

Les salaires, dimensions conjoncturelle et territoriale



Réalisé par :

GAABOURI Reda

MISSINHOUN Johannes

NOUNAGNON Vianou Didier

SAIDI EI OUALI Amal

Table des matières :

| | |
|--|-----------|
| INTRODUCTION | 3 |
| Objectifs de l'étude..... | 4 |
| PARTIE 1 : AJUSTEMENT DES SALAIRES NOMINAUX AUX PRIX A LA CONSOMMATION | 5 |
| 1.1 Revue de littérature..... | 6 |
| 1.2 Données et Sources..... | 7 |
| 1.3 Analyse descriptive..... | 8 |
| 1.4 Analyse de la stationnarité..... | 10 |
| 1.5 Test de cointégration de Johansen | 12 |
| 1.6 Estimation du modèle vectoriel à correction d'erreur | 14 |
| 1.7 Test de validation du modèle..... | 15 |
| 1.8 Fonctions de réponses impulsionnelles | 16 |
| Synthèse | 18 |
| PARTIE 2 : LES DÉTERMINANTS DU NIVEAU DE RÉMUNÉRATION DANS LES ZONES D'EMPLOI..... | 20 |
| 2.1 Contexte de l'étude | 21 |
| 2. 1 Revue de littérature..... | 22 |
| 2. 2 Analyse descriptive..... | 23 |
| 2.4 Analyse du lien entre le salaire net moyen horaire et les parts des secteurs et catégories professionnelles | 28 |
| 2.5 Analyse économétrique | 30 |
| 2.6 Estimations du modèle SDM et résultats obtenus | 34 |
| 2.7 Validation du modèle | 37 |
| 2.8 Contribution des variables explicatives..... | 37 |
| 2.9 Etude des résidus du Centre val de Loire | 41 |
| Synthèse | 43 |
| CONCLUSION | 44 |

Figures :

| | |
|---|----|
| Figure 1 : Matrice de corrélation | 28 |
| Figure 2 : Cercle de corrélation | 29 |

Tableaux :

| | |
|--|----|
| Tableau 1: Statistiques descriptives de l'IPC et du salaire nominal..... | 10 |
| Tableau 2 : Test de Dickey-Fuller | 11 |
| Tableau 3 : Test de trace de Johansen sur les variations de salaire et l'inflation | 12 |
| Tableau 4 : Résultats de l'estimation..... | 14 |
| Tableau 5 : Test de validation du modèle..... | 15 |
| Tableau 6 : Statistiques descriptives..... | 26 |
| Tableau 7 : Test de spécificité | 34 |
| Tableau 8 : Tableau des estimations du modèle SDM | 35 |
| Tableau 9 : Contribution des variables explicatives..... | 37 |
| Tableau 10 : Top 10 des zones dans lesquelles le salaire attendu par notre modèle s'écarte plus du salaire observé | 39 |
| Tableau 11 : Top 10 des résidus des zones du centre Val de Loire..... | 41 |

Graphiques :

| | |
|---|----|
| Graphique 1 : Evolution de l'IPC et des salaires sur la période 2000-2023..... | 8 |
| Graphique 2 : Evolution des différences premières de l'IPC et du salaire nominal..... | 12 |
| Graphique 3 : Evolution des secteurs d'activités en France de 1998 à 2020 | 25 |
| Graphique 4 : Répartition (en %) des salariés par catégorie socio-professionnelle | 42 |

Cartes :

| | |
|---|----|
| Carte 1 : Répartition des salariés par zone d'emploi en 2021 | 23 |
| Carte 2 : Disparité des salaires nets moyens horaires par zone d'emploi..... | 24 |
| Carte 3 : Répartition des salariés du secteur tertiaire marchand..... | 27 |
| Carte 4 : Répartition des résidus (écart entre les valeurs observées et les valeurs ajustées) | 39 |
| Carte 5 : Carte des résidus dans le centre val de Loire..... | 41 |

INTRODUCTION

Dans le contexte économique actuel de la France, il est essentiel de comprendre les ajustements des salaires nominaux par rapport à l'inflation ainsi que les facteurs déterminant les variations du salaire net moyen horaire dans différentes zones d'emploi. Cette compréhension approfondie est cruciale pour les décideurs politiques, les entreprises et les acteurs sociaux. Notre étude se divise en deux parties distinctes pour traiter ces deux axes de recherche.

La première partie de l'étude vise à analyser comment les salaires nominaux s'ajustent par rapport aux indices de prix à la consommation (IPC). Cette analyse permettra de déterminer si la croissance salariale suit ou dépasse l'inflation, un indicateur clé de la santé économique d'une société. En étudiant ce lien, nous pourrions également évaluer l'efficacité des mécanismes de protection du revenu, comme les indexations salariales et les négociations collectives, qui visent à préserver le standard de vie des travailleurs. L'approche méthodologique inclura l'analyse des données salariales récentes et l'application de modèles économétriques.

La deuxième partie se concentrera sur l'analyse des déterminants des salaires nets moyens horaires, en mettant l'accent sur le rôle des grands secteurs économiques et des différentes catégories socio-professionnelles. Nous explorerons comment des facteurs tels que l'appartenance sectorielle et la catégorie professionnelle influencent les niveaux de rémunération dans diverses zones d'emploi. Cette section inclura une modélisation détaillée pour identifier les disparités régionales et sectorielles et proposera des recommandations pour adresser les éventuelles inégalités structurelles au sein du marché du travail.

Dans chacune de ces parties, nous utiliserons des techniques d'analyse statistique et économétrique avancées sur les données de salaires et d'IPC de 2000 à 2023 pour la première partie et les données de 2021 sur les salaires nets moyens horaires pour la deuxième partie, avec une préparation minutieuse des données et une exploration descriptive préliminaire. Ce découpage en deux parties permettra une compréhension plus claire et structurée des dynamiques économiques en jeu, facilitant ainsi la formulation de politiques plus ciblées et efficaces pour améliorer les conditions de vie et de travail des salariés en France.

Objectifs de l'étude

Objectif 1 : Analyser l'ajustement des salaires nominaux aux prix en prenant en compte les délais de cet ajustement.

Objectif 2 : Construire un modèle économétrique pour expliquer les niveaux de rémunération dans les zones d'emploi en France.

**PARTIE 1 : AJUSTEMENT DES SALAIRES NOMINAUX AUX
PRIX A LA CONSOMMATION**

1.1 Revue de littérature

Depuis le début des années 1980, la croissance des salaires réels en France a considérablement ralenti par rapport aux décennies précédentes, marquant l'entrée de l'économie dans une période de modération salariale.

Cette phase est principalement attribuable à la montée du chômage, à la désindexation des salaires, ainsi qu'à des revalorisations moins fréquentes du Smic. Notre exploration se concentre sur la dynamique complexe entre les prix et les salaires, en utilisant des modèles économétriques pour décrypter les mécanismes sous-jacents à cette relation.

Au cœur de l'analyse des dynamiques macroéconomiques actuelles, la théorie de Phillips (1958) joue un rôle crucial, mettant en lumière l'interdépendance entre les niveaux des prix et des salaires. Les recherches de Vladimir Passeron et Fabrice Romans sur les périodes 1971-1981 et 1982-2001 illustrent une indexation presque unitaire des salaires aux prix dans un contexte d'inflation élevée, suivie d'une indexation plus faible et retardée, où les évolutions de prix du trimestre en cours n'influencent plus celles des salaires de manière significative. Cette modification des dynamiques de prix-salaire reflète les politiques de blocage des prix et des revenus, ainsi que les changements dans les pratiques de négociation salariale au sein des entreprises.

Les études menées par Rault et L'Horly (2002) approfondissent cette analyse, révélant l'impact de l'inflation sur la progression du Smic et des salaires. Leur utilisation d'un modèle vectoriel à correction d'erreur (VECM) souligne une interaction dynamique entre le Smic, l'inflation, et la formation des salaires, montrant que les hausses du Smic avant 1981 avaient un effet modérateur sur les salaires qui devient moins significatif par la suite. En approfondissant l'analyse, Heyer et al. (2000) ont employé la méthode de Johansen pour estimer une boucle prix-salaire, mettant en évidence la non-stationnarité des séries et les interactions simultanées entre les prix et les salaires. Cette approche a permis d'identifier des équations de court terme pour les prix et les salaires, offrant une vue structurelle de la courbe de Phillips comme une relation de cointégration.

D'autre part, Lucotte et Pradines-Jobet (2023) ont analysé l'effet d'un choc d'inflation sur la dynamique des prix en utilisant un modèle autorégressif vectoriel (VAR) en panel,

examinant un large échantillon de pays de l'OCDE et non-OCDE sur une période étendue de 1960 à 2019.

Leur recherche illustre la manière dont un choc initial d'inflation peut engendrer une inflation élevée et persistante, contribuant à une meilleure compréhension de la propagation de l'inflation.

Les variables clés telles que le taux de chômage, le salaire nominal par tête, l'indice des prix, et la productivité, mises en avant dans l'analyse de Heyer et al. (2000), offrent un cadre précieux pour comprendre les mécanismes de transmission au sein de la boucle prix-salaire. Notamment, le taux de chômage sert d'indicateur des tensions sur le marché du travail, influençant la capacité des salariés à négocier des salaires plus élevés qui, à leur tour, peuvent se répercuter sur les niveaux de prix à travers les coûts de production.

Cette revue met en lumière la complexité des relations entre les salaires et les prix en France, dévoilant les multiples couches d'interaction économique qui façonnent ces dynamiques. En traversant différentes époques et en utilisant des méthodologies variées, les recherches examinées fournissent une compréhension enrichie de la façon dont les politiques économiques et les conditions de marché influencent la formation des salaires et la structure des prix, offrant des insights précieux pour les décideurs et les analystes économiques.

1.2 Données et Sources

Les données utilisées pour l'étude sont des séries trimestrielles de l'indice des prix à la consommation et des salaires nominaux produites par l'Institut National de la statistique et des études économiques (INSEE).

Les séries des salaires exploitées sont issues de l'enquête trimestrielle Activité et conditions d'emploi de la main-d'œuvre (Acemo) réalisée par la Dares. L'enquête couvre les établissements des entreprises de 10 salariés ou plus, situés en France hors Mayotte.

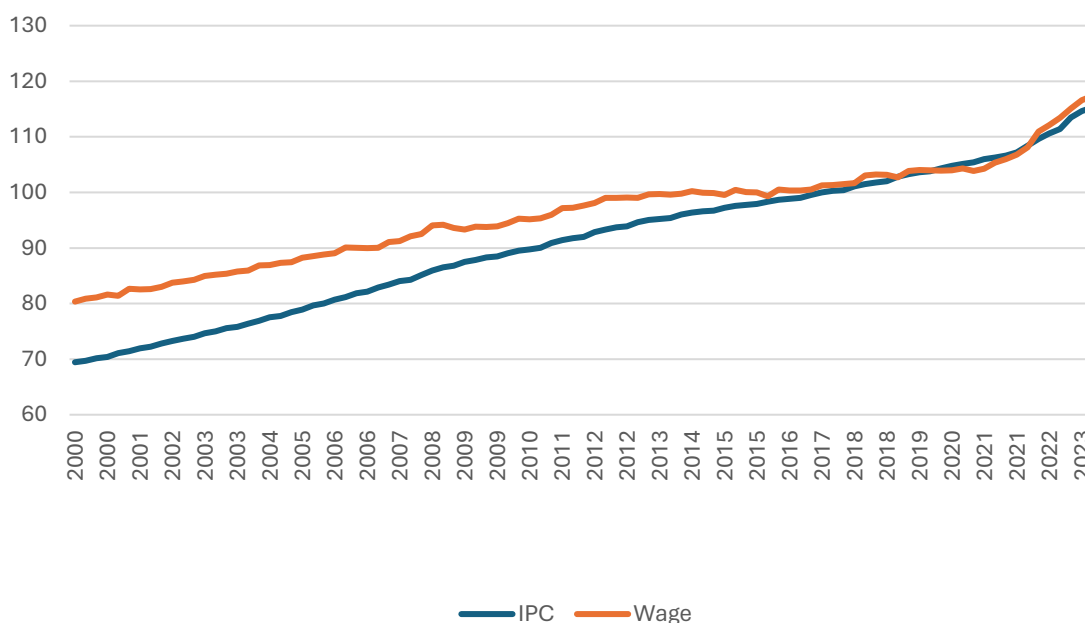
Tous les employeurs sont concernés, à l'exception de quatre catégories d'entre eux : employeurs agricoles, administration publique, activités des ménages, activités extraterritoriales. Les établissements de 250 salariés ou plus sont interrogés de façon exhaustive. Ceux de moins de 250 salariés sont interrogés par sondage et l'échantillon est renouvelé par quart chaque année.

Les données proposées par la Dares détaillent les indices de salaires pour les quatre grandes catégories de salariés : les ouvriers, les employés, les professions intermédiaires et les cadres.

1.3 Analyse descriptive

▪ Description de l'évolution des indices

Le graphique suivant présente l'évolution de l'indice des prix sur la période de 2000 à 2023.



Graphique 1: Evolution de l'IPC et des salaires sur la période 2000-2023

En observant la trajectoire des deux séries, on constate que les indices des prix et des salaires affichent une tendance générale à la hausse.

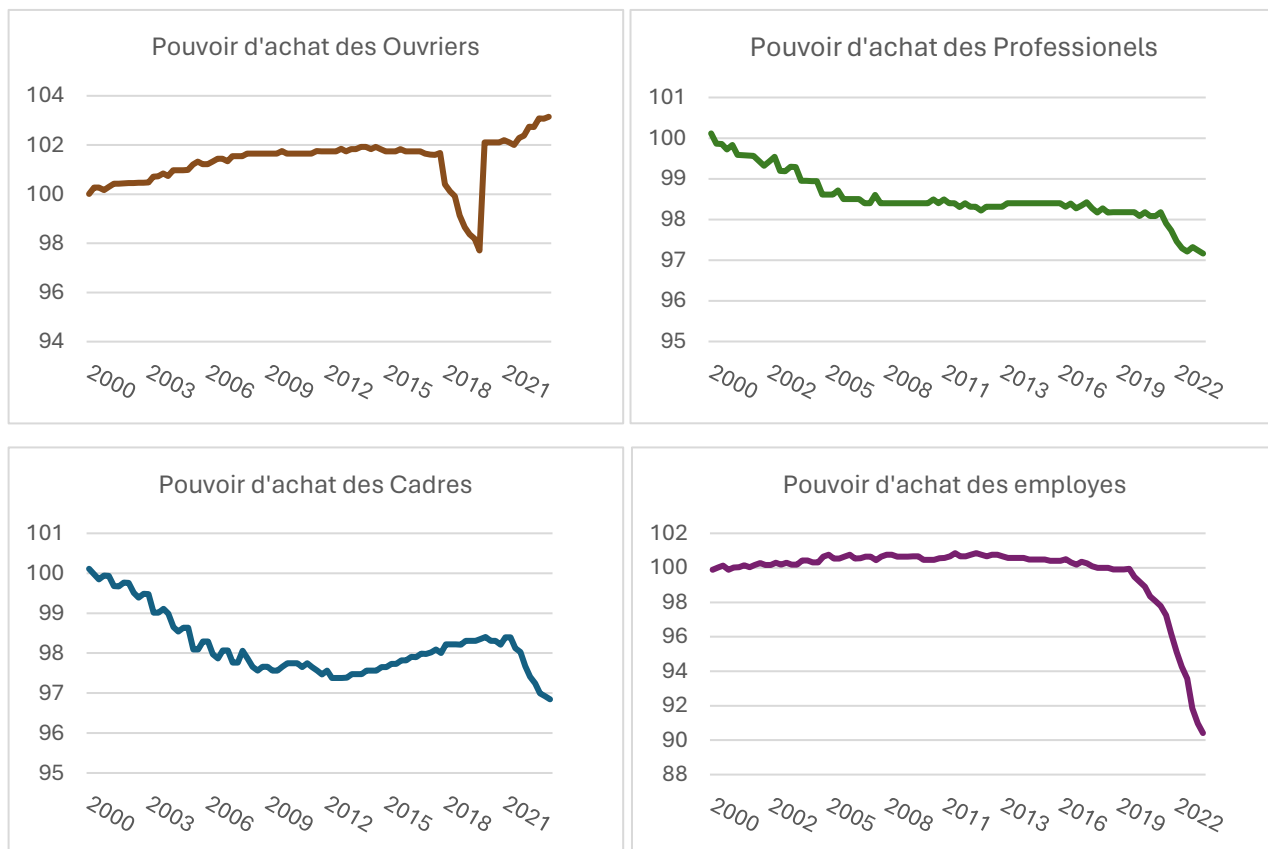
Les deux séries semblent avoir un parcours similaire indiquant que les salaires et les prix pourraient évoluer ensemble.

Cette évolution suggère une certaine relation de long terme entre les salaires et l'inflation traduisant l'existence d'une éventuelle relation de cointégration. Le pic observé dans le graphique pour l'année 2022 coïncide avec la crise en Ukraine, qui a entraîné des répercussions économiques mondiales considérables, notamment en raison des perturbations des chaînes d'approvisionnement et de l'augmentation des coûts énergétiques.

Ces facteurs ont exercé une pression inflationniste substantielle, ce qui est reflété par l'accroissement prononcé de l'indice des prix à la consommation.

▪ Description de l'évolution du pouvoir d'achat par catégorie socioprofessionnelle.

Figure 2 : Evolution du pouvoir d'achat sur la période 2000-2023.



Après une période de légère augmentation entre 2000 et 2018 on observe une chute importante du pouvoir d'achat suivi d'un rétablissement de la dynamique après la Covid. Ce phénomène suggère que les salaires nominaux des ouvriers ne s'ajustent pas instantanément aux changements des prix, mais qu'il peut y avoir un mécanisme de correction qui intervient après un certain délai.

Quant aux Cadres et Professionnels, on observe une diminution du pouvoir d'achat, indiquant que les salaires nominaux n'augmentent pas au même rythme que des prix, ce qui entraîne une dégradation du pouvoir d'achat au fil du temps. En ce qui concerne la catégorie des professionnels le graphique montre un pouvoir d'achat stable qui chute

brusquement vers la fin de la période observée. L'évolution du pouvoir d'achat sur la période montre que les salaires ne s'ajustent pas instantanément aux prix.

▪ Statistiques descriptives

La description des données est une étape importante de la démarche d'analyse. Elle permet de se familiariser avec les variables, d'en étudier les tendances et l'évolution.

Le tableau suivant décrit les principales caractéristiques (Moyenne, Ecart-type, Minimum, Maximum) de l'indice des prix et des salaires.

Tableau 1: Statistiques descriptives de l'IPC et du salaire nominal

| Caractéristiques | | | | |
|------------------|---------|---------|---------|------------|
| | Moyenne | Minimum | Maximum | Écart-type |
| IPC | 90.97 | 69.44 | 115.30 | 12.37 |
| Salaire | 96.07 | 80.35 | 117.37 | 8.75 |

Ce tableau révèle que la variance des prix est supérieure à celle des salaires, cela peut suggérer que les prix subissent de plus grandes fluctuations que les salaires sur la période étudiée.

Quels impacts sur l'ajustement ?

Une forte volatilité des prix peut rendre l'ajustement des salaires plus complexe pour les employeurs et les travailleurs. Les entreprises peuvent être réticentes à augmenter les salaires si elles anticipent une instabilité des prix à court terme, ce qui peut conduire à des négociations salariales plus difficiles. De plus, les travailleurs peuvent être moins enclins à accepter des hausses de salaire si l'incertitude économique les rend préoccupés quant à la durabilité de ces augmentations.

1.4 Analyse de la stationnarité

L'analyse de stationnarité des deux séries est essentielle afin d'analyser leur dynamique. Cette démarche est déterminante pour identifier si les séries sont stationnaires ou non et si elles sont intégrées du même ordre afin de se prononcer sur l'hypothèse émise plus haut.

Tableau 2 : Test de Dickey-Fuller

| Test de Dickey-Fuller | | | | | |
|---------------------------|-----------------------|-------------|---------|-------------------------------|---------------------------------------|
| | Lags (Critère AIC) | Statistique | Pvalue | pvalue (test de Ljung-Box) | Ordre d'intégration |
| $\Delta^2 \log_{IPC}$ | 4 | -5.6345 | <0.0001 | 0.06295 | I (0) sans tendance et sans constante |
| $\Delta^2 \log_{salaire}$ | 4 | -1.1963 | <0.0001 | 0.9025 | I (0) sans tendance et sans constante |

Le test de Dickey-Fuller réalisé sur les deux séries montrent qu'elles ne sont pas stationnaires en niveau et sont toutes deux intégrées d'ordre 2 sans tendance et sans constante.

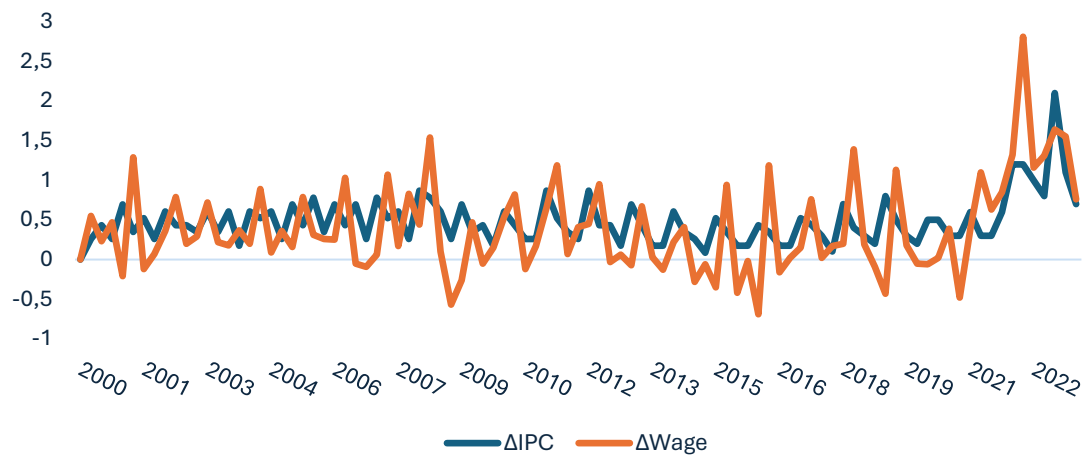
L'interprétation économique de ce résultat est assez complexe mais s'aligne avec les résultats de la plupart des études menées sur ces données (*Heyer Eric, Le Bihan Hervé, Lerais Frédéric. Relation de Phillips, boucle prix-salaire : une estimation par la méthode de Johansen*). Le fait que les deux séries soient intégrées du même ordre renforce la possibilité de la cointégration. Cela pourrait impliquer que les ajustements des salaires nominaux suivent ceux des prix à la consommation sur le long terme, un point important pour comprendre la dynamique entre inflation et salaire.

Le résultat obtenu oriente l'étude sur une nouvelle piste : l'existence d'une relation entre les variations des salaires et l'inflation¹.

¹ Il s'agit de la différence première des deux séries.

▪ Analyse de la relation entre l'inflation et la variation des salaires.

Le graphique suivant présente l'évolution des deux séries en différences sur la période.



Graphique 2 : Evolution des différences premières de l'IPC et du salaire nominal

L'analyse du graphique nous permet de constater que malgré la différenciation des deux séries, elles présentent toujours une volatilité assez élevée caractéristique de la non-stationnarité.

Etant donné que les deux séries différenciées sont intégrées d'ordre 1, un test de cointégration de Johansen a été réalisé afin de conclure quant à l'existence d'une relation de cointégration entre elles.

1.5 Test de cointégration de Johansen

Avant de procéder au test de Johansen, il convient de déterminer l'ordre de décalage des séries dans le modèle VAR.

Pour cela, un test LR (log-likelihood ratio) a été réalisé. Les critères d'informations convergent tous pour le retard $p=4$.²

Tableau 3 : Test de trace de Johansen sur les variations de salaire et l'inflation

| Test de trace de Johansen | | | |
|---------------------------|------------|-------------|-----------------|
| Valeurs propres | H_0 | Statistique | Valeur critique |
| 0.0285059 | $r = 0$ | 53.2 | 19.96 |
| 0.4300308 | $r \leq 1$ | 2.89 | 9.24 |

²Les résultats du test sont reportés en annexe.

La statistique de trace associée à la 1^{ère} valeur propre (53.2) excède largement la valeur critique de 19.96 au seuil de 5%. Ceci permet de rejeter l'hypothèse nulle (H_0) de l'absence de relations de cointégration ($r=0$), suggérant donc l'existence d'au moins une relation de cointégration entre les variables étudiées.

La statistique associée à la seconde valeur propre (2.89), est en dessous de la valeur critique de 9.24, ce qui nous amène à ne pas rejeter l'hypothèse nulle de l'existence au plus d'une relation de cointégration ($r \leq 1$). Cela indique qu'il y a exactement une relation de cointégration entre les variables.

La mise en œuvre de ce test est particulièrement pertinente dans ce cadre puisque l'identification d'une relation de cointégration nous permet de conclure qu'il existe une relation d'équilibre stable sur le long terme entre les variations des prix (l'inflation) et les variations des salaires nominaux. Cette découverte est essentielle car elle peut indiquer que les variations des salaires s'ajustent aux changements des prix à la consommation sur le long terme.

Des résultats obtenus, un modèle vectoriel à correction d'erreur semble donc adéquat pour étudier la dynamique entre les deux séries.

▪ Spécification du modèle

La conclusion précédente indiquant l'existence d'une relation de cointégration entre l'inflation et les variations des salaires nominaux, justifie l'adoption d'un modèle vectoriel de correction d'erreur avec trois retards VECM (3).

Le VECM est particulièrement pertinent dans ce contexte car il permet de modéliser la dynamique de court terme tout en intégrant l'équilibre de long terme mis en évidence par la cointégration. C'est un moyen efficace d'analyser comment les variables se corrigent et convergent vers l'équilibre de long terme après un écart dû à un choc.

Dans le contexte d'étude, le modèle VECM est représenté par le système d'équations suivant :

$$\Delta^2 IPC_t = \alpha_1(\Delta IPC_{t-1} - \beta_2 \Delta Wage_{t-1}) + \gamma_{11}^{(1)} \Delta^2 IPC_{t-1} + \gamma_{12}^{(1)} \Delta^2 Wage_{t-1} + \gamma_{11}^{(2)} \Delta^2 IPC_{t-2} + \gamma_{12}^{(2)} \Delta^2 Wage_{t-2} + \gamma_{11}^{(3)} \Delta^2 IPC_{t-3} + \gamma_{12}^{(3)} \Delta^2 Wage_{t-3} + \varepsilon_{1t}$$

$$\Delta^2 Wage_t = \alpha_2(\Delta IPC_{t-1} - \beta_2 \Delta Wage_{t-1}) + \gamma_{21}^{(1)} \Delta^2 IPC_{t-1} + \gamma_{22}^{(1)} \Delta^2 Wage_{t-1} + \gamma_{21}^{(2)} \Delta^2 IPC_{t-2} + \gamma_{22}^{(2)} \Delta^2 Wage_{t-2} + \gamma_{21}^{(3)} \Delta^2 IPC_{t-3} + \gamma_{22}^{(3)} \Delta^2 Wage_{t-3} + \varepsilon_{2t}$$

Chaque terme dans les équations a une signification spécifique :

- α_1 et α_2 : Ce sont les coefficients de correction d'erreur (force de rappel à l'équilibre), qui mesurent la vitesse à laquelle les variables dépendantes réagissent pour corriger tout déséquilibre en direction de l'équilibre de long terme.
- γ_{11} , γ_{12} , γ_{21} , γ_{22} : Ce sont les coefficients autorégressifs qui mesurent l'impact à court terme des valeurs passées des variables sur les valeurs actuelles.

1.6 Estimation du modèle vectoriel à correction d'erreur

Les résultats de l'estimation³ sont présentés dans le tableau qui suit :

Tableau 4 : Résultats de l'estimation

| Variables | Paramètres estimés | Ecart-type | t-test | Pr(> t) |
|---------------|--------------------|------------|---------|---------------|
| Ect1 | 0.725979 | 0.2932842 | 2.4753 | 0.0153705 * |
| Covid_19 | 0.00339011 | 0.00159813 | 2.1213 | 0.0369171 * |
| Diff_wage.dl1 | 1.1295018 | 0.27141070 | 4.1616 | 7.749e-05 *** |
| Diff_IPC.dl1 | -0.94875153 | 0.12195428 | -7.7796 | 1.906e-11 *** |
| Diff_wage.dl2 | 0.45665153 | 0.32820957 | 1.3913 | 0.1678836 |
| Diff_IPC.dl2 | -0.8032744 | 0.17016186 | -4.7206 | 9.569-06 *** |
| Diff_wage.dl3 | 0.31533125 | 0.35046055 | 0.8998 | 0.3708814 |
| Diff_IPC.dl3 | -0.750616 | 0.20800880 | -3.6086 | 0.0005282 *** |

Note : *** '1%', ** '5%', * '10%'

³ L'équation associée à l'inflation est reportée en annexe.

Dans le modèle, la variable $ect1$ représente le terme de correction d'erreur, avec un coefficient estimé de 0,7259. Ce coefficient est significativement différent de zéro ($p\text{-value} < 0,05$), ce qui indique que les variations des salaires s'ajustent à 73% à un déséquilibre entre elles et l'inflation. Nous pouvons donc conclure que les salaires s'ajustent aux prix.

L'indexation des salaires aux prix est-elle unitaire ?

Nous avons réalisé un test de Student pour vérifier si le coefficient associé au terme de correction d'erreur est égal à l'unité.

Les hypothèses du test :

$$\begin{cases} H_0 : \alpha = 1 \\ H_1 : \alpha < 1 \end{cases}$$

La statistique de test est de -0.9343, inférieur à la valeur critique de 1.64 pour un seuil de significativité de 5%. On conclut donc que l'ajustement n'est pas parfait.

1.7 Test de validation du modèle

Après estimation du modèle, il convient de juger de sa validité au travers des propriétés des résidus.

Nous présentons dans le tableau qui suit les tests d'autocorrélation de Box-Pierce et d'hétéroscédasticité réalisés sur les résidus de l'estimation.

Tableau 5 : Test de validation du modèle

| Tests | Statistiques | Pvalues associées | Décision |
|---------------------|----------------------|-------------------|-----------|
| Non-Autocorrélation | $\chi^2(6) = 11.8$ | 0.06658 | Non Rejet |
| Homoscédasticité | $\chi^2(7) = 16.979$ | 0.01754 | Rejet |

Les résultats des tests indiquent que les résidus ne présentent pas d'autocorrélation. Dans le cadre de la validation de notre modèle VECM, un test d'hétéroscédasticité a été également réalisé afin d'examiner la constance de la variance des résidus à différents lags.

Les résultats, présentés dans le Tableau 5, indiquent que jusqu'au lag 6, le modèle ne présente pas d'hétéroscédasticité, comme en témoignent les p-values associées aux statistiques de test qui sont supérieures à 0,05. Cependant, au lag 7, la p-value de 0.01754 suggère une possible hétéroscédasticité, indiquant une non-constance de la variance des

erreurs pour ce lag. Pour remédier à ce problème et garantir l'efficacité de notre estimation, nous avons adopté la méthode HAC (Heteroscedasticity and Autocorrelation Consistent). Cette approche ajuste les estimations de la matrice de covariance pour tenir compte de l'hétéroscédasticité possible, assurant ainsi que notre modèle reste robuste même en présence de variations de variance à différents lags.

1.8 Fonctions de réponses impulsionnelles

Les fonctions de réponse impulsionnelles sont un outil d'analyse post-estimation, qui illustrent comment une variable au sein du système réagit à des chocs uniques (ou impulsions) sur une autre variable au cours du temps. Ces fonctions permettent de comprendre la dynamique temporelle et la vitesse d'ajustement des variables en réponse à ces chocs.

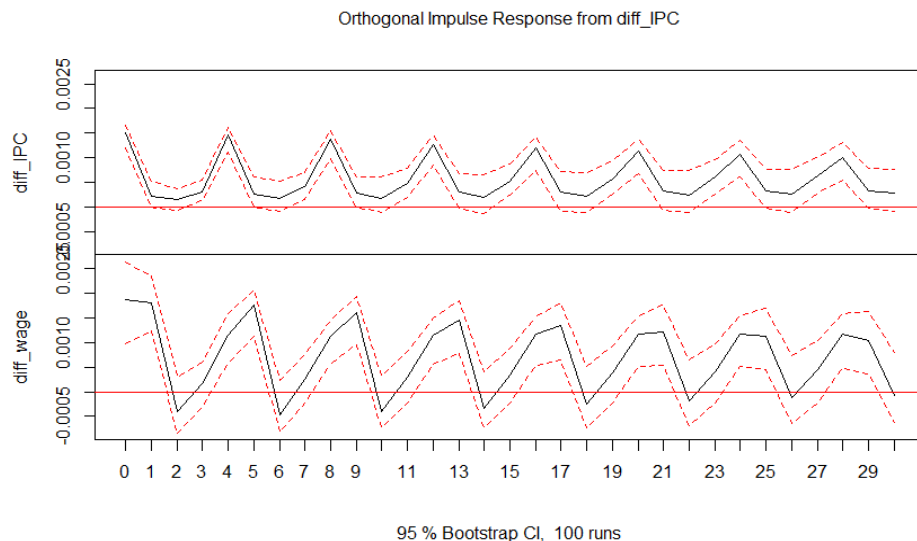


Figure 1 : Les fonctions de réponses orthogonalisées

Le graphique ci-dessus montre les réponses impulsionnelles orthogonales de diff_wage (variations de salaire nominal) et diff_IPC (inflation) à un choc d'un écart-type dans diff_IPC.

En d'autres termes, il indique comment les variations des salaires nominaux et l'inflation évoluent en réponse à une modification inattendue des prix à la consommation. Les fluctuations qui suivent le choc initial et les oscillations qui se produisent avant que la réponse

ne se stabilise à nouveau au niveau zéro varie de 3 à 4 trimestres ce qui indiquerait que les variations des salaires s'ajustent à un choc sur l'inflation au bout de 3-4 trimestres.

Synthèse

L'étude présentée a abordé de manière approfondie l'évolution des salaires nominaux par rapport à l'inflation de 2000 à 2023.

Cette analyse révèle une hausse générale des deux indices sur la période, suggérant une possible relation de long terme entre elles. L'année 2022, marquée par la crise en Ukraine, illustre comment des chocs externes peuvent influencer de manière significative l'inflation. L'analyse de la stationnarité des séries indique que ni les prix ni les salaires ne sont stationnaires en niveau mais le sont après double différenciation.

Ces résultats sont importants car ils nous ont orienté sur une nouvelle piste : l'existence d'une relation de long terme entre les variations des salaires et l'inflation. La présence d'une relation de cointégration entre les variations des salaires et l'inflation a été confirmée au moyen du test de Johansen, justifiant ainsi l'emploi d'un modèle vectoriel à correction d'erreur (VECM) avec trois retards. Ce modèle permet d'examiner la dynamique de court terme et l'ajustement à long terme entre les salaires et les prix, révélant des ajustements significatifs des salaires en réponse aux fluctuations des prix. Le VECM est validé par des tests d'autocorrélation et d'hétéroscédasticité, indiquant que les résidus du modèle sont non-auto-corrélés mais présentent un problème d'hétéroscédasticité qui a été corrigé par la méthode HAC. Les fonctions de réponse impulsionnelles ont montré que les salaires s'ajustent aux chocs sur l'inflation entre trois et quatre trimestres.

Dans le régime d'inflation qui caractérise la France, on n'observe pas l'indexation immédiate et presque unitaire des salaires sur les prix. Les évolutions de prix du trimestre en cours n'influencent pas significativement celles des salaires. Cette influence est retardée, puisqu'elle s'exerce avec un délai allant jusqu'à trois-quatre trimestres.

Le délai constaté entre une hausse de prix et son impact sur le salaire nominal traduit sans doute le fait, d'un point de vue microéconomique, que les revalorisations de salaires se font dans les entreprises sur la base d'accords couvrant des périodes plus longues qu'auparavant. Le retard d'ajustement du salaire de base sur les prix entraîne des conséquences sur l'évolution du pouvoir d'achat du salaire de base à court terme : une hausse (respectivement une baisse) de l'inflation n'est pas répercutée instantanément sur les salaires, ce qui entraîne une baisse (respectivement une hausse) transitoire du pouvoir d'achat.

Ces résultats offrent des perspectives cruciales pour les décideurs, qui peuvent ainsi mieux comprendre les mécanismes d'ajustement du marché du travail français et concevoir des interventions ciblées pour stabiliser le pouvoir d'achat.

**PARTIE 2 : LES DÉTERMINANTS DU NIVEAU DE
RÉMUNÉRATION DANS LES ZONES D'EMPLOI**

2.1 Contexte de l'étude

Dans le contexte économique actuel en France, comprendre les variations du salaire net moyen horaire entre différentes zones d'emploi est essentiel pour les décideurs politiques, les entreprises et les acteurs sociaux. Les disparités salariales peuvent refléter une multitude de facteurs économiques, sociaux et géographiques qui affectent la qualité de vie et la compétitivité des régions. Dans cette étude, il s'agira d'analyser ces disparités en mettant l'accent sur le rôle des grands secteurs économiques et des différentes catégories socio-professionnelles.

Cette étude vise à évaluer l'impact des différents secteurs économiques et catégories socio-professionnelles sur le salaire net moyen horaire au sein des zones d'emploi en France. Nous nous concentrerons sur les relations entre les secteurs tels que l'agriculture, l'industrie, la construction, ainsi que les secteurs du tertiaire marchand (le commerce, les transports et services divers) et non marchand (l'administration publique, enseignement, santé et action sociale), de même l'influence des catégories socio-professionnelles, depuis les chefs d'entreprise jusqu'aux ouvriers est prise en compte.

Notre démarche analytique se déroule en trois phases : nous commencerons par identifier les facteurs clés qui expliquent les salaires nets horaires moyens par zone, puis nous ferons une analyse des résidus de notre modèle afin de juger de sa validité, pour enfin déceler les zones où il y a une divergence marquée entre les salaires prédits par notre modèle et les salaires réellement observés accompagné d'une analyse plus poussée pour la région du Centre-Val-de-Loire.

Afin d'atteindre cet objectif, des techniques et méthodes d'analyse statistiques et économétriques ont été utilisées sur les données salariales de 2021. Tout d'abord, nous avons soigneusement préparé nos données pour l'analyse, puis réalisé une exploration descriptive afin de synthétiser l'information essentielle contenue dans notre base de données. Ces étapes préliminaires nous ont permis de se préparer à la modélisation économétrique.

2. 1 Revue de littérature

Depuis quelques années, l'étude des déterminants des salaires nets horaires est devenue un sujet de débat prépondérant dans la littérature économique. Divers chercheurs ont abordé cette question sous plusieurs angles, ce qui enrichit notre compréhension de la complexité des structures salariales. Cette revue de littérature se propose d'explorer les impacts des grands secteurs économiques et des catégories professionnelles sur les rémunérations des salariés. Nous débuterons par une discussion sur le cadre théorique avant de nous pencher sur des études empiriques pertinentes.

La théorie du capital humain, formulée par Becker en 1964, postule que les investissements en éducation et formation sont directement liés aux revenus futurs. Elle explique que le salaire d'un individu est influencé par son niveau d'étude et ses compétences. Peu après, Doeringer et Piore en 1971 ont introduit la théorie de la segmentation du marché du travail, qui n'est pas radicalement différente de la théorie du capital humain. Selon eux, le marché du travail est divisé en segments distincts, chacun caractérisé par des conditions d'emploi et des salaires qui varient, souvent sous l'influence de facteurs institutionnels ou corporatifs.

Des études de cas concrets ont été menées pour examiner ces théories, par exemple, l'analyse des salaires dans les industries technologiques en Californie a révélé que les régions à haute densité technologique offrent des rémunérations supérieures, dues à la demande accrue de compétences spécifiques et à l'innovation constante (Smith, 2010). De même, Ferguson, en 2012, a observé que dans l'industrie énergétique au Texas, les salaires sont en moyenne 20% plus élevés que le reste du pays, une différence attribuée à la spécialisation technique et aux risques professionnels inhérents. En Allemagne, Muller a trouvé en 2017 que les managers et les professionnels techniques gagnent en moyenne 30% de plus que les employés des services ou de la manufacture. Ces études illustrent bien la complexité et la variabilité des salaires, mais elles présentent également certaines limitations méthodologiques. La dépendance aux données auto-déclarées, par exemple, peut introduire des biais significatifs. De plus, les disparités régionales dans la collecte des données peuvent fausser les comparaisons entre différentes régions.

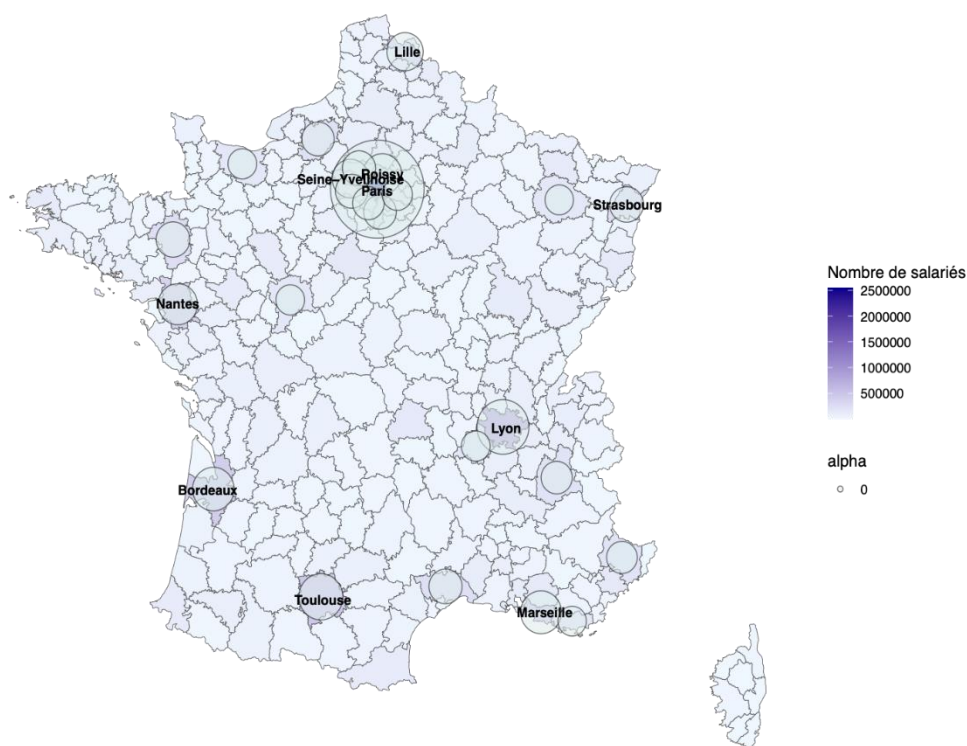
Notre présente étude sur les déterminants des rémunérations nets horaires moyens par zone d'emploi, nous permettra de tester les différentes théories et d'explorer le marché de travail en France.

2. 2 Analyse descriptive

▪ Répartition du nombre de salariés et des salaires par zone d'emploi en France

Dans le cadre de notre analyse des déterminants du niveau de rémunération à l'échelle locale, il est essentiel d'abord de visualiser la répartition géographique du nombre de salariés ainsi que la répartition du salaire net moyen horaire par zone d'emploi. Cela nous permet de déceler des tendances spatiales et sectorielles qui pourraient influencer les niveaux de salaire observés.

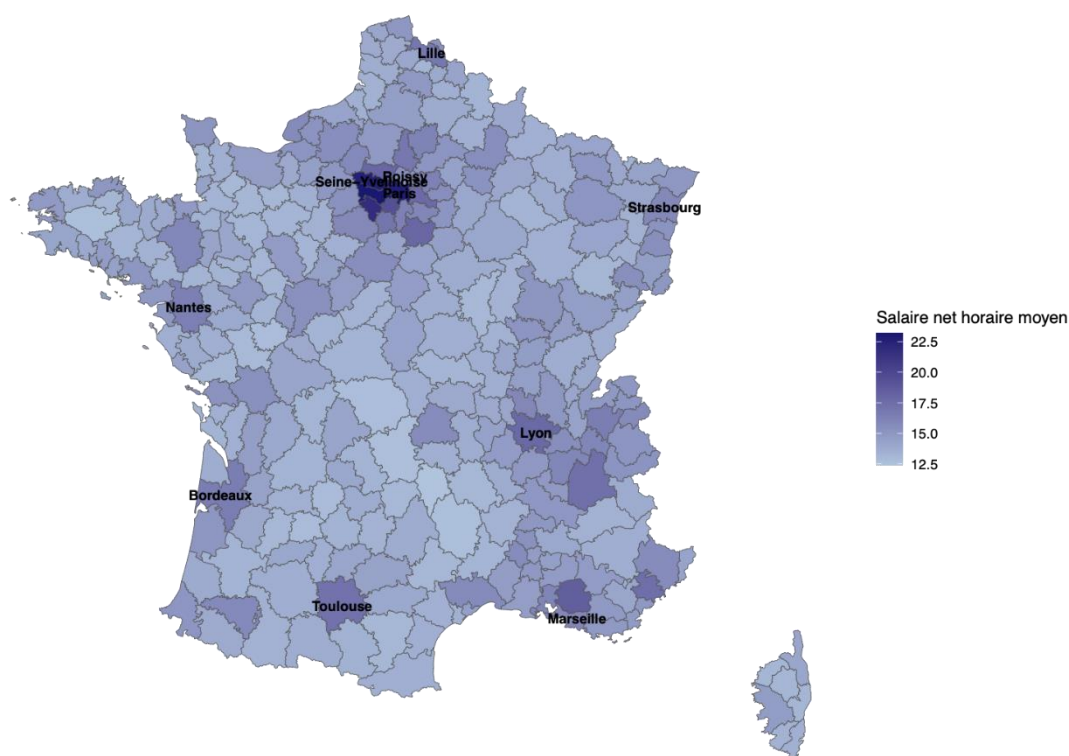
Carte 1 : Répartition des salariés par zone d'emploi en 2021



L'analyse de cette carte nous permet de dire qu'il y a une forte concentration de salariés en Île-de-France. En effet, nous constatons que le nombre de salariés qu'il y a en Île-de-France est plus élevé que celui constaté dans les autres régions de la France métropolitaine. Cette conclusion confirme son statut de région métropolitaine par excellence, où le dynamisme économique se traduit par un vaste marché de l'emploi.

Cette agglomération dense souligne l'attractivité de la capitale, non seulement en termes d'opportunités d'emploi mais aussi de diversité des secteurs d'activité. Un autre constat assez évident est celui du nombre relativement faible de salariés dans les zones rurales ou celles éloignées des grands centres urbains signalant des marchés de l'emploi plus restreints ou des spécialisations économiques qui pourraient bénéficier de mesures de soutien ciblées.

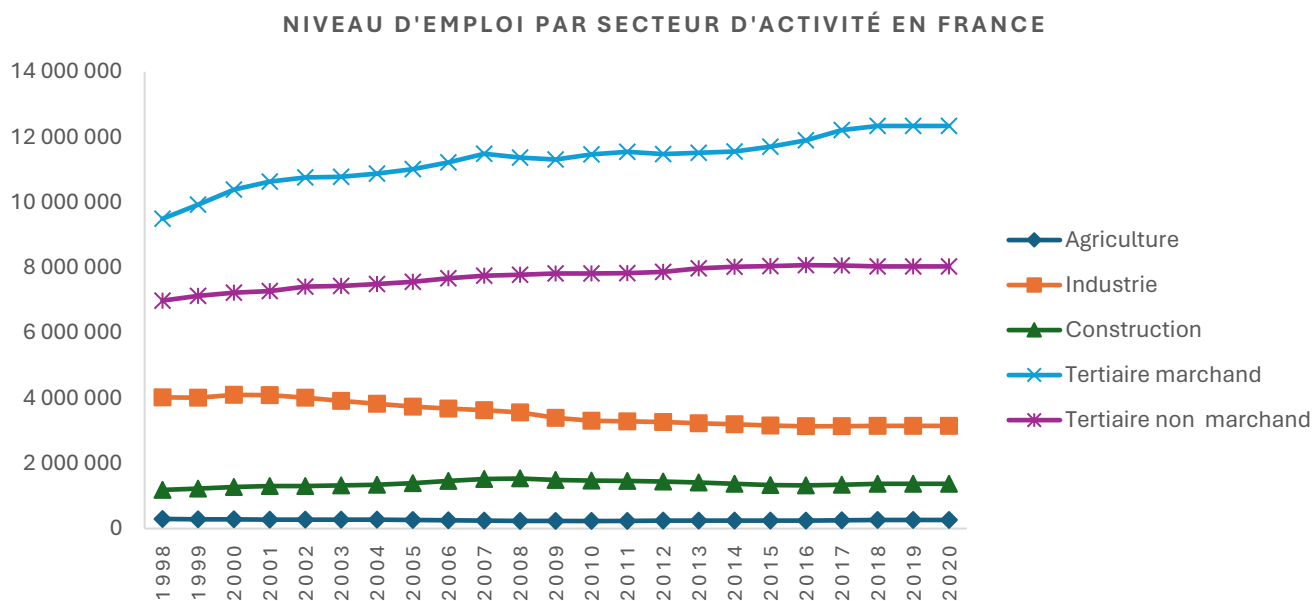
Carte 2 : Disparité des salaires nets moyens horaires par zone d'emploi



En analysant cette carte, nous constatons que l'Île-de-France se distingue par des salaires élevés, grâce à son statut de centre économique majeur, riche en compétences et en entreprises innovantes. De même, les régions proches de la Suisse, comme la Bourgogne-Franche-Comté et l'Auvergne-Rhône-Alpes, bénéficient de salaires supérieurs, influencés par la proximité avec ce pays où les rémunérations sont généreuses. En outre, des variations notables de salaires sont observées selon les zones, avec des villes comme Lyon ou Nice, où les secteurs des services avancés boostent les rémunérations, tandis que les zones à forte présence industrielle, comme autour de Saint-Étienne ou Clermont-Ferrand, affichent des salaires moins compétitifs par rapport au secteur des services.

Passant de l'analyse générale de la répartition des salariés et des salaires par zone d'emploi à une exploration plus détaillée, nous nous penchons maintenant sur les dynamiques spécifiques des secteurs d'activités et des catégories socio-professionnelles.

▪ Etude sur les secteurs d'activités et les catégories socio-professionnelles



Graphique 3 : Evolution des secteurs d'activités en France de 1998 à 2020

Le graphique ci-dessus illustre les tendances sectorielles en France de 1998 à 2020, mettant en lumière une augmentation notable de l'emploi dans les services, aussi bien marchands que non marchands, et un recul dans l'industrie, signe d'une transition vers une économie principalement orientée vers les services. Alors que le secteur de la construction a connu une légère hausse, l'emploi agricole est resté peu élevé mais stable.

Après avoir examiné la progression du nombre de salariés par zone d'activité de 1998 à 2020, nous allons maintenant étudier les interactions entre les parts des différents secteurs d'activité et des catégories socio-professionnelles en 2021.

Tableau 6 : Statistiques descriptives

| Variables | Moyenne | Minimum | Maximum | Ecart type |
|--|---------|---------|---------|------------|
| Salaire net horaire moyen | 14,47 | 12,39 | 23,17 | 1,47 |
| Part de l'agriculture | 0,06 | 0,00 | 1,47 | 0,16 |
| Part de l'Industrie | 19,08 | 5,52 | 40,24 | 7,04 |
| Part de la construction | 8,81 | 5,68 | 18,31 | 1,80 |
| Part du secteur tertiaire marchand | 59,05 | 43,84 | 78,13 | 6,74 |
| Part du secteur tertiaire non marchand | 12,67 | 4,47 | 30,04 | 3,21 |
| Part des chefs d'entreprise | 1,41 | 0,45 | 3,31 | 0,52 |
| Part des cadres et profession intellectuelle | 11,53 | 5,23 | 41,17 | 5,32 |
| Part des professions intermédiaires | 18,92 | 9,82 | 25,88 | 2,21 |
| Part des employés | 30,57 | 19,24 | 50,56 | 5,18 |
| Part Ouvriers | 37,57 | 15,11 | 59,51 | 7,42 |

Ce tableau confirme que le secteur tertiaire marchand domine le marché de l'emploi avec une moyenne significative de 59,05%, parallèlement, le secteur tertiaire non marchand, avec une moyenne de 12,67%, joue un rôle clé dans l'emploi réparti uniformément à travers le pays, contribuant à la stabilité économique. Le secteur industriel quant à lui, bien que moins prédominant (19,08%), conserve une présence significative dans certaines régions, soulignant l'importance de l'innovation pour rester compétitif. En revanche, l'agriculture présente la part la plus faible (0,06%) en moyenne, ce qui reflète sa moindre prédominance dans la structure de l'emploi en France.

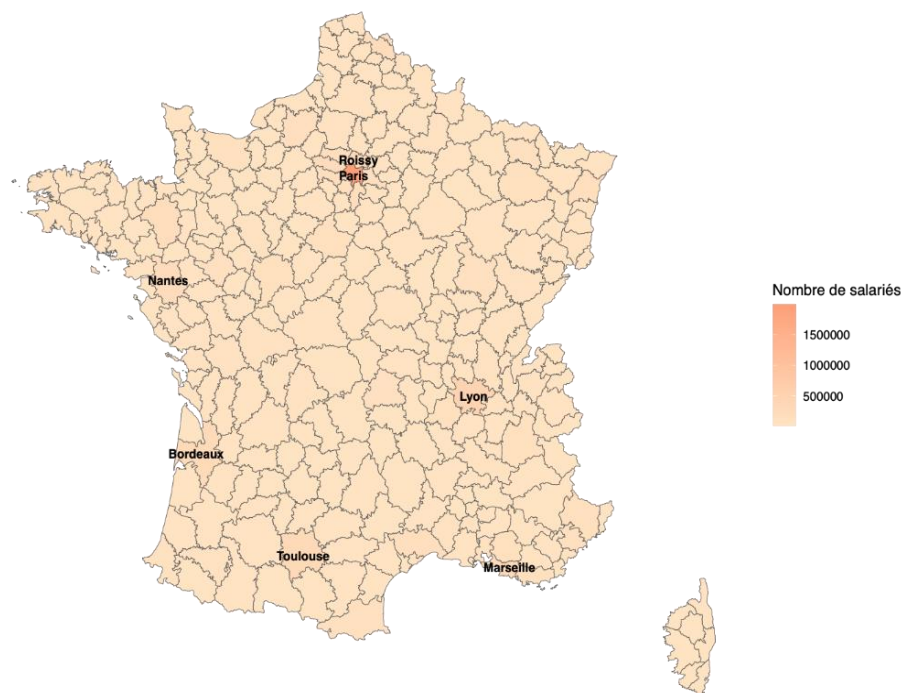
De plus la structure détaillée de l'emploi révèle également des nuances importantes entre les différentes catégories socio-professionnelles. Les ouvriers et les employés, avec des moyennes respectives de 37,57% et 30,57%, constituent les segments les plus volumineux de la main-d'œuvre, reflétant le rôle essentiel de ces professions dans le maintien des opérations quotidiennes des secteurs dominants, tertiaire marchand et industriel. Les professions intermédiaires, qui représentent 18,92%, ainsi que les cadres et professions intellectuelles supérieures, à 11,53%, indiquent une demande significative pour des

compétences spécialisées, surtout dans les régions où le secteur tertiaire marchand et non marchand est florissant.

En contraste, les chefs d'entreprises, bien que moins nombreux avec seulement 1,41%, jouent un rôle disproportionné dans la dynamisation et l'orientation de l'économie, en stimulant l'innovation et en adaptant les entreprises aux évolutions du marché. Ces distributions de l'emploi mettent en évidence non seulement la diversité des rôles au sein de l'économie française mais aussi leur interdépendance, chaque catégorie contribuant à des aspects spécifiques du dynamisme économique national.

Nous approfondirons spécifiquement le développement du secteur tertiaire marchand à travers une carte détaillée qui exposera la distribution des emplois de ce secteur en 2021, soulignant des zones de concentration spécifiques à travers le pays.

Carte 3 : Répartition des salariés du secteur tertiaire marchand



La carte ci-dessus nous montre la répartition des salariés du secteur tertiaire marchand. Elle révèle que ce secteur est particulièrement concentré dans les grandes métropoles comme Paris, Lyon, et Marseille, où il y a une forte densité de services à haute valeur ajoutée.

Ces résultats permettent de formuler des hypothèses sur les déterminants des disparités salariales, essentielles pour cibler des politiques visant à rectifier les déséquilibres économiques et stimuler le développement régional.

2.4 Analyse du lien entre le salaire net moyen horaire et les parts des secteurs et catégories professionnelles

■ Matrice de corrélation

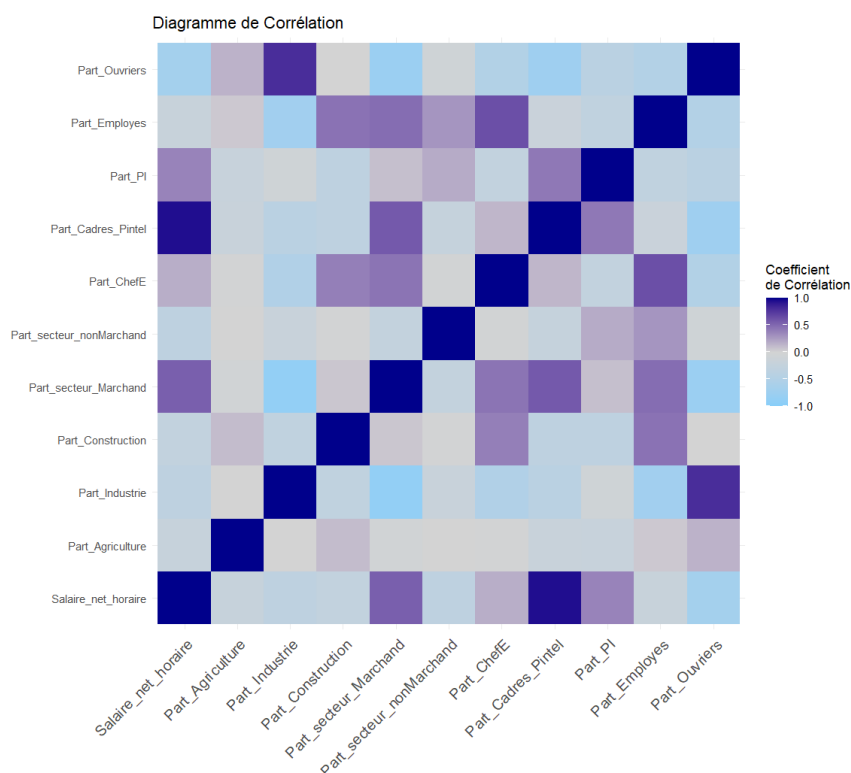


Figure 1: Matrice de corrélation

Cette matrice de corrélation présente la force et la direction des relations linéaires entre diverses variables économiques et sociales. Selon cette matrice, on note que la proportion des cadres et professions intellectuelles, ainsi que celle des ouvriers, est fortement corrélée au salaire net moyen horaire, ce qui pourrait indiquer des tendances générales dans la structure de l'emploi ou des rémunérations au sein du marché du travail français. De même, on note que la part de l'industrie et la part du secteur tertiaire non marchand figurent parmi les secteurs les plus corrélés au salaire net moyen horaire.

Afin de confirmer les résultats de la matrice de corrélation, nous allons réaliser une analyse en composantes principales.

▪ Analyse en Composante Principale

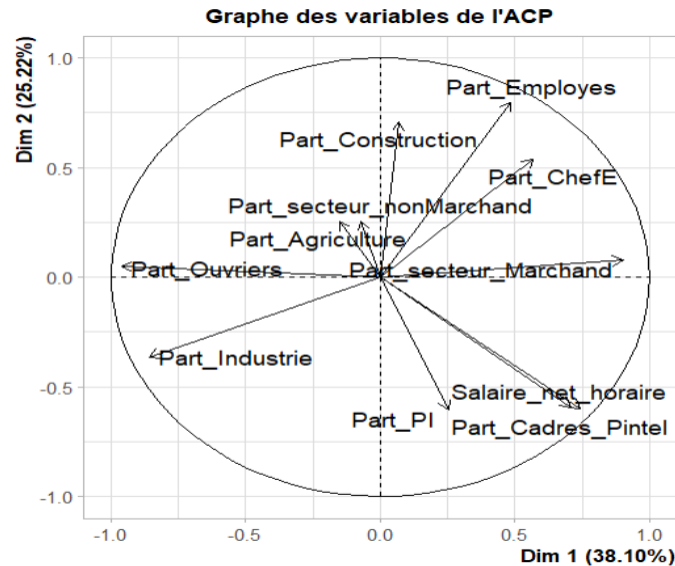


Figure 2 : Cercle de corrélation

Le cercle de corrélation dérivé de l'analyse en composantes principales montre comment les différentes variables du marché du travail français sont interconnectées. Les variables comme la part des cadres et le salaire net horaire moyen sont fortement liées à la première composante principale, indiquant que les zones avec plus de cadres ont tendance à avoir des salaires horaires plus élevés. En revanche, les ouvriers et l'industrie sont moins influents dans le modèle d'ACP, suggérant une moindre association avec les salaires et autres variables. Les positions relatives des variables nous informent sur leurs relations positives ou négatives potentielles, offrant un aperçu de la structure sous-jacente du marché du travail.

Cette analyse met en évidence une économie française où les régions avec un grand nombre de cadres tendent à avoir des salaires plus élevés, soulignant la concentration de secteurs à forte valeur ajoutée dans certaines zones urbaines. Cela peut indiquer une polarisation croissante du marché du travail et nécessiter des politiques de développement régional ciblées pour soutenir les secteurs moins représentés et favoriser une répartition plus équilibrée de la croissance économique.

▪ Espace des individus

Le graphe des individus de l'ACP (en annexe) montre la distribution des zones d'emploi en France selon les deux principales composantes. Les points situés loin du centre, connus comme des outliers, tels que "Ghisonaccia", "Saint Florent", "Ussel", "Cherbourg en Cotentin" et "Paris", peuvent indiquer des caractéristiques uniques ou des conditions économiques qui diffèrent significativement de la moyenne nationale.

Ces outliers pourraient représenter des zones avec des dynamiques de marché du travail particulières, nécessitant une attention spécifique pour les stratégies de développement régional ou les politiques d'emploi.

Après avoir posé les bases avec une analyse descriptive détaillée, nous allons maintenant passer à une analyse économétrique rigoureuse. Cette phase approfondie nous permettra d'aborder l'objectif principal de répondre à la question 2, en évaluant les relations quantitatives et les influences spécifiques des secteurs d'activité et des catégories socio-professionnelles sur le salaire net horaire moyen par zone d'emploi.

2.5 Analyse économétrique

▪ Compréhension de la méthodologie adoptée

L'analyse descriptive nous a permis de retenir les variables suivantes : la part du secteur tertiaire marchand, la part du secteur tertiaire non marchand, la part des chefs d'entreprise, la part des cadres et professions intellectuelles supérieures, la part des professions intermédiaires et la part des employés.

La structure des données, nous fait soupçonner que le salaire net moyen horaire dans une région peut être influencé par les salaires des régions voisines. Ce qui nous a permis d'opter pour l'utilisation d'un modèle d'économétrie spatiale. Ce modèle est spécifiquement conçu pour analyser les données avec des interactions ou des dépendances géographiques, permettant d'explorer les autocorrélations spatiales entre les salaires. Ainsi, il est possible que les variations de salaires dans une région soient partiellement expliquées par les niveaux de salaire des régions adjacentes.

Afin de retenir un bon modèle d'économétrie spatiale, nous avons fait les tests de spécificité tels que le test d'autocorrélation spatiale de Moran et le test du multiplicateur de Lagrange.

▪ Test de Moran

Le test d'autocorrélation spatiale de Moran est un outil statistique essentiel utilisé pour détecter la présence d'autocorrélation dans une variable mesurée à travers différentes localisations géographiques. Ce test examine si des valeurs similaires de la variable tendent à se regrouper spatialement ou si elles sont dispersées de manière aléatoire sur le territoire étudié.

Une valeur d'indice de Moran positive et significative indique que les zones géographiquement proches présentent des valeurs similaires, suggérant une autocorrélation positive. À l'inverse, une valeur négative signale une autocorrélation négative, où des valeurs similaires tendent à être séparées spatialement.

Règle de décision : si la valeur de l'indice est significativement différente de zéro (selon un seuil de signification prédéfini comme 0.05), l'hypothèse nulle est rejetée, indiquant la présence d'une autocorrélation spatiale.

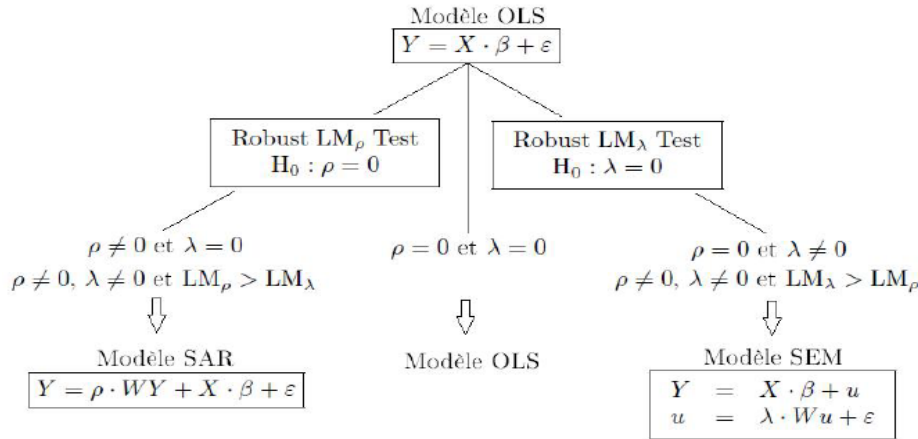
▪ Test du multiplicateur de Lagrange pour l'autocorrélation spatiale

Les tests LM-lag et LM-error sont conçus pour détecter l'omission d'une dépendance spatiale structurelle dans les modèles de régression. L'hypothèse nulle pour le LM-lag est qu'il n'y a pas de retard spatial (lag) nécessaire dans le modèle, tandis que pour le LM-error, l'hypothèse nulle est qu'il n'y a pas de corrélation d'erreur spatiale requise.

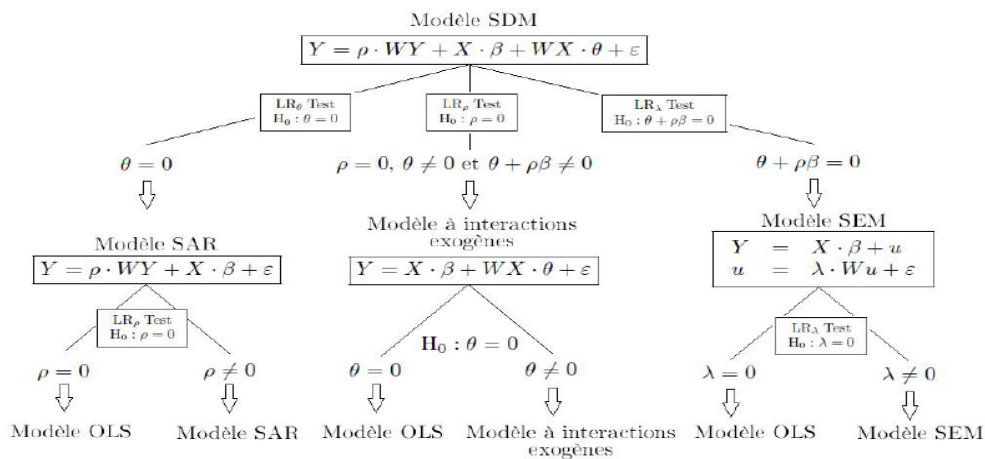
La règle de décision : elle implique de calculer la valeur du test LM et de la comparer à une valeur critique de la distribution Chi-carré avec un degré de liberté, généralement à un niveau de significativité de 5%. Si la valeur du test LM dépasse la valeur critique, l'hypothèse nulle est rejetée, suggérant que le modèle serait amélioré par l'ajout d'un terme spatial correspondant, soit dans les retards (LM-lag) soit dans les erreurs (LM-error). Cette conclusion nous conduit à réviser le modèle initial pour incorporer les effets spatiaux appropriés, améliorant ainsi la fiabilité et la précision des estimations du modèle.

■ Méthodologie pour la sélection du modèle

Pour sélectionner le bon modèle parmi les modèles d'économétrie spatiale potentiels, nous allons suivre les deux approches suivantes, présentées dans le document de l'INSEE



– Approche *bottom-up*



– Approche *top-down*

Source : LESAGE *et al.* 2009

L'idée de ces deux approches est que l'on part d'abord d'une estimation par les MCO pour aboutir au modèle le plus adapté aux données après avoir constaté l'autocorrélation spatiale entre les données.

Après le test de Moran pour l'autocorrélation spatiale, si l'on conclut à une présence d'autocorrélation spatiale, on estime un modèle de régression linéaire par les MCO. Avec le modèle MCO estimé, on fait les tests de multiplicateur de Lagrange pour détecter d'une part la dépendance spatiale dans le modèle (captée par ρ) et d'autre part, la dépendance spatiale entre les termes d'erreurs (captée par θ). Si les deux coefficients permettant de

détecter les deux formes d'autocorrélations sont significatifs, on retient le modèle d'erreur spatiale de Durbin Watson SDM (Spatial Durbin Model).

Avant de parler des résultats que nous avons obtenus, nous allons aborder un concept que représente le cœur de l'analyse spatiale, la matrice de voisinage.

▪ Concept de matrice de voisinage

Dans le contexte de l'analyse spatiale, la matrice de voisinage est centrale pour comprendre les interactions entre unités géographiques. Il postule que des unités proches (voisines) dans l'espace ont tendance à influencer les uns les autres plus fortement que celles éloignées. Ce voisinage peut être défini en termes de distance physique, mais aussi par des relations économiques, sociales, ou environnementales. Dans la modélisation spatiale, la matrice de pondération spatiale (W) capte ces relations de voisinage, influençant les valeurs d'une variable d'intérêt à travers le voisinage spatial. Ainsi, le voisinage est un principe clé pour analyser la dépendance spatiale et les effets d'entraînement entre les régions ou les points dans l'espace.

Plusieurs méthodes permettent de construire cette matrice dont la méthode de contiguïté qui est une approche commune, simple et claire pour construire cette matrice, où des unités sont considérées comme voisines si elles partagent une frontière (méthode "rook") ou un point (méthode "Queen"). D'autres méthodes incluent la distance euclidienne, où le poids diminue avec la distance, ou la méthode k-plus proches voisins, où seuls les k voisins les plus proches sont considérés. Ces matrices sont essentielles pour capturer les interdépendances spatiales et sont utilisées pour analyser comment les caractéristiques d'une région peuvent influencer et être influencées par ses voisins. Dans notre étude, nous avons choisi la méthode de contiguïté (l'approche rook) grâce à sa simplicité et nous avons jugé que cette méthode cadre beaucoup plus avec nos données étant donné que nous avons à faire avec des zones d'emploi.

2.6 Estimations du modèle SDM et résultats obtenus

i) Présentation du modèle

Le modèle Spatial Durbin Model (SDM) est une forme avancée de régression spatiale qui intègre les interactions spatiales en incluant non seulement les termes retardés de la variable dépendante mais aussi ceux des variables explicatives. En formule, cela se présente comme :

$$Y = \rho.WY + X.\beta + WX.\theta + \varepsilon$$

Où :

- Y est un vecteur (N, 1), vecteur des variables expliquées
- X est la matrice de taille (N, k) des variables explicatives
- W est la matrice de voisinage (N, N)

Les termes ρ , β , W , θ et ε sont respectivement le coefficient auto régressif mesurant l'autocorrélation spatiale dans le modèle, le vecteur des paramètres à estimer, la matrice de voisinage, les effets spatiaux lagués WX des variables explicatives et le terme d'erreur du modèle.

Soulignons que dans notre étude, le vecteur Y est constitué par les salaires nets moyens horaires par zone d'emploi en France et la matrice X est constituée des variables suivantes : la part du secteur non marchand, la part du secteur marchand, la part des chefs d'entreprises, la part des cadres et professions intellectuelles, la part des professions intermédiaires et la part des employés.

ii) Présentation des résultats

- Les tests de spécificité (Test de Moran et test LM)

Tableau 7 : Test de spécificité

| Test | Statistique de Test | P-value | Décision |
|--|---------------------|------------|----------|
| Moran (Absence d'autocorrélation spatiale) | 13,775 | 2,2.10e-16 | Rejet |
| Absence d'autocorrélation spatiale dans le modèle | 74,736 | 2,2.10e-16 | Rejet |
| Absence d'autocorrélation spatiale entre les termes d'erreur | 72,174 | 2,2.10e-16 | Rejet |

Les résultats des tests de spécification que nous avons réalisés nous ont montré que la spécification SDM est la plus adaptée à nos données.

En effet, le test de Moran nous a révélé la présence d'autocorrélation spatiale, ce qui nous a permis de continuer les autres tests : test LM d'autocorrélation spatiale dans le modèle et le test LM d'autocorrélation spatiale entre les termes d'erreurs. En analysant les résultats de ces derniers, nous constatons que tous les deux coefficients captant les deux formes d'autocorrélation spatiale sont significatifs, ce qui nous conduit à l'estimations du modèle spatial de Durbin Watson.

Le tableau suivant présente les résultats que nous avons obtenus pour les estimations :

Tableau 8 : Tableau des estimations du modèle SDM

| | Coefficients estimés | Erreur standard | Z score | Pr(> z) |
|---|-------------------------|-----------------|---------|--------------|
| (Intercept) | 5,502 | 0,912 | 6,032 | 0,000*** |
| Part du secteur tertiaire non-Marchand | -0,055 | 0,010 | -5,288 | 0,000*** |
| Part du secteur tertiaire Marchand | -0,029 | 0,007 | -4,174 | 0,000*** |
| Part des chefs d'entreprise | 0,251 | 0,079 | 3,170 | 0,002*** |
| Part des cadres et professions intellectuelles | 0,240 | 0,007 | 32,783 | < 2,2e-16*** |
| Part des professions intermédiaires | 0,029 | 0,013 | 2,306 | 0,021** |
| Part des employés | 0,015 | 0,010 | 1,466 | 0,143 |
| Lag.Part du secteur tertiaire non-Marchand | -0,002 | 0,019 | -0,100 | 0,920 |
| Lag.Part du secteur tertiaire Marchand | 0,018 | 0,012 | 1,475 | 0,140 |
| Lag. Part des chefs d'entreprise | -0,215 | 0,103 | -2,083 | 0,037** |
| Lag. Part des cadres et professions intellectuelles | -0,102 | 0,020 | -4,970 | 0,000*** |
| Lag. Part des professions intermédiaires | -0,011 | 0,023 | -0,465 | 0,642 |
| Lag. Part des employés | -0,001 | 0,018 | -0,040 | 0,968 |

Les estimations des paramètres fournissent une vue approfondie des effets directs et spatiaux des différentes variables explicatives.

Les coefficients estimés pour les variables comme la part des chefs d'entreprise, la part des cadres et professions intellectuelles, et la part des professions Intermédiaires sont significatifs et positifs, indiquant qu'une augmentation de ces parts d'emploi par secteur dans une zone d'emploi est associée à une augmentation des salaires nets moyens horaires.

Cela pourrait suggérer que les zones d'emploi avec une proportion plus élevée de cadres ou d'individus dans des postes à responsabilité tendent à afficher des salaires moyens plus élevés, peut-être en raison de la concentration de secteurs d'activité à forte valeur ajoutée ou de la nécessité d'attirer des talents qualifiés avec des salaires compétitifs.

Les effets retardés (lag) des variables indiquent l'influence des caractéristiques des zones d'emploi voisines. Par exemple, les coefficients négatifs pour lag chef d'entreprise et Lag. Part des cadres et professions intellectuelles suggèrent que la concentration de ces catégories professionnelles dans les zones environnantes pourrait avoir un impact concurrentiel sur les salaires des zones d'emploi.

En outre, des variables comme la part du secteur tertiaire marchand et la part du secteur tertiaire non marchand montrent des coefficients négatifs, ce qui pourrait refléter une saturation de ces secteurs dans certaines zones ou une concurrence accrue pouvant conduire à une pression à la baisse sur les salaires.

▪ Quelle implication économique ?

Les coefficients positifs pour des variables telles que la part des chefs d'entreprise et la part des cadres soulignent l'importance de ces professions hautement qualifiées dans la détermination des niveaux de salaires locaux. En économie, ces professions sont souvent associées à des secteurs à haute productivité et à forte valeur ajoutée, tels que la technologie, la finance ou les services professionnels. Une concentration élevée de ces professions dans une zone particulière pourrait indiquer un environnement dynamique et innovant, qui attire et récompense les compétences avec des salaires plus élevés.

Les effets retardés négatifs (ou l'impact spatial lag) montrent que les caractéristiques économiques des régions voisines peuvent exercer une influence concurrentielle sur les salaires d'une zone d'emploi donnée. Par exemple, une forte présence de chefs d'entreprise ou de cadres dans les zones adjacentes pourrait introduire une concurrence pour les mêmes talents, ce qui peut conduire à une homogénéisation ou à une diminution relative des salaires.

2.7 Validation du modèle

Le test de Moran sur les résidus (en annexe) : En analysant le test de Moran (présenté dans l'annexe), nous constatons que la statistique de Moran I est de -0.015327738, avec une déviation standard de -0.31744 et une valeur p de 0.6245, suggérant qu'il n'y a pas de preuve d'autocorrélation spatiale significative dans les résidus.

Le test d'autocorrélation de Durbin Watson : Il vérifie l'autocorrélation des résidus dans les modèles de régression. Un DW proche de 2 suggère qu'il n'y a pas d'autocorrélation série des résidus, un DW inférieur à 2 suggère une autocorrélation positive, et un DW supérieur à 2 indique une autocorrélation négative. Le résultat présenté en annexe montre que la statistique de test pour ce test est 2.0056 avec une valeur p de 0.5191, indiquant l'absence d'autocorrélation série dans les résidus. Aussi, notons que le modèle est globalement significatif (avec une pvalue de la statistique de Fisher inférieure à 1%).

En revanche, pour le test de normalité des résidus, nous notons que le W de Shapiro-Wilk est 0.94651 avec une pvalue extrêmement basse, ce qui signifie que nous rejetons l'hypothèse de normalité des résidus¹.

Soulignons que les résultats des paramètres estimés que nous avons présentés sont robustes à l'hétéroscédasticité. De ce fait, nous n'avons pas tester l'hétéroscédasticité des résidus. Etant donné que nos résidus sont non-auto corrélés, nous pouvons valider le modèle.

2.8 Contribution des variables explicatives

Tableau 9 : Contribution des variables explicatives

| | Impacts directs | Impacts indirects | Total |
|--|-----------------|-------------------|-------------|
| Part du secteur tertiaire non-Marchand | -0.05933487 | -0.065386029 | -0.12472090 |
| Part du secteur tertiaire Marchand | -0.02890717 | 0.004036031 | -0.02487114 |
| Part des chefs d'entreprise | 0.23896015 | -0.161017793 | 0.07794236 |
| Part des cadres et professions intellectuelles | 0.24429482 | 0.061115893 | 0.30541071 |
| Part des professions intermédiaires | 0.03004306 | 0.011218464 | 0.04126152 |
| Part des employés | 0.01646364 | 0.015743226 | 0.03220686 |

Impacts directs : Représentent l'effet immédiat de la variable indépendante sur la variable dépendante dans la zone d'emploi donnée. Par exemple, la part des cadres a un impact direct positif, ce qui signifie que l'augmentation de la part des cadres dans une zone est associée à une augmentation du salaire net moyen dans cette même zone.

Impacts Indirects : Aussi connus sous le nom d'effets de débordement ou effets spatiaux, ils représentent l'effet de la variable indépendante dans une zone d'emploi sur la variable dépendante dans une autre zone. Un impact indirect négatif de la part des administrations suggère que l'augmentation de la part des administrations dans une zone peut avoir un effet négatif sur les salaires dans les zones voisines.

Impact total : Combinaison des effets directs et indirects, reflétant l'influence globale d'une variable indépendante sur la variable dépendante sur l'ensemble du système spatial. Pour la part des chefs d'entreprise, l'impact total est positif, indiquant que, dans l'ensemble, une augmentation de la part des chefs d'entreprise est associée à une augmentation des salaires nets moyens à la fois localement et dans les zones avoisinantes.

Pour déterminer les variables les plus importantes pour expliquer le salaire net horaire moyen par zone d'emploi en France, nous avons utilisé la valeur des impacts totaux de chaque variable.

De ce fait, nous trouvons que :

La Part des cadres et professions intellectuelles est la variable la plus explicative avec un impact positif. À l'opposé, la Part des administrations montre l'impact total négatif le plus significatif, ce qui suggère que sa hausse est associée à une baisse des salaires nets moyens horaires. Ces résultats peuvent indiquer où les efforts de développement économique et de formation professionnelle pourraient être ciblés pour améliorer les niveaux de salaire moyen dans les différentes zones d'emploi en France.

- Les zones d'emploi de la région dans lesquelles le salaire attendu s'écarte le plus du salaire observé

Carte 4 : Répartition des résidus (écart entre les valeurs observées et les valeurs ajustées)

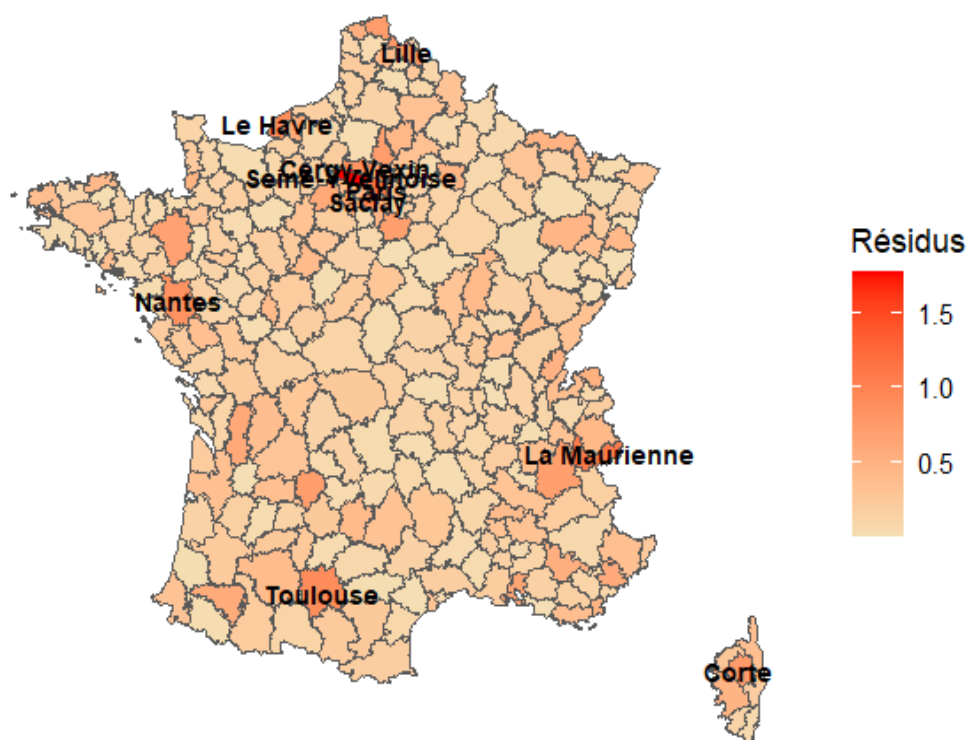


Tableau 10 : Top 10 des zones dans lesquelles le salaire attendu par notre modèle s'écarte plus du salaire observé

| Rang | Zone | Résidus | Résidus en valeur absolue |
|------|------------------|------------|---------------------------|
| 1 | Seine-Yvelinoise | 1,7594611 | 1,7594611 |
| 2 | Paris | 1,1450547 | 1,1450547 |
| 3 | La Maurienne | 1,1018344 | 1,1018344 |
| 4 | Saclay | -0,9736523 | 0,9736523 |
| 5 | Le Havre | 0,942297 | 0,942297 |
| 6 | Toulouse | -0,9070074 | 0,9070074 |
| 7 | Nantes | -0,8962729 | 0,8962729 |
| 8 | Cergy-Vexin | -0,8422297 | 0,8422297 |
| 9 | Lille | -0,7684322 | 0,7684322 |
| 10 | Dunkerque | 0,7538062 | 0,7538062 |

Ce tableau révèle l'écart entre les salaires attendus, tels que prédits par notre modèle (SDM), et les salaires réellement observés dans différentes zones géographiques de la France métropolitaine. Des résidus élevés dans certaines zones d'emploi suggèrent que le modèle n'englobe peut-être pas tous les facteurs influençant les salaires dans ces zones, révélant ainsi des disparités économiques régionales. Les zones avec des écarts significatifs comme Seine-Yvelinoise, Paris, La Maurienne, Saclay et Le Havre (qui font partie des villes les plus attractives de la France) pourraient indiquer des conditions de marché du travail particulières ou des politiques économiques qui méritent une analyse plus approfondie. De tels résultats peuvent être un indicateur pour les décideurs politiques sur les régions nécessitant une attention spécifique, qu'il s'agisse de corriger des surévaluations salariales qui pourraient mener à des tensions inflationnistes, ou d'adresser des salaires inférieurs aux prévisions, signalant peut-être des problèmes de chômage ou de sous-emploi.

Des résidus positifs élevés dans certaines zones pourraient indiquer un avantage économique dû à des facteurs tels que l'attraction de talents ou de technologies avancées grâce à une plus grande intégration dans l'économie mondiale. Inversement, des résidus négatifs importants pourraient pointer vers une compétitivité réduite due aux pressions salariales de la concurrence internationale.

2.9 Etude des résidus du Centre val de Loire

Carte 5 : Carte des résidus dans le centre val de Loire

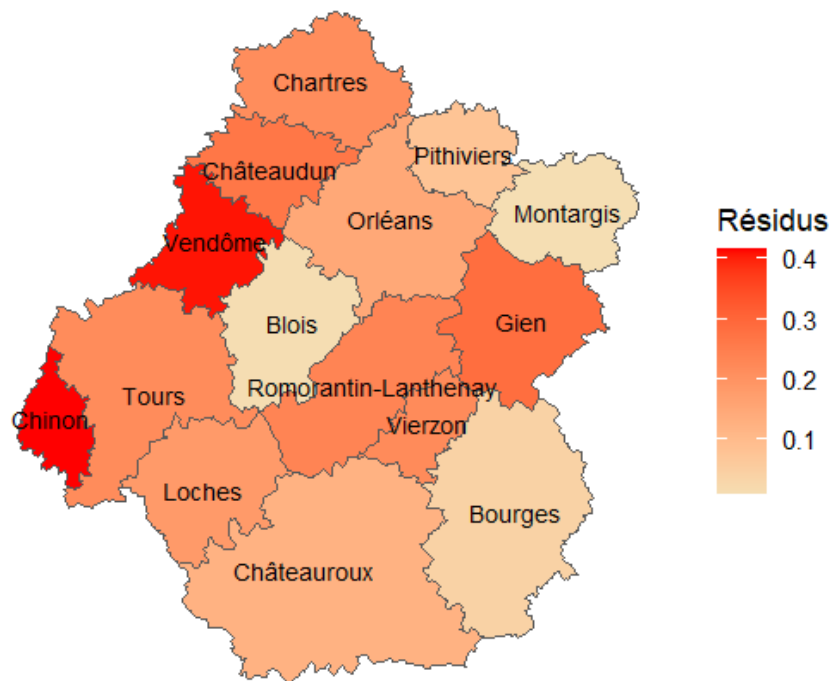
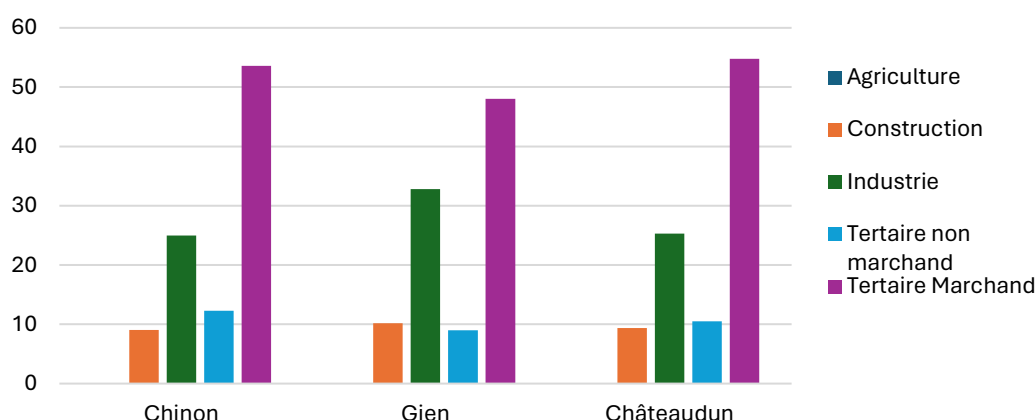


Tableau 11 : Top 10 des résidus des zones du centre Val de Loire

| Rang | Zone | Résidus | Valeur absolue des résidus |
|------|--------------------|--------------|----------------------------|
| 1 | Chinon | 0,415187291 | 0,415187291 |
| 2 | Gien | 0,28477446 | 0,28477446 |
| 3 | Châteaudun | -0,265808294 | 0,265808294 |
| 4 | Chartres | -0,214752538 | 0,214752538 |
| 5 | Loches | -0,184164885 | 0,184164885 |
| 6 | Orléans | -0,148949301 | 0,148949301 |
| 7 | Châteauroux | -0,124093654 | 0,124093654 |
| 8 | Bourges | -0,040106793 | 0,040106793 |
| 9 | Blois | 0,011531587 | 0,011531587 |
| 10 | Montargis | -0,008590048 | 0,008590048 |

L'analyse de la carte des résidus et du tableau ci-dessus nous montre que notre modèle a bien réussi à ajuster le salaire net moyen horaire dans la région du Centre Val de Loire. Ce tableau des résidus révèle non seulement des écarts significatifs mais aussi des tendances qui peuvent signaler des opportunités ou des contraintes économiques spécifiques à la région.

Les zones d'emploi de Chinon, Gien et Châteaudun présentent des écarts significatifs entre les salaires observés et les salaires prédits. Ces écarts peuvent être attribués à la forte concentration d'industries dans ces régions, ce qui se traduit par une prédominance d'ouvriers. Il est à noter que ces variables n'ont pas été incluses dans le modèle économétrique, c'est pourquoi le modèle n'arrive pas à bien ajuster le salaire net horaire moyen dans ces zones.



Graphique 4: Répartition (en %) des salariés par secteur d'activité

La région Centre-Val de Loire, où se trouvent ces zones d'emploi, est caractérisée par une forte présence industrielle, abritant des multinationales telles que *Deret*, *Amazon*, et *Cargill*. Cette particularité explique en partie les difficultés du modèle à capturer efficacement les dynamiques salariales dans ces régions.

Ces différences doivent inciter les décideurs politiques à une analyse plus fine des conditions de marché régionales, afin de formuler des politiques susceptibles de corriger les déséquilibres, qu'il s'agisse d'ajuster les prévisions salariales pour éviter des distorsions inflationnistes ou de combler les lacunes en matière d'emploi et de compétences.

En somme, les résidus de notre modèle économétrique représentent des indicateurs de la dynamique économique à l'œuvre, nécessitant une interprétation minutieuse pour orienter les politiques régionales vers une croissance inclusive et soutenue.

Synthèse

Pour expliquer le salaire net moyen horaire par zone d'emploi, nous avons opté pour le choix d'un modèle spatial de Durbin. Ce modèle nous a permis de prendre en compte les différentes formes de dépendance qui pourraient exister entre les données. Ce modèle a évalué l'impact des parts sectorielles du commerce, des administrations, des chefs d'entreprise, des cadres et professions intellectuelles supérieures, des professions intermédiaires et des employés sur les salaires. Les résultats montrent que les zones avec une plus grande proportion de cadres ou de professions intellectuelles affichent des salaires plus élevés, suggérant une concentration de secteurs à haute valeur ajoutée.

En ce qui concerne les variables qui contribuent beaucoup plus à la variance expliquée, nous notons que la part des cadres et professions intellectuelles et la part du secteur tertiaire non-marchand sont les variables qui contribuent le plus significativement à la variance des salaires. Alors que la présence accrue des cadres et professions intellectuelles est associée à des salaires plus élevés, une plus grande part des administrations est liée à des salaires plus bas. Ceci indique l'importance de ces professions hautement qualifiées et souligne le besoin de politiques de développement économique et de formation ciblées.

En outre, l'analyse des résidus révèle que le modèle ne capture pas entièrement tous les facteurs déterminant les salaires dans certaines zones. Des résidus élevés indiquent des zones où les salaires observés s'écartent significativement de ceux prévus par le modèle. Ces écarts suggèrent des spécificités régionales qui pourraient être liées à des conditions de marché du travail uniques ou à des politiques économiques locales. Les zones d'emploi avec des résidus significatifs tels que Seine-Yvelinoise, Paris, La Maurienne, Saclay et Le Havre pourraient nécessiter une analyse plus approfondie pour comprendre ces disparités.

CONCLUSION

Cette étude approfondie sur l'ajustement des salaires nominaux face à l'inflation et les facteurs déterminant les salaires nets moyens horaires dans différentes zones d'emploi en France a révélé des résultats intéressants pour la compréhension des dynamiques salariales et économiques actuelles. En divisant l'analyse en deux parties distinctes, nous avons pu adresser spécifiquement les mécanismes d'ajustement salarial face aux prix et les disparités salariales selon les secteurs économiques et les catégories socio-professionnelles.

La première partie de notre étude a mis en lumière comment les salaires s'adaptent aux variations de l'indice des prix à la consommation. Cette analyse a confirmé que les ajustements salariaux ne sont pas immédiats, ce qui reflète les conditions macroéconomiques et les politiques d'emploi.

La deuxième partie a exploré les déterminants des salaires nets moyens par zone d'emploi, en mettant en évidence l'influence significative des grands secteurs économiques et des catégories professionnelles. Les résultats ont révélé que les zones avec une concentration plus élevée de cadres ou de professions hautement qualifiées tendent à présenter des salaires supérieurs, ce qui reflète la demande pour des compétences spécialisées et la valeur ajoutée générée par ces emplois.

En conclusion, cette étude n'offre pas seulement une analyse détaillée de la situation salariale en France mais également une base solide pour la formulation de politiques ciblées visant à améliorer les conditions de vie et de travail des salariés. Ces résultats obtenus sont essentiels pour les décideurs politiques, les entreprises et les acteurs sociaux qui cherchent à comprendre et à répondre aux défis économiques complexes de notre temps.

Bibliographie :

- Becker, G. S. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.
- Doeringer, P. B., & Piore, M. J. (1971). Internal Labor Markets and Manpower Analysis.
- Azariadis, C. (1975). Implicit Contracts and Underemployment Equilibria.
- Fujita, M., & Thisse, J.-F. (2002). Economics of Agglomeration.
- Smith, J. (2010). "Technological Industries and Wage Levels in California."
- Ferguson, R. (2012). "Energy Sector's Impact on Wages in Texas."
- Müller, V. (2017). "The Role of Professional Categories in Wage Determination in Germany."
- Johansen S. (1995b). « A Statistical Analysis of Cointegration for 1(2) Variables », *Econometric Theory*, vol. 11, pp 25-59.
- Heyer Eric, Le Bihan Hervé, Lerais Frédéric. « Relation de Phillips, boucle prix-salaire : une estimation par la méthode de Johansen » In : *Économie & prévision*, n°146, 2000-5. pp. 43-60.
- Vladimir PASSERON, Fabrice ROMANS « Prévoir l'évolution des salaires en France ». Note de conjoncture INSEE, Mars 2002.
- Floch, J.-M., & Le Saout, R. (date non spécifiée). *Économétrie spatiale : modèles courants* (Chapitre 6, pp. 154-175). INSEE, ENSAI.

Annexe Partie I

1. Test de Dickey-Fuller

Les différents retards retenus pour le test de Dickey-Fuller ont été choisis en se basant sur le critère d'information AIC.

L'optimalité des retards peut être évaluée au travers du test de Box-Pierce effectué sur les résidus.

a. La variable IPC en niveau

La spécification retenue est celle du modèle avec constante.

```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####
```

Test regression drift

Call:

```
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
```

Residuals:

| | Min | 1Q | Median | 3Q | Max |
|--|------------|------------|------------|-----------|-----------|
| | -0.0040222 | -0.0021780 | -0.0007077 | 0.0020260 | 0.0134134 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|-----------|------------|---------|----------|
| (Intercept) | 0.015628 | 0.011058 | 1.413 | 0.1610 |
| z.lag.1 | -0.002305 | 0.002258 | -1.021 | 0.3101 |
| z.diff.lag | 0.185768 | 0.103082 | 1.802 | 0.0749 |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.002915 on 90 degrees of freedom

Multiple R-squared: 0.04938, Adjusted R-squared: 0.02826

F-statistic: 2.338 on 2 and 90 DF, p-value: 0.1024

Value of test-statistic is: -1.0208 24.4434

Critical values for test statistics:

| | 1pct | 5pct | 10pct |
|------|-------|-------|-------|
| tau2 | -3.51 | -2.89 | -2.58 |
| phi1 | 6.70 | 4.71 | 3.86 |

Le de Box-Pierce sur les résidus du modèle montre que les résidus sont non auto-corrélés.

Box-Pierce test

```
data: df_test@testreg$residuals
```

```
X-squared = 0.19143, df = 1, p-value = 0.6617
```

b. La variable IPC en différence première

Une fois différenciée, la constante présente dans la spécification en niveau disparaît. Le modèle pour le test en différence première ne comporte ni tendance, ni constante.

```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
  
Test regression none  
  
Call:  
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
  
Residuals:  
      Min       1Q   Median       3Q      Max   
-0.0040081 -0.0022566 -0.0006265  0.0021747  0.0127123  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)      
z.lag.1      0.0008811  0.0001294   6.810 1.02e-09 ***  
z.diff.lag  0.2066394  0.1025761   2.014  0.0469 *    
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.002931 on 91 degrees of freedom  
Multiple R-squared:  0.7785,    Adjusted R-squared:  0.7736  
F-statistic: 159.9 on 2 and 91 DF,  p-value: < 2.2e-16  
  
Value of test-statistic is: 6.8104  
  
Critical values for test statistics:  
      1pct   5pct 10pct  
tau1 -2.6 -1.95 -1.61
```

Le de Box-Pierce sur les résidus du modèle montre que les résidus sont non auto-corrélés.

Box-Pierce test

```
data: df_test@testreg$residuals  
X-squared = 0.25967, df = 1, p-value = 0.6103
```


c. La variable IPC en différence seconde

```
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.0039943 -0.0009857 -0.0000801  0.0009522  0.0070533

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1         -2.2853     0.4056  -5.635 2.29e-07 ***
z.diff.lag1      0.6750     0.3163   2.134 0.035746 *
z.diff.lag2      0.1072     0.2177   0.492 0.623712
z.diff.lag3     -0.3834     0.1101  -3.481 0.000794 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.001863 on 84 degrees of freedom
Multiple R-squared:  0.9221,    Adjusted R-squared:  0.9184
F-statistic: 248.5 on 4 and 84 DF,  p-value: < 2.2e-16

Value of test-statistic is: -5.6346

Critical values for test statistics:
    1pct    5pct   10pct
tau1 -2.6 -1.95 -1.61

>
> Box.test(df_test@testreg$residuals)

Box-Pierce test

data:  df_test@testreg$residuals
X-squared = 0.20047, df = 1, p-value = 0.6543
```

La variable en différence seconde est stationnaire sans tendance et sans constante.

d. La variable salaire (wage) en niveau.

La spécification retenue est celle du modèle avec constante.

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression drift

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.013764 -0.003413 -0.000770  0.002488  0.016432

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.009535   0.029619  -0.322   0.748
z.lag.1      0.002267   0.006213   0.365   0.716
z.diff.lag1  0.088890   0.091305   0.974   0.333
z.diff.lag2  0.071459   0.092130   0.776   0.440
z.diff.lag3  0.009687   0.093434   0.104   0.918
z.diff.lag4  0.564077   0.093699   6.020 4.42e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.004826 on 84 degrees of freedom
Multiple R-squared:  0.3311,    Adjusted R-squared:  0.2912
F-statistic: 8.315 on 5 and 84 DF,  p-value: 2.026e-06

Value of test-statistic is: 0.3649 1.3523

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.51 -2.89 -2.58
phi1  6.70  4.71  3.86
```

Le test de Box-Pierce montre que les résidus sont non auto corrélés.

Box-Pierce test

```
data: df_test@testreg$residuals
X-squared = 1.8558, df = 1, p-value = 0.1731
```

e. Test sur le salaire en différence première

```
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.0136829 -0.0028355  0.0004129  0.0030965  0.0152711

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1        -0.1224    0.1023  -1.196  0.23496
z.diff.lag1    -0.6048    0.1435  -4.214 6.29e-05 ***
z.diff.lag2    -0.4962    0.1417  -3.503 0.00074 ***
z.diff.lag3    -0.4243    0.1307  -3.246 0.00168 **
z.diff.lag4     0.1871    0.1069   1.751 0.08365 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.004665 on 84 degrees of freedom
Multiple R-squared:  0.6161,    Adjusted R-squared:  0.5933
F-statistic: 26.97 on 5 and 84 DF,  p-value: 3.63e-16

Value of test-statistic is: -1.1963

Critical values for test statistics:
    1pct  5pct 10pct
tau1 -2.6 -1.95 -1.61

> Box.test(df_test@testreg$residuals)

Box-Pierce test

data:  df_test@testreg$residuals
X-squared = 0.015018, df = 1, p-value = 0.9025
```

f. Test sur la variable salaire en différence seconde

```
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-0.0142190 -0.0029218  0.0000498  0.0025972  0.0141613

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1        -2.5525     0.3928  -6.499 5.46e-09 ***
z.diff.lag1     0.8696     0.3068   2.834 0.00575 **
z.diff.lag2     0.2988     0.2080   1.436 0.15462
z.diff.lag3    -0.1708     0.1034  -1.652 0.10234
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.004641 on 84 degrees of freedom
Multiple R-squared:  0.8669,    Adjusted R-squared:  0.8606
F-statistic: 136.8 on 4 and 84 DF,  p-value: < 2.2e-16

Value of test-statistic is: -6.4987

Critical values for test statistics:
      1pct   5pct  10pct
tau1 -2.6 -1.95 -1.61

> Box.test(df_test@testreg$residuals)

      Box-Pierce test

data:  df_test@testreg$residuals
X-squared = 0.13016, df = 1, p-value = 0.7183
```

La variable en différence seconde est stationnaire sans tendance et sans constante.

2. Test de cointégration de Johansen

- Sélection du nombre de retards optimal pour le VAR

```
> diff_IPC=diff(base$log_IPC)
> diff_wage=diff(base$log_wage)
> x=data.frame(diff_IPC,diff_wage)
> selection <- VARselect(x, lag.max=10, type="const")
> print(selection)
$selection
AIC(n)  HQ(n)  SC(n)  FPE(n)
      4      4      4      4

$Criteria
      1      2      3      4      5      6      7
AIC(n) -2.259994e+01 -2.277526e+01 -2.343686e+01 -2.389671e+01 -2.381803e+01 -2.386233e+01 -2.381015e+01
HQ(n)   -2.253014e+01 -2.265893e+01 -2.327400e+01 -2.368731e+01 -2.356210e+01 -2.355988e+01 -2.346116e+01
SC(n)   -2.242631e+01 -2.248587e+01 -2.303172e+01 -2.337582e+01 -2.318138e+01 -2.310994e+01 -2.294200e+01
FPE(n)  1.531078e-10  1.285151e-10  6.634961e-11  4.192851e-11  4.542306e-11  4.354091e-11  4.599866e-11
      8      9     10
AIC(n) -2.384732e+01 -2.385476e+01 -2.378270e+01
HQ(n)   -2.345180e+01 -2.341271e+01 -2.329412e+01
SC(n)   -2.286342e+01 -2.275511e+01 -2.256729e+01
FPE(n)  4.447963e-11  4.435288e-11  4.794106e-11
```

3. Test de Trace de Johansen

```
#####
# Johansen-Procedure #
#####

Test type: trace statistic , without linear trend and const:

Eigenvalues (lambda):
[1] 4.204898e-01 3.157012e-02 1.860695e-18

Values of teststatistic and critical values of test:

      test 10pct  5pct  1pct
r <= 1 |   2.89   7.52   9.24 12.97
r = 0  |  51.99 17.85 19.96 24.60

Eigenvectors, normalised to first column:
(These are the cointegration relations)

      diff_IPC.l4 diff_wage.l4 constant
diff_IPC.l4  1.000000000 1.000000000 1.000000000
diff_wage.l4 -0.865345711 -0.141480853 -0.27443526
constant    -0.002147181 -0.004294392 -0.00790902

Weights w:
(This is the loading matrix)

      diff_IPC.l4 diff_wage.l4 constant
diff_IPC.d  -0.5410382  -0.1408554 1.963047e-16
diff_wage.d   0.7259793  -0.4438047 7.079953e-16
```

4. Estimation du modèle VECM

a. Equation associée à l'inflation

Response diff_IPC.d :

Call:

```
lm(formula = diff_IPC.d ~ ect1 + C + diff_IPC.dl1 + diff_wage.dl1 +
    diff_IPC.dl2 + diff_wage.dl2 + diff_IPC.dl3 + diff_wage.dl3 -
    1, data = data.mat)
```

Residuals:

| | Min | 1Q | Median | 3Q | Max |
|--|------------|------------|-----------|-----------|-----------|
| | -0.0037073 | -0.0009249 | 0.0000939 | 0.0007716 | 0.0042356 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) | |
|---------------|------------|------------|---------|----------|-----|
| ect1 | -0.5410382 | 0.1079082 | -5.014 | 3.03e-06 | *** |
| C | 0.0010247 | 0.0004794 | 2.138 | 0.03551 | * |
| diff_IPC.dl1 | -1.0557036 | 0.0952378 | -11.085 | < 2e-16 | *** |
| diff_wage.dl1 | 0.1514705 | 0.0484505 | 3.126 | 0.00245 | ** |
| diff_IPC.dl2 | -1.2428254 | 0.1367061 | -9.091 | 4.74e-14 | *** |
| diff_wage.dl2 | 0.2370235 | 0.0700308 | 3.385 | 0.00110 | ** |
| diff_IPC.dl3 | -1.2639253 | 0.1379936 | -9.159 | 3.47e-14 | *** |
| diff_wage.dl3 | 0.3905852 | 0.0759276 | 5.144 | 1.80e-06 | *** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.00157 on 82 degrees of freedom
 Multiple R-squared: 0.8377, Adjusted R-squared: 0.8219
 F-statistic: 52.91 on 8 and 82 DF, p-value: < 2.2e-16

b. Equation associée au salaire

Response diff_wage.d :

Call:

```
lm(formula = diff_wage.d ~ ect1 + C + diff_IPC.dl1 + diff_wage.dl1 +
    diff_IPC.dl2 + diff_wage.dl2 + diff_IPC.dl3 + diff_wage.dl3 -
    1, data = data.mat)
```

Residuals:

| | Min | 1Q | Median | 3Q | Max |
|--|------------|------------|-----------|-----------|-----------|
| | -0.0106551 | -0.0020074 | 0.0003129 | 0.0022889 | 0.0108720 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) | |
|---------------|-----------|------------|---------|----------|-----|
| ect1 | 0.725979 | 0.274980 | 2.640 | 0.00992 | ** |
| C | 0.003390 | 0.001222 | 2.775 | 0.00683 | ** |
| diff_IPC.dl1 | 1.129501 | 0.242692 | 4.654 | 1.24e-05 | *** |
| diff_wage.dl1 | -0.948752 | 0.123465 | -7.684 | 2.94e-11 | *** |
| diff_IPC.dl2 | 0.456653 | 0.348365 | 1.311 | 0.19357 | |
| diff_wage.dl2 | -0.803274 | 0.178458 | -4.501 | 2.21e-05 | *** |
| diff_IPC.dl3 | 0.315331 | 0.351646 | 0.897 | 0.37249 | |
| diff_wage.dl3 | -0.750616 | 0.193484 | -3.879 | 0.00021 | *** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.004002 on 82 degrees of freedom

Multiple R-squared: 0.7424, Adjusted R-squared: 0.7173

F-statistic: 29.54 on 8 and 82 DF, p-value: < 2.2e-16

c. Le vecteur de cointégration

| | ect1 |
|--------------|--------------|
| diff_IPC.l4 | 1.000000000 |
| diff_wage.l4 | -0.865345711 |
| constant | -0.002147181 |

d. Estimation des écarts-types des paramètres tenant compte de l'hétéroscédasticité

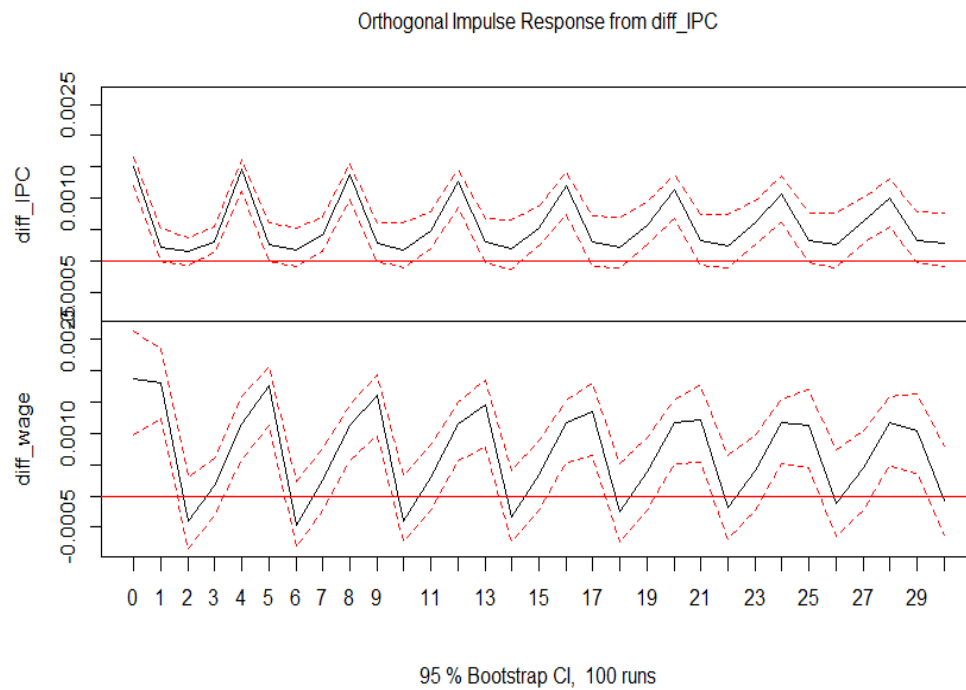
t test of coefficients:

| | Estimate | Std. Error | t value | Pr(> t) | |
|---------------------------|-------------|------------|----------|-----------|-----|
| diff_IPC.d:ect1 | -0.54103819 | 0.13952312 | -3.8778 | 0.0002117 | *** |
| diff_IPC.d:C | 0.00102475 | 0.00066611 | 1.5384 | 0.1277978 | |
| diff_IPC.d:diff_IPC.dl1 | -1.05570356 | 0.10495375 | -10.0587 | 5.715e-16 | *** |
| diff_IPC.d:diff_wage.dl1 | 0.15147050 | 0.05050142 | 2.9993 | 0.0035819 | ** |
| diff_IPC.d:diff_IPC.dl2 | -1.24282536 | 0.15796315 | -7.8678 | 1.276e-11 | *** |
| diff_IPC.d:diff_wage.dl2 | 0.23702353 | 0.07876210 | 3.0094 | 0.0034773 | ** |
| diff_IPC.d:diff_IPC.dl3 | -1.26392532 | 0.14612834 | -8.6494 | 3.587e-13 | *** |
| diff_IPC.d:diff_wage.dl3 | 0.39058522 | 0.09567189 | 4.0825 | 0.0001029 | *** |
| diff_wage.d:ect1 | 0.72597926 | 0.29328421 | 2.4753 | 0.0153705 | * |
| diff_wage.d:C | 0.00339011 | 0.00159813 | 2.1213 | 0.0369171 | * |
| diff_wage.d:diff_IPC.dl1 | 1.12950108 | 0.27141070 | 4.1616 | 7.749e-05 | *** |
| diff_wage.d:diff_wage.dl1 | -0.94875153 | 0.12195428 | -7.7796 | 1.906e-11 | *** |
| diff_wage.d:diff_IPC.dl2 | 0.45665342 | 0.32820957 | 1.3913 | 0.1678836 | |
| diff_wage.d:diff_wage.dl2 | -0.80327444 | 0.17016186 | -4.7206 | 9.569e-06 | *** |
| diff_wage.d:diff_IPC.dl3 | 0.31533125 | 0.35046055 | 0.8998 | 0.3708814 | |
| diff_wage.d:diff_wage.dl3 | -0.75061590 | 0.20800880 | -3.6086 | 0.0005282 | *** |

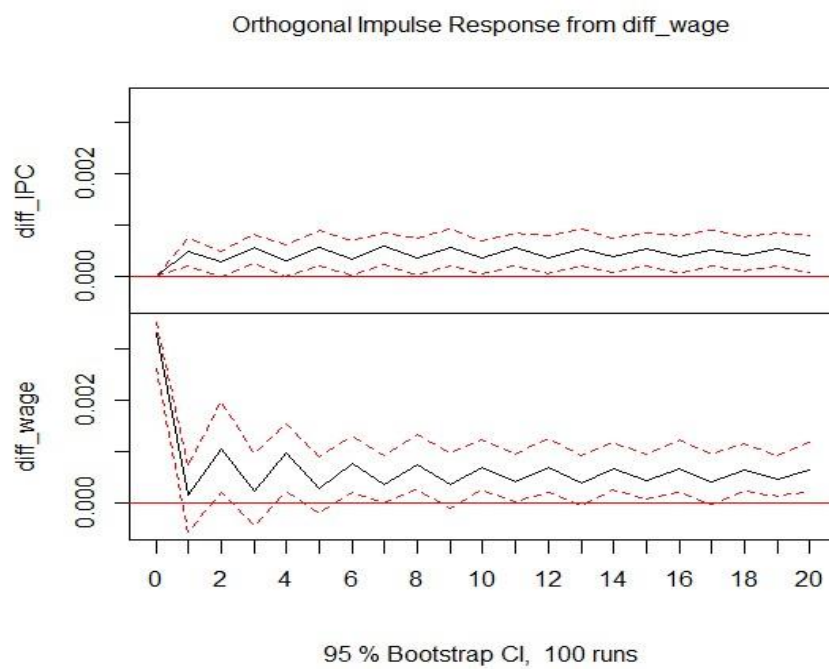
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5. Les fonctions de réponses impulsionnelles orthogonales

a. Réponses par suite d'un choc sur l'inflation

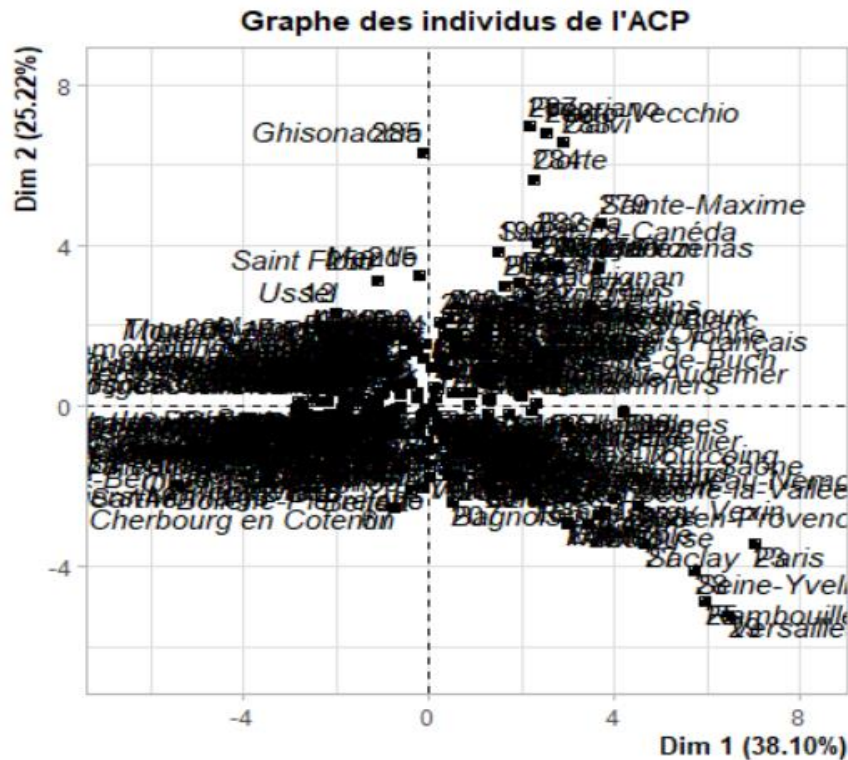


b. Réponses par suite d'un choc sur l'inflation



Annexe Partie 2

- Espace des individus de l'ACP



- Test de Moran pour vérifier l'autocorrélation spatiale entre les salaires nets moyens horaires

Moran I test under normality

data: Salaire_net_horaire
weights: listw_voisinage

Moran I statistic standard deviate = 13.775, p-value < 2.2e-16

alternative hypothesis: greater

sample estimates:

| Moran I statistic | Expectation | Variance |
|-------------------|--------------|-------------|
| 0.513516102 | -0.003496503 | 0.001408612 |

- Le modèle MCO

Call:

```
lm(formula = model, data = base)
```

Residuals:

| Min | 1Q | Median | 3Q | Max |
|----------|----------|----------|---------|---------|
| -1.67507 | -0.23187 | -0.03648 | 0.23554 | 1.95370 |

Coefficients:

| | Estimate | Std. Error | t value | Pr(> t) | |
|--------------------------|-----------|------------|---------|----------|-----|
| (Intercept) | 12.646519 | 0.359652 | 35.163 | < 2e-16 | *** |
| Part_secteur_nonMarchand | -0.070058 | 0.011610 | -6.034 | 5.05e-09 | *** |
| Part_secteur_Marchand | -0.015390 | 0.007866 | -1.956 | 0.051407 | . |
| Part_ChefE | 0.235754 | 0.066993 | 3.519 | 0.000505 | *** |
| Part_Cadres_Pintel | 0.257920 | 0.008299 | 31.080 | < 2e-16 | *** |
| Part_PI | 0.018488 | 0.014395 | 1.284 | 0.200072 | |
| Part_Employes | -0.001103 | 0.011459 | -0.096 | 0.923380 | |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4134 on 280 degrees of freedom

Multiple R-squared: 0.9232, Adjusted R-squared: 0.9215

F-statistic: 560.9 on 6 and 280 DF, p-value: < 2.2e-16

- Les tests pré-modélisation

Rao's score (a.k.a Lagrange multiplier) diagnostics for spatial dependence

data:

```
model: lm(formula = model, data = base)
```

```
test weights: listw_voisinage
```

RSerr = 72.174, df = 1, p-value < 2.2e-16

Rao's score (a.k.a Lagrange multiplier) diagnostics for spatial dependence

data:

```
model: lm(formula = model, data = base)
```

```
test weights: listw_voisinage
```

RSlag = 74.736, df = 1, p-value < 2.2e-16

Rao's score (a.k.a Lagrange multiplier) diagnostics for spatial dependence

data:

```
model: lm(formula = model, data = base)
```

```
test weights: listw_voisinage
```

adjRSerr = 35.747, df = 1, p-value = 2.247e-09

Rao's score (a.k.a Lagrange multiplier) diagnostics for spatial dependence

data:

model: lm(formula = model, data = base)

test weights: listw_voisinage

adjRSlag = 38.309, df = 1, p-value = 6.039e-10

Rao's score (a.k.a Lagrange multiplier) diagnostics for spatial dependence

- Modèle SDM

Call: lagsarlm(formula = model, data = base, listw = listw_voisinage,
type = "mixed", zero.policy = T)

Residuals:

| | Min | 1Q | Median | 3Q | Max |
|--|-----------|-----------|-----------|----------|----------|
| | -0.987025 | -0.173915 | -0.030751 | 0.149000 | 1.767778 |

Type: mixed

Coefficients: (asymptotic standard errors)

| | Estimate | Std. Error | z value | Pr(> z) |
|------------------------------|-------------|------------|---------|-----------|
| (Intercept) | 5.50183658 | 0.91212953 | 6.0319 | 1.621e-09 |
| Part_secteur_nonMarchand | -0.05459493 | 0.01032398 | -5.2882 | 1.235e-07 |
| Part_secteur_Marchand | -0.02919975 | 0.00699610 | -4.1737 | 2.997e-05 |
| Part_ChefE | 0.25063260 | 0.07906873 | 3.1698 | 0.001525 |
| Part_Cadres_Pintel | 0.23986443 | 0.00731672 | 32.7830 | < 2.2e-16 |
| Part_PI | 0.02922981 | 0.01267491 | 2.3061 | 0.021104 |
| Part_Employes | 0.01532238 | 0.01045257 | 1.4659 | 0.142677 |
| lag.Part_secteur_nonMarchand | -0.00189901 | 0.01893432 | -0.1003 | 0.920111 |
| lag.Part_secteur_Marchand | 0.01793405 | 0.01215794 | 1.4751 | 0.140188 |
| lag.Part_ChefE | -0.21532760 | 0.10335309 | -2.0834 | 0.037213 |
| lag.Part_Cadres_Pintel | -0.10152472 | 0.02042857 | -4.9697 | 6.704e-07 |
| lag.Part_PI | -0.01053988 | 0.02267534 | -0.4648 | 0.642063 |
| lag.Part_Employes | -0.00073387 | 0.01848119 | -0.0397 | 0.968325 |

Rho: 0.54704, LR test value: 66.462, p-value: 3.3307e-16

Asymptotic standard error: 0.060847

z-value: 8.9904, p-value: < 2.22e-16

Wald statistic: 80.827, p-value: < 2.22e-16

Log likelihood: -91.86586 for mixed model

ML residual variance (sigma squared): 0.10349, (sigma: 0.3217)

Number of observations: 287

Number of parameters estimated: 15

AIC: 213.73, (AIC for lm: 278.19)

LM test for residual autocorrelation

test value: 1.0143, p-value: 0.31388

- Test de Moran sur les résidus et test d'autocorrélation de Durbin Watson

Moran I test under randomisation

```
data: residuals
weights: listw_voisinage
```

```
Moran I statistic standard deviate = -0.28793, p-value = 0.6133
alternative hypothesis: greater
sample estimates:
```

| Moran I statistic | Expectation | Variance |
|-------------------|--------------|-------------|
| -0.014227441 | -0.003496503 | 0.001389035 |

Durbin-Watson test

```
data: residuals ~ 1
DW = 1.9875, p-value = 0.4578
alternative hypothesis: true autocorrelation is greater than 0
```

Shapiro-Wilk normality test

```
data: ze.sdm$residuals
W = 0.94721, p-value = 1.225e-08
```