat} + 1} (D_s^{0a2} - D_s^{3a5})}{\frac{N_{s,assmat} + 1}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} + 2 \beta_2 + 2 \beta_2 \frac{N_{s,enfants}^{3a5} \alpha_s^{3a5}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}}} \\
& - \frac{2 \beta_2 \beta_8 te_s N_{s,enfants}^{3a5} \alpha_s^{3a5} \frac{\sum_i age_{i,s}^{p_4}}{N_{s,assmat} + 1}}{\frac{N_{s,assmat} + 1}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} + 2 \beta_2 + 2 \beta_2 \frac{N_{s,enfants}^{3a5} \alpha_s^{3a5}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}}}
\end{aligned}$$

#### F.a.vi. Analyse des prédictions du modèle

Le modèle économique auquel cette annexe aboutit est le suivant :

$$\begin{aligned}
q_s^{0a2} = & \frac{\phi_1 N_{s,assmat} (tt_s + traj_s)^{r_1} - N_{s,assmat} D_s^{0a2} - \sum_i g_{i,s} - \beta_8 \sum_i age_{i,s}^{p_4} (tt_s + traj_s)}{\frac{N_{s,assmat} + 1}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} + 2 \beta_2 + 2 \beta_2 \frac{N_{s,enfants}^{3a5} \alpha_s^{3a5}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}}} \\
& - \frac{2 \beta_2 N_{s,enfants}^{3a5} \alpha_s^{3a5} \frac{N_{s,assmat}}{N_{s,assmat} + 1} (D_s^{0a2} - D_s^{3a5})}{\frac{N_{s,assmat} + 1}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} + 2 \beta_2 + 2 \beta_2 \frac{N_{s,enfants}^{3a5} \alpha_s^{3a5}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}}} \\
& - \frac{2 \beta_2 \beta_8 te_s N_{s,enfants}^{3a5} \alpha_s^{3a5} \frac{\sum_i age_{i,s}^{p_4}}{N_{s,assmat} + 1}}{\frac{N_{s,assmat} + 1}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} + 2 \beta_2 + 2 \beta_2 \frac{N_{s,enfants}^{3a5} \alpha_s^{3a5}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}}}
\end{aligned}$$

$$\frac{q_s^{3a5}}{N_{s,enfants}^{3a5} \alpha_s^{3a5}} - \frac{q_s^{0a2}}{N_{s,enfants}^{0a2} \alpha_s^{0a2}} = \frac{N_{s,assmat}}{N_{s,assmat} + 1} (D_s^{0a2} - D_s^{3a5}) + \beta_8 te_s \frac{\sum_i age_{i,s}^{p_4}}{N_{s,assmat} + 1}$$

$$w_s^{0a2} = \phi_1 (tt_s + traj_s)^{r_1-1} - \frac{q_s^{0a2}}{N_{s,enfants}^{0a2} \alpha_s^{0a2} (tt_s + traj_s)} - \frac{D_s^{0a2}}{tt_s + traj_s}$$

$$w_s^{3a5} = \phi_1 \frac{(tt_s + traj_s)^{r_1}}{tt_s + traj_s - te_s} - \frac{q_s^{3a5}}{N_{s,enfants}^{3a5} \alpha_s^{3a5} (tt_s + traj_s - te_s)} - \frac{D_s^{3a5}}{tt_s + traj_s - te_s}$$

$$h_s^{0a2} = (tt_s + traj_s) q_s^{0a2}$$

$$h_s^{3a5} = (tt_s + traj_s - te_s) q_s^{3a5}$$

Il est composé de 6 équations. Transposé en système économétrique à équations simultanées, il sera identifiable. En effet, il s'agira d'un système à équations simultanées triangulaire supérieur et pour chaque variable endogène, il y a au moins une variable exogène présente dans l'équation qui la détermine et qui est exclue des équations dans lesquelles la variable endogène est explicative.

#### *Variables exogènes à influence ambiguë*

Tout d'abord, la positivité de  $q_s^{0a2}$  est supposée acquise, ce qui n'est pas évident au vu de sa formule. Certaines variables interviennent à plusieurs reprises dans la formule de  $q_s^{0a2}$  et dans des sens opposés. Il n'est donc pas possible d'anticiper leur influence sur le nombre d'enfants de moins de 3 ans gardés par les assistantes maternelles. Ainsi, l'influence des évolutions du temps de travail  $tt_s$  et du temps de trajet  $traj_s$  sur le nombre d'enfants de moins de 3 ans dépend pleinement de la valeur des paramètres et des variables exogènes. De même, le nombre d'assistantes maternelles  $N_{s,assmat}$  intervient dans divers termes et dans des sens opposés dans la formule du nombre d'enfants de moins de 3 ans gardés par des assistantes maternelles  $q_s^{0a2}$ .

$q_s^{0a2}$  dépend négativement de  $D_s^{0a2}$ . Donc quand le seuil de participation des parents des enfants de moins de 3 ans augmente, le nombre d'enfants de moins de 3 ans gardés par des assistantes maternelles diminue. Cependant, concrètement, cela se vérifie difficilement par l'économétrie. En effet,  $D_s^{3a5}$  intervient dans le sens opposé de  $D_s^{0a2}$ . Comme ces deux variables exogènes sont des combinaisons linéaires des mêmes variables exogènes (taux d'activité sur le bassin de vie, taux de chômage et salaire médian), il n'est pas possible de prévoir l'influence de l'évolution d'une de ces variables sur la garde d'enfants de moins de 3 ans.

La demande potentielle en garde pour les enfants de 3 à 5 ans  $N_{s,enfants}^{3a5} \alpha_s^{3a5}$  a un effet ambigu sur la garde des enfants de moins de 3 ans  $q_s^{0a2}$ . En effet,  $D_s^{0a2} - D_s^{3a5}$  peut aussi bien être positif que négatif. Pourtant, il avait été argumenté dans le paragraphe « Seuil de participation  $k_{j,s}$  des employeurs » de la section F.a.iii que le seuil de participation est *a priori* plus élevé pour les employeurs parents d'enfants de 3 à 5 ans que pour les parents d'enfants de moins de 3 ans. Donc  $D_s^{0a2} - D_s^{3a5}$  est négatif et le terme qui lui est lié est donc de signe opposé par rapport aux autres termes dans lesquels  $N_{s,enfants}^{3a5} \alpha_s^{3a5}$  intervient. Il est donc impossible d'anticiper l'effet d'une hausse du nombre d'enfants de 3 à 5 ans sur la garde des enfants de moins de 3 ans.

### Variables exogènes à influence explicite

Certaines variables exogènes font toujours varier la garde des enfants de moins de 3 ans dans un sens déterminé. C'est le cas des coûts des assistantes maternelles  $g_{i,s}$ . Lorsqu'ils croissent, la garde des enfants de moins de 3 ans dans le bassin de vie diminue. Donc quand la concurrence augmente (à nombre d'assistantes maternelles constant), quand la densité diminue, quand les assistantes maternelles sont peu expérimentées, quand elles sont plus âgées (cela intervient aussi à travers le terme lié à  $\sum_i age_{i,s}^{p_4}$ ) et quand les loyers du bassin de vie sont plus élevés, le nombre d'enfants de moins de 3 ans gardés diminue. Ces constats peuvent être nuancés : il paraît logique lorsqu'une assistante maternelle isolée est étudiée mais quand la garde des assistantes maternelles est agrégée, le raisonnement n'est pas évident. L'économétrie permet donc de conforter ou infirmer ces prédictions du modèle et donc les fonctions de coûts qui ont été choisies.

Le nombre d'heures d'école  $te_s$  joue positivement sur le nombre d'enfants de 3 à 5 ans gardés : plus la durée de la semaine scolaire est élevée pour les 3 à 5 ans, plus les assistantes maternelles vont souhaiter garder ces enfants. Là encore, cette prévision (non vérifiable empiriquement) est surprenante car les assistantes maternelles ont tout de même besoin de faire un certain nombre d'heures pour que leur métier soit rentable. Même si dans le même temps leur salaire pour les enfants de 3 à 5 ans augmente. Cette prévision provient de l'introduction dans la fonction de profit des assistantes maternelles un coût de fatigue horaire  $\beta_8 age_{i,s}^{p_4} h_{i,s}$ . Ainsi, dans le cadre du modèle, une heure de travail engendre déjà de la fatigue. En réalité, il avait été initialement suggéré un coût horaire de la forme  $\beta_1 \mathbb{1} \left\{ \frac{\int_s h_{i,j,s,t} q_{i,j,s,t} dj}{\int_s jours_{i,j,s,t} q_{i,j,s,t} dj} \geq 8 \right\} \left( \frac{\int_s h_{i,j,s,t} q_{i,j,s,t} dj}{\int_s jours_{i,j,s,t} q_{i,j,s,t} dj} - 8 \right)^{p_1}$  dans l'annexe C.a.i. L'idée était que l'assistante maternelle ne ressent la fatigue qu'au-delà d'un certain nombre d'heures par jour travaillées par enfant. Cette hypothèse, plus réaliste que celle d'un coût de fatigue horaire dès la première heure de travail, ne permettait cependant pas de résoudre le modèle. Avec cette hypothèse, en-deçà d'un certain nombre d'heures de garde, le coût de fatigue lié à la garde des enfants de 3 à 5 ans et des enfants de moins de 3 ans n'interviendrait pas dans les choix de l'assistante maternelle. L'influence du temps d'école sur le choix de garder des enfants de moins de 3 ans ou des enfants de 3 à 5 ans est donc probablement moins forte que ce que prédit le modèle, du fait de la simplification qui a été opérée dans la modélisation, même si cela ne sera pas testable empiriquement.

Lorsque la préférence pour le mode de garde d'assistante maternelle pour les enfants de moins de 3 ans augmente ou lorsque le nombre d'enfants de moins de 3 ans augmente, la demande potentielle de garde  $N_{s,enfants}^{0a2} \alpha_s^{0a2}$  augmente. En conséquence, le nombre d'enfants de moins de 3 ans gardés par les assistantes maternelles augmente logiquement.

La modélisation avec deux catégories d'enfants est donc également plus riche que celle à un enfant dans le sens où cela permet de tester des hypothèses faites dans cette étude comme celles qui ont abouti aux coûts formulés en équation (1). Cette piste n'a cependant pas été poursuivie et pourrait faire l'objet d'une étude ultérieure.

#### F.b. Résolution dans le cas où les employeurs conservent du temps de loisir

La discussion du paragraphe sur le « Choix du nombre d'heures de garde par les parents » dans la section 1.3.2 indique qu'il n'y a pas qu'un maximum possible pour les parents d'après la figure 4. Il y en a trois en tout. Par commodité, les parents sont toujours supposés symétriques pour leur choix optimal. Choisir un temps de garde nul n'a alors aucun sens puisque le métier d'assistante maternelle n'existerait pas. Les parents peuvent aussi choisir de faire garder leurs enfants durant leur temps de travail cumulé à leur temps de trajet. Cela a déjà été traité en 1.3.2. Reste la dernière possibilité : les employeurs décident de conserver du temps de loisir car leur utilité marginale à le faire est supérieure au salaire versé à l'assistante maternelle, même une fois qu'ils sont revenus du travail. Cette hypothèse est cohérente. Elle peut par exemple traduire la situation d'une famille monoparentale où le parent conserve du temps une fin d'après-midi par semaine pour une activité extraprofessionnelle.

En supposant que chaque parent ne faisait garder qu'un enfant, d'après l'équation (3) en section 1.3.2, la fonction de bénéfice des employeurs était égale à :

$$\begin{aligned} u_{j,s} - h_{j,s} w_s &= \phi_1 \mathbb{1}\{h_{j,s} \leq tt_{j,s} + traj_{j,s}\} h_{j,s}^{r_1} \\ &+ \mathbb{1}\{h_{j,s} > tt_{j,s} + traj_{j,s}\} [\phi_1 (tt_{j,s} + traj_{j,s})^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{r_2}] \\ &+ \phi_2 \mathbb{1}\{h_{j,s,t} > tt_{j,s} + traj_{j,s}\} h_{j,s}^{r_2} \\ &- h_{j,s} w_s \end{aligned}$$

Le maximum est alors atteint en la deuxième partie de la fonction d'utilité. Il suffit donc de maximiser  $\phi_2 h_{j,s}^{r_2} - h_{j,s} w_s$  en  $h_{j,s}$ . Donc  $\phi_2 r_2 h_{j,s}^{r_2-1} = w_s$ . Or  $0 < r_2 < 1$ . Il sera supposé  $r_2 = \frac{1}{2}$  qui sera la valeur la plus simple pour la résolution. Donc  $h_{j,s} = \frac{\phi_2^2}{4 w_s^2} \forall j$ .

Les parents comparent ensuite leur bénéfice à faire garder leur enfant à leur seuil de participation. En notant  $D_s = \eta_0 + \eta_1 \lambda_s + \eta_2 \mu_s + \eta_3 med_s$  et en utilisant l'hypothèse de répartition uniforme sur  $\varepsilon_{j,s}$  :

$$\begin{aligned} &P(u_{j,s} - h_{j,s} w_s \geq k_{j,s}) \\ &= P\left(\phi_1 (tt_s + traj_s)^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{\frac{1}{2}} + \phi_2 h_{j,s}^{\frac{1}{2}} - h_{j,s} w_s \geq D_s + \varepsilon_{j,s}\right) \\ &= P\left(\phi_1 (tt_s + traj_s)^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{\frac{1}{2}} + \frac{\phi_2^2}{2 w_s} - \frac{\phi_2^2}{4 w_s} \geq D_s + \varepsilon_{j,s}\right) \\ &= P\left(\phi_1 (tt_s + traj_s)^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{\frac{1}{2}} + \frac{\phi_2^2}{4 w_s} \geq D_s + \varepsilon_{j,s}\right) \\ &= P\left(\varepsilon_{j,s} \leq \phi_1 (tt_s + traj_s)^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{\frac{1}{2}} + \frac{\phi_2^2}{4 w_s} - D_s\right) \\ &= \phi_1 (tt_s + traj_s)^{r_1} - \phi_2 (tt_{j,s} + traj_{j,s})^{\frac{1}{2}} + \frac{\phi_2^2}{4 w_s} - D_s \end{aligned}$$

Donc, en estimant que les parents sont symétriques,  $traj_{j,s} = traj_s$  et  $tt_{j,s} = tt_s \forall j$ , ce qui donne l'équation de demande de garde :

$q_s = N_{s,enfants} \alpha_s \left[ \phi_1 (tt_s + traj_s)^{r_1} - \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + \frac{\phi_2^2}{4 w_s} - D_s \right]$  bien que l'utilisation de la loi des grands nombres ne soit pas rigoureuse dans tous les bassins de vie. Ceci permet d'écrire :

$$w_s = \frac{\phi_2^2}{4 \left( \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s \right)}$$

Les assistantes maternelles tiennent compte de cette relation au moment de maximiser leur profit. En notant  $g_{i,s} = \beta_6 conc_s + \beta_7 age_{i,s}^{p_3} + \beta_9 dens_{i,s} + \beta_{10} exp_{i,s} + \beta_{11} loyer_{i,s}$ , les assistantes maternelles maximisent comme en section 1.3.2 :

$$\Pi_{i,s} = q_{i,s} h_{i,s} w_s - \beta_1 h_{i,s}^{p_1} - \beta_2 q_{i,s}^{p_2} - g_{i,s} q_{i,s} - \beta_8 age_{i,s}^{p_4} h_{i,s} - F_{i,s}$$

Or d'après le jeu détaillé en section 1.2.2, le nombre d'heures  $h_{i,s}$  est imposé par les employeurs. Comme chaque parent n'emploie qu'une assistante maternelle par hypothèse et comme les parents sont symétriques,  $h_{i,s} = h_{j,s} = \frac{\phi_2^2}{4 w_s^2}$ . Donc :

$$\begin{aligned}
\Pi_{i,s} &= \frac{\phi_2^2}{4 w_s} q_{i,s} - \beta_1 \left( \frac{\phi_2^2}{4 w_s^2} \right)^{p_1} - \beta_2 q_{i,s}^{p_2} - g_{i,s} q_{i,s} - \beta_8 age_{i,s}^{p_4} \frac{\phi_2^2}{4 w_s^2} - F_{i,s} \\
&= \left( \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s \right) q_{i,s} \\
&\quad - \beta_1 \left( \frac{4}{\phi_2^2} \right)^{p_1} \left( \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s \right)^{2p_1} \\
&\quad - \beta_2 q_{i,s}^{p_2} - g_{i,s} q_{i,s} - F_{i,s} \\
&\quad - \beta_8 age_{i,s}^{p_4} \frac{4}{\phi_2^2} \left( \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s \right)^2
\end{aligned}$$

Dans les annexes C.a.i et C.a.ii,  $p_1$  et  $p_2$  avaient été définis tels que  $p_1 > 1$  et  $p_2 > 1$ . Pour des fins de résolution et pour sommer sur les assistantes maternelles  $i$  le nombre d'enfants gardés  $q_{i,s}$ , il est supposé que  $p_2 = 2$ . Pour la même raison, et bien que  $p_1 > 1$  en théorie,  $p_1 = 1$  est posé. Sinon, il est compliqué de sommer les  $q_{i,s}$  dans la condition du premier ordre afin d'obtenir  $q_s$ .

$$\begin{aligned}
\Pi_{i,s} &= \left( \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s \right) q_{i,s} \\
&\quad - \beta_1 \frac{4}{\phi_2^2} \left( \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s \right)^2 \\
&\quad - \beta_2 q_{i,s}^2 - g_{i,s} q_{i,s} - F_{i,s} \\
&\quad - \beta_8 age_{i,s}^{p_4} \frac{4}{\phi_2^2} \left( \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s \right)^2
\end{aligned}$$

Donc, quand les assistantes maternelles maximisent leur profit selon le nombre d'enfants gardés :

$$\begin{aligned}
\frac{\partial \Pi_{i,s}}{\partial q_{i,s}} &= 0 \\
\Leftrightarrow \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s + \frac{q_{i,s}}{N_{s,enfants} \alpha_s} \\
&= \frac{8 (\beta_1 + \beta_8 age_{i,s}^{p_4})}{N_{s,enfants} \alpha_s \phi_2^2} \left( \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s \right) \\
&\quad + 2 \beta_2 q_{i,s} + g_{i,s}
\end{aligned}$$

En sommant :

$$\begin{aligned}
&\frac{(N_{s,assmat} + 1)}{N_{s,enfants} \alpha_s} q_s - \phi_1 N_{s,assmat} (tt_s + traj_s)^{r_1} + \phi_2 N_{s,assmat} (tt_s + traj_s)^{\frac{1}{2}} + N_{s,assmat} D_s \\
&= \frac{8 (\beta_1 N_{s,assmat} + \beta_8 \sum_i age_{i,s}^{p_4})}{N_{s,enfants} \alpha_s \phi_2^2} \left( \frac{q_s}{N_{s,enfants} \alpha_s} - \phi_1 (tt_s + traj_s)^{r_1} + \phi_2 (tt_s + traj_s)^{\frac{1}{2}} + D_s \right) \\
&\quad + 2 \beta_2 q_s + \sum_i g_{i,s}
\end{aligned}$$

$q_s$  intervient positivement de chaque côté de l'égalité. Il est difficile de comparer les termes entre eux pour savoir lequel est plus grand que l'autre. Il n'est donc pas possible d'analyser le sens de variation de  $q_s$  en fonction des variables exogènes ici. Comme il s'agit d'un modèle secondaire de l'étude, le développement de cette résolution n'a pas été poursuivi. Ce modèle dépend des mêmes variables que le modèle résolu en section 1.3.2. Sans mener une analyse plus fine, il semble possible

que les variables exogènes puissent donner un résultat opposé au modèle de référence. Par exemple, une augmentation de la demande contenue dans le terme  $D_s$  pourrait conduire à une réduction du nombre d'enfants gardés par assistante maternelle, bien que cela soit contre-intuitif.

## G. Résultats des différents modèles

### G.a. Tableaux de résultats

Tableau 5 : Résultats des régressions sur la variable du nombre d'enfants gardés par assistante maternelle

Nombre d'enfants gardés par assistante maternelle $\log\left(\frac{q}{N_{assmat}}\right)$				
Variables explicatives	OLS	Équations simultanées	SESLX	SESAR
Constante	2,0567*** [1,35 ; 2,77]	2,0567*** [1,35 ; 2,77]	3,5600*** [2,49 ; 4,63]	0,0051
Taux d'offre $\log\left(\frac{N_{assmat}}{N_{enfants \alpha}}\right)$	-0,1573*** [-0,18 ; -0,13]	-0,1573*** [-0,18 ; -0,13]	-0,1624*** [-0,19 ; -0,13]	-0,1310
Taux d'actifs occupés à temps partiel $tt$	0,3868*** [0,19 ; 0,58]	0,3868*** [0,19 ; 0,58]	-0,1709 [-0,51 ; 0,17]	0,1246
Taux d'actifs à faible temps de trajet $traj$	0,1318** [0,01 ; 0,25]	0,1318** [0,01 ; 0,25]	0,1585** [0,00 ; 0,32]	-0,0053
Taux d'actifs de 25 à 54 ans $\lambda$	1,5702*** [1,27 ; 1,87]	1,5702*** [1,27 ; 1,87]	0,7142*** [0,38 ; 1,05]	0,7698
Taux de chômage de 25 à 54 ans $\mu$	-0,6818*** [-0,97 ; -0,40]	-0,6818*** [-0,97 ; -0,40]	-0,5855*** [-0,95 ; -0,22]	-0,4612
Salaires médians $med$	-0,2705*** [-0,34 ; -0,20]	-0,2705*** [-0,34 ; -0,20]	0,0066 [-0,10 ; 0,12]	-0,0564
Densité $\log\left(\frac{\sum_i dens_i}{N_{assmat}}\right)$	0,0020 [-0,004 ; 0,01]	0,0020 [-0,004 ; 0,01]	-0,0055 [-0,01 ; 0,001]	-0,0093
Expérience $\frac{\sum_i exp_i}{N_{assmat}}$	0,0775*** [0,07 ; 0,09]	0,0775*** [0,07 ; 0,09]	0,0548*** [0,04 ; 0,07]	0,0467
Âge $\frac{\sum_i age_i}{N_{assmat}}$	-0,0091*** [-0,012 ; -0,006]	-0,0091*** [-0,012 ; -0,006]	-0,0048*** [-0,008 ; -0,002]	-0,0024
$W \frac{q}{N_{assmat}}$				0,5668
$W \log\left(\frac{N_{assmat}}{N_{enfants \alpha}}\right)$			0,0015 [-0,04 ; 0,04]	
$Wtt$			0,6403*** [0,22 ; 1,06]	
$Wtraj$			0,1203 [-0,12 ; 0,36]	
$W\lambda$			1,9815*** [1,43 ; 2,53]	
$W\mu$			0,2236 [-0,31 ; 0,76]	
$Wmed$			-0,5063*** [-0,66 ; -0,35]	
$W \log\left(\frac{\sum_i dens_i}{N_{assmat}}\right)$			0,0271*** [0,02 ; 0,04]	

$W \frac{\sum_i exp_i}{N_{assmat}}$			0,0340*** [0,01 ; 0,06]	
$W \frac{\sum_i age_i}{N_{assmat}}$			-0,0151*** [-0,02 ; -0,01]	

Note : \*\*\*p < 0,01, \*\*p < 0,05, \*p < 0,1

En première ligne de chaque case figure le coefficient. En seconde ligne figure entre crochets l'intervalle de confiance à 95%.

Les intervalles de confiance ont été calculés à partir des écarts types en supposant que les résidus suivent une loi normale pour les modèles OLS, à équations simultanées et SESLX. Ils n'ont pas été calculés pour le modèle SESAR, faute d'une méthode permettant une estimation convergente de la variance des coefficients.

Tableau 6 : Résultats des régressions sur la variable du salaire horaire des assistantes maternelles

Nombre d'enfants gardés par assistante maternelle $\log(w)$				
Variables explicatives	OLS	Équations simultanées	SESLX	SESAR
Constante	-6,4275*** [-7,06 ; -5,79]	-5,4636*** [-6,25 ; -4,67]	-8,6754*** [-9,57 ; -7,78]	-2,3827
Taux d'offre $\log\left(\frac{N_{assmat}}{N_{enfants} \alpha}\right)$	-0,2368*** [-0,26 ; -0,21]	-0,3052*** [-0,34 ; -0,27]	-0,0634*** [-0,10 ; -0,03]	-0,0957
Taux d'actifs occupés à temps partiel $tt$	0,7585*** [0,58 ; 0,94]	0,9149*** [0,70 ; 1,13]	0,1182 [-0,17 ; 0,40]	0,2502
Taux d'actifs à faible temps de trajet $traj$	0,1856*** [0,08 ; 0,30]	0,2326*** [0,11 ; 0,35]	0,0259 [-0,11 ; 0,16]	0,0827
Taux d'actifs de 25 à 54 ans $\lambda$	0,0466 [-0,24 ; 0,33]	0,8697*** [0,43 ; 1,31]	0,3620** [0,05 ; 0,68]	0,4041
Taux de chômage de 25 à 54 ans $\mu$	0,4079*** [0,16 ; 0,66]	0,0355 [-0,28 ; 0,35]	0,4573*** [0,16 ; 0,75]	0,0350
Salaire médian $med$	0,7093*** [0,65 ; 0,77]	0,5699*** [0,49 ; 0,65]	0,3239*** [0,23 ; 0,41]	0,2055
Nombre d'enfants gardés par assistante maternelle $\log\left(\frac{q}{N_{assmat}}\right)$	-0,1902*** [-0,23 ; -0,15]	-0,6007*** [-0,76 ; -0,44]	-0,1673*** [-0,31 ; -0,03]	-0,5573
$W \log\left(\frac{q}{N_{assmat}}\right)$				0,4536
$Ww$				0,8511
$W \log\left(\frac{N_{assmat}}{N_{enfants} \alpha}\right)$			-0,3086*** [-0,34 ; -0,27]	
$Wtt$			1,1781*** [0,82 ; 1,54]	
$Wtraj$			0,1543 [-0,05 ; 0,36]	
$W\lambda$			-0,4947* [-1,07 ; 0,08]	
$W\mu$			-0,4392** [-0,85 ; -0,03]	
$Wmed$			0,5895*** [0,44 ; 0,74]	

Note : \*\*\*p < 0,01, \*\*p < 0,05, \*p < 0,1

En première ligne de chaque case figure le coefficient. En seconde ligne figure entre crochets l'intervalle de confiance à 95%.

Les intervalles de confiance ont été calculés à partir des écarts types en supposant que les résidus suivent une loi normale pour les modèles OLS, à équations simultanées et SESLX. Ils n'ont pas été calculés pour le modèle SESAR, faute d'une méthode permettant une estimation convergente de la variance des coefficients.

Tableau 7 : Résultats des régressions sur la variable du nombre d'heures de garde par assistante maternelle

Nombre d'enfants gardés par assistante maternelle $\log\left(\frac{h}{N_{assmat}}\right)$				
Variables explicatives	OLS	Équations simultanées	SESLX	SESAR
Constante	5,0535*** [4,99 ; 5,12]	4,6693*** [4,56 ; 4,77]	4,6684*** [4,57 ; 4,77]	0,5834
Taux d'actifs occupés à temps partiel <i>tt</i>	-1,4916 *** [-1,70 ; -1,29]	-1,5331*** [-1,75 ; -1,31]	-0,6159*** [-1,01 ; -0,22]	-0,4660
Taux d'actifs à faible temps de trajet <i>traj</i>	-0,8176*** [-0,93 ; -0,71]	-0,9645*** [-1,08 ; -0,85]	-0,5838*** [-0,76 ; -0,41]	-0,3335
Nombre d'enfants gardés par assistante maternelle $\log\left(\frac{q}{N_{assmat}}\right)$	0,9021*** [0,86 ; 0,95]	1,2407*** [1,15 ; 1,33]	1,3602*** [1,28 ; 1,44]	0,8246
$W \log\left(\frac{q}{N_{assmat}}\right)$				-0,6116
$W \log\left(\frac{h}{N_{assmat}}\right)$				0,8757
<i>Wtt</i>			-1,3933*** [-1,89 ; -0,90]	
<i>Wtraj</i>			-0,8373*** [-1,08 ; -0,59]	

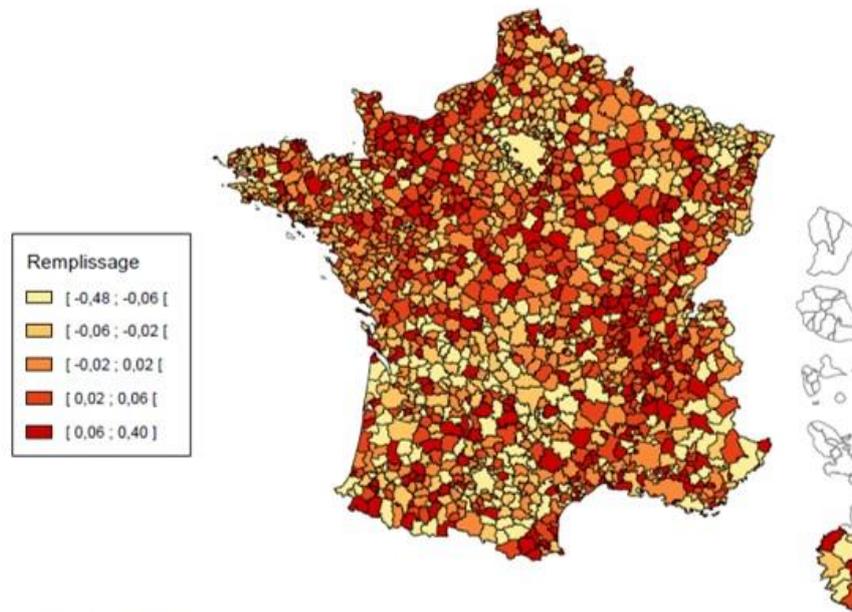
Note : \*\*\*p < 0,01, \*\*p < 0,05, \*p < 0,1

En première ligne de chaque case figure le coefficient. En seconde ligne figure entre crochets l'intervalle de confiance à 95%.

Les intervalles de confiance ont été calculés à partir des écarts types en supposant que les résidus suivent une loi normale pour les modèles OLS, à équations simultanées et SESLX. Ils n'ont pas été calculés pour le modèle SESAR, faute d'une méthode permettant une estimation convergente de la variance des coefficients.

G.b. Résultats du modèles SESAR

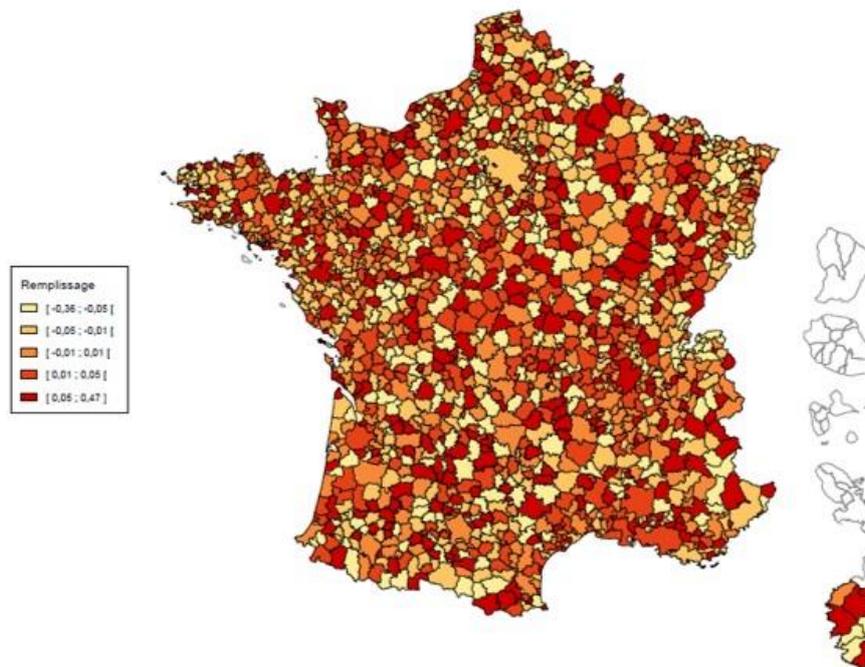
Résidus de l'équation du nombre d'enfants – Modèle SESAR



Carte réalisée avec Cartes & Données - © Artique

Figure 14 - Résidus de l'équation 14 du nombre d'enfants gardés par assistante maternelle dans le modèle SESAR

Résidus de l'équation de salaire – Modèle SESAR



Carte réalisée avec Cartes & Données - © Artique

Figure 15 - Résidus de l'équation 15 du salaire des assistantes maternelles dans le modèle SESAR

## Résidus de l'équation du nombre d'heures de garde – Modèle SESAR

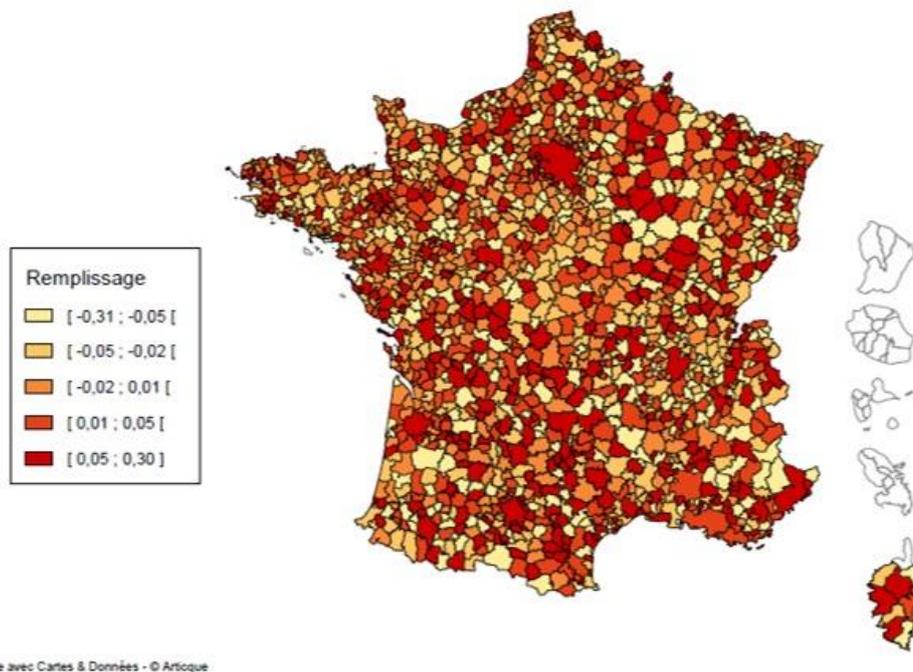


Figure 16 - Résidus de l'équation 16 du nombre d'heures de garde par assistante maternelle dans le modèle SESAR