

Economies d'Agglomération et Rentabilité des PME françaises

Document de travail

N° 2026-01 - Janvier 2026



Résumé

Cette étude examine si les économies d'agglomération – les gains de productivité liés à la concentration géographique de l'activité – se traduisent par une meilleure rentabilité réelle des Petites et Moyennes Entreprises françaises. En s'appuyant sur les données financières d'un panel de 675 000 PME françaises entre 2017 et 2022 ainsi que sur les caractéristiques des zones d'emploi, l'analyse montre que, si la densité économique accroît bien la richesse produite, elle augmente tout autant les charges de personnel, en cohérence avec les résultats établis par la littérature traitant des effets d'agglomération sur la productivité. Ces deux effets se compensent, de sorte que la rentabilité réelle, mesurée par l'Excédent Brut d'Exploitation par travailleur, n'est pas plus élevée dans les territoires économiquement denses. Ce résultat est conforté par l'analyse de l'hétérogénéité sectorielle de l'effet qui montre qu'à secteur donné, la valeur ajoutée au coût des facteurs et les charges de personnel sont affectées de manière similaire par les variables d'agglomération, aboutissant à une compensation au moins partielle des effets. Cette dernière analyse permet toutefois d'identifier quelques secteurs pour lesquels la compensation n'est pas totale.

Mots clés : Economies d'agglomération, Densité économique, Performance de l'entreprise, Petites et Moyennes Entreprises, Productivité des facteurs, Différentiels de salaire

Codes JEL : R11, R12, L25, D24, J31

Abstract

This paper examines whether agglomeration economies – productivity gains arising from the spatial concentration of economic activity – translate into higher real profitability for French Small and Medium-Sized Enterprises. Relying on financial data from a panel of 675,000 French SMEs between 2017 and 2022 and on the characteristics of French employment areas, the analysis shows that while economic density increases value added, it raises labor costs to a similar extent, consistent with the literature on agglomeration effects on productivity. These two effects offset each other, so that real profitability, measured by gross operating surplus per worker, is no higher in economically denser areas. This result is supported by the analysis of sectoral heterogeneity, which shows that, within a given sector, both value added at factor cost and labor costs respond similarly to agglomeration variables, leading to at least partial compensation of the effects. The latter analysis nevertheless identifies a few sectors where this compensation is not complete.

Keywords: Agglomeration economies, Economic density, Firm performance, Small and Medium-Sized Enterprises, Factor Productivity, Wage differentials

JEL codes: R11, R12, L25, D24, J31

Ce document de travail ne reflète pas la position de l'Insee et n'engage que son auteur.

REMERCIEMENTS

Je tiens tout d'abord à exprimer ma gratitude à Xavier Bonnet et Christian Calzada, pour la confiance qu'ils m'ont témoignée tout au long de ce travail ainsi que pour leurs retours toujours pertinents.

Je souhaite ensuite adresser mes remerciements à toutes celles et ceux avec qui j'ai eu l'occasion d'échanger au fil de cette recherche, sans qui ce travail n'aurait pas pris la même forme. Je pense en particulier à Laurent Gobillon et Olivier Meslin, dont les conseils éclairés, les discussions exigeantes et la grande disponibilité ont été déterminants dans l'élaboration de cette analyse. Je remercie également Gabrielle Gambuli et Claire Dutey pour leurs suggestions, qui ont profondément enrichi la réflexion. Mes remerciements vont aussi à Jean Dupin, Swann Maillefert et Hervé Bacheré, pour leurs éclairages méthodologiques. Enfin, je suis reconnaissante envers l'ensemble des personnes, collègues du Département de l'Action Régionale, de la Direction Interrégionale de l'INSEE Antilles-Guyane et de mon entourage, qui ont pris le temps d'échanger, de relire ou simplement d'encourager la réalisation de ce travail.

1 INTRODUCTION

La concentration spatiale de l'activité économique interroge depuis longtemps les économistes sur l'existence d'économies d'agglomération, qui désignent des économies d'échelle externes, c'est-à-dire tout effet qui accroît les revenus des travailleurs et des entreprises lorsque la taille de l'économie locale augmente.

Sur le plan théorique, les économies d'agglomération sont notamment attribuées aux travaux de Marshall (1890), qui met en avant trois canaux par lesquels les agents économiques peuvent bénéficier d'une forte concentration de l'activité. Premièrement, les effets de débordement technologiques – en particulier les externalités de connaissance – apparaissent lorsqu'un employé change d'entreprise ou lors d'échanges informels entre travailleurs de firmes voisines au sein d'un même marché local. Deuxièmement, la concentration de l'activité, surtout lorsqu'elle est spécialisée, conduit à la formation d'un marché commun du travail plus qualifié, favorisant ainsi un meilleur appariement entre emplois et travailleurs. Enfin, l'agglomération peut accroître la diversité des biens intermédiaires disponibles pour les entreprises, encourageant la spécialisation et stimulant la productivité. Une typologie plus récente et également très utilisée distingue, de son côté, trois grands mécanismes à l'origine des effets d'agglomération : les effets de partage (notamment des facteurs de production), d'appariement (en particulier de la main-d'œuvre) et d'apprentissage (comme les effets de débordement de connaissance) (Duranton et Puga, 2004).

Bien que les outils empiriques ne permettent pas d'identifier de manière causale lesquels de ces mécanismes sont à l'œuvre et leur importance respective (Combes et Gobillon, 2015), la littérature empirique a désormais largement établi l'existence d'un effet net¹ positif et causal d'agglomération sur la *productivité* dans des contextes économiques variés. La productivité des agents économiques bénéficie tout d'abord d'économies inter-sectorielles, aussi appelées *économies d'urbanisation*. Ainsi, l'élasticité de la productivité totale des facteurs à la densité en emploi locale s'établit en moyenne autour de 6%², et celle de la productivité du travail, mesurée par les salaires, à près de 4%³ (voir Ahlfeldt et Pietrostefani (2019) pour une méta-analyse de l'effet de la densité en emploi). La productivité bénéficie également souvent d'économies d'agglomération infra-sectorielles, aussi appelées *économies de localisation*, généralement appréciées par la spécialisation sectorielle locale.

L'existence d'économies d'agglomération n'est pas spécifique au cas américain, auquel la littérature s'est d'abord intéressée; de nombreux articles traitent de la France et mettent en exergue l'existence d'effets d'agglomération semblables. S'intéressant aux variables d'agglomération des zones d'emploi françaises, Combes et al. (2010) et Barbesol et Briant (2008) soulignent l'impact positif et significatif de la densité économique ainsi que de celle des zones alentours sur la productivité totale des facteurs dans la zone, bien que, comme le montrent Barbesol et Briant (2008), cet effet ne concerne qu'un petit nombre d'industries manufacturières et de services. Combes et al. (2008) montrent que les différences géographiques de salaire s'expliquent partiellement par l'existence d'économies d'urbanisation qui augmentent la productivité du travail là où l'activité

1. L'effet est dit net au sens qu'il confond tous les mécanismes à l'œuvre, qui opèrent possiblement dans des directions opposées sur la productivité.

2. Moyenne des effets estimés de 15 articles dont 13% portent sur des pays à niveau de revenus faible à moyen et pondérée par le nombre de citations des articles, calculée par Ahlfeldt et Pietrostefani (2019)

3. Moyenne des effets estimés de 47 articles dont 19% portent sur des pays à niveau de revenus faible à moyen et pondérée par le nombre de citations des articles, calculée par Ahlfeldt et Pietrostefani (2019).

économique est plus concentrée.

Une question se pose alors : pour une même entreprise, les bénéfices d'agglomération en termes de *productivité* totale des facteurs l'emportent-ils sur la *hausse du coût du travail* qui découle également de la concentration spatiale de l'activité? Cette étude vise à déterminer si les différences spatiales de rentabilité réelle des Petites et Moyennes Entreprises (PME) françaises, appréciée par l'Excédent Brut d'Exploitation (EBE) par travailleur, peuvent s'expliquer par la présence d'économies d'agglomération. En comparant la rentabilité financière des PME françaises hexagonales et dromiennes, Bilonière et al. (2025) ont établi qu'il subsiste dans certains cas des écarts résiduels de rentabilité financière après appariement au plus proche voisin sur la base des caractéristiques individuelles des entreprises. La présente étude se distingue de Bilonière et al. (2025) par son approche; au delà de considérer une variable expliquée indépendante des politiques de financement des entreprises, elle cherche à établir si les économies d'agglomération peuvent expliquer les différentiels spatiaux de rentabilité nets des caractéristiques individuelles, en tenant également compte de l'hétérogénéité sectorielle des économies d'agglomération.

Cette étude montre que les variables d'agglomération ne permettent pas d'expliquer les écarts inter-zones d'emploi d'EBE par travailleur des PME françaises entre 2017 et 2022 au global sur tous les secteurs. Afin de comprendre l'origine de cette absence d'effet, cette étude examine comment les mêmes variables d'agglomération agissent sur la valeur ajoutée au coût des facteurs, croissante de la productivité totale des facteurs, et sur les charges de personnel, croissantes de la productivité du travail. Ces deux variables, dont la différence est égale à l'EBE, sont toutes deux affectées positivement, significativement et dans une magnitude semblable par la densité en emploi. Les économies d'agglomération qui augmentent la productivité se compensent donc avec les économies d'agglomération sur les salaires pour les PME, de sorte que la richesse créée par travailleur n'est pas plus élevée dans les zones économiquement concentrées une fois nette du coût du travail. Ce résultat est conforté par l'analyse de l'hétérogénéité sectorielle de l'effet, qui montre qu'à secteur donné, la valeur ajoutée au coût des facteurs et les charges de personnel sont affectées de manière similaire par les variables d'agglomération, aboutissant à une compensation au moins partielle des effets. Cette dernière analyse permet toutefois d'identifier quelques secteurs pour lesquels la compensation n'est pas totale.

La suite du document est structurée comme suit : la section 2 passe en revue la littérature pertinente; la section 3 détaille la méthodologie empirique; la section 4 présente les résultats; puis la section 5 conclut.

2 REVUE DE LITTÉRATURE

Il est désormais bien établi que la densité locale accroît la productivité des entreprises et des travailleurs (Combes et Gobillon, 2015), et qu'il s'agit de la principale variable d'agglomération avec un effet positif et significatif sur la productivité. La riche littérature sur le sujet a d'abord rencontré de nombreuses limites méthodologiques, concluant dans un premier temps à la seule présence d'économies de localisation. C'est par exemple le cas de Henderson et al. (1995) qui mesurent la spécialisation par le logarithme de la part de l'emploi sectoriel local, en contrôlant également pour le logarithme de l'emploi sectoriel local. Or, comme le

soulignent Combes (2000), cela équivaut à conclure à un effet positif de la petite taille d'un marché local sur la productivité. L'effet disparaît lorsque le contrôle de l'emploi sectoriel local est remplacé par le contrôle de l'emploi local. Henderson (2003) conclut à l'absence d'économies d'urbanisation significatives sur la productivité de firmes américaines mais recourt à des effets fixes individuels qui restreignent l'identification de l'effet de la densité locale à ses seules variations temporelles qui sont très faibles.

La littérature de la fin des années 1990 et des années 2000 s'est donc particulièrement attachée à développer un cadre méthodologique rigoureux et également crédible du point de vue de l'identification causale. Deux biais d'endogénéité affectant l'estimation de l'effet de la densité locale sont distingués par Combes et al. (2011). Il y a tout d'abord le biais de la quantité de travail endogène selon lequel l'activité se concentre là où les entreprises sont déjà les plus productives, ou bien là où il existe des facteurs externes qui attirent la main d'œuvre et favorisent la productivité. Les travaux fondateurs sur la correction de ce biais sont ceux de Ciccone et Hall (1996), Rosenthal et Strange (2008) et Combes et al. (2010). Le premier instrumente la variable de densité par la même variable retardée de nombreuses années. Cet instrument n'est toutefois pas parfaitement exogène s'il existe des variables corrélées avec la densité historique et la productivité contemporaine, comme certaines caractéristiques naturelles. Pour cette raison, les travaux de Rosenthal et Strange (2008) et Combes et al. (2010) instrumentent la densité locale par des variables géologiques, des instruments bien plus exogènes mais également trop faibles. La correction du biais de la quantité de travail endogène n'affecte toutefois que peu les estimations. Le second biais d'endogénéité désigné comme le biais de la qualité endogène de travail est lui bien plus important. La main d'œuvre la plus qualifiée étant attirée par les zones les plus denses, il y a un effet de composition qui biaise les estimations de l'effet de la densité à la hausse, si bien qu'une fois corrigées, les estimations sont divisées par deux comme le montrent Combes et al. (2008). Ces derniers contrôlent pour l'hétérogénéité individuelle de la productivité du travail en introduisant un effet fixe individuel dans l'esprit de Glaeser et Maré (2001) dans une spécification qui régresse le salaire par individu sur des variables d'agglomération et estimée sur un panel de plusieurs millions de salariés français. L'élasticité de la productivité du travail à la densité en emploi qui en découle est de 0,02. Ces contributions fondatrices en matière d'identification causale des économies d'agglomération ont établi des nouveaux standards de publication. Comme le notent en effet Ahlfeldt et Pietrostefani (2019) dans leur méta-analyse de l'effet de la densité locale, les travaux publiés après 2008, l'an médian de publication sur le sujet, marquent une nette progression vers l'usage de contrôles adéquats et l'exploitation appropriée de la variation spatio-temporelle par le recours à des effets fixes. De plus, l'effet estimé des articles non publiés dans des revues académiques est en moyenne plus élevé de 0,4 écart-type.

Si les méthodologies empiriques ont beaucoup évolué, les variables d'agglomération étudiées sont restées très similaires au cours du temps, notamment parce qu'on sait désormais, au moins pour la densité locale, corriger les biais d'endogénéité. Les variables d'agglomération couramment étudiées sont donc la densité locale de manière quasi systématique, le potentiel de marché qui capture l'effet de la densité locale des marchés locaux à proximité et la diversité sectorielle à la suite des travaux de Jacobs (1969) pour les économies d'urbanisation, et la spécialisation sectorielle pour les économies de localisation. Si les économies de localisation ont la plupart du temps un effet positif ou nul sur la productivité, son pouvoir explicatif reste faible relativement à l'effet de la densité, tandis que l'effet de la diversité sectorielle n'est pas robuste à travers les

études (Combes et Gobillon, 2015). L'effet du potentiel de marché sur la productivité est positif et significatif dans les études portant sur des données françaises (par exemple Combes et al. (2010); Barbesol et Briant (2008)). L'effet estimé de toutes ces variables n'est toutefois qu'un effet net, qui ne distingue pas dans quelle mesure ils proviennent d'effet de partage, d'appariement ou d'apprentissage pour reprendre la typologie de Duranton et Puga (2004) évoquée en introduction. Il y a eu quelques tentatives de distinction de mécanisme dont Rosenthal et Strange (2001) constitue le principal exemple, qui recourent à des variables de proxy censées adéquatement capturer les partages de facteurs de production, de connaissance et de main d'œuvre. Les résultats de ce type d'analyse restent encore très descriptifs et l'identification crédible des mécanismes à l'origine des économies d'agglomération constitue un des défis de cette littérature (Combes et Gobillon, 2015).

Cette littérature a également tenté de quantifier l'étendue spatiale des effets et, bien qu'il y ait divergence sur l'étendue exacte, il y a consensus sur le fait que les effets décroissent rapidement avec la distance (Combes et Gobillon, 2015). Duranton et Overman (2005) établissent par exemple que les effets ne vont pas au-delà de 50km. C'est une des raisons pour lesquelles la littérature portant sur les économies d'agglomération en France s'est intéressée aux caractéristiques des zones d'emploi en tant que marché local, un découpage réalisé par l'INSEE qui décrit l'espace au sein duquel la plupart des trajets domicile-travail s'opèrent, d'un rayon moyen de 34km, dont la densité en emploi a un effet positif et significatif sur la productivité totale des facteurs et sur les salaires.

L'existence d'économies d'agglomération n'est pas spécifique aux pays développés comme la France ou les Etats-Unis; en moyenne, l'effet estimé de la densité dans les articles qui s'intéressent aux pays à faible et moyen revenu est même un peu plus élevé (Ahlfeldt et Pietrostefani, 2019). Combes et al. (2015) investiguent par exemple l'effet de la densité sur les salaires dans 87 grandes villes chinoises et établissent une élasticité de 0,10-0,12 tandis qu'un même ordre de grandeur est trouvé pour l'effet de la densité sur les revenus annuels individuels en Inde par Chauvin et al. (2017).

Si la nature des variables d'agglomération étudiées est restée assez stable dans le temps, davantage de variables dépendantes ont quant à elles été explorées ces dernières années. Ainsi, des études se sont par exemple intéressées à l'effet de l'agglomération sur la santé mentale (Melis et al., 2015), la végétalisation de l'espace (Lin et al., 2015) ou encore la criminalité (Raleigh et Galster, 2015). Les performances économiques des entreprises, au-delà de la productivité totale des facteurs et du dépôt de brevets, n'ont jamais été étudiées, certainement pour des raisons de confidentialité des données. Il s'agit de l'objectif de cette étude qui, en étudiant les effets d'agglomération sur la rentabilité réelle d'une entreprise, s'inscrit dans la continuité de cette littérature.

3 MÉTHODOLOGIE EMPIRIQUE

A/ Données : champ, choix des variables et sources

A.1 - L'échelle géographique d'observation

Le choix de l'aire géographique au sein de laquelle les variables d'agglomération sont mesurées est important; il relève d'une hypothèse implicite sur l'ampleur géographique des effets d'agglomération. Comme expliqué en section 2, les effets d'agglomération ont été établis par la littérature comme étant des effets assez localisés, et c'est une des raisons pour lesquelles les articles portant sur le cas français se sont intéressés aux zones d'emplois. Ce découpage réalisé par l'INSEE et dont le dernier millésime date de 2020 est défini comme "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". Composé de 306 unités en 2020, ce zonage, plus fin qu'un département, reflète donc les contours des zones où l'activité économique se concentre. Comme les articles portant sur les économies d'agglomération en France ont établi l'existence d'effets positifs sur les salaires (Combes et al., 2008) et sur la productivité totale des facteurs (Combes et al. (2010); Barbesol et Briant (2008)) en s'intéressant aux variables d'agglomération des **zones d'emploi**, cette étude fait de même afin d'établir si le net de ces effets est toujours positif et significatif.

A.2 - Les variables dépendantes et leur unité d'observation

L'étude de la performance économique des entreprises nécessite de retenir une variable permettant de l'apprécier. En analyse financière, de nombreux indicateurs rendent compte de la capacité de l'entreprise à rentabiliser ses capitaux propres et d'exploitation, à générer des recettes fiscales ou à assurer sa solvabilité, détaillés dans Mordant (1999).

Cette étude, portant en premier lieu sur les écarts spatiaux de rentabilité, retient l'Excédent Brut d'Exploitation (EBE) comme indicateur d'analyse. L'EBE correspond au solde des opérations d'exploitation directement liées à la production et présente l'avantage majeur d'être indépendant des politiques d'investissement, de financement ou d'amortissement, contrairement au bénéfice comptable. Il représente le résultat économique brut de l'activité opérationnelle, c'est-à-dire le montant net des ressources régulièrement générées par l'exploitation, et indique donc la rentabilité réelle de l'entreprise. Cet indicateur se rapproche fortement de la notion de profit en théorie microéconomique et se calcule en retranchant le coût du travail de la valeur ajoutée brute au coût des facteurs (VACF)⁴. L'EBE est ainsi pertinent pour évaluer si la hausse du coût du travail liée à la concentration spatiale de l'activité économique est plus que compensée par les gains de richesse créée par travailleur. Il est également utile d'examiner séparément ces deux composantes afin de mieux comprendre l'effet des variables d'agglomération sur l'EBE. Si les mécanismes théorisés et empiriquement identifiés existent, un effet positif et significatif de la densité locale sur la VACF d'une part et sur les charges de personnel d'autre part devraient être observés, justifiant l'analyse de ces deux agrégats comptables.

Ces variables sont issues des données du Fichier Approché des Résultats d'Esane (Esane FARE) qui contient

4. La VACF correspond à la valeur ajoutée aux prix de base après impôts et subventions, représentant la part de la valeur ajoutée distribuée entre travail et capital.

les données financières des entreprises françaises. Cette source présente deux niveaux d'observation possibles : au niveau de l'entreprise au sens économique, et au sens de l'unité légale. Le concept d'entreprise au sens économique a été introduit par la loi de modernisation de l'économie (LME) de 2008 (décret n° 2008-1354) qui la définit à partir de critères économiques et fournit donc une meilleure vision du tissu productif. Elle y est définie comme "la plus petite combinaison d'unités légales qui constitue une unité organisationnelle de production de biens et services jouissant d'une certaine autonomie de décision, notamment pour l'affectation de ses ressources courantes". L'observation des variables comptables au niveau de l'unité légale est donc le plus désagrégé des deux niveaux, et correspond pour cette raison à l'échelle d'observation retenue dans cette étude. L'étude des déterminants des écarts géographiques de performance économique requiert en effet de situer spatialement les variables dépendantes. Or une entreprise au sens économique peut être constituée de plusieurs unités légales géographiquement éloignées, regroupant elles-mêmes plusieurs établissements implantés dans des marchés locaux aux caractéristiques d'agglomération hétérogènes. De ce fait, l'échantillon doit être restreint aux entreprises mono-établissements, qui constituent le champ de cette étude comme détaillé en section 3.A.5. Une conception de l'entreprise au sens de l'unité légale permet donc d'avoir l'échantillon le plus grand, puisque davantage d'unités légales sont mono-établissement que d'entreprises au sens économique ne le sont. De plus, l'approche en entreprise au sens économique classe toutes les unités légales d'une entreprise dans le secteur de celle-ci (Béguin et Hecquet, 2015). L'approche en unité légale permet donc d'avoir une classification sectorielle locale plus précise. L'**unité d'observation** est donc **établissement-année**, chaque établissement étant l'unique établissement de son unité légale dans les régressions principales.

L'analyse de variables telles que l'EBE, la VACF ou les charges de personnel se heurte à un problème de dimension : leur valeur absolue dépend fortement du volume de moyens de production mobilisés. Comparer directement ces variables entre les entreprises sans tenir compte de leur différence de taille risquerait de refléter un simple effet de taille plutôt que leur performance relative. Pour cette raison, les variables dépendantes de cette étude sont rapportées au nombre de travailleurs⁵ en équivalent temps plein de l'année précédente.⁶ Plusieurs raisons motivent ce choix de dénominateur, au détriment du capital d'exploitation, qui aurait permis d'étudier le ratio de rentabilité économique. Tout d'abord, la théorie des économies d'agglomération montre que les effets transitent essentiellement par la main d'œuvre : les entreprises avec davantage d'employés sont donc potentiellement les plus susceptibles de profiter des effets d'agglomération. Ensuite, l'information sur les EQTP est disponible au niveau de l'établissement, ce qui n'est pas le cas des variables comptables, et cela importe lorsque le champ est étendu aux unités légales pluri-établissement. Enfin, le retardement de la variable d'un an permet d'éviter toute simultanéité entre l'EBE et le nombre de travailleurs, par exemple si un plus grand EBE conduit l'unité légale à recruter davantage la même année. Les principales variables dépendantes sont donc l'**EBE par travailleur** que l'on pourrait approcher d'une notion de "rentabilité apparente du travail" pour faire un parallèle avec les intitulés de mesures de productivité⁷, la **valeur ajoutée brute au coût des facteurs par travailleur**, et les **charges de personnel par travailleur**.

5. Dans la littérature sur les économies d'agglomération, les variables dépendantes sont souvent exprimées en logarithme, permettant d'estimer des élasticités plutôt que des effets de niveau. Comme l'EBE peut être nul ou négatif, cela n'est pas possible ici.

6. Les résultats sont qualitativement robustes à des formes alternatives de la variable dépendante, comme montré en Annexe C.

7. La productivité apparente du travail mesure le rapport entre la valeur ajoutée et la quantité de travail utilisée pour la générer (Kremp et Sklénard, 2019).

Leur répartition géographique, après restriction au champ décrit en section 3.A.5, est représentée dans les Figures 1 à 3. Les changements de couleur des cartes sont codés selon des seuils correspondant aux déciles de la variable, ce qui permet également d'illustrer leurs statistiques descriptives.

Figure 1 - $\frac{EBE_{it}}{EQTP_{it-1}}$ moyen par ZE 2020, en 2022 (en milliers euros par EQTP)

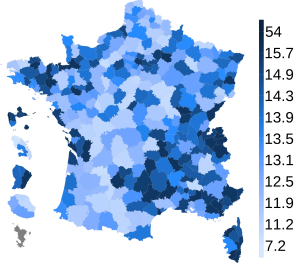


Figure 2 - $\frac{VACF_{it}}{EQTP_{it-1}}$ moyenne par ZE 2020, en 2022 (en milliers euros par EQTP)

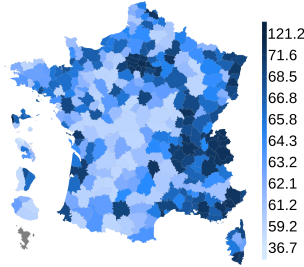
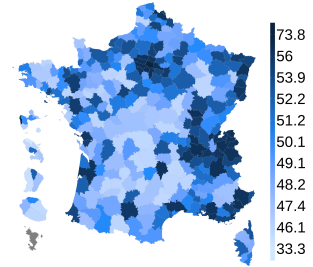


Figure 3 - $\frac{\text{Charges de personnel}_{it}}{EQTP_{it-1}}$ moyennes par ZE 2020, en 2022 (en milliers euros par EQTP)



A.3 - Les variables d'agglomération

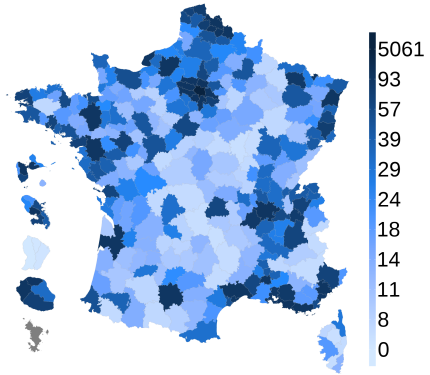
Ce document étudie l'effet de quatre variables standards dans la littérature sur les économies d'agglomération (voir Combes et Gobillon (2015) pour une revue de la littérature empirique) : la densité en emploi, le potentiel de marché, la diversité sectorielle et la spécialisation sectorielle de la zone. Ces variables sont construites comme suit :

Densité en emploi : Principale variable d'intérêt des articles portant sur les économies d'agglomération, cette variable d'urbanisation mesure, tous secteurs confondus, la concentration de l'activité économique de la zone d'emploi. Elle est construite à partir des Fichiers de Localisation des Rémunérations et de l'Emploi Salarié (FLORES), qui couvrent les statistiques démographiques de tous les établissements employeurs de France hors Mayotte, de la manière suivante :

$$\text{Densité}_{z,t} = \frac{\sum_{i \in Z_t} EQTP_{it}}{\text{Surface}_z}$$

où z désigne la zone d'emploi dans laquelle l'établissement considéré est situé, t l'année d'observation, i un établissement et Z_t l'ensemble des établissements localisés dans la zone d'emploi z l'année t .

Figure 4 - Densité en emploi des zones d'emploi 2020, en 2022



Potentiel de marché : Cette variable a pour objectif de tenir compte de la possibilité d'effets de débordements ("spillovers") des zones d'emploi alentours. Il est calculé à partir de la précédente mesure de densité en emploi et des distances, deux-à-deux, entre les barycentres des zones d'emploi :

$$\text{Potentiel de marché}_{z,t} = \sum_{z' \neq z} \frac{\text{Densité}_{z',t}}{\text{Distance}_{z'z}}$$

où z' désigne une autre zone d'emploi que la zone considérée.

Spécialisation sectorielle : Unique variable d'économie de localisation de cette étude, elle mesure la spécialisation relative de la zone d'emploi de l'établissement considéré dans le secteur d'activité de cet établissement comme la part de l'emploi sectoriel. Cette variable est également calculée à partir de la base FLORES :

$$\text{Spécialisation}_{z,s,t} = \frac{\sum_{i \in Z_t} \text{EQTP}_{i,t} \times 1\{s(i) = s\}}{\sum_{i \in Z_t} \text{EQTP}_{i,t}}$$

où s désigne le secteur de l'établissement considéré.

Diversité sectorielle : Cette variable d'économie d'urbanisation est parfois appelée variable d'économie *jacobienne* du fait des travaux de Jacobs (1969) sur les effets de "fertilisation croisée" entre secteurs. Si Jacobs n'est pas à l'origine de la manière dont cette variable est construite, elle est cependant très couramment spécifiée comme suit :

$$\text{Diversité}_{z,t} = \frac{1}{\sum_s (\text{Spécialisation}_{z,s,t})^2}$$

Cette variable est en fait l'inverse d'un indice de Herfindahl-Hirschman de concentration sectorielle; il est maximal lorsqu'il y a équi-répartition de l'emploi entre les secteurs de la zone d'emploi, et est dans ce cas égal au nombre de secteurs actifs dans la zone; il vaut 1 lorsque tout l'emploi est concentré dans un seul secteur d'activité.

Figure 5 - Potentiel de marché des zones d'emploi 2020, en 2022

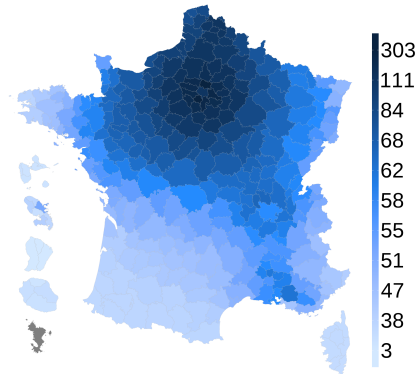
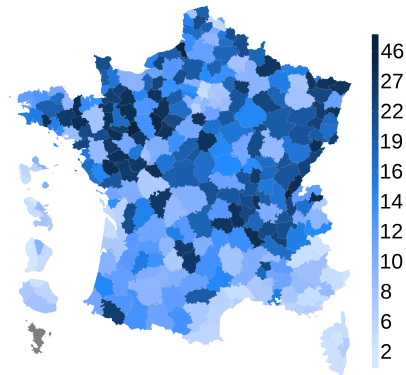


Figure 6 - Spécialisation sectorielle des zones d'emploi 2020 dans les industries manufacturières⁸, en 2022 (en %)



8. Codées "C" dans la NAF rév.2 au niveau 21.

Pour ces deux dernières variables, la nomenclature sectorielle NAF rév. 2 en 88 niveaux est utilisée pour les spécifications principales. Il est toutefois montré en Annexe C que les résultats sont robustes au recours à un autre niveau de granularité sectorielle, à savoir la NAF rév. 2 en 38 niveaux.

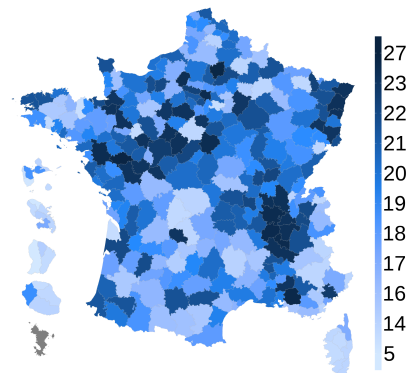
Il est important de noter que toutes ces variables sont calculées à partir de l'échantillon complet d'établissements employeurs. Ce n'est qu'après calcul des variables d'agglomération que l'échantillon est éventuellement restreint à certaines unités légales.

A.4 - Les variables instrumentales et de contrôle

Comme expliqué en section 2, la littérature empirique s'est particulièrement attachée ces dernières années à inférer la relation causale des variables d'agglomération – essentiellement la densité en emploi – sur un certain nombre de variables dépendantes de la manière la plus rigoureuse possible. Il est désormais standard et nécessaire de tenir compte de l'existence de deux biais détaillés dans Combes et al. (2011) :

Le biais de la quantité de travail endogène : Dans cette étude de la rentabilité des entreprises, il y a un biais de la quantité de travail endogène lorsqu'il y a causalité inverse, c'est-à-dire si les zones les plus rentables attirent davantage de travailleurs, ou lorsqu'il existe des variables omises qui attirent davantage de travailleurs et favorisent la rentabilité par ailleurs, comme cela pourrait être le cas des aménités (infrastructures de transport, proximité à un cours d'eau ...) par exemple. Depuis Ciccone et Hall (1996), il est courant de corriger ce biais par l'instrumentation des variables de densité en emploi et de potentiel de marché par les mêmes variables retardées de plusieurs décennies. Cette étude utilise le recensement de la population de 1876, disponible sur le site internet de l'INSEE, pour construire ces instruments qui sont donc les **densités de population** et le **potentiel de marché⁹ historiques** des zones d'emploi. La validité de ces instruments repose sur l'hypothèse que les déterminants de la localisation de la population en 1876 ne sont pas liés aux déterminants actuels des performances économiques des entreprises (exogénéité des instruments) mais que la densité de population d'il y a 150 ans est corrélée avec la concentration de l'activité économique actuelle (pertinence des instruments). L'hypothèse de pertinence tient à l'inertie de la répartition spatiale des personnes et de l'activité dans le temps, particulièrement marquée en France (Eaton et Eckstein, 1997) et qui s'explique notamment du fait de la durabilité des bâtis. L'exogénéité de l'instrument peut raisonnablement s'argumenter par le fait que la structure de l'économie française en 1876 était très différente de celle d'aujourd'hui. En 1876, la France était en effet en pleine seconde révolution industrielle, alors qu'on observe ces dernières décennies une désindustrialisation de l'activité économique. De plus, deux guerres mondiales séparent les années d'observation de cette étude de l'année des instruments. Ces variables ne

Figure 7 - Diversité sectorielle des zones d'emploi 2020, en 2022



9. Construit à partir de la densité de population historique et donc exprimé en nombre d'habitants et non de travailleurs.

sont malheureusement pas disponibles pour les zones d'emploi dromiennes ou corses. Les spécifications qui instrumentent les variables de densité économique et de potentiel marchand seront donc restreintes au champ des établissements de la France hexagonale et continentale.

Cependant, un biais de la quantité de main d'œuvre persiste après instrumentation s'il existe des variables qui expliquaient la localisation de la population en 1876 et qui favoriseraient la rentabilité des entreprises aujourd'hui. Cela pourrait être le cas des variables de géographie de première nature ("first-nature geography") comme par exemple la présence de reliefs ou la proximité avec des cours d'eau. Pour cette raison, les spécifications de cette étude contrôlent pour les caractéristiques des zones d'emploi suivantes :

Part de la surface classée en zone montagne : Cette variable est construite en sommant les surfaces des communes de la zone d'emploi classées en zones de montagne au titre du droit de l'urbanisme et en rapportant cette somme à la surface de la zone d'emploi. La classification en zone montagne au titre du droit de l'urbanisme est mise à disposition par l'Observatoire National des Territoires.

Part de la surface couverte par de l'eau : Cette variable calcule la part de la surface de la zone d'emploi couverte par de l'eau d'origine naturelle (lacs, fleuves et rivières naturels). Les contours des éléments hydrographiques sont obtenus par requête OpenStreetMap.

Distance avec le littoral : Cette variable est calculée à partir des données géospatiales de la limite terre-mer pour la France métropolitaine et du trait de côte HISTOLITT pour les DROMs ¹⁰, tous deux mis à disposition par le Shom. Elle est égale à la distance à vol d'oiseau la plus courte entre le barycentre de la zone d'emploi et le littoral.

Distance avec la frontière : Cette variable est motivée par le fait que les limites administratives des pays étaient globalement les mêmes en 1876, et que la variable de potentiel de marché ne tient pas compte du potentiel représenté par les régions étrangères à proximité. Elle est calculée à partir des données géospatiales sur les limites administratives des pays. Elle est égale à la distance à vol d'oiseau la plus courte entre le barycentre de la zone d'emploi et une frontière d'un pays étranger.

De plus, afin de faciliter les interprétations, les spécifications contrôlent également pour la **superficie de la zone d'emploi** afin que l'effet de la densité estimé reflète réellement un effet d'agglomération de l'activité et non de la taille du zonage.

Le biais de la qualité du travail endogène : Il y a un biais de la qualité du travail endogène ¹¹, lorsque la main d'œuvre la plus qualifiée est attirée par les zones les plus concentrées économiquement (par exemple parce qu'elles sont en général caractérisées par une plus grande offre culturelle), un phénomène étudié en détail dans Combes et al. (2008) qui montrent que l'effet de la densité économique est divisé par deux

10. Les DROMs sont en effet inclus en test de robustesse en Annexe C

11. Cet intitulé est la traduction littérale du concept d'"endogenous quality of labor bias". Il ne désigne pas la qualité du travail au sens normatif du terme mais plutôt les qualifications.

lorsque ce biais est corrigé. Dans ce cas, on ne voudrait pas que les effets d'agglomération soient confondus avec les effets de composition liés à une qualité de la main d'œuvre acquise autrement que par les économies d'agglomération (capital humain des parents par exemple). Depuis Glaeser et Maré (2001), il est courant de tirer profit de la longitudinalité des données pour absorber la part invariante de la qualité du travail avec un effet fixe travailleur lorsque la variable dépendante est le salaire. Comme cette étude porte sur les performances économiques des établissements, il n'est pas possible de faire cela. Alternativement, et en adéquation avec les articles portant sur la productivité totale des facteurs, cette étude approxime la qualité de la main d'œuvre de l'établissement par la **part en équivalent temps plein des emplois très qualifiés** en recourant à la nomenclature PCS 2020, à savoir la part des équivalents temps plein appartenant à la catégorie des "cadres et professions intellectuelles supérieures". Cette information est calculée à partir de la Base Tous Salariés (BTS) qui contient les variables socio-démographiques des salariés en France.

D'autres effets de composition que celui de la qualité du travail pourraient émerger et biaiser les estimations des effets d'agglomération si les caractéristiques des entreprises sont corrélées avec les variables d'agglomération sans en être la conséquence, et qu'elles expliquent par ailleurs la rentabilité de l'entreprise. Il pourrait en être ainsi pour l'âge de l'entreprise, le type de groupe auquel elle appartient, si elle est profilée, si elle reçoit des subventions d'exploitation ou encore sa taille. Cette dernière variable est particulièrement importante, puisque les variables comptables étant normalisées par la taille de l'entreprise comme expliqué en section 3.A.2, il pourrait y avoir un effet significatif estimé des variables d'agglomération sur la variable dépendante si les entreprises des zones les plus concentrées sont en moyenne composées d'entreprises avec un petit nombre de salariés par exemple. Il est donc nécessaire que les résultats soient robustes à l'inclusion de l'**âge de l'entreprise** au sens économique, de l'**origine du groupe** auquel elle appartient, des **subventions d'exploitation** reçues par l'unité légale et de la **taille de l'établissement** identique au dénominateur de la variable dépendante, à savoir retardée d'un an et exprimée pour les spécifications principales en équivalent temps plein.

A.5 - Champ de l'étude

L'étude des effets d'agglomération sur la performance économique des entreprises requiert de restreindre le champ des données à des entreprises spécifiques. Celles-ci doivent n'avoir, comme évoqué dans la section 3.A.2, qu'un seul établissement ¹², afin d'être en mesure de situer spatialement la performance économique qui n'est comptabilisée, au niveau le plus fin, qu'à l'échelle de l'unité légale. Cette restriction est standard dans la littérature sur les économies d'agglomération, mais ce choix n'est pas anodin car il pourrait induire un biais de sélection d'échantillonnage si le compte d'établissements par unité légale n'est pas aléatoire. Pour cette raison, il est vérifié en Annexe C que les résultats restent qualitativement inchangés lorsque l'on considère également les entreprises pluri-établissements et que l'EBE est ventilé parmi les établissements au prorata de la rémunération brute ¹³, la même méthode qui est retenue par l'INSEE pour le calcul de la valeur ajoutée. Afin de retenir des unités légales comparables et ainsi de réduire le bruit

12. En pratique, le compte des établissements par unité légale est réalisé dans cette étude sur l'échantillon des établissements employeurs, étant ceux qui produisent l'essentiel de la richesse ; un établissement non-employeur d'une unité légale qui emploie au moins 3 salariés désignant par exemple un entrepôt ou une boîte aux lettres.

13. Obtenue à partir de la Base Tous Salariés.

des estimations, le champ est également restreint aux unités légales employant au moins deux Équivalents Temps Plein (EQTP), et qui font partie d'entreprises qui au sens économique sont des PME avec plus de 10 EQTP¹⁴. Cette restriction permet notamment d'exclure toutes les unités légales qui relèvent du régime de micro-entrepreneur qui couvre des réalités économiques très diverses¹⁵ et dont les statistiques sont pour la plupart imputées. Le champ de cette étude est également contraint par les champs statistiques des différentes sources mobilisées. Les données d'entreprise s'appuient comme expliqué en section 3.A.2 sur le dispositif Esane (Élaboration des Statistiques ANnuelles d'Entreprises) qui ne couvre exhaustivement les unités légales que d'un nombre restreint de secteurs d'activité. Le secteur agricole n'est pas couvert par la source, tandis que ceux de la santé humaine-action sociale et de l'enseignement intègrent de nombreux établissements publics dont le poids économique est significatif; l'observation de ces activités uniquement à travers les entreprises privées offre donc une perspective incomplète, ce qui vaut également pour les organisations associatives, dont l'activité échappe souvent à la logique marchande. De même, la comptabilité spécifique des entreprises financières et d'assurance n'est pas homogène avec celle des autres secteurs qui ne sont de toute manière couvertes que partiellement par la source. Ces secteurs sont donc exclus de cette étude pour se restreindre au champ des secteurs principalement marchands non agricoles et non financiers, comme c'est le cas de la plupart des études portant sur les entreprises en France (voir par exemple la publication Insee Références "Les entreprises en France" (2023)). En définitive, le champ de cette étude se réfère aux **établissements qui emploient au moins 2 équivalents temps plein, qui sont l'unique établissement employeur de leur unité légale, qui appartiennent à une PME d'au moins 10 EQTP au sens de l'entreprise économique, qui sont présents dans la base au moins deux années consécutivement¹⁶, localisés en France hexagonale et continentale¹⁷, et dont le secteur d'activité est principalement marchand, non agricole et non financier**. Une troncature des valeurs extrêmes est également effectuée : les unités légales dont la valeur de la variable dépendante se situe dans les 5% les plus faibles ou les 5% les plus élevées, au sein de chaque secteur et année, sont exclues de l'échantillon. Il en est de même pour les unités légales percevant des subventions d'exploitation parmi les 1% plus élevées de leur secteur et de leur année. L'échantillon n'est restreint qu'après avoir calculé les variables d'agglomération. Une fois restreinte, la base est constituée de près de 675 000 observations établissement-année. La robustesse des résultats à des champs alternatifs, notamment à l'inclusion des unités légales pluri-établissement est testée en Annexe C.

14. Cela exclut donc les micro-entreprises ainsi que les PME de moins de 10 EQTP puisque ces dernières réalisent nécessairement plus de 2 millions d'euros de chiffre d'affaires pour être catégorisées comme des PME et non comme des micro-entreprises. Les variables comptables de cette étude étant en effet rapportées au nombre de travailleurs, conserver les établissements des PME de moins de 10 EQTP pourrait biaiser les estimations si leur localisation était corrélée avec les variables d'agglomération.

15. Voir par exemple l'Insee Références "Emploi et revenus des indépendants" (2025).

16. Du fait de l'emploi de variables retardées d'un an.

17. Du fait de l'indisponibilité des instruments pour la Corse et les DROMs : voir la discussion en section 3.A.4. Comme l'instrumentation affecte peu les résultats, les spécifications principales sont re-conduites sans instrumentation afin de vérifier la robustesse des résultats à l'inclusion des zones d'emploi corses et dromiennes (hors Mayotte) en Annexe C.

B/ Spécifications économétriques

B.1 - Estimation d'effets homogènes entre secteurs d'activité

La spécification de base de cette étude repose sur une procédure en deux étapes détaillée dans Gobillon (2004) estimée par Moindres Carrés Ordinaires :

$$\begin{aligned} \text{1ère étape :} \quad & \frac{Y_{it}}{\text{EQTP}_{it-1}} = X_{it}\delta + \beta_{zst} + \varepsilon_{it} \\ \text{2ème étape :} \quad & \widehat{\beta}_{zst} = \theta_{st} + X_{zt} \times \varphi + X_{zst} \times \gamma + X_z \times \sigma + \epsilon_{zst} \end{aligned} \quad (1)$$

où $Y \in \{\text{EBE, VACF, Charges de personnel}\}$. X_{it} est la matrice des variables de contrôle au niveau établissement-année, qui ont pour objectif de limiter les biais liés aux effets de composition, en particulier celui de la qualité endogène de la main d'oeuvre comme discuté en section 3.A.4. Les effets fixes zone d'emploi-secteur-année β_{zst} capturent donc les écarts moyens de la variable dépendante entre clusters zone d'emploi-secteur-année nets des caractéristiques individuelles. Une fois estimés, ces effets fixes sont régressés en seconde étape sur les variables d'agglomération ; les variables d'économies d'urbanisation, possiblement préalablement instrumentées, correspondent à la matrice X_{zt} ; l'unique variable d'économie de localisation est contenue dans la matrice X_{zst} ; φ et γ sont donc respectivement les vecteurs des coefficients d'économie d'urbanisation et de localisation. X_z est la matrice des caractéristiques des zones d'emploi invariantes dans le temps. Elles correspondent aux variables de contrôle telles que les caractéristiques de géographie de première nature ou la superficie des zones d'emploi 2020 évoquées en section 3.A.4. θ_{st} est un effet fixe secteur-année qui capture la part de la variable dépendante invariante au sein de chaque cluster secteur-année. Ceci inclut par exemple des chocs de demande sur le bien final ou d'offre sur les biens intermédiaires. Ces chocs peuvent affecter certains secteurs en particulier, en raison des spécificités de leur fonction de production ou de leur marché. Ils pourraient alors être corrélés avec certaines variables d'agglomération si la composition sectorielle varie entre zones.

Cette approche en deux étapes a l'avantage d'être plus rigoureuse vis-à-vis des potentielles variables individuelles omises qui seraient à l'origine d'effets de composition. En effet, dans une approche en une seule étape, le terme d'erreur individuel ε_{it} doit être orthogonal à X_{zst} et X_{zt} , ce qui n'est pas le cas si des variables individuelles omises sont corrélées avec certaines variables d'agglomération. Certaines approches managériales pourraient par exemple être davantage répandues dans les zones concentrées et impacter par ailleurs les performances des entreprises. Une approche en deux étapes évite ce biais lié à l'hétérogénéité individuelle, car l'estimation des effets fixes β_{zst} n'impose pas qu'ils soient orthogonaux au terme d'erreur individuel ε_{it} .

Il convient toutefois de faire preuve de prudence avec cette approche, car elle introduit de l'hétéroscédasticité (Gobillon, 2004). En effet, les effets fixes β_{zst} ne sont pas estimés avec la même précision selon la taille des clusters zone d'emploi-secteur-année : les coefficients issus de petits clusters présentent une variance plus élevée. Afin de corriger ce problème et de garantir la validité de l'hypothèse d'homoscédasticité entre clusters, la régression de deuxième étape est pondérée par le nombre d'observations dans chaque cluster.

Une telle spécification facilite également l'analyse des résultats en termes de forces et faiblesses écono-

miques des zones d'emploi françaises; une fois les deux spécifications estimées, l'espérance conditionnelle de la seconde étape fournit la prédiction de la performance économique moyenne des entreprises de chaque zone d'emploi du fait des caractéristiques des *zones*. Une telle analyse est proposée en fin de section 4.

B.2 - Hétérogénéité sectorielle des effets

Il n'y a aucune raison pour que les performances économiques des établissements de secteurs différents soient affectées de manière homogène par les variables d'agglomération. Henderson (2003) montre par exemple que les externalités de localisation sur la productivité sont plus fortes dans les secteurs de haute technologie que dans les secteurs plus traditionnels aux Etats-Unis. Sur le plan théorique, cette hétérogénéité sectorielle pourrait en effet se justifier par l'importance de connaissances spécifiques ou encore d'intensités capitalistiques différentes entre les secteurs puisque les économies d'agglomération transitent entre autres par des effets de débordement de connaissance et par la qualité de l'appariement entre la main d'œuvre et les emplois. Il est donc également intéressant d'estimer (1) séparément pour chaque secteur. Ainsi, dans une seconde partie de l'analyse, les deux étapes suivantes sont conduites pour chaque secteur d'activité :

$$\begin{aligned} \text{1ère étape :} \quad & \frac{Y_{it}}{\text{EQTP}_{it-1}} = X_{it}\delta + \beta_{zt} + \varepsilon_{it} \\ \text{2ème étape :} \quad & \widehat{\beta}_{zt} = \theta_t + X_{zt} \times \varphi + X_{zst} \times \gamma + X_z \times \sigma + \epsilon_{zt} \end{aligned} \quad (2)$$

La principale différence avec (1) est que les vecteurs de coefficients δ , φ , γ et σ peuvent désormais différer selon les secteurs. Cette procédure requiert qu'il y ait suffisamment d'établissements dans chaque secteur d'activité et que ces établissements soient localisés dans suffisamment de zones d'emploi différentes. Les secteurs retenus sont donc ceux pour lesquels il y a au moins 1500 observations établissement-année et que ces observations correspondent à au moins 30 zones d'emploi différentes. Comme pour construire les variables d'agglomération des spécifications principales, la nomenclature sectorielle NAF rév.2 en 88 niveaux est utilisée, mais il est également montré en Annexe C que les résultats restent qualitativement inchangés lorsque une nomenclature sectorielle plus agrégée est utilisée.

4 RÉSULTATS

Les estimations d'effets d'agglomération homogènes entre secteurs sont présentées dans un premier temps dans les tables 1 à 3. L'hétérogénéité sectorielle des effets est ensuite explorée et les résultats sont présentés en Figure 8.

La table 1 présente les résultats de l'estimation d'effets d'agglomération sur l'EBE par travailleur, la principale variable d'intérêt de cette étude. Il s'agit de déterminer notamment si la richesse créée par le personnel, en complémentarité avec le capital physique, est plus élevée là où l'activité est plus concentrée et ce même après rémunération du personnel, alors que l'existence d'économie d'agglomération positives et significatives sur les salaires a largement été démontrée par la littérature. Les différentes spécifications suivent la logique suivante; elles testent alternativement pour la présence d'économies d'urbanisation uniquement,

Table 1 - Effets des variables d'agglomération sur l'EBE par EQTP

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					1.188*** (0.099)	1.188*** (0.099)	1.162*** (0.097)	1.162*** (0.097)
log(EQTP _{t-1})							-1.084*** (0.183)	-1.084*** (0.183)
log(Age)							1.209*** (0.175)	1.209*** (0.175)
Subventions d'exploitation							0.713*** (0.152)	0.713*** (0.152)
Groupe franco-français							0.540*** (0.170)	0.540*** (0.170)
Multinationale sous contrôle français							-0.124 (0.336)	-0.124 (0.336)
Multinationale sous contrôle étranger							1.965*** (0.431)	1.965*** (0.431)
Pas de groupe, nationalité étrangère							1.756* (1.004)	1.756* (1.004)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	674296	674296	674296	674296	674206	674206	673720	673720
R2	0.260	0.260	0.260	0.260	0.261	0.261	0.265	0.265
R2 intra-groupes	NA	NA	NA	NA	0.002	0.002	0.007	0.007
R2 ajusté	0.187	0.187	0.187	0.187	0.188	0.188	0.192	0.192
R2 ajusté intra-groupes	NA	NA	NA	NA	0.002	0.002	0.007	0.007

Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	0.123 (0.125)	0.129 (0.130)	0.124 (0.124)	0.124 (0.132)	0.013 (0.123)	0.013 (0.134)	0.038 (0.123)	0.038 (0.133)
log(Potentiel de marché)	-0.161 (0.238)	-0.141 (0.224)	-0.117 (0.270)	-0.098 (0.250)	-0.130 (0.258)	-0.122 (0.246)	-0.120 (0.255)	-0.103 (0.236)
log(Diversité sectorielle)	0.427 (0.462)	0.317 (0.447)	0.717 (0.468)	0.630 (0.472)	0.614 (0.439)	0.578 (0.454)	0.607 (0.443)	0.533 (0.451)
log(Superficie)	-0.443** (0.196)	-0.388** (0.197)	-0.548*** (0.192)	-0.500** (0.198)	-0.563*** (0.189)	-0.543*** (0.195)	-0.549*** (0.188)	-0.507*** (0.193)
log(Spécialisation sectorielle)		0.356 (0.277)		0.308 (0.277)		0.126 (0.272)		0.261 (0.270)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	✓	✓	✓	✓	✓	✓
Observations	61154	61154	61154	61154	61115	61115	61103	61103
R2	0.561	0.561	0.781	0.781	0.778	0.778	0.771	0.771
R2 intra-groupes	0.004	0.005	0.503	0.502	0.502	0.502	0.501	0.501
R2 ajusté	0.558	0.559	0.779	0.779	0.777	0.776	0.769	0.769
R2 ajusté intra-groupes	0.004	0.005	0.503	0.502	0.502	0.502	0.501	0.501

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe 8. La modalité de référence de la nature du groupe est 'Pas de groupe, nationalité française'. Le champ et les sources sont décrits en section 3. Les R2 des colonnes (1) et (2) ne sont pas directement comparables avec ceux des colonnes suivantes, car pour ces dernières, la régression d'une variable dépendante débruitée sur des régresseurs eux-mêmes débruités (via l'instrumentation) tend mécaniquement à augmenter les R2.

puis d'économie d'urbanisation et de localisation, et corrigent progressivement les biais d'endogénéité. Les colonnes (1) et (2) régressent l'EBE par travailleur sur les variables d'agglomération présentées en section 3, sans correction ni de biais de la quantité de travail endogène, ni de celui de la qualité du travail endogène. Les estimations des effets des quatre variables d'intérêt; la densité en emploi, le potentiel de marché, la diversité sectorielle et la spécialisation sectorielle, ne sont pas significatives. Cela indique qu'il n'y a pas de corrélation entre l'EBE par travailleur et ces caractéristiques d'agglomération, ce qui pouvait d'ailleurs déjà s'observer en comparant la répartition géographique de l'EBE par travailleur en Figure 1 avec les répartitions géographiques des variables d'agglomération en Figures 4 à 7. Ce n'est pas pour autant qu'il ne puisse pas y

avoir d'effet causal. Ces estimations sont en effet biaisées. Il y a tout d'abord le biais de la quantité de travail endogène, que les colonnes (3) et (4) visent à corriger. Comme expliqué en section 3, ce biais désigne à la fois la possible causalité inverse qu'il peut y avoir de l'attractivité d'une zone particulièrement performante économiquement pour les travailleurs qui sont donc plus nombreux dans ces zones, et l'existence de variables omises qui expliquent à la fois les performances économiques des entreprises et la localisation des travailleurs, telles que les caractéristiques géographiques. Les coefficients de la densité et du potentiel de marché pourraient alors être biaisés à la hausse, et l'effet causal sur l'EBE pourrait être négatif et significatif. Les colonnes (3) et (4) instrumentent donc les variables de densité et de potentiel de marché par les mêmes variables retardées de près de 150 ans et contrôlent également pour les caractéristiques géographiques des zones d'emploi décrites en section 3 afin de corriger ce biais. Les coefficients restent non significatifs : l'absence d'effet n'est donc pas due au biais de la quantité de travail endogène. Toutefois, il subsiste encore le biais de la qualité du travail endogène. Particulièrement mis en évidence dans Combes et al. (2008), il désigne le fait que les travailleurs les plus qualifiés soient attirés par les zones les plus denses, or une large part de ces qualifications n'est pas le produit d'effets d'agglomération, comme le capital social des parents par exemple. Le coût du travail représenté par le personnel très qualifié étant plus élevé, les charges de personnel sont donc possiblement plus élevées dans les zones plus denses du simple fait de cette différence d'intensité en travail très qualifié des unités légales qui composent les zones d'emploi. Une fois ce biais contrôlé, les estimations des effets des économies d'agglomération sur la productivité sont divisées par deux (Combes et al., 2008). Ce biais pourrait être à l'origine de la non significativité des variables d'agglomération sur l'EBE par travailleur alors même qu'il pourrait y avoir un effet causal significatif, positif comme négatif. La première étape des colonnes (5) et (6) contrôle donc pour la part des emplois très qualifiés des établissements afin que les effets fixes zone-secteur-année estimés n'absorbent pas ces effets de composition de qualité de la main d'oeuvre. Les coefficients estimés restent cependant non significatifs, et diminuent même en magnitude, suggérant plutôt l'existence d'un petit biais positif de la qualité de la main d'oeuvre sur la richesse créée plus que sur le coût additionnel qu'il représente. Les colonnes (7) et (8) contrôlent pour d'autres effets de composition potentiels. $\log(EQTP_{t-1})$ permet par exemple de comparer l'EBE par travailleur d'établissements dont le nombre de travailleurs est identique. En effet, s'il y a des rendements d'échelle internes suffisamment importants, l'EBE serait susceptible d'augmenter plus que proportionnellement aux heures travaillées, ce qui entraînerait un biais positif du coefficient de la densité en emploi du fait de la construction de la variable dépendante si les établissements dans les zones les plus denses ont en moyenne plus d'EQTP. Certaines collectivités locales peuvent aussi accorder davantage de subventions d'exploitation, qui augmentent l'EBE. Les résultats ne sont pas affectés par ces éventuels effets de composition.

En moyenne sur tous les secteurs, il n'y a donc pas d'effets significatifs d'agglomération sur l'EBE par travailleur. Ce résultat peut aussi bien être le produit d'une absence d'économies d'agglomération sur les variables dont la différence forme l'EBE, que la résultante d'effets positifs et significatifs qui se compensent, ce qui serait en adéquation avec les résultats précédemment établis par la littérature. L'objectif des régressions présentées dans les tables 2 et 3 est donc d'investiguer les raisons de cette absence d'effet, en testant les effets d'agglomération sur la valeur ajoutée brute au coût des facteurs et sur les charges de personnel.

La table 2 présente les résultats de l'estimation des effets d'agglomération sur la valeur ajoutée brute au

coût des facteurs. Les colonnes (1) et (2) indiquent une forte corrélation entre la densité en emploi et la VACF : une hausse de 1% de la densité en emploi est associée à une hausse de la VACF par équivalent temps plein de plus de 20 euros. Les autres variables d'agglomération ne sont toutefois pas significatives de manière robuste. Après correction du biais de la quantité de travail endogène, les résultats restent semblables, bien que l'effet de la densité diminue un peu, en adéquation avec l'existence d'un biais positif mais relativement faible de la quantité de travail établi par Combes et al. (2010). Le biais de la qualité du travail endogène, contrôlé dans les colonnes (5) et (6), est lui bien plus important, en adéquation de nouveau avec la littérature. En effet, après correction de ce biais, on estime désormais qu'une hausse de 1% de la densité en emploi est associée à une hausse d'environ 10 euros de VACF par travailleur. Cet effet est robuste au contrôle des autres effets de composition potentiels. Les autres variables d'agglomération n'ont quant à elles toujours pas d'effet significatif robuste.

Il y a donc un effet positif et significatif de la principale variable d'économie d'urbanisation, la densité en emploi, sur la valeur ajoutée au coût des facteurs, qui témoigne de l'existence d'économies d'agglomération. L'absence d'effet d'agglomération sur l'EBE par EQTP des entreprises ne peut donc être expliquée par l'absence d'économies d'agglomération sur les composantes de l'EBE. En principe, les économies d'agglomération sur les charges de personnel devraient être de nature semblable pour expliquer l'absence d'effet sur l'EBE par EQTP. Les résultats présentés en table 3 permettent de vérifier cela.

La table 3 présente les résultats de l'estimation des effets d'agglomération sur les charges de personnel, qui sont retranchées à la VACF pour obtenir l'EBE. Elles comprennent les charges sur les salaires et les charges sociales. Les effets d'agglomération sur les charges de personnel sont, comme attendu, très similaires à ceux sur la VACF. En effet, l'effet estimé biaisé est approximativement de 20 euros de charges supplémentaires par EQTP pour une hausse de 1% de la densité en emploi, tandis qu'il est pratiquement divisé par deux lorsque le biais de la qualité du travail endogène est contrôlé. De nouveau, seule la densité en emploi a un effet significatif et robuste à travers les différentes spécifications, qui s'établit à environ 10 euros de charges additionnelles par EQTP lorsque la densité en emploi augmente de 1%.

Ces résultats confirment donc que l'absence d'effet d'agglomération sur l'EBE tous secteurs confondus s'explique par la présence d'économies d'agglomération qui augmentent à la fois la richesse créée et les salaires là où l'activité économique est plus concentrée. Il est toutefois possible que, de par leurs différentes technologies de production, les secteurs soient affectés différemment par ces variables d'agglomération et qu'il existe des secteurs pour lesquels l'effet sur la VACF et les charges de personnel ne se compensent pas exactement, voire qu'il n'y ait d'effet sur aucune de ces variables. Il est donc également intéressant d'étudier l'hétérogénéité sectorielle des effets d'agglomération, dont les résultats sont présentés en Figure 8.

On peut d'abord observer que, quelle que soit la variable dépendante, les effets d'agglomération semblent concerner un faible nombre de secteurs, bien que cela puisse être en partie expliqué par la perte de précision des estimateurs du fait de la réduction de la taille de l'échantillon. Ce constat avait déjà été fait par Barbesol et Briant (2008) dans leur étude portant sur l'effet des économies d'agglomération sur la productivité des entreprises françaises. De plus, les économies d'agglomération n'apparaissent pas concerner davantage de secteurs industriels que de services, ce qui avait également été observé par Barbesol et Briant (2008). Les effets des variables d'économies d'urbanisation que sont la densité en emploi et le potentiel de marché sur

Table 2 - Effets des variables d'agglomération sur la VACF par EQTP

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					8.779*** (0.539)	8.779*** (0.539)	8.509*** (0.536)	8.509*** (0.536)
log(EQTP _{t-1})							-7.319*** (0.519)	-7.319*** (0.519)
log(Age)							2.400*** (0.295)	2.400*** (0.295)
Subventions d'exploitation							0.523** (0.261)	0.523** (0.261)
Groupe franco-français							-0.826 (0.512)	-0.826 (0.512)
Multinationale sous contrôle français							1.156* (0.689)	1.156* (0.689)
Multinationale sous contrôle étranger							12.143*** (0.947)	12.143*** (0.947)
Pas de groupe, nationalité étrangère							23.055*** (2.385)	23.055*** (2.385)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	674296	674296	674296	674296	674215	674215	673827	673827
R2	0.413	0.413	0.413	0.413	0.427	0.427	0.439	0.439
R2 intra-groupes	NA	NA	NA	NA	0.023	0.023	0.043	0.043
R2 ajusté	0.355	0.355	0.355	0.355	0.370	0.370	0.383	0.383
R2 ajusté intra-groupes	NA	NA	NA	NA	0.023	0.023	0.043	0.043

Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	2.171*** (0.334)	2.200*** (0.315)	1.980*** (0.289)	1.979*** (0.263)	1.164*** (0.222)	1.164*** (0.224)	1.211*** (0.242)	1.210*** (0.230)
log(Potentiel de marché)	0.624 (0.740)	0.726 (0.694)	0.203 (0.811)	0.305 (0.739)	0.098 (0.652)	0.114 (0.631)	0.179 (0.624)	0.253 (0.560)
log(Diversité sectorielle)	1.019 (1.680)	0.464 (1.546)	2.866* (1.597)	2.422 (1.557)	2.152* (1.285)	2.084 (1.305)	1.986 (1.267)	1.666 (1.235)
log(Superficie)	-0.594 (0.511)	-0.315 (0.504)	-1.484*** (0.485)	-1.232** (0.525)	-1.601*** (0.446)	-1.562*** (0.502)	-1.530*** (0.444)	-1.349*** (0.487)
log(Spécialisation sectorielle)		1.811 (1.362)		1.589 (1.407)		0.243 (1.233)		1.144 (1.214)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	✓	✓	✓	✓	✓	✓
Observations	61094	61094	61094	61094	61057	61057	61049	61049
R2	0.637	0.638	0.845	0.845	0.819	0.819	0.815	0.815
R2 intra-groupes	0.048	0.052	0.594	0.594	0.575	0.575	0.574	0.574
R2 ajusté	0.635	0.636	0.844	0.844	0.818	0.818	0.813	0.814
R2 ajusté intra-groupes	0.048	0.051	0.593	0.594	0.575	0.575	0.574	0.574

***p < 0.01; **p < 0.05; *p < 0.1. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe 8. La modalité de référence de la nature du groupe est 'Pas de groupe, nationalité française'. Le champ et les sources sont décrits en section 3. Les R2 des colonnes (1) et (2) ne sont pas directement comparables avec ceux des colonnes suivantes, car pour ces dernières, la régression d'une variable dépendante débruitée sur des régresseurs eux-mêmes débruités (via l'instrumentation) tend mécaniquement à augmenter les R2.

la VACF et les charges de personnel sont positifs pour la plupart des secteurs, tandis que l'effet de la spécialisation sectorielle est plus hétérogène. Ceci pourrait possiblement être attribué à un effet concurrentiel d'intensité variable selon les secteurs, bien que ces estimations ne permettent pas d'identifier les mécanismes à l'œuvre. En s'appuyant sur la nomenclature sectorielle disponible en Annexe A, on observe que les secteurs les plus exposés aux économies d'agglomération liées à la concentration de l'activité économique sur les charges de personnel sont les activités spécialisées, scientifiques et techniques, tandis que la rémunération moyenne des EQTP est plus sensible à la spécialisation sectorielle pour les secteurs des transports et de l'entreposage.

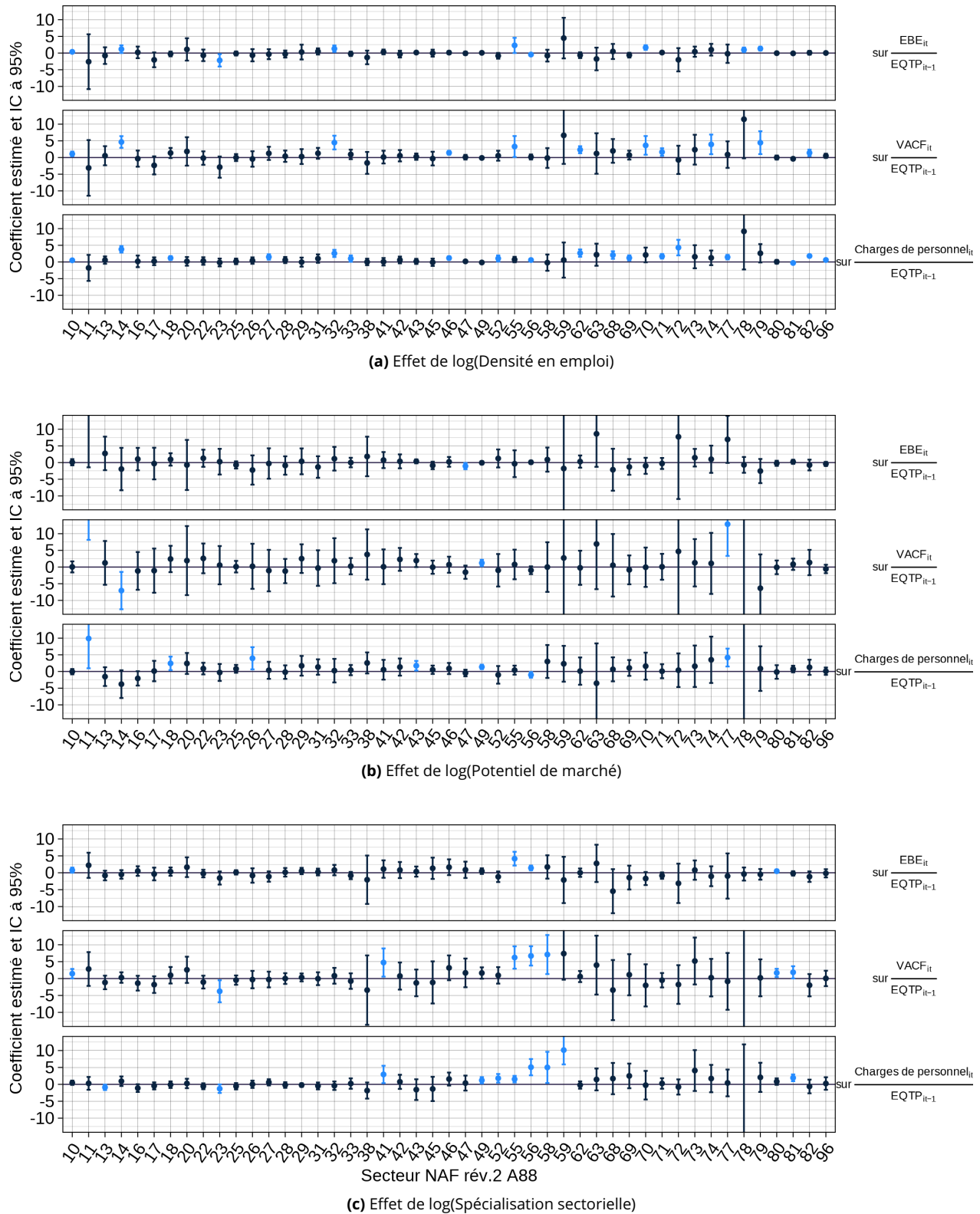
Table 3 - Effets des variables d'agglomération sur les charges de personnel par EQTP

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					7.189*** (0.534)	7.189*** (0.534)	7.005*** (0.539)	7.005*** (0.539)
log(EQTP _{t-1})							-5.634*** (0.373)	-5.634*** (0.373)
log(Age)							0.910*** (0.115)	0.910*** (0.115)
Subventions d'exploitation							-0.273** (0.131)	-0.273** (0.131)
Groupe franco-français							-1.533** (0.668)	-1.533** (0.668)
Multinationale sous contrôle français							1.191** (0.535)	1.191** (0.535)
Multinationale sous contrôle étranger							9.763*** (0.570)	9.763*** (0.570)
Pas de groupe, nationalité étrangère							16.778*** (1.754)	16.778*** (1.754)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	674296	674296	674296	674296	674216	674216	673773	673773
R2	0.476	0.476	0.476	0.476	0.489	0.489	0.501	0.501
R2 intra-groupes	NA	NA	NA	NA	0.026	0.026	0.047	0.047
R2 ajusté	0.424	0.424	0.424	0.424	0.439	0.439	0.451	0.451
R2 ajusté intra-groupes	NA	NA	NA	NA	0.026	0.026	0.047	0.047
Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	1.949** (0.290)	1.972*** (0.268)	1.768*** (0.246)	1.768*** (0.225)	1.125*** (0.178)	1.125*** (0.178)	1.138*** (0.190)	1.138*** (0.180)
log(Potentiel de marché)	0.940 (0.584)	1.016* (0.558)	0.481 (0.617)	0.557 (0.569)	0.407 (0.499)	0.413 (0.485)	0.489 (0.466)	0.539 (0.427)
log(Diversité sectorielle)	0.430 (1.270)	0.027 (1.167)	1.922 (1.205)	1.606 (1.186)	1.427 (0.932)	1.402 (0.944)	1.300 (0.919)	1.093 (0.897)
log(Superficie)	-0.302 (0.388)	-0.092 (0.396)	-1.071*** (0.365)	-0.885** (0.423)	-1.167*** (0.333)	-1.153*** (0.413)	-1.102*** (0.330)	-0.980** (0.399)
log(Spécialisation sectorielle)		1.334 (1.215)		1.159 (1.262)		0.093 (1.125)		0.759 (1.107)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	✓	✓	✓	✓	✓	✓
Observations	60677	60677	60677	60677	60639	60639	60628	60628
R2	0.645	0.646	0.852	0.852	0.827	0.827	0.825	0.826
R2 intra-groupes	0.051	0.053	0.604	0.605	0.588	0.588	0.587	0.587
R2 ajusté	0.643	0.644	0.851	0.851	0.826	0.826	0.824	0.824
R2 ajusté intra-groupes	0.051	0.053	0.604	0.605	0.587	0.587	0.587	0.587

***p < 0.01; **p < 0.05; *p < 0.1. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe 8. La modalité de référence de la nature du groupe est 'Pas de groupe, nationalité française'. Le champ et les sources sont décrits en section 3. Les R2 des colonnes (1) et (2) ne sont pas directement comparables avec ceux des colonnes suivantes, car pour ces dernières, la régression d'une variable dépendante débruitée sur des régresseurs eux-mêmes débruités (via l'instrumentation) tend mécaniquement à augmenter les R2.

L'interprétation de ces coefficients se heurte à la principale limite de l'étude empirique des économies d'agglomération; l'effet estimé n'est qu'un effet net résultant de multiples mécanismes opérant possiblement dans des directions différentes, non identifiés par les estimateurs. Toutefois, cette analyse permet de faire un constat éclairant pour cette étude : la direction des effets d'agglomération sur la VACF et sur les charges de personnel généralement identique pour un même secteur. Ceci tend à confirmer de nouveau que la faiblesse des effets d'agglomération sur l'EBE provient d'économies d'agglomération sur la VACF et les charges de personnel qui se compensent au moins partiellement. Cette analyse permet toutefois d'identifier quelques secteurs pour lesquels cette compensation n'est pas totale.

Figure 8 - Étude de l'hétérogénéité sectorielle des effets d'agglomération sur les performances économiques -
Spécification de la colonne (8)



Note : Les nomenclatures sectorielles sont disponibles en Annexe A. Le champ et les sources sont décrits en section 3. Les coefficients significatifs au seuil de 5% ainsi que leur intervalle de confiance à 95% sont en bleu.

Lecture : Une hausse de 1% de la densité en emploi augmente significativement l'EBE par équivalent temps plein des unités légales du secteur 14 (Industrie de l'habillement) de 12,50€.

Dans les secteurs des industries alimentaires, des industries de l'habillement ou encore de l'hébergement, il persiste par exemple un effet net positif et significatif de la densité en emploi sur l'EBE par travailleur. Cela pourrait par exemple se produire si le pouvoir de négociation salariale était inférieur dans ces secteurs, auquel cas les charges de personnel augmentent moins vite que la productivité. De nouveau, cette étude n'explique pas quels sont les mécanismes à l'origine de cet effet.

Ces procédures d'estimation sectorielles permettent de produire une analyse des forces et faiblesses des zones d'emploi françaises en termes d'économies d'agglomération dans chacun des secteurs. En effet, il est possible de prédire, au sens économétrique du terme, la contribution des effets d'agglomération à la valeur moyenne par zone d'emploi de l'EBE, de la VACF et des charges de personnel par travailleur, à partir des coefficients estimés lors de la régression de deuxième étape. Cette étude propose d'illustrer cela en s'intéressant au cas spécifique des industries alimentaires, les analyses conduites en figure 8 montrant qu'il est particulièrement sujet aux économies d'agglomération. La table 4 détaille les résultats des régressions de deuxième étape pour chacune des variables dépendantes.

Table 4 - Régressions de deuxième étape - Spécification de la colonne (8) - Secteur des industries alimentaires

	$\frac{EBE_{it}}{EQTP_{it-1}}$	$\frac{VACF_{it}}{EQTP_{it-1}}$	$\frac{\text{Charges de personnel}_{it}}{EQTP_{it-1}}$
log(Densité en emploi)	0.37** (0.12)	1.09** (0.27)	0.50*** (0.12)
log(Potentiel de marché)	0.11 (0.36)	0.05 (0.65)	-0.05 (0.31)
log(Diversité sectorielle)	-0.64 (0.91)	-1.45 (1.47)	0.45 (0.71)
log(Superficie)	0.01 (0.26)	-0.14 (0.45)	-0.30 (0.25)
log(Spécialisation sectorielle)	0.77** (0.30)	1.45** (0.54)	0.47 (0.25)
Part de la superficie couverte par de l'eau	95.37 (72.97)	-50.77 (117.38)	-127.67** (44.29)
Part de la superficie classée en zone montagne	2.84*** (0.66)	5.98*** (1.03)	2.77*** (0.50)
Distance à la frontière	0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Distance au littoral	-0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Effets fixes année	✓	✓	✓
Densité et potentiel de marché instrumentés	✓	✓	✓
Observations	1659	1658	1661
R2	0.86	0.88	0.91
R2 intra-groupes	0.84	0.85	0.87
R2 ajusté	0.85	0.88	0.91
R2 ajusté intra-groupes	0.84	0.85	0.87

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc à une année. Aucune variable explicative n'est ici centrée-réduite lors du processus d'estimation afin de pouvoir ensuite calculer l'espérance de la variable dépendante conditionnelle aux caractéristiques observées. Les ordres de grandeur des coefficients ne sont donc pas comparables. Le champ et les sources sont décrits en section 3.

On retrouve, pour les variables de densité, de potentiel de marché et de spécialisation, les coefficients représentés en figure 8. Il est possible d'isoler la contribution des effets d'agglomération *uniquement*, i.e. en excluant les caractéristiques de géographie de première nature de la prédiction¹⁸. L'écart entre la performance moyenne par zone d'emploi des PME observée et prédite réside alors dans la contribution des caractéristiques naturelles qui, comme elles sont corrélées avec la densité en emploi instrumentée, sont incluses dans le modèle, et dans la contribution de toute autre variable locale non corrélée avec les variables de spécialisation et de diversité sectorielles par hypothèse de la spécification économétrique. On peut également décomposer les effets parmi les variables d'urbanisation et de localisation afin de mieux comprendre les déterminants des performances relatives prédites par le modèle. La contribution des effets d'agglomération à la prédiction des performances locales des PME est représentée en figure 9.

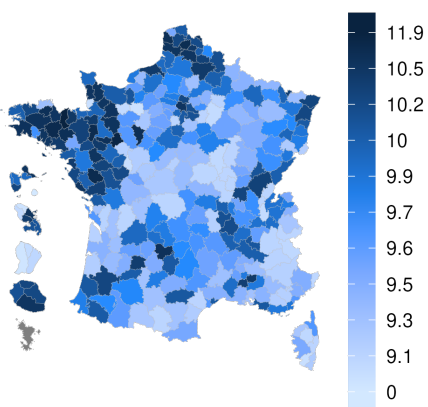
Cette analyse permet d'observer que la rentabilité des PME du secteur des industries alimentaires localisées de l'ouest de la région Auvergne Rhône-Alpes au sud de la région Nouvelle-Aquitaine en passant par la région Occitanie est susceptible de bénéficier de la spécialisation locale dans le secteur des industries alimentaires relativement plus élevée qu'ailleurs en France, tout comme pour les zones d'emploi du nord-ouest de la France. Les externalités de localisation, qu'il s'agisse d'un meilleur partage des facteurs de production, d'une plus grande diversité de ces derniers, d'une main d'œuvre plus spécialisée ou encore d'externalités de connaissances, y sont donc plus élevées dans ces régions, et sont surtout telles qu'elles peuvent contribuer à une plus grande rentabilité des PME du secteur des industries alimentaires.

Si des PME du secteur des industries alimentaires choisissent de se localiser à proximité de Paris ou bien dans les îles de Guadeloupe, Martinique et de La Réunion par exemple, alors celles-ci pourront bénéficier d'une plus grande concentration de l'activité économique tous secteurs confondus à l'origine d'externalités d'urbanisation qui favorisent la rentabilité.

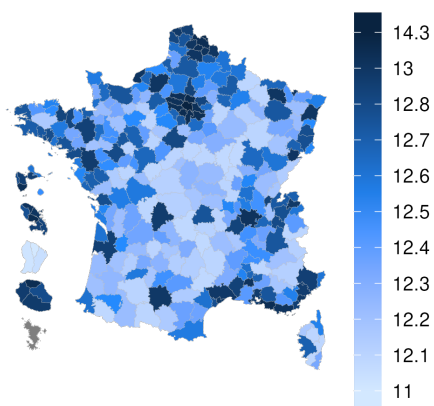
Combinées, les économies d'urbanisation et de localisation prédisent les plus fortes externalités d'agglomération sur la rentabilité des PME du secteur des industries alimentaires dans le nord et le nord-ouest de la France, avec des disparités régionales prédites qui proviennent essentiellement des différentiels de spécialisation relative des zones d'emploi dans le secteur des industries alimentaires.

18. Les effets fixes année sont conservés puisqu'ils absorbent une partie du niveau moyen de l'EBE par travailleur, toutes zones d'emploi confondues, dans le secteur des industries alimentaires.

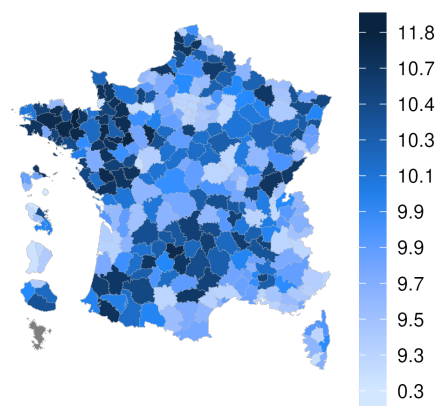
Figure 9 - Contribution des effets d'agglomération à la valeur moyenne par zone d'emploi prédite de $\frac{EBE_{it}}{EQTP_{it-1}}$ des établissements du secteur des industries alimentaires



(a) Contribution des économies d'agglomération



(b) Contribution des économies d'urbanisation



(c) Contribution des économies de localisation

Note : Le champ et les sources sont décrits en section 3. La spécification économétrique sur laquelle le modèle est estimé est celle de la table 4. Certaines variables étant exclues du conditionnement lors de l'étape de prédiction, les valeurs en légende ne peuvent pas être directement interprétées comme la valeur moyenne par zone d'emploi de l'EBE par travailleur prédite par le modèle. Seuls les écarts entre les valeurs sont informatifs, en l'occurrence de la disparité de performance prédite par les économies d'agglomération.

Lecture : 10% des zones d'emploi françaises ont des caractéristiques d'agglomération telles que la moyenne prédite de l'EBE par travailleur des établissements du secteur des industries alimentaires est de plus de 119 euros supérieure, toutes choses égales par ailleurs.

5 CONCLUSION

Cette étude se propose de déterminer si les disparités spatiales de performance économique des PME françaises peuvent s'expliquer par l'existence d'effets d'agglomération. Le cadre méthodologique, qui répond à une logique d'identification causale, est le même que celui qui a permis d'établir l'existence d'économies d'agglomération sur la productivité et les salaires dans les travaux fondateurs sur le sujet. Cette étude montre que les disparités spatiales d'Excédent Brut d'Exploitation, qui mesure la rentabilité réelle des entreprises et qui correspond à la différence entre la valeur ajoutée brute au coût des facteurs et les charges de personnel, ne sont pas expliquées par les variables d'agglomération tous secteurs confondus. L'étude de l'effet de ces mêmes variables sur la valeur ajoutée au coût des facteurs puis sur les charges de personnel montre un effet positif, significatif et de magnitude semblable de la densité en emploi sur ces deux composantes de l'EBE. Les économies d'agglomération qui augmentent la productivité se compensent donc avec les économies d'agglomération sur les salaires pour les PME, de sorte que la richesse créée par travailleur n'est pas plus élevée dans les zones économiquement concentrées une fois nette du coût du travail. Ce résultat est appuyé par l'étude de l'hétérogénéité sectorielle des effets qui montre que pour un même secteur, la valeur ajoutée au coût des facteurs et les charges de personnel sont affectées de façon similaire par les variables d'agglomération, résultant en une compensation au moins partielle des économies d'agglomération. Cette analyse identifie toutefois quelques secteurs pour lesquels cette compensation n'est pas totale et où il subsiste des effets d'agglomération significatifs sur la rentabilité. L'étude causale des mécanismes à l'œuvre, qui constitue encore aujourd'hui un des défis de la littérature empirique sur les économies d'agglomération, pourrait éclairer les raisons de cette hétérogénéité sectorielle.

RÉFÉRENCES

- Ahlfeldt, G. M., & Pietrostefani, E. (2019). The economic effects of density : A synthesis. *Journal of Urban Economics*, 111, 93-107.
- Barbesol, Y., & Briant, A. (2008). Économies d'agglomération et productivité des entreprises : estimation sur données individuelles françaises. *Economie et statistique*, 419(1), 31-54.
- Béguin, J.-M., & Hecquet, V. (2015). *Avec la définition économique des entreprises, une meilleure vision du tissu productif* (INSEE Références. Les entreprises en France, Edition 2015).
- Bilionière, M., Clarenc, P., & Salibekyan-Rosain, Z. (2025). *Comparaison de la rentabilité financière des PME dromiennes et hexagonales* (Document de travail INSEE n°2025-17).
- Chauvin, J. P., Glaeser, E., Ma, Y., & Tobio, K. (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics*, 98, 17-49.

- Ciccone, A., & Hall, R. E. (1996). Productivity and the Density of Economic Activity. *The American Economic Review*, 86(1), 54-70.
- Combes, P.-P. (2000). Marshall-Arrow-Romer externalities and city growth. *CERAS wp*, 99(06).
- Combes, P.-P., Démurger, S., & Li, S. (2015). Migration externalities in Chinese cities. *European Economic Review*, 76, 152-167.
- Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities : Sorting matters! *Journal of Urban Economics*, 63(2), 723-742.
- Combes, P.-P., Duranton, G., & Gobillon, L. (2011). The identification of agglomeration economies. *Journal of economic geography*, 11(2), 253-266.
- Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2010, avril). Estimating Agglomeration Economies with History, Geology, and Worker Effects. In *Agglomeration Economics* (p. 15-66). University of Chicago Press.
- Combes, P.-P., & Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of regional and urban economics* (p. 247-348, T. 5). Elsevier.
- Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *The Review of Economic Studies*, 72(4), 1077-1106.
- Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics* (p. 2063-2117, T. 4). Elsevier.
- Eaton, J., & Eckstein, Z. (1997). Cities and growth : Theory and evidence from France and Japan. *Regional science and urban Economics*, 27(4-5), 443-474.
- Glaeser, E. L., & Maré, D. C. (2001). Cities and Skills. *Journal of Labor Economics*, 19(2), 316-342.
- Gobillon, L. (2004). The estimation of cluster effects in linear panel models. *Processed, ined.*
- Henderson, V. (2003). Marshall's scale economies. *Journal of urban economics*, 53(1), 1-28.
- Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial Development in Cities. *Journal of Political Economy*, 103(5), 1067-1090.

- Jacobs, J. (1969). *The Economy of Cities*. Knopf Doubleday Publishing Group.
- Kremp, E., & Sklénard, G. (2019). *Productivité du travail et du capital : une mesure renouvelée au niveau de l'entreprise* (INSEE Références. Les entreprises en France, Edition 2019).
- Lin, B., Meyers, J., & Barnett, G. (2015). Understanding the potential loss and inequities of green space distribution with urban densification. *Urban forestry & urban greening*, 14(4), 952-958.
- Marshall, A. (1890). Principles of economics. *London : Mac-Millan*, 1-627.
- Melis, G., Gelormino, E., Marra, G., Ferracin, E., & Costa, G. (2015). The effects of the urban built environment on mental health : A cohort study in a large northern Italian city. *International journal of environmental research and public health*, 12(11), 14898-14915.
- Mordant, G. (1999). *Méthodologie d'analyse financière* (Document de travail INSEE n°E9907).
- Raleigh, E., & Galster, G. (2015). Neighborhood Disinvestment, Abandonment, and Crime Dynamics. *Journal of Urban Affairs*, 37(4), 367-396.
- Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of urban economics*, 50(2), 191-229.
- Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, 64(2), 373-389.

Annexe A - Nomenclature Sectorielle NAF rév. 2

Table 5 - NAF rév.2 - Classifications en 21, 38 et 88 niveaux

A21	Intitulé	A38	Intitulé	A88	Intitulé
A	Agriculture, Sylviculture Et Pêche	AZ	Agriculture, Sylviculture Et Pêche	01	Culture Et Production Animale, Chasse Et Services Annexes
				02	Sylviculture Et Exploitation Forestière
				03	Pêche Et Aquaculture
B	Industries Extractives	BZ	Industries Extractives	05	Extraction De Houille Et De Lignite
				06	Extraction D'Hydrocarbures
				07	Extraction De Minerais Métalliques
				08	Autres Industries Extractives
				09	Services De Soutien Aux Industries Ex- tractives
C	Industrie Manufacturière	CA	Fabrication De Denrées Ali- mentaires, De Boissons Et De Produits À Base De Tabac	10	Industries Alimentaires
				11	Fabrication De Boissons
				12	Fabrication De Produits À Base De Ta- bac
		CB	Fabrication De Textiles, Indus- tries De L'Habillement, Indus- trie Du Cuir Et De La Chaussure	13	Fabrication De Textiles
				14	Industrie De L'Habillement
				15	Industrie Du Cuir Et De La Chaussure
		CC	Travail Du Bois, Industries Du Papier Et Imprimerie	16	Travail Du Bois Et Fabrication D'Ar- ticles En Bois Et En Liège, À L'Exception Des Meubles ; Fabrication D'Articles En Vannerie Et Sparterie
				17	Industrie Du Papier Et Du Carton
				18	Imprimerie Et Reproduction D'Enregis- trements
		CD	Cokéfaction Et Raffinage	19	Cokéfaction Et Raffinage
		CE	Industrie Chimique	20	Industrie Chimique
		CF	Industrie Pharmaceutique	21	Industrie Pharmaceutique
		CG	Fabrication De Produits En Ca- outchouc Et En Plastique Ainsi Que D'Autres Produits Miné- raux Non Métalliques	22	Fabrication De Produits En Caou- tchouc Et En Plastique
				23	Fabrication D'Autres Produits Miné- raux Non Métalliques
		CH	Métallurgie Et Fabrication De Produits Métalliques À L'Ex- ception Des Machines Et Des Équipements	24	Métallurgie
				25	Fabrication De Produits Métalliques, À L'Exception Des Machines Et Des Équi- pements

A21	Intitulé	A38	Intitulé	A88	Intitulé
		CI	Fabrication De Produits Informatiques, Électroniques Et Optiques	26	Fabrication De Produits Informatiques, Électroniques Et Optiques
		CJ	Fabrication D'Équipements Électriques	27	Fabrication D'Équipements Électriques
		CK	Fabrication De Machines Et Équipements N.c.a.	28	Fabrication De Machines Et Équipements N.c.a.
		CL	Fabrication De Matériels De Transport	29	Industrie Automobile
				30	Fabrication D'Autres Matériels De Transport
		CM	Autres Industries Manufacturières; Réparation Et Installation De Machines Et D'Équipements	31	Fabrication De Meubles
				32	Autres Industries Manufacturières
				33	Réparation Et Installation De Machines Et D'Équipements
D	Production Et Distribution D'Électricité, De Gaz, De Vapeur Et D'Air Conditionné	DZ	Production Et Distribution D'Électricité, De Gaz, De Vapeur Et D'Air Conditionné	35	Production Et Distribution D'Électricité, De Gaz, De Vapeur Et D'Air Conditionné
E	Production Et Distribution D'Eau; Assainissement, Gestion Des Déchets Et Dépollution	EZ	Production Et Distribution D'Eau; Assainissement, Gestion Des Déchets Et Dépollution	36	Captage, Traitement Et Distribution D'Eau
				37	Collecte Et Traitement Des Eaux Usées
				38	Collecte, Traitement Et Élimination Des Déchets; Récupération
				39	Dépollution Et Autres Services De Gestion Des Déchets
F	Construction	FZ	Construction	41	Construction De Bâtiments
				42	Génie Civil
				43	Travaux De Construction Spécialisés
G	Commerce; Réparation D'Automobiles Et De Motocycles	GZ	Commerce; Réparation D'Automobiles Et De Motocycles	45	Commerce Et Réparation D'Automobiles Et De Motocycles
				46	Commerce De Gros, À L'Exception Des Automobiles Et Des Motocycles
				47	Commerce De Détail, À L'Exception Des Automobiles Et Des Motocycles
H	Transports Et Entreposage	HZ	Transports Et Entreposage	49	Transports Terrestres Et Transport Par Conduites
				50	Transports Par Eau
				51	Transports Aériens
				52	Entreposage Et Services Auxiliaires Des Transports
				53	Activités De Poste Et De Courrier
I	Hébergement Et Restauration	IZ	Hébergement Et Restauration	55	Hébergement
				56	Restauration
J	Information Et Communication	JA	Edition, Audiovisuel Et Diffusion	58	Édition

A21	Intitulé	A38	Intitulé	A88	Intitulé
				59	Production De Films Cinématographiques, De Vidéo Et De Programmes De Télévision; Enregistrement Sonore Et Édition Musicale
				60	Programmation Et Diffusion
		JB	Télécommunications	61	Télécommunications
		JC	Activités Informatiques Et Services D'Information	62	Programmation, Conseil Et Autres Activités Informatiques
				63	Services D'Information
K	Activités Financières Et D'Assurance	KZ	Activités Financières Et D'Assurance	64	Activités Des Services Financiers, Hors Assurance Et Caisses De Retraite
				65	Assurance
				66	Activités Auxiliaires De Services Financiers Et D'Assurance
L	Activités Immobilières	LZ	Activités Immobilières	68	Activités Immobilières
M	Activités Spécialisées, Scientifiques Et Techniques	MA	Activités Juridiques, Comptables, De Gestion, D'Architecture, D'Ingénierie, De Contrôle Et D'Analyses Techniques	69	Activités Juridiques Et Comptables
				70	Activités Des Sièges Sociaux; Conseil De Gestion
				71	Activités D'Architecture Et D'Ingénierie; Activités De Contrôle Et Analyses Techniques
		MB	Recherche-Développement Scientifique	72	Recherche-Développement Scientifique
		MC	Autres Activités Spécialisées, Scientifiques Et Techniques	73	Publicité Et Études De Marché
				74	Autres Activités Spécialisées, Scientifiques Et Techniques
				75	Activités Vétérinaires
N	Activités De Services Administratifs Et De Soutien	NZ	Activités De Services Administratifs Et De Soutien	77	Activités De Location Et Location-Bail
				78	Activités Liées À L'Emploi
				79	Activités Des Agences De Voyage, Voyagistes, Services De Réservation Et Activités Connexes
				80	Enquêtes Et Sécurité
				81	Services Relatifs Aux Bâtiments Et Aménagement Paysager
				82	Activités Administratives Et Autres Activités De Soutien Aux Entreprises
O	Administration Publique	OZ	Administration Publique	84	Administration Publique Et Défense; Sécurité Sociale Obligatoire
P	Enseignement	PZ	Enseignement	85	Enseignement
Q	Santé Humaine Et Action Sociale	QA	Activités Pour La Santé Humaine	86	Activités Pour La Santé Humaine
		QB	Hébergement Médico-Social Et Social Et Action Sociale Sans Hébergement	87	Hébergement Médico-Social Et Social

A21	Intitulé	A38	Intitulé	A88	Intitulé
R	Arts, Spectacles Et Activités Ré- créatives	RZ	Arts, Spectacles Et Activités Ré- créatives	88	Action Sociale Sans Hébergement
				90	Activités Créatives, Artistiques Et De Spectacle
				91	Bibliothèques, Archives, Musées Et Autres Activités Culturelles
				92	Organisation De Jeux De Hasard Et D'Argent
				93	Activités Sportives, Récréatives Et De Loisirs
S	Autres Activités De Services	SZ	Autres Activités De Services	94	Activités Des Organisations Associa- tives
				95	Réparation D'Ordinateurs Et De Biens Personnels Et Domestiques
				96	Autres Services Personnels
T	Activités Des Ménages En Tant Qu'Employeurs; Activités Indifférenciées Des Ménages En Tant Que Producteurs De Biens Et Services Pour Usage Propre	TZ	Activités Des Ménages En Tant Qu'Employeurs; Activités Indifférenciées Des Ménages En Tant Que Producteurs De Biens Et Services Pour Usage Propre	97	Activités Des Ménages En Tant Qu'Em- ployeurs De Personnel Domestique
				98	Activités Indifférenciées Des Ménages En Tant Que Producteurs De Biens Et Services Pour Usage Propre
U	Activités Extra-Territoriales	UZ	Activités Extra-Territoriales	99	Activités Des Organisations Et Orga- nismes Extraterritoriaux

Annexe B - Statistiques descriptives

Table 6 - Statistiques descriptives des principales variables d'intérêt

Variable	Min	Q1	Médiane	Q3	Max	Moyenne	SD
Échantillon zone d'emploi-année							
<i>Économies d'urbanisation</i>							
Densité en emploi	3.36	12.37	22.41	42.04	5060.51	59.87	287.64
Potentiel de marché	27.91	48.07	55.86	73.25	302.51	68.6	38.95
Diversité sectorielle	8.08	17.36	19.38	21.27	27.56	19.21	3.2
Échantillon zone d'emploi-secteur-année							
<i>Économies de localisation</i>							
Spécialisation sectorielle	0	0	0	0.01	0.38	0.01	0.02
Échantillon établissement-année							
<i>Économies d'urbanisation</i>							
Densité en emploi	3.5	27.19	64.55	268.79	5060.51	770.32	1641.25
Potentiel de marché	28.19	49.15	57.33	107.51	302.51	80.14	51.42
Diversité sectorielle	8.34	18.58	21.11	23.87	27.56	21.11	3.55
<i>Économies de localisation</i>							
Spécialisation sectorielle	0	0.01	0.03	0.05	0.31	0.03	0.03
<i>Caractéristiques individuelles</i>							
EBE	-91885	21.36	97.78	265.31	37116.09	222.91	653.83
$\frac{EBE}{EQTP_{t-1}}$	-549.06	1.89	7.88	17.98	486.08	12.11	20.32
VACF	-46122.34	462.47	785.96	1381.96	59158.4	1214.05	1569.77
$\frac{VACF}{EQTP_{t-1}}$	-327.06	41.44	55.21	76.56	1608.57	65.84	49.43
Charges de personnel	0.77	402.27	661.95	1126.94	58636.43	988.49	1210.77
$\frac{\text{Charges de personnel}}{EQTP_{t-1}}$	0.09	36.58	46.21	59.61	1516.26	53.18	40.31
Part emplois très qualifiés	0	0	0.05	0.16	1	0.14	0.23
$EQTP_{t-1}$	0.03	9	13.66	22.34	602.32	19.44	20.78
Age	0	9	17	29	122	21.13	17.48
Subventions d'exploitation	-882.47	0	0	5.19	4881.14	12.09	66.58
Entreprise profilée	Inf	NA	NA	NA	-Inf	NA	NA
Groupe franco-français	0	0	1	1	1	0.56	0.5
Multinationale sous contrôle français	0	0	0	0	1	0.04	0.2
Multinationale sous contrôle étranger	0	0	0	0	1	0.04	0.19
Pas de groupe, nationalité étrangère	0	0	0	0	1	0	0.05
Pas de groupe, nationalité française	0	0	0	1	1	0.36	0.48

Les valeurs sont arrondies à 2 décimales près. Le champ de l'échantillon établissement-année est le même que celui décrit en section 3 avec le filtre à valeurs aberrantes appliqué sur l'EBE par EQTP, sauf pour les statistiques portant sur la VACF et les charges de personnel, pour lesquelles le filtre est respectivement appliqué à la VACF par EQTP et aux charges personnel par EQTP. La nomenclature sectorielle utilisée est la NAF rév.2 en 88 niveaux.

Annexe C - Tests de robustesse

Table 7 - Effets des variables d'agglomération sur l'EBE par EQTP - Robustesse à l'inclusion des unités légales pluri-établissement

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					1.449*** (0.135)	1.449*** (0.135)	1.472*** (0.128)	1.472*** (0.128)
log(EQTP _{t-1})							1.039*** (0.073)	1.039*** (0.073)
log(Age)							0.520** (0.211)	0.520** (0.211)
Subventions d'exploitation							0.133 (0.181)	0.133 (0.181)
Groupe franco-français							1.707*** (0.257)	1.707*** (0.257)
Multinationale sous contrôle français							0.449 (0.577)	0.449 (0.577)
Multinationale sous contrôle étranger							0.160 (0.652)	0.160 (0.652)
Pas de groupe, nationalité étrangère							2.291 (1.913)	2.291 (1.913)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	888124	888124	888124	888124	888032	888032	749654	749654
R2	0.221	0.221	0.221	0.221	0.222	0.222	0.227	0.227
R2 intra-groupes	NA	NA	NA	NA	0.002	0.002	0.004	0.004
R2 ajusté	0.157	0.157	0.157	0.157	0.159	0.159	0.165	0.165
R2 ajusté intra-groupes	NA	NA	NA	NA	0.002	0.002	0.004	0.004
Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	0.337** (0.156)	0.349** (0.159)	0.342*** (0.100)	0.349*** (0.127)	0.217** (0.095)	0.221* (0.129)	0.227** (0.100)	0.229 (0.140)
log(Potentiel de marché)	-0.213 (0.260)	-0.178 (0.245)	-0.221 (0.310)	-0.188 (0.288)	-0.250 (0.298)	-0.229 (0.283)	-0.344 (0.304)	-0.333 (0.296)
log(Diversité sectorielle)	0.414 (0.507)	0.229 (0.469)	0.780 (0.589)	0.599 (0.557)	0.582 (0.537)	0.467 (0.517)	0.504 (0.533)	0.438 (0.520)
log(Superficie)	-0.389** (0.196)	-0.301 (0.196)	-0.511*** (0.176)	-0.425** (0.181)	-0.542*** (0.173)	-0.487*** (0.180)	-0.559*** (0.180)	-0.528*** (0.188)
log(Spécialisation sectorielle)		0.641* (0.337)		0.594* (0.338)		0.378 (0.325)		0.215 (0.341)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	✓	✓	✓	✓	✓	✓
Observations	66755	66755	66755	66755	66720	66720	55756	55756
R2	0.511	0.512	0.770	0.770	0.765	0.765	0.772	0.772
R2 intra-groupes	0.007	0.008	0.533	0.532	0.530	0.530	0.541	0.541
R2 ajusté	0.509	0.509	0.769	0.769	0.764	0.764	0.770	0.770
R2 ajusté intra-groupes	0.006	0.008	0.532	0.532	0.530	0.530	0.541	0.541

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe B. Les sources sont décrites en section 3. Le champ est identique à celui décrit en section 3 à l'exception qu'il est désormais étendu aux unités légales pluri-établissement employeur. L'EBE est ventilé parmi les établissements au prorata de la rémunération brute puis rapporté aux EQTP de l'établissement l'année précédente. Les EQTP et la part des emplois très qualifiés sont les seules variables individuelles observées au niveau de l'établissement : l'âge est celui de l'entreprise économique et les subventions sont celles perçues par l'unité légale. Les R2 des colonnes (1) et (2) ne sont pas directement comparables avec ceux des colonnes suivantes, car pour ces dernières, la régression d'une variable dépendante débruitée sur des régresseurs eux-mêmes débruités (via l'instrumentation) tend mécaniquement à augmenter les R2.

L'objectif de ces tests de robustesse est de s'assurer que les effets significatifs établis dans les régressions principales le sont toujours et du même ordre face à des choix alternatifs, en particulier de champ statistique. En effet, on veut avant tout éviter les erreurs de type I – les faux positifs –, autrement dit conclure à l'existence d'un effet lorsqu'il n'y en a pas. On vérifiera donc surtout la cohérence des estimations du coefficient de la densité lorsque les variables dépendantes sont la VACF ou les charges de personnel par travailleur. Le fait d'observer ici un petit effet positif et significatif de la densité en emploi locale sur l'EBE par travailleur ne permet pas de conclure à l'existence d'un effet puisque ceux-ci ne sont significatifs ni dans les régressions principales ni en colonne (8). De plus, une partie de la variance expliquée par le modèle provient de la variance de la variable

de ventilation entre établissement; la part de l'établissement dans la rémunération brute totale de l'unité légale. Or, comme la table 3 le montre, la densité en emploi a un effet sur les charges de personnel. Il est donc probable qu'une partie de la significativité des coefficients proviennent en fait du pouvoir explicatif de la densité sur la variable de ventilation. On préférera donc conclure à l'absence d'effet au risque de commettre une erreur de type II. On peut toutefois raisonnablement conclure de ces analyses dans leur ensemble que la hausse des salaires n'excède pas celle de la productivité.

Table 8 - Effets des variables d'agglomération sur la VACF par EQTP - Robustesse à l'inclusion des unités légales pluri-établissement

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					8.724*** (0.423)	8.724*** (0.423)	8.694*** (0.422)	8.694*** (0.422)
log(EQTP _{t-1})							-6.201*** (0.516)	-6.201*** (0.516)
log(Age)							0.470*** (0.160)	0.470*** (0.160)
Subventions d'exploitation							-0.081 (0.231)	-0.081 (0.231)
Groupe franco-français							1.074** (0.418)	1.074** (0.418)
Multinationale sous contrôle français							1.272 (0.867)	1.272 (0.867)
Multinationale sous contrôle étranger							7.997*** (1.169)	7.997*** (1.169)
Pas de groupe, nationalité étrangère							16.532*** (2.949)	16.532*** (2.949)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	748094	748094	748094	748094	748022	748022	748017	748017
R2	0.415	0.415	0.415	0.415	0.429	0.429	0.439	0.439
R2 intra-groupes	NA	NA	NA	NA	0.023	0.023	0.040	0.040
R2 ajusté	0.368	0.368	0.368	0.368	0.383	0.383	0.394	0.394
R2 ajusté intra-groupes	NA	NA	NA	NA	0.023	0.023	0.040	0.040

Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	2.499*** (0.423)	2.539*** (0.388)	2.137*** (0.297)	2.160*** (0.274)	1.389*** (0.246)	1.396*** (0.243)	1.418*** (0.272)	1.436*** (0.263)
log(Potentiel de marché)	0.609 (0.709)	0.725 (0.645)	0.354 (0.827)	0.458 (0.750)	0.138 (0.693)	0.167 (0.647)	0.424 (0.695)	0.506 (0.616)
log(Diversité sectorielle)	0.894 (1.518)	0.286 (1.364)	3.441* (1.780)	2.863* (1.670)	2.308* (1.397)	2.144 (1.323)	2.151 (1.357)	1.695 (1.249)
log(Superficie)	-0.421 (0.466)	-0.131 (0.429)	-1.359*** (0.467)	-1.086** (0.500)	-1.539*** (0.417)	-1.462*** (0.464)	-1.575*** (0.411)	-1.359*** (0.450)
log(Spécialisation sectorielle)		2.090 (1.550)		1.888 (1.593)		0.534 (1.409)		1.489 (1.394)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	✓	✓	✓	✓	✓	✓
Observations	55758	55758	55758	55758	55724	55724	55724	55724
R2	0.677	0.678	0.878	0.878	0.862	0.862	0.862	0.862
R2 intra-groupes	0.057	0.061	0.645	0.645	0.628	0.628	0.628	0.628
R2 ajusté	0.675	0.677	0.878	0.878	0.861	0.861	0.861	0.861
R2 ajusté intra-groupes	0.057	0.061	0.645	0.645	0.628	0.627	0.628	0.628

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe B. Les sources sont décrites en section 3. Le champ est identique à celui décrit en section 3 à l'exception qu'il est désormais étendu aux unités légales pluri-établissement employeur. La VACF est ventilée parmi les établissements au prorata de la rémunération brute puis rapportée aux EQTP de l'établissement l'année précédente. Les EQTP et la part des emplois très qualifiés sont les seules variables individuelles observées au niveau de l'établissement : l'âge est celui de l'entreprise économique et les subventions sont celles perçues par l'unité légale. Les R2 des colonnes (1) et (2) ne sont pas directement comparables avec ceux des colonnes suivantes, car pour ces dernières, la régression d'une variable dépendante débruitée sur des régresseurs eux-mêmes débruités (via l'instrumentation) tend mécaniquement à augmenter les R2.

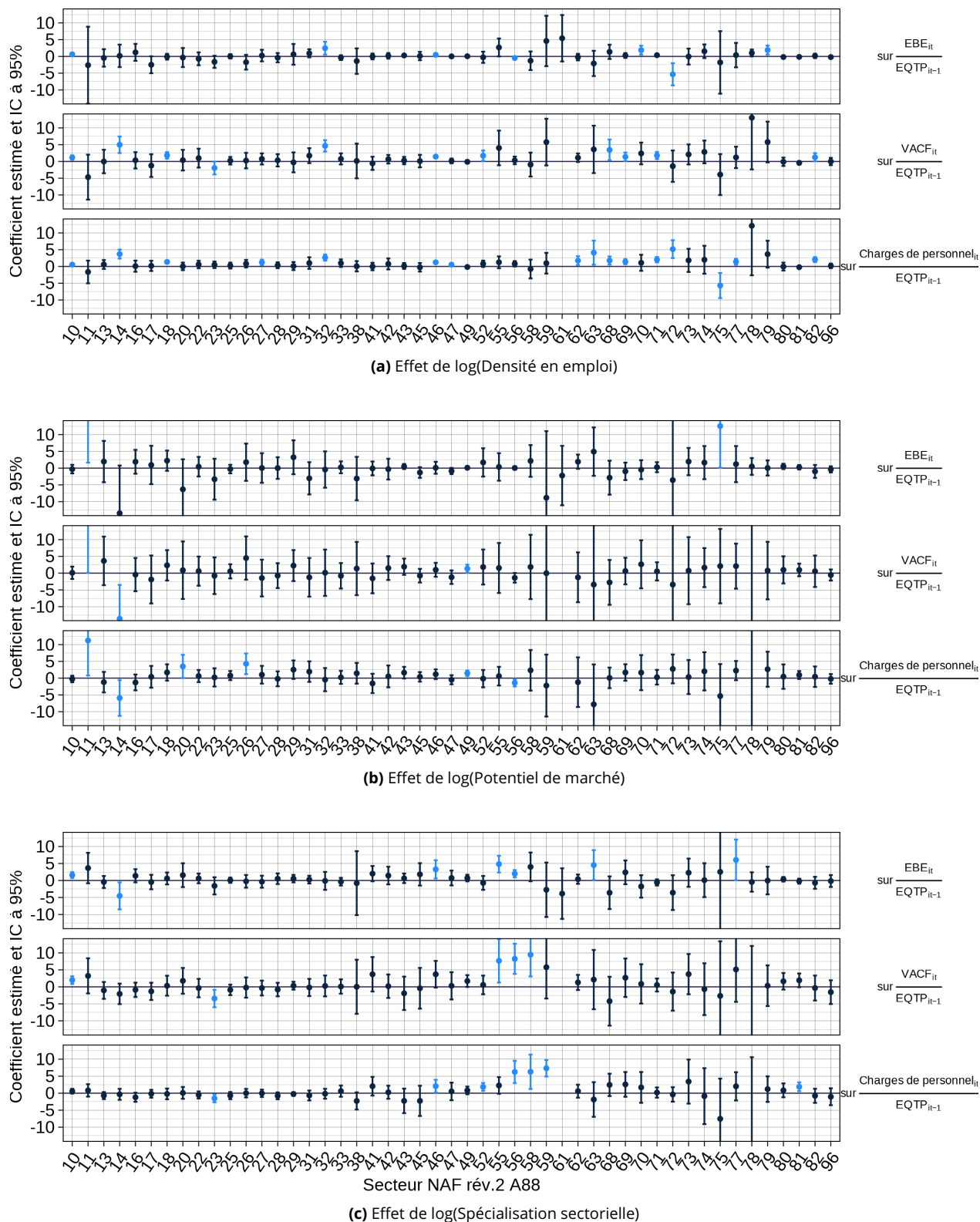
Table 9 - Effets des variables d'agglomération sur les charges de personnel par EQTP - Robustesse à l'inclusion des unités légales pluri-établissement

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					7.320*** (0.431)	7.320*** (0.431)	7.279*** (0.427)	7.279*** (0.427)
log(EQTP _{t-1})							-5.400*** (0.400)	-5.400*** (0.400)
log(Age)							-0.265 (0.205)	-0.265 (0.205)
Subventions d'exploitation							-0.319*** (0.061)	-0.319*** (0.061)
Groupe franco-français							-0.361 (0.380)	-0.361 (0.380)
Multinationale sous contrôle français							1.403** (0.570)	1.403** (0.570)
Multinationale sous contrôle étranger							7.293*** (0.941)	7.293*** (0.941)
Pas de groupe, nationalité étrangère							14.648*** (2.318)	14.648*** (2.318)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	748094	748094	748094	748094	748016	748016	748012	748012
R2	0.485	0.485	0.485	0.485	0.498	0.498	0.510	0.510
R2 intra-groupes	NA	NA	NA	NA	0.026	0.026	0.048	0.048
R2 ajusté	0.443	0.443	0.443	0.443	0.458	0.458	0.470	0.470
R2 ajusté intra-groupes	NA	NA	NA	NA	0.026	0.026	0.048	0.048

Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	2.240*** (0.335)	2.271*** (0.301)	1.927*** (0.245)	1.944*** (0.218)	1.314*** (0.186)	1.318*** (0.178)	1.311*** (0.211)	1.325*** (0.195)
log(Potentiel de marché)	0.897 (0.587)	0.983* (0.541)	0.541 (0.684)	0.619 (0.630)	0.381 (0.575)	0.396 (0.542)	0.631 (0.569)	0.692 (0.514)
log(Diversité sectorielle)	0.186 (1.198)	-0.259 (1.088)	2.178* (1.315)	1.764 (1.242)	1.327 (1.020)	1.246 (0.955)	1.203 (0.973)	0.881 (0.890)
log(Superficie)	-0.071 (0.376)	0.148 (0.365)	-0.921** (0.372)	-0.718* (0.434)	-1.069*** (0.342)	-1.029** (0.419)	-1.077*** (0.333)	-0.919** (0.403)
log(Spécialisation sectorielle)		1.547 (1.376)		1.378 (1.419)		0.271 (1.277)		1.072 (1.262)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	✓	✓	✓	✓	✓	✓
Observations	55534	55534	55534	55534	55499	55499	55499	55499
R2	0.694	0.695	0.885	0.885	0.871	0.871	0.872	0.872
R2 intra-groupes	0.058	0.061	0.646	0.646	0.631	0.631	0.631	0.631
R2 ajusté	0.692	0.693	0.884	0.884	0.870	0.870	0.871	0.871
R2 ajusté intra-groupes	0.058	0.061	0.646	0.646	0.631	0.631	0.631	0.631

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe 8. Les sources sont décrites en section 3. Le champ est identique à celui décrit en section 3 à l'exception qu'il est désormais étendu aux unités légales pluri-établissement employeur. Les charges de personnel sont ventilées parmi les établissements au prorata de la rémunération brute puis rapportées aux EQTP de l'établissement l'année précédente. Les EQTP et la part des emplois très qualifiés sont les seules variables individuelles observées au niveau de l'établissement : l'âge est celui de l'entreprise économique et les subventions sont celles perçues par l'unité légale. Les R2 des colonnes (1) et (2) ne sont pas directement comparables avec ceux des colonnes suivantes, car pour ces dernières, la régression d'une variable dépendante débruitée sur des régresseurs eux-mêmes débruités (via l'instrumentation) tend mécaniquement à augmenter les R2.

Figure 10 - Étude de l'hétérogénéité sectorielle des effets d'agglomération sur les performances économiques -
 Spécification de la colonne (8) – Robustesse à l'inclusion des unités légales pluri-établissement



Note : Les nomenclatures sectorielles sont disponibles en Annexe A. Les sources sont décrites en section 3. Le champ est identique à celui décrit en section 3 à l'exception qu'il est désormais étendu aux unités légales pluri-établissement employeur. Les coefficients significatifs au seuil de 5% ainsi que leur intervalle de confiance à 95% sont en bleu.

Table 10 - Effets des variables d'agglomération sur l'EBE par EQTP - Robustesse au recours à la NAF rév.2 en 38 niveaux

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					1.036*** (0.139)	1.036*** (0.139)	0.973*** (0.124)	0.973*** (0.124)
log(EQTP _{t-1})							-1.135*** (0.188)	-1.135*** (0.188)
log(Age)							1.526*** (0.200)	1.526*** (0.200)
Subventions d'exploitation							0.446*** (0.163)	0.446*** (0.163)
Groupe franco-français							0.260* (0.149)	0.260* (0.149)
Multinationale sous contrôle français							-0.525 (0.406)	-0.525 (0.406)
Multinationale sous contrôle étranger							2.290*** (0.501)	2.290*** (0.501)
Pas de groupe, nationalité étrangère							1.303* (0.776)	1.303* (0.776)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	674508	674508	674508	674508	674417	674417	673928	673928
R2	0.178	0.178	0.178	0.178	0.179	0.179	0.185	0.185
R2 intra-groupes	NA	NA	NA	NA	0.002	0.002	0.009	0.009
R2 ajusté	0.135	0.135	0.135	0.135	0.137	0.137	0.143	0.143
R2 ajusté intra-groupes	NA	NA	NA	NA	0.002	0.002	0.009	0.009
Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	0.082 (0.114)	0.081 (0.116)	0.094 (0.062)	0.091 (0.073)	-0.009 (0.070)	-0.009 (0.077)	0.025 (0.068)	0.023 (0.074)
log(Potentiel de marché)	-0.343 (0.210)	-0.334* (0.195)	-0.290 (0.251)	-0.282 (0.237)	-0.315 (0.250)	-0.317 (0.248)	-0.306 (0.257)	-0.301 (0.249)
log(Diversité sectorielle)	-1.323** (0.600)	-1.356** (0.599)	-1.204 (0.755)	-1.224 (0.764)	-1.271 (0.775)	-1.267 (0.781)	-1.246 (0.779)	-1.259 (0.784)
log(Superficie)	-0.478** (0.215)	-0.453** (0.208)	-0.590*** (0.214)	-0.570*** (0.211)	-0.607*** (0.211)	-0.611*** (0.210)	-0.559*** (0.205)	-0.547*** (0.204)
log(Spécialisation sectorielle)		0.207 (0.324)		0.154 (0.310)		-0.029 (0.339)		0.099 (0.309)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	✓	✓	✓	✓	✓	✓
Observations	33292	33292	33292	33292	33281	33281	33278	33278
R2	0.552	0.552	0.784	0.784	0.778	0.778	0.772	0.772
R2 intra-groupes	0.005	0.006	0.520	0.520	0.519	0.519	0.518	0.518
R2 ajusté	0.550	0.550	0.783	0.782	0.777	0.777	0.771	0.771
R2 ajusté intra-groupes	0.005	0.005	0.520	0.519	0.519	0.519	0.518	0.517

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe B. La modalité de référence de la nature du groupe est 'Pas de groupe, nationalité française'. Le champ et les sources sont décrits en section 3. Les R2 des colonnes (1) et (2) ne sont pas directement comparables avec ceux des colonnes suivantes, car pour ces dernières, la régression d'une variable dépendante débruitée sur des régresseurs eux-mêmes débruités (via l'instrumentation) tend mécaniquement à augmenter les R2.

Table 11 - Effets des variables d'agglomération sur la VACF par EQTP - Robustesse au recours à la NAF rév.2 en 38 niveaux

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					10.248*** (0.414)	10.248*** (0.414)	9.819*** (0.384)	9.819*** (0.384)
log(EQTP _{t-1})							-5.824*** (0.623)	-5.824*** (0.623)
log(Age)							2.911*** (0.305)	2.911*** (0.305)
Subventions d'exploitation							0.446* (0.267)	0.446* (0.267)
Groupe franco-français							-2.908*** (0.273)	-2.908*** (0.273)
Multinationale sous contrôle français							-0.476 (0.522)	-0.476 (0.522)
Multinationale sous contrôle étranger							10.562*** (1.102)	10.562*** (1.102)
Pas de groupe, nationalité étrangère							21.628*** (2.618)	21.628*** (2.618)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	674508	674508	674508	674508	674424	674424	674035	674035
R2	0.289	0.289	0.289	0.289	0.331	0.331	0.350	0.350
R2 intra-groupes	NA	NA	NA	NA	0.059	0.059	0.085	0.085
R2 ajusté	0.253	0.253	0.253	0.253	0.297	0.297	0.316	0.316
R2 ajusté intra-groupes	NA	NA	NA	NA	0.059	0.059	0.085	0.085

Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	2.227*** (0.276)	2.208*** (0.264)	2.169*** (0.314)	2.112*** (0.240)	1.157*** (0.188)	1.134*** (0.165)	1.174*** (0.208)	1.139*** (0.169)
log(Potentiel de marché)	0.888 (0.794)	1.026 (0.645)	0.371 (1.006)	0.532 (0.863)	0.112 (0.809)	0.179 (0.725)	0.161 (0.759)	0.260 (0.639)
log(Diversité sectorielle)	-2.108 (1.582)	-2.627* (1.449)	-1.230 (1.487)	-1.626 (1.606)	-1.959 (1.568)	-2.122 (1.600)	-1.471 (1.493)	-1.713 (1.544)
log(Superficie)	-0.568 (0.595)	-0.165 (0.546)	-1.323** (0.566)	-0.922* (0.521)	-1.472*** (0.510)	-1.306*** (0.504)	-1.332*** (0.497)	-1.086** (0.483)
log(Spécialisation sectorielle)		3.278*** (0.996)		3.103*** (1.020)		1.285 (0.799)		1.900** (0.751)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	✓	✓	✓	✓	✓	✓
Observations	33253	33253	33253	33253	33243	33243	33241	33241
R2	0.809	0.817	0.938	0.938	0.898	0.897	0.891	0.891
R2 intra-groupes	0.215	0.248	0.746	0.747	0.684	0.683	0.678	0.678
R2 ajusté	0.808	0.816	0.938	0.938	0.897	0.897	0.891	0.891
R2 ajusté intra-groupes	0.215	0.248	0.746	0.747	0.683	0.683	0.678	0.678

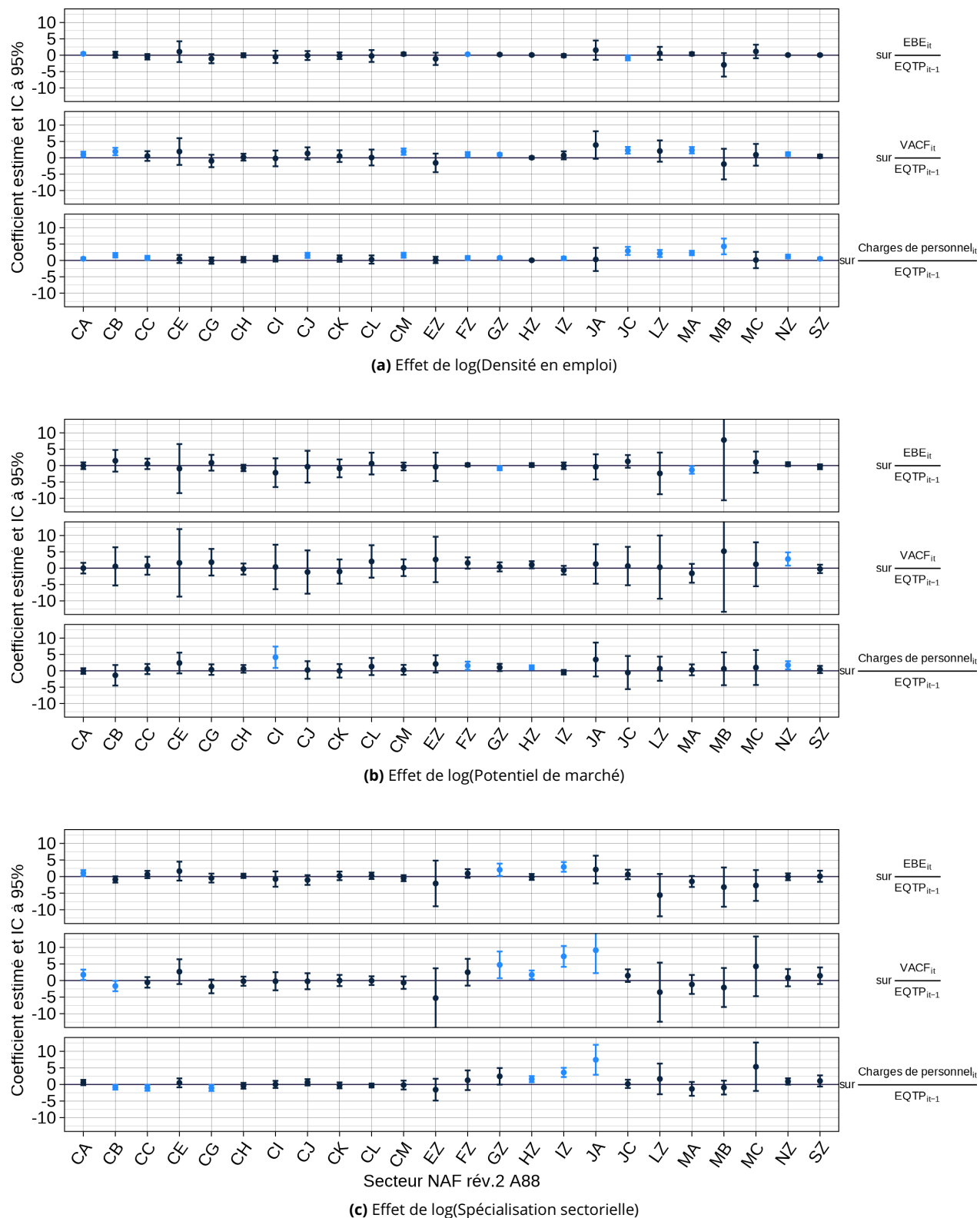
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe B. La modalité de référence de la nature du groupe est 'Pas de groupe, nationalité française'. Le champ et les sources sont décrits en section 3. Les R2 des colonnes (1) et (2) ne sont pas directement comparables avec ceux des colonnes suivantes, car pour ces dernières, la régression d'une variable dépendante débruitée sur des régresseurs eux-mêmes débruités (via l'instrumentation) tend mécaniquement à augmenter les R2.

Table 12 - Effets des variables d'agglomération sur les charges de personnel par EQTP - Robustesse au recours à la NAF
rév.2 en 38 niveaux

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					9.020*** (0.277)	9.020*** (0.277)	8.731*** (0.276)	8.731*** (0.276)
log(EQTP _{t-1})							-4.030*** (0.434)	-4.030*** (0.434)
log(Age)							1.028*** (0.107)	1.028*** (0.107)
Subventions d'exploitation							-0.106 (0.084)	-0.106 (0.084)
Groupe franco-français							-3.338*** (0.448)	-3.338*** (0.448)
Multinationale sous contrôle français							-0.034 (0.286)	-0.034 (0.286)
Multinationale sous contrôle étranger							7.619*** (0.505)	7.619*** (0.505)
Pas de groupe, nationalité étrangère							15.172*** (1.706)	15.172*** (1.706)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	674508	674508	674508	674508	674428	674428	673976	673976
R2	0.391	0.391	0.391	0.391	0.461	0.461	0.482	0.482
R2 intra-groupes	NA	NA	NA	NA	0.115	0.115	0.149	0.149
R2 ajusté	0.359	0.359	0.359	0.359	0.433	0.433	0.455	0.455
R2 ajusté intra-groupes	NA	NA	NA	NA	0.115	0.115	0.149	0.149
Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	2.070*** (0.225)	2.056*** (0.190)	2.023*** (0.309)	1.973*** (0.247)	1.159*** (0.160)	1.138*** (0.139)	1.131*** (0.175)	1.102*** (0.143)
log(Potentiel de marché)	1.166** (0.578)	1.288*** (0.467)	0.613 (0.717)	0.761 (0.621)	0.458 (0.560)	0.521 (0.496)	0.525 (0.512)	0.609 (0.426)
log(Diversité sectorielle)	-0.805 (1.192)	-1.236 (1.012)	-0.249 (0.884)	-0.561 (1.016)	-0.775 (0.882)	-0.910 (0.919)	-0.282 (0.832)	-0.460 (0.892)
log(Superficie)	-0.154 (0.396)	0.207 (0.355)	-0.752** (0.354)	-0.392 (0.306)	-0.897*** (0.303)	-0.741** (0.296)	-0.809*** (0.294)	-0.603** (0.280)
log(Spécialisation sectorielle)		2.869*** (0.817)		2.748*** (0.852)		1.185** (0.504)		1.567*** (0.493)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	✓	✓	✓	✓	✓	✓
Observations	33059	33059	33059	33059	33049	33049	33047	33047
R2	0.887	0.897	0.974	0.975	0.952	0.953	0.951	0.952
R2 intra-groupes	0.395	0.448	0.860	0.865	0.809	0.810	0.803	0.807
R2 ajusté	0.887	0.897	0.974	0.975	0.952	0.953	0.950	0.951
R2 ajusté intra-groupes	0.395	0.448	0.860	0.865	0.809	0.810	0.803	0.807

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe B. La modalité de référence de la nature du groupe est 'Pas de groupe, nationalité française'. Le champ et les sources sont décrits en section 3. Les R2 des colonnes (1) et (2) ne sont pas directement comparables avec ceux des colonnes suivantes, car pour ces dernières, la régression d'une variable dépendante débruitée sur des régresseurs eux-mêmes débruités (via l'instrumentation) tend mécaniquement à augmenter les R2.

Figure 11 - Étude de l'hétérogénéité sectorielle des effets d'agglomération sur les performances économiques -
 Spécification de la colonne (8) – Robustesse au recours à la NAF rév.2 en 38 niveaux



Note : Les nomenclatures sectorielles sont disponibles en Annexe A. Le champ et les sources sont décrits en section 3. Les coefficients significatifs au seuil de 5% ainsi que leur intervalle de confiance à 95% sont en bleu.

Table 13 - Effets des variables d'agglomération sur les variables dépendantes sous des formes alternatives

Étape 1 : Régression au niveau établissement-année						
	EBE	VACF	log(VACF)	log(VACF)	Charges de personnel	log(Charges de personnel)
Part emplois très qualifiés	19.261*** (1.491)	148.102*** (2.903)	0.120*** (0.004)	0.134*** (0.003)	135.570*** (3.105)	0.120*** (0.004)
log(EQTP _{t-1})				0.699*** (0.016)		0.149*** (0.003)
EQTP _{t-1}	5.127*** (0.226)	41.438*** (1.252)			38.251*** (0.885)	0.722*** (0.018)
log(Age)	27.109*** (1.564)	52.975*** (3.134)	0.194*** (0.003)	0.038*** (0.004)	17.251*** (2.121)	0.018*** (0.002)
Subventions d'exploitation	14.294*** (3.183)	12.680*** (4.018)	0.066*** (0.004)	0.012*** (0.003)	-0.268 (2.168)	0.003* (0.004)
Groupe franco-français	-4.201 (3.474)	-53.973*** (2.613)	-0.199*** (0.010)	-0.056*** (0.005)	-57.770*** (3.597)	-0.089*** (0.009)
Multinationale sous contrôle français	3.545 (6.471)	16.868* (9.051)	-0.066*** (0.017)	-0.033*** (0.008)	22.138*** (6.838)	-0.032*** (0.015)
Multinationale sous contrôle étranger	62.938*** (7.804)	243.996*** (17.931)	0.204*** (0.013)	0.091*** (0.007)	220.282*** (10.106)	0.111*** (0.006)
Pas de groupe, nationalité étrangère	9.716 (20.265)	431.542*** (52.055)	0.419*** (0.027)	0.190*** (0.024)	346.302*** (36.287)	0.188*** (0.020)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓
Observations	673747	673778	672722	672722	673728	673728
R2	0.334	0.631	0.340	0.664	0.751	0.763
R2 intra-groupes	0.083	0.487	0.098	0.540	0.642	0.660
R2 ajusté	0.267	0.594	0.275	0.630	0.726	0.739
R2 ajusté intra-groupes	0.083	0.487	0.098	0.540	0.642	0.660

Étape 2 : Régression au niveau zone-secteur-année						
	EBE	VACF	log(VACF)	log(VACF)	Charges de personnel	log(Charges de personnel)
log(Densité en emploi)	-0.785 (1.709)	19.617*** (3.614)	0.011** (0.005)	0.012*** (0.003)	20.792*** (2.878)	0.013*** (0.004)
log(Potentiel de marché)	1.616 (3.355)	6.501 (9.357)	0.026*** (0.006)	0.007 (0.010)	8.493 (7.895)	0.015* (0.009)
log(Diversité sectorielle)	5.749 (7.951)	39.552*** (18.392)	0.014 (0.015)	0.046*** (0.017)	31.468** (14.064)	0.045*** (0.017)
log(Superficie)	-5.817** (2.509)	-17.756*** (6.647)	-0.007 (0.006)	-0.023*** (0.007)	-11.344** (5.051)	-0.017*** (0.006)
log(Spécialisation sectorielle)	4.000 (5.345)	33.811*** (12.764)	0.096*** (0.011)	0.026*** (0.009)	27.762*** (10.658)	0.024*** (0.008)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	✓	✓	✓	✓	✓	✓
Observations	61058	60961	60894	60894	60644	60644
R2	0.687	0.755	0.812	0.827	0.826	0.875
R2 intra-groupes	0.425	0.537	0.601	0.597	0.641	0.625
R2 ajusté	0.685	0.754	0.811	0.826	0.825	0.875
R2 ajusté intra-groupes	0.425	0.537	0.600	0.597	0.641	0.625

*** p < 0.01; ** p < 0.05; * p < 0.1. Les étiquettes des colonnes réfèrent à la variable dépendante de la régression. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe B. La modalité de référence de la nature du groupe est "Pas de groupe, nationalité française". Le champ et les sources sont décrits en section 3. Le nombre d'observations peut varier du fait de la restriction à des valeurs strictement positives lorsque les variables dépendantes sont exprimées en log, et du fait du filtre à valeurs aberrantes qui conduit à exclure un nombre différent d'observations selon les variables dépendantes.

La variable dépendante en colonne (8) est intéressante : elle permet de s'assurer que la VACF et les charges de personnel croissent au même rythme en présence d'économies d'agglomération lorsque les effets d'agglomération sont non significatifs, et permet également de conclure que le taux de marge n'est pas impacté par les économies d'agglomération, puisque

$$\text{Charges de personnel} = \frac{\text{VACF}}{1 - \frac{\text{EBE}}{\text{VACF}}} = \frac{1}{1 - \text{Taux de marge}}$$

Table 14 - Effets des variables d'agglomération sur l'EBE par EQTP - Robustesse à l'inclusion de la Corse et des DROMs (hors Mayotte)

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					1.188*** (0.097)	1.188*** (0.097)	1.164*** (0.095)	1.164*** (0.095)
log(EQTP _{t-1})							-1.090*** (0.183)	-1.090*** (0.183)
log(Age)							1.188*** (0.176)	1.188*** (0.176)
Subventions d'exploitation							0.704*** (0.154)	0.704*** (0.154)
Groupe franco-français							0.541*** (0.163)	0.541*** (0.163)
Multinationale sous contrôle français							-0.149 (0.322)	-0.149 (0.322)
Multinationale sous contrôle étranger							1.925*** (0.429)	1.925*** (0.429)
Pas de groupe, nationalité étrangère							1.751* (0.998)	1.751* (0.998)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	697030	697030	697030	697030	696936	696936	696436	696436
R2	0.266	0.266	0.266	0.266	0.267	0.267	0.271	0.271
R2 intra-groupes	NA	NA	NA	NA	0.002	0.002	0.007	0.007
R2 ajusté	0.191	0.191	0.191	0.191	0.192	0.192	0.196	0.196
R2 ajusté intra-groupes	NA	NA	NA	NA	0.002	0.002	0.007	0.007

Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	0.122 (0.125)	0.131 (0.132)	0.139 (0.122)	0.147 (0.131)	0.030 (0.122)	0.034 (0.131)	0.068 (0.120)	0.075 (0.130)
log(Potentiel de marché)	-0.076 (0.179)	-0.070 (0.172)	0.094 (0.196)	0.087 (0.189)	0.038 (0.187)	0.035 (0.184)	0.025 (0.184)	0.019 (0.178)
log(Diversité sectorielle)	0.428 (0.398)	0.280 (0.390)	0.693* (0.405)	0.538 (0.392)	0.604 (0.390)	0.524 (0.384)	0.506 (0.406)	0.371 (0.395)
log(Superficie)	-0.293* (0.153)	-0.241 (0.159)	-0.224 (0.149)	-0.178 (0.155)	-0.268* (0.147)	-0.244 (0.152)	-0.257* (0.143)	-0.217 (0.149)
log(Spécialisation sectorielle)		0.406 (0.282)		0.384 (0.283)		0.198 (0.281)		0.334 (0.277)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	×	×	×	×	×	×
Observations	64588	64588	64588	64588	64548	64548	64535	64535
R2	0.548	0.548	0.548	0.549	0.542	0.543	0.529	0.530
R2 intra-groupes	0.003	0.004	0.004	0.005	0.002	0.003	0.003	0.003
R2 ajusté	0.545	0.546	0.546	0.546	0.540	0.540	0.527	0.527
R2 ajusté intra-groupes	0.003	0.004	0.004	0.005	0.002	0.002	0.003	0.003

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe B. La modalité de référence de la nature du groupe est 'Pas de groupe, nationalité française'. Les sources sont décrites en section 3. Le champ est identique à celui décrit en section 3 à l'exception qu'il est désormais étendu aux unités légales corses et dromiennes hors Mayotte.

Table 15 - Effets des variables d'agglomération sur la VACF par EQTP - Robustesse à l'inclusion de la Corse et des DROMs (hors Mayotte)

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					8.706*** (0.522)	8.706*** (0.522)	8.439*** (0.520)	8.439*** (0.520)
log(EQTP _{t-1})							-7.213*** (0.512)	-7.213*** (0.512)
log(Age)							2.404*** (0.291)	2.404*** (0.291)
Subventions d'exploitation							0.528** (0.259)	0.528** (0.259)
Groupe franco-français							-0.869* (0.499)	-0.869* (0.499)
Multinationale sous contrôle français							1.081 (0.672)	1.081 (0.672)
Multinationale sous contrôle étranger							11.989*** (0.900)	11.989*** (0.900)
Pas de groupe, nationalité étrangère							22.771*** (2.299)	22.771*** (2.299)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	697030	697030	697030	697030	696944	696944	696540	696540
R2	0.414	0.414	0.414	0.414	0.427	0.427	0.439	0.439
R2 intra-groupes	NA	NA	NA	NA	0.024	0.024	0.043	0.043
R2 ajusté	0.354	0.354	0.354	0.354	0.369	0.369	0.382	0.382
R2 ajusté intra-groupes	NA	NA	NA	NA	0.024	0.024	0.043	0.043

Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	2.157*** (0.331)	2.201*** (0.312)	2.269*** (0.370)	2.304*** (0.350)	1.469*** (0.308)	1.479*** (0.299)	1.473*** (0.308)	1.499*** (0.297)
log(Potentiel de marché)	1.026 (0.632)	1.057* (0.594)	1.374* (0.705)	1.342** (0.663)	0.966* (0.560)	0.957* (0.548)	1.094** (0.511)	1.070** (0.479)
log(Diversité sectorielle)	1.005 (1.533)	0.286 (1.445)	1.344 (1.454)	0.595 (1.364)	0.747 (1.276)	0.546 (1.213)	0.804 (1.296)	0.246 (1.195)
log(Superficie)	-0.023 (0.420)	0.233 (0.401)	0.024 (0.408)	0.244 (0.397)	-0.295 (0.360)	-0.236 (0.366)	-0.247 (0.355)	-0.083 (0.351)
log(Spécialisation sectorielle)		1.996 (1.322)		1.864 (1.355)		0.500 (1.190)		1.388 (1.167)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	×	×	×	×	×	×
Observations	64515	64515	64515	64515	64477	64477	64468	64468
R2	0.631	0.632	0.632	0.634	0.578	0.579	0.570	0.571
R2 intra-groupes	0.046	0.050	0.050	0.053	0.025	0.025	0.025	0.027
R2 ajusté	0.629	0.630	0.630	0.632	0.576	0.576	0.567	0.568
R2 ajusté intra-groupes	0.046	0.050	0.050	0.053	0.024	0.025	0.025	0.027

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe B. La modalité de référence de la nature du groupe est 'Pas de groupe, nationalité française'. Les sources sont décrites en section 3. Le champ est identique à celui décrit en section 3 à l'exception qu'il est désormais étendu aux unités légales corses et dromiennes hors Mayotte.

La significativité du potentiel de marché est à considérer avec précaution puisque cette variable fait beaucoup moins sens pour les zones d'emploi corses et dromiennes, dans la mesure où seules les densités des zones d'emploi alentour françaises sont considérées. Cette variable prend donc une valeur très faible pour ces zones d'emploi dont les unités légales ont par ailleurs des performances un peu inférieures relativement aux autres zones d'emploi. L'objectif de cette table est surtout de vérifier que la magnitude et significativité de la densité locale est inchangée afin d'éviter, là encore, les erreurs de type I.

Table 16 - Effets des variables d'agglomération sur les charges de personnel par EQTP - Robustesse à l'inclusion de la Corse et des DROMs (hors Mayotte)

Étape 1 : Régression au niveau établissement-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Part emplois très qualifiés					7.147*** (0.521)	7.147*** (0.521)	6.964*** (0.528)	6.964*** (0.528)
log(EQTP _{t-1})							-5.521*** (0.367)	-5.521*** (0.367)
log(Age)							0.938*** (0.118)	0.938*** (0.118)
Subventions d'exploitation							-0.243* (0.133)	-0.243* (0.133)
Groupe franco-français							-1.571** (0.660)	-1.571** (0.660)
Multinationale sous contrôle français							1.113** (0.551)	1.113** (0.551)
Multinationale sous contrôle étranger							9.630*** (0.536)	9.630*** (0.536)
Pas de groupe, nationalité étrangère							16.496*** (1.787)	16.496*** (1.787)
Effets fixes Zone-Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Observations	697030	697030	697030	697030	696947	696947	696490	696490
R2	0.477	0.477	0.477	0.477	0.491	0.491	0.502	0.502
R2 intra-groupes	NA	NA	NA	NA	0.026	0.026	0.047	0.047
R2 ajusté	0.424	0.424	0.424	0.424	0.439	0.439	0.451	0.451
R2 ajusté intra-groupes	NA	NA	NA	NA	0.026	0.026	0.047	0.047

Étape 2 : Régression au niveau zone-secteur-année								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(Densité en emploi)	1.943*** (0.287)	1.977*** (0.262)	2.039*** (0.324)	2.066*** (0.300)	1.406*** (0.267)	1.411*** (0.254)	1.365*** (0.268)	1.383*** (0.251)
log(Potentiel de marché)	1.275*** (0.495)	1.301*** (0.466)	1.437*** (0.532)	1.417*** (0.504)	1.141*** (0.414)	1.137*** (0.407)	1.286*** (0.371)	1.272*** (0.350)
log(Diversité sectorielle)	0.398 (1.146)	-0.124 (1.096)	0.441 (1.038)	-0.100 (0.970)	0.065 (0.913)	-0.048 (0.839)	0.244 (0.920)	-0.127 (0.820)
log(Superficie)	0.137 (0.316)	0.330 (0.303)	0.113 (0.308)	0.278 (0.308)	-0.132 (0.268)	-0.098 (0.290)	-0.090 (0.266)	0.023 (0.277)
log(Spécialisation sectorielle)		1.475 (1.168)		1.366 (1.203)		0.286 (1.065)		0.938 (1.047)
Effets fixes Secteur-Année	✓	✓	✓	✓	✓	✓	✓	✓
Caractéristiques géographiques	×	×	✓	✓	✓	✓	✓	✓
Densité et potentiel de marché instrumentés	×	×	×	×	×	×	×	×
Observations	64086	64086	64086	64086	64049	64049	64037	64037
R2	0.641	0.642	0.643	0.644	0.589	0.589	0.586	0.586
R2 intra-groupes	0.050	0.053	0.055	0.057	0.030	0.030	0.030	0.031
R2 ajusté	0.639	0.640	0.641	0.641	0.587	0.587	0.583	0.584
R2 ajusté intra-groupes	0.050	0.053	0.055	0.057	0.030	0.030	0.030	0.031

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Les erreurs standards sont reportées entre parenthèses. La matrice de variance-covariance des erreurs est supposée bloc-diagonale, chaque bloc correspondant à un couple établissement-année pour l'étape 1, et à un couple secteur-année pour l'étape 2. La part des emplois très qualifiés et les subventions d'exploitation pouvant être nulles, ces variables ne sont pas exprimées en log mais centrées et réduites. Les statistiques descriptives des variables sont disponibles en Annexe B. La modalité de référence de la nature du groupe est 'Pas de groupe, nationalité française'. Les sources sont décrites en section 3. Le champ est identique à celui décrit en section 3 à l'exception qu'il est désormais étendu aux unités légales corses et dromiennes hors Mayotte.

Série des Documents de Travail

« Action régionale »

H2011/01

Panorama des villes moyennes
Jean-Michel Floch et Bernard Morel

H2012/01

Les déterminants régionaux de l'innovation
Benoît Buisson, Lionel Doisneau, Claire Kubrak, Michelle Mongo et Corinne Autant-Bernard

H2012/02

Estimations locales du handicap dans l'enquête Handicap-Santé 2008
Josiane Le Guennec

H2012/03

Estimations régionales dans l'enquête nationale Transport et Déplacements 2007-2008
Josiane Le Guennec

H2012/04

Détection des disparités socio-économiques
L'apport de la statistique spatiale
Jean-Michel Floch

H2013/01

Concentration et spécialisation des activités économiques : des outils pour analyser les tissus productifs locaux
Claire Kubrak

H2014/01

La ségrégation spatiale dans les grandes unités urbaines de France métropolitaine : une approche par les revenus
Gaëlle Dabet Jean-Michel Floch

H2014/02

Rapport du groupe de travail DREES/INSEE/DRJSCS sur les indicateurs locaux de suivi du plan pluriannuel de lutte contre la pauvreté et pour l'inclusion sociale (PPLPIS)

H2016/01

Élaboration des Produits Intérieurs Bruts régionaux en base 2010 : Principes et méthodologie
Benoît Hurpeau

H2018/01

Structurel, résiduel, géographique : Principe et mise en oeuvre des approches comptable et économétrique
Claire Kubrak

H2018/02

Les dynamiques de la qualité de vie dans les territoires
Robert Reynard Pascal Vialette

H2018/03

Guide de sémiologie graphique

H2018/04

Guide de sémiologie cartographique

H2019/01

Les entreprises régionales Concepts, usages, mesure statistique
Lionel Doisneau

H2021/01

Des projections nationales aux projections locales
Laure Crusson Jérôme Fabre

H2021/02

Décomposition de l'évolution de l'emploi par catégorie d'entreprise
Hervé Bacheré Benoît Mirouse Zoé Brassier

2022-18

La grille communale de densité à 7 niveaux
Simon Beck, Marie-Pierre De Bellefon, Jocelyn Forest, Mathilde Gerardin, David Levy,

2023-01

Cas d'utilisation des indices de Theil pour mesurer les disparités spatiales selon les revenus à l'intercommunal
Mathilde Gerardin Julien Pramil

2023-11

La rétopolation en 2010 du zonage en aires d'attraction des villes de 2020
Simon Beck Olivier Pégaz-Blanc Adem Khamallah

2023-15

Utilisation d'une méthode de scoring pour des estimations locales à partir d'une enquête nationale : exemple avec l'enquête TIC-ménages
Aline Branche-Seigeot

2025-01

Une estimation du taux de vulnérabilité énergétique des ménages pour le logement
Simon Beck Kendal Masson Virginie Mora Simon Prusse

2026-01

Economies d'agglomération et rentabilité des PME françaises
Marine LESUR