QUELS SONT LES DETERMINANTS DES DISPARITES GEOGRAPHIQUES DE TAUX D'INVALIDITE EN BELGIQUE ?[1]

PAR | **OCTAVE DE BROUWER ET ILAN TOJEROW**
DULBEA, SBS – EM, Université libre de Bruxelles

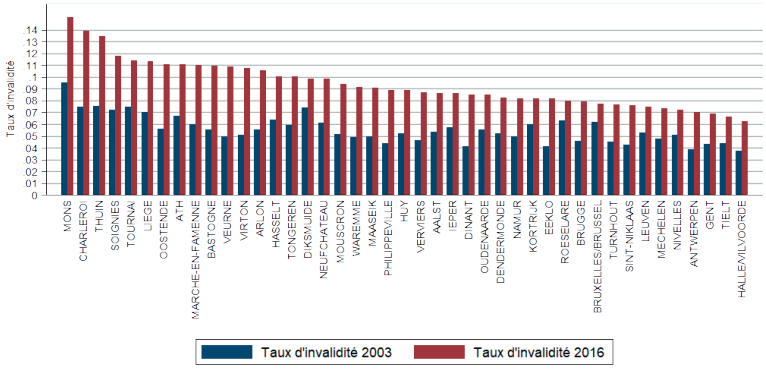
1. INTRODUCTION

Avec un taux d'invalidité[2] de 8,87 %, la Belgique a atteint un taux historiquement élevé en 2016. Son augmentation depuis 15 ans est un phénomène préoccupant tant en raison des coûts économiques associés que des questions sociétales qu'il soulève. Alors que de nombreuses mesures ont été récemment prises pour limiter ce phénomène, les causes de cette rapide augmentation sont encore mal connues. La présente étude propose de mettre en lumière une dimension particulièrement importante de l'invalidité en Belgique, à savoir son hétérogénéité géographique. En effet, comme le montre la figure 1, les disparités géographiques du taux d'invalidité sont élevées et ne montrent pas de trajectoires convergentes ces dix dernières années.

(1) Cet article a été réalisé avec le soutien financier du Service public fédéral Sécurité sociale, dans le cadre de la convention de recherche « Recherche sur les motifs des différences entre arrondissements en matière de reconnaissance de l'invalidité ».

(2) Tout au long de ce travail, le taux d'invalidité sera défini comme le nombre d'invalides divisé par le nombre de titulaires indemnisables duquel on a retiré les prépensionnés (TI-prép). En effet, les prépensionnés sont des titulaires indemnisables dont la probabilité de tomber en invalidité est nulle étant donné que ceux-ci bénéficient déjà d'un régime financier plus favorable.

FIGURE 1 : DISPERSION GEOGRAPHIQUE DU TAUX D'INVALIDITE 2003-2016 – SALARIES



Sources : données INAMI et calculs propres.

Dans cette perspective, il s'agira dans un premier temps de discerner les différents facteurs à l'origine des disparités géographiques du taux d'invalidité en Belgique. Une fois ces facteurs identifiés, il s'agira dans un deuxième temps d'évaluer leur capacité à expliquer la variation géographique du taux d'invalidité. Dans cette étude, nous nous concentrerons sur le régime des salariés et des demandeurs d'emploi, excluant donc les indépendants. En effet, bien que des disparités géographiques de taux d'invalidité existent chez les indépendants, celles-ci sont restées stables dans le temps. De plus, les taux d'invalidité chez les indépendants sont relativement faibles dans presque tous les arrondissements alors que de grandes disparités de niveaux et de trajectoires sont observées chez les salariés.

Dans cette étude, nous menons des analyses statistiques à deux niveaux. Premièrement, une analyse macroéconomique au niveau des communes est effectuée dans le but de déterminer les facteurs socio-économiques affectant le taux d'invalidité local. Deuxièmement, une analyse au niveau individuel est réalisée à partir d'un échantillon représentatif de la population belge, dans le but d'évaluer l'effet de résider dans un certain arrondissement sur la probabilité d'entrer en invalidité en contrôlant pour un grand nombre de caractéristiques individuelles.

Les deux analyses sont complémentaires dans le sens où l'analyse macroéconomique (au niveau communal) sera effectuée sur des stocks (ici le taux d'invalidité), tandis que l'analyse microéconomique (au niveau individuel) sera effectuée

sur des flux (ici, la probabilité individuelle d'entrer en invalidité). De plus, les méthodes économétriques utilisées dans chaque analyse seront substantiellement différentes. Dans l'analyse macroéconomique, nous utiliserons les méthodes de l'économétrie des données de panel (modèle « pooled OLS » et modèle à « effet fixe »), en analysant l'effet d'un ensemble de variables socio-économiques et d'un effet fixe d'arrondissement sur le taux d'invalidité communal. Premièrement, cette analyse permettra de déterminer quelles caractéristiques socio-économiques affectent le taux d'invalidité local. Deuxièmement, nous étudierons comment les effets d'arrondissement évoluent lorsqu'on inclut progressivement les variables explicatives dans le modèle. De cette manière, nous pourrons mesurer la part de l'hétérogénéité des effets d'arrondissement qui est due à une composition différente des populations en termes de caractéristiques observables et la part de l'hétérogénéité restant inexpliquée par le modèle. Dans l'analyse microéconomique, un modèle de probabilité linéaire sera utilisé dans le but d'estimer la probabilité individuelle d'entrée en invalidité au cours d'un laps de temps de deux années. Cette probabilité sera estimée séparément pour chaque arrondissement. L'inclusion des caractéristiques individuelles dans le modèle permettra d'estimer la probabilité d'entrée en invalidité dans chaque arrondissement en éliminant les effets de composition de population (c'est-à-dire en assumant que les caractéristiques individuelles observables sont réparties de manière homogène entre tous les arrondissements).

Plusieurs résultats intéressants ressortent de cette étude. Premièrement, nous montrons que les disparités géographiques ne peuvent être expliquées de manière satisfaisante par des différences en termes de structure d'âge, de ratio hommes/femmes ou de ratio employés/ouvriers. Deuxièmement, il existe une corrélation conditionnelle négative entre l'évolution temporelle du taux de chômage et l'évolution temporelle du taux d'invalidité. En effet, toutes choses égales par ailleurs, une augmentation (diminution) d'un point de pourcentage du taux de chômage entraîne une diminution (augmentation) de 0,247 points de pourcentage du taux d'invalidité. Ce résultat, combiné avec l'observation que les personnes au chômage ont connu une forte augmentation de leur taux d'entrée en invalidité, indique qu'un transfert d'une partie des personnes au chômage vers le système d'invalidité a eu lieu au cours de la période couverte par cette étude. Troisièmement, au niveau individuel, la situation vis-à-vis du marché du travail (notamment le fait d'être au chômage, le nombre de trimestres précédemment passés au chômage ou encore le revenu annuel) représente autant la cause la plus importante de l'entrée en invalidité que le facteur observable le plus important pour expliquer les disparités géographiques de la probabilité d'entrer en invalidité. Enfin, bien que notre modèle puisse expliquer relativement bien les disparités géographiques de

la probabilité d'entrée en invalidité pour les arrondissements entre le premier et le troisième quartile, celui-ci ne peut pas correctement expliquer les valeurs extrêmes en haut en en bas de la distribution géographique. De plus, alors que notre modèle explique bien les disparités géographiques de taux d'entrée en invalidité pour les maladies excluant les troubles mentaux et les maladies du système locomoteur, celui-ci ne permet pas d'expliquer correctement les différences géographiques observées pour ces deux catégories de maladies. Ce résultat indique que d'autres facteurs non observés influencent substantiellement les taux d'entrée en invalidité pour ces deux catégories de maladies.

La structure de l'article est la suivante. Premièrement, nous présentons une description de l'état de l'invalidité en Belgique afin d'avoir une vision d'ensemble sur ce sujet. Deuxièmement, nous discutons les facteurs explicatifs du taux d'invalidité ayant déjà été étudiés dans la littérature scientifique. Troisièmement, nous décrivons les données de nos échantillons et établissons un certain nombre de faits relatifs à la dispersion géographique du taux d'invalidité. Quatrièmement, nous décrivons les résultats de l'analyse macroéconomique et microéconomique. Enfin, nous établissons nos conclusions.

2. ETAT DE L'INVALIDITE EN BELGIQUE

Avant d'entamer des analyses statistiques plus approfondies, il est important de cerner les aspects principaux de l'invalidité en Belgique. Pour cela, nous fournissons un certain nombre de statistiques descriptives construites à partir des données de l'Institut national d'assurance maladie-invalidité (INAMI).[3] Nous nous concentrons ici sur 4 dimensions, à savoir le type de maladie, l'âge, le sexe et le statut employé/ouvrier.

Deux catégories de maladies constituent la grande majorité du stock de personnes en invalidité, à savoir les troubles mentaux et les maladies du système locomoteur et du tissu conjonctif.[4] A elles seules, ces deux catégories de pathologies regroupaient en 2015, 65 % du total des personnes en invalidité et enregistraient également les taux d'augmentation les plus rapides sur la période 2006-2015, avec respectivement +73 % pour les troubles mentaux et +89 % pour les maladies du système locomoteur et du tissu conjonctif. De manière

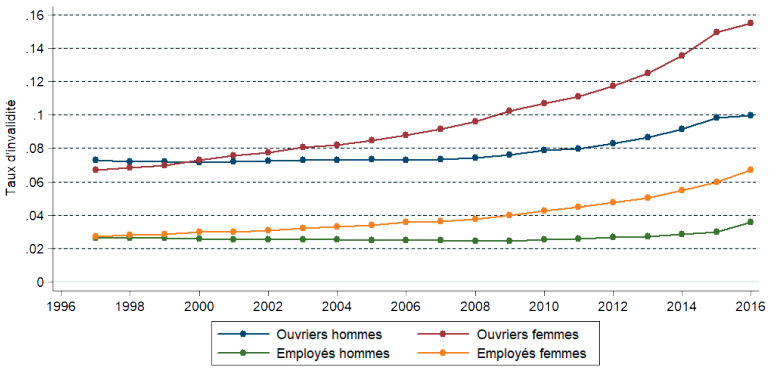
(3) Notons que d'autres analyses descriptives ont également été menées dans les rapports annuels de l'INAMI et par la Banque nationale de Belgique (BNB) (Saks, 2017).

(4) L'INAMI classe les personnes en invalidité selon 17 catégories de pathologies.

intéressante, ces deux catégories de pathologies sont également prépondérantes dans la plupart des autres pays de l'Organisation de Coopération et de Développement Economiques (OCDE) (OCDE, 2010).

Ensuite, comme le montre la figure 2, l'augmentation du taux d'invalidité ces quinze dernières années est en grande majorité le résultat d'une augmentation des taux d'invalidité des femmes et en particulier des femmes ayant le statut d'ouvrières.[5] La catégorie ayant connu l'augmentation la plus faible est celle des employés hommes, bien qu'une légère augmentation soit observée depuis 2008.

FIGURE 2 : EVOLUTION DU TAUX D'INVALIDITE PAR SEXE ET STATUT SOCIAL



Sources : données INAMI et calculs propres. Notes : le taux d'invalidité est défini comme le nombre d'invalides divisé par le nombre de titulaires indemnisables duquel on a retiré les prépensionnés (TI-prép).

Afin de calculer la part de l'augmentation du taux d'invalidité total étant associée à l'augmentation du taux d'invalidité des ouvrières, nous avons utilisé la méthode du « shift-share ». Cette méthode permet de décomposer l'évolution du taux d'invalidité total en un effet intra-catégories (évolution du taux d'invalidité à

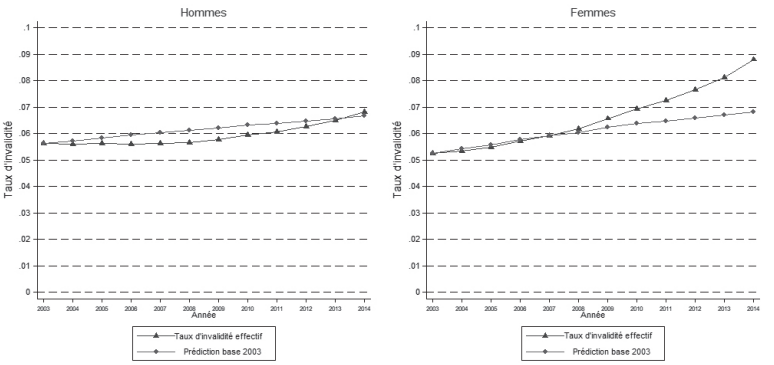
(5) Notons que cette classification a été établie par l'INAMI et ne distingue pas les demandeurs d'emplois des personnes en emploi. Les personnes au chômage sont donc classées en fonction de la dernière profession occupée avant le chômage.

l'intérieur de chaque catégorie de travailleurs) et inter-catégories (évolution de la part relative de chaque catégorie de travailleurs dans la population active).[6] Cette décomposition montre que 80 % de l'augmentation du taux d'invalidité entre 2003 et 2014 est due à l'augmentation du taux d'invalidité chez les femmes, dont 50 % pour les ouvrières et 30 % pour les employées. En d'autres mots, ce n'est pas l'augmentation de la proportion de femmes (ayant des taux d'invalidité en moyenne plus élevés) qui a provoqué l'augmentation du taux d'invalidité national, mais bien la forte augmentation du taux d'invalidité chez celles-ci et en particulier chez les ouvrières.

Une composante importante du risque d'entrée en invalidité est bien évidemment l'âge. D'une part, l'âge est un déterminant important du risque de rencontrer des soucis de santé et d'autre part, l'âge exacerbe les disparités du niveau de santé entre différentes populations ayant eu des conditions de vie différentes. Afin d'étudier le rôle du vieillissement de la population sur le taux d'invalidité pour différentes catégories de travailleurs, nous avons eu recours à un exercice, consistant à comparer l'évolution du taux d'invalidité effectif avec le taux d'invalidité qui aurait été observé si les taux d'invalidité à l'intérieur de chaque catégorie d'âge étaient restés constants. En d'autres termes, nous construisons l'évolution hypothétique du taux d'invalidité dans le cas où seule la composition de la structure d'âge aurait changé au cours de la période étudiée. La figure 3 montre le résultat de cet exercice, en prenant l'année 2003 comme référence. Nous observons à nouveau d'importantes différences entre les hommes et les femmes. En effet, chez les hommes, nous voyons que le taux d'invalidité est presque identique en 2014 au taux hypothétique construit sur base de l'année 2003. Un simple calcul nous montre que l'évolution de la structure d'âge peut expliquer 88,2 % de l'augmentation du taux d'invalidité sur la période considérée chez les hommes. En d'autres termes, le vieillissement démographique constitue l'essentiel de l'augmentation du taux d'invalidité chez les hommes. Chez les femmes en revanche, nous voyons qu'un décrochage du taux d'invalidité effectif a eu lieu entre 2007 et 2008. Nos calculs indiquent cette fois que seul 43,4 % de l'augmentation du taux d'invalidité peut être expliqué par l'évolution de la structure d'âge. En d'autres termes, le vieillissement de la population ne constitue qu'une explication partielle de l'augmentation du taux d'invalidité chez les femmes.

(6) Nous détaillons les résultats obtenus par cette méthode en annexe (Annexe 1).

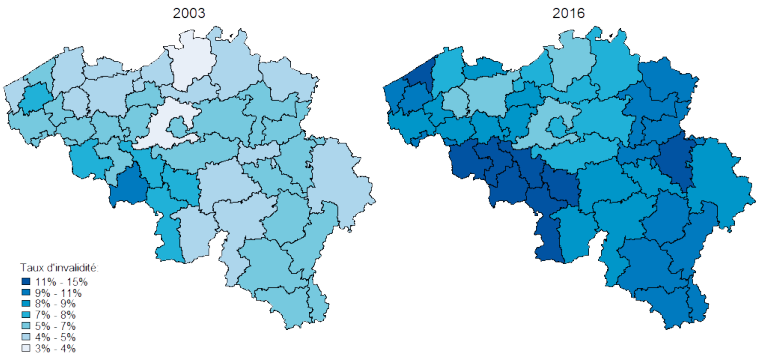
FIGURE 3 : COMPARAISON DU TAUX D'INVALIDITE TAUX D'INVALIDITE EFFECTIF ET PREDIT SUR BASE DE L'ANNEE 2003



Sources : données INAMI et calculs propres.

Afin d'avoir une idée visuelle des disparités géographiques de taux d'invalidité en Belgique, la figure 4 présente une cartographie des taux d'invalidité en Belgique en 2003 et 2016. Nous observons sur cette carte que même si des disparités géographiques existaient déjà en 2003, celles-ci se sont accentuées durant la période 2003-2016. Les taux d'invalidité les plus faibles sont observés dans les arrondissements d'Hal-Vilvoorde, de Tielt et de Gand. Sur la période 2003-2016, ces arrondissements ont de plus connu une augmentation limitée du taux d'invalidité. Les arrondissements du Hainaut, de Liège et dans une moindre mesure d'Ostende enregistrent les taux d'invalidité les plus élevés. La figure 1 montre en particulier que les arrondissements de Mons (15 %), Charleroi (13,9 %) et Thuin (13,5 %) possèdent des taux d'invalidité largement au-dessus de la moyenne nationale.

FIGURE 4 : CARTOGRAPHIE DU TAUX D'INVALIDITE EN 2003 ET 2016



Sources : données INAMI et calculs propres.

Nous analysons à présent comment les facteurs étudiés précédemment (âge, statut social, ratio hommes/femmes) influencent les disparités de taux d'invalidité entre les arrondissements. Pour cela, nous procédons d'une part à une analyse de corrélations entre les taux d'invalidité par arrondissements pour les différents groupes de travailleurs (par sexe-statut social et type de maladie) et d'autre part à une standardisation du taux d'invalidité par âge, sexe et statut social.[7]

La figure 5 montre que les taux d'invalidité par arrondissement dans les différentes catégories de travailleurs (par sexe et statut social) sont fortement corrélés entre eux. Seule la relation entre les employés hommes et les ouvrières semble plus faible (r=0,28), indiquant que certains arrondissements peuvent avoir des taux d'invalidité élevés chez les premiers et faibles chez les secondes. La corrélation élevée dans les autres catégories indique qu'il existe un effet d'arrondissement commun aux différentes catégories de travailleurs.

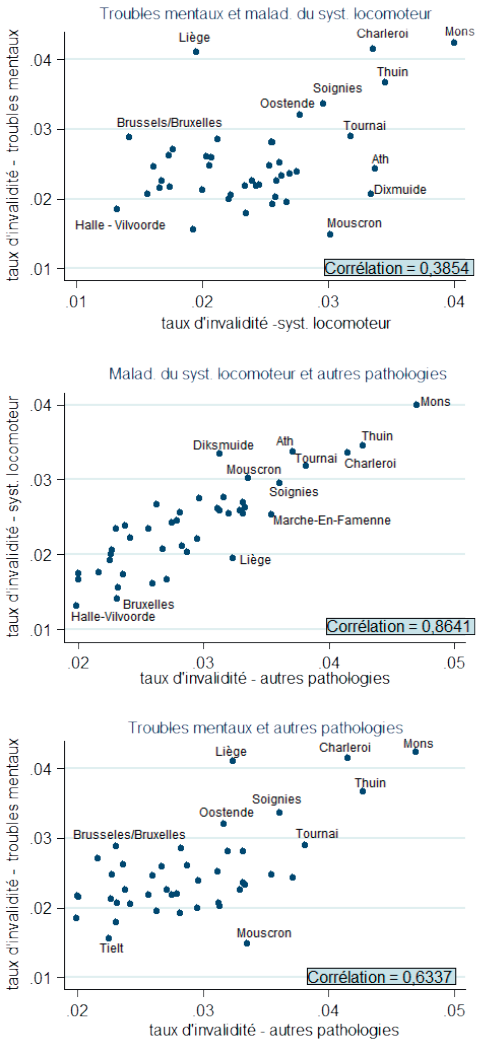
(7) La standardisation du taux d'invalidité consiste à comparer les taux d'invalidité entre les arrondissements en faisant l'hypothèse que ceux-ci contiennent une proportion égale de travailleurs de chaque groupe. Ici, la standardisation consiste donc à considérer des proportions égales en termes de catégories d'âge, de sexe et de statut social.

FIGURE 5 : CORRELOGRAMME DES TAUX D'INVALIDITE PAR SEXE ET TYPE DE CONTRAT

Sources : données INAMI et calculs propres.

En effectuant le même exercice de corrélations, pour les différentes catégories de maladies, nous obtenons des résultats similaires, à savoir que les taux d'invalidité locaux pour différentes catégories de maladies sont fortement corrélés entre eux. Seules quelques observations semblent se distinguer dans la catégorie des troubles mentaux, à savoir Liège et (dans une moindre mesure) Bruxelles, dont les taux d'invalidité dans la catégorie des troubles mentaux sont particulièrement élevés comparativement aux autres catégories de maladies et Mouscron, pour lequel celui-ci est particulièrement faible.

FIGURE 6 : CORRELATIONS DES TAUX D'INVALIDITE POUR DIFFERENTES CATEGORIES DE MALADIES



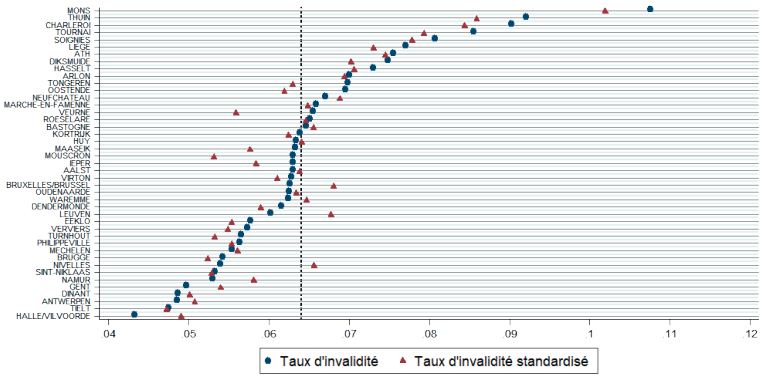
Sources : données INAMI et calculs propres.

La figure 7 montre la variation inter-arrondissement des taux d'invalidité observés et standardisés.[8] Pour ne pas nous concentrer sur une année en particulier, nous avons représenté ici la moyenne des taux effectifs et standardisés sur la période 2005-2013. La ligne verticale en pointillés représente la moyenne générale. Ce graphique montre une hétérogénéité dans la manière dont les arrondissements réagissent à la standardisation. En effet, pour certains arrondissements (Nivelles, Louvain et Bruxelles), le taux d'invalidité standardisé est largement au-dessus du taux d'invalidité effectif. Pour d'autres arrondissements (Ypres, Mouscron, Maaseik, Furnes, Ostende et Tongres), on observe des taux d'invalidité bien en-dessous des taux effectifs. Pour ces arrondissements, la composition de la population active en termes de structure d'âge, de ratios hommes/femmes et employés/ouvriers joue donc un rôle important sur le taux d'invalidité local. Cependant, pour la grande majorité des arrondissements, la standardisation ne réduit que partiellement les écarts par rapport à la moyenne.[9] Cela montre que pour la majorité des arrondissements, les différences en termes de structure d'âge, de ratios hommes/femmes et employés/ouvriers n'expliquent qu'une faible part de leur écart par rapport à la moyenne nationale. D'autres facteurs restent donc à être identifiés.

(8) De manière plus précise, la méthode de standardisation directe consiste à supprimer l'effet des différences entre arrondissements en termes de structure d'âge de la population. Pour cela, un taux d'invalidité hypothétique est calculé pour chaque arrondissement, en supposant que la structure d'âge, le nombre d'hommes et femmes ainsi que le nombre d'employés et ouvriers soient les mêmes dans tous les arrondissements. Formellement, nous avons : $Taux.inv_{standard_k} = \frac{\sum_{i=1}^{n} taux_{inv_{ik}} * \overline{TI}_i}{\overline{TI}}$ où i représente la classe d'âge et la catégorie de travailleurs (par ex. ouvrier homme 44-40 ans), k représente l'arrondissement et $\overline{TI}_i$ représente le nombre moyen de titulaires indemnisables du groupe par arrondissement en Belgique. Autrement dit, cette formule suppose que chaque arrondissement possède un nombre similaire de titulaires indemnisables dans chaque sexe, statut social et tranche d'âge.

(9) Par exemple, le taux d'invalidité de Mons passe de 10,8 % à 10,2 %, laissant l'écart par rapport à la moyenne tout aussi important.

FIGURE 7 : TAUX D'INVALIDITE EFFECTIF ET STANDARDISE PAR AGE, SEXE ET STATUT SOCIAL



En résumé, cette partie descriptive a permis de montrer plusieurs aspects importants de l'invalidité en Belgique. Premièrement, une tendance à la hausse existe dans tous les arrondissements. Cette tendance est surtout prononcée chez les femmes et en particulier les femmes ayant le statut d'ouvrières. De plus, alors que le vieillissement de population explique 88 % de l'augmentation du taux d'invalidité chez les hommes entre 2003 et 2014, celui-ci n'explique que 43 % de l'augmentation chez les femmes. Ensuite, parmi les différentes maladies invalidantes, les troubles mentaux et les maladies du système locomoteur sont les plus largement représentés, constituant 65 % du total des invalidités. Concernant la dimension géographique, nous avons observé que des taux d'invalidité élevés sont concentrés dans certaines zones géographiques, à savoir le Hainaut et dans une moindre mesure, Liège, le Limbourg et la Flandre Occidentale. Nous avons ensuite observé que les taux d'invalidité des différentes catégories de travailleurs dans chaque arrondissement sont fortement corrélés entre eux (à l'exception des employés hommes et ouvrières). Les taux d'invalidité par type de maladie sont également corrélés entre eux, à certaines exceptions près. Enfin, en appliquant la méthode de standardisation, nous avons observé que la composition de la population active en termes d'âge, de ratio hommes/femmes et de statut employé/ouvrier joue un rôle important pour une minorité d'arrondissements. Pour la grande majorité des arrondissements, cet effet de composition ne joue qu'un rôle mineur. D'autres facteurs à identifier jouent donc un rôle important sur le taux d'invalidité local.

3. REVUE DE LA LITTERATURE

Avant de décrire nos analyses économétriques, nous discutons dans cette section des connaissances scientifiques actuelles sur les déterminants de l'invalidité. L'objectif est ici de fournir un cadre pour comprendre comment des disparités géographiques de taux d'invalidité peuvent survenir et ainsi d'identifier des variables à inclure dans nos analyses économétriques. Notons que le nombre d'études ayant utilisé des données belges étant très faible, nous nous concentrerons donc sur des études menées dans d'autres pays, en particulier aux Etats-Unis, en Norvège, en Suède et aux Pays-Bas.

Nous distinguons ici quatre grandes catégories de facteurs pouvant expliquer l'évolution du taux d'invalidité sur un territoire donné. Ceux-ci sont les suivants : la structure démographique locale, le marché du travail local, le tissu social local et l'organisation du système de santé local. Ces quatre catégories regroupent bien évidemment une multiplicité de facteurs qui interagissent les uns avec les autres. Notons également que la mobilité géographique des personnes constitue un cinquième facteur dynamique, agissant à travers le regroupement géographique de personnes ayant des niveaux de santé similaires. Enfin, nous discutons comment certaines études ont tenté de quantifier le rôle de différents effets incitatifs dans le processus menant à l'invalidité, ou encore d'estimer la capacité de travail restante des personnes en invalidité.

3.1. STRUCTURE DEMOGRAPHIQUE LOCALE

Dans la section précédente, nous avons observé que l'évolution récente de la structure démographique ne peut expliquer que partiellement l'évolution du taux d'invalidité en Belgique et encore moins son hétérogénéité entre arrondissements. Plusieurs études ont obtenu des résultats similaires pour d'autres pays. Aux Etats-Unis, Duggan et Imberman (2008) montrent que l'évolution de la structure d'âge entre 1984 et 2003 ne peut expliquer que 6 % de l'augmentation du taux d'invalidité, dont 15,5 % pour les hommes et seulement 3,6 % pour les femmes. Des résultats similaires ont été obtenus pour de nombreux pays de l'OCDE (OCDE, 2010). Une autre explication souvent évoquée pour expliquer l'évolution du nombre d'invalides est l'augmentation de la participation des femmes sur le marché du travail, ayant entraîné une augmentation du nombre d'assurées et ainsi du nombre d'invalides. Notons que cette explication est valable uniquement pour expliquer l'augmentation du nombre d'invalides et non l'augmentation du taux d'invalidité, celle-ci ayant eu lieu en grande majorité à l'intérieur de chaque catégorie de travailleurs (par sexe-catégorie sociale). Enfin, notons que si la structure démographique ne peut

expliquer qu'une faible part des disparités géographiques de taux d'invalidité, une possibilité est que celle-ci interagisse avec d'autres variables. Par exemple, dans nos données, nous observons que l'écart de taux d'invalidité entre les deux arrondissements aux extrémités de la distribution géographique du taux d'invalidité (Max.-Min.) était de 2,9 points de pourcentage pour les 30-34 ans et de 14,7 points de pourcentage pour les 55-59 ans en 2013. Autrement dit, le vieillissement de population est susceptible d'exacerber les disparités géographiques de taux d'invalidité en augmentant la proportion de personnes âgées, pour lesquelles les disparités géographiques de taux d'invalidité sont les plus importantes.

3.2. MARCHE DU TRAVAIL LOCAL

Le marché du travail local peut influencer la santé et le risque d'invalidité à travers de nombreux mécanismes. Nous nous contenterons donc de discuter des résultats d'un certain nombre d'études ayant tenté de mesurer (1) les effets du cycle économique et (2) l'effet des conditions de travail sur le risque d'invalidité. Une régularité observée dans tous les pays étudiés est que le taux d'invalidité est négativement corrélé avec le revenu local par habitant (OCDE, 2010). Cette corrélation est cependant difficilement interprétable en raison du nombre d'explications potentielles (causalité inverse, tendances pré-existantes, accès local aux soins de santé, etc…). Afin d'isoler la relation causale entre la conjoncture économique locale et le taux d'invalidité local, plusieurs études ont donc utilisé des méthodes de variables instrumentales, consistant à exploiter des variations exogènes du revenu local par habitant et à mesurer leur effet sur le taux d'invalidité local (Black et al., 2002 ; Charles et al., 2017). En utilisant l'interaction entre les ressources locales (au niveau du district) en pétrole et charbon et les prix de ces matières premières comme instruments, Black et al. (2002) et Charles et al. (2017) estiment ainsi pour les Etats-Unis qu'une augmentation exogène de 1 % du revenu local par habitant entraîne une diminution des dépenses en invalidité par habitant entre 0,3 % et 0,4 %. Malgré un effet causal bien établi, la conjoncture économique peut affecter le taux d'invalidité local à travers un grand nombre de mécanismes. Pour identifier ceux-ci, d'autres études ont donc utilisé des données microéconomiques permettant une analyse plus fine des déterminants de l'entrée et sortie en/de l'invalidité. Bratsberg et al. (2013) montrent pour la Norvège qu'une perte d'emploi augmente la probabilité de bénéficier du programme d'invalidité permanente dans les quatre années de 100 % (50 %) chez les hommes (femmes). Selon leurs calculs, les pertes d'emploi pourraient ainsi expliquer 28 % (13 %) du nombre total d'entrées en invalidité chez les

hommes (femmes) en Norvège. Bien que la relation causale sur l'invalidité n'ait pas été montrée directement, d'autres études montrent qu'une période de récession économique peut également affecter la santé des travailleurs via une diminution des investissements de sécurité des entreprises (Boone et al., 2011 ; Mc Manus et Schaur, 2016) ou une augmentation du stress des travailleurs lié au risque de la perte d'emploi (Colantone et al., 2015 ; Caroli et Godard, 2016). Les conditions économiques peuvent également influencer le taux d'invalidité à travers le taux de sortie de ce statut. Plusieurs études ont montré l'importance du rôle de l'employeur (Fevang et al., 2014) ainsi que des incitants financiers sur le retour à l'emploi de personnes en invalidité (Fevang et al., 2017 ; Kostol et Mogstad, 2014 ; Campolieti et Ridell, 2012 ; Borghans et al., 2014). Il est en effet possible qu'un risque élevé de tomber au chômage une fois retourné sur le marché du travail décourage certaines personnes en invalidité de retourner au travail ou de chercher un nouvel emploi. Enfin, un aspect plus dynamique n'ayant pas été exploré par la littérature fait référence à la mobilité géographique des entreprises. Il est possible que des conditions économiques défavorables entraînent à long terme le départ de certaines entreprises offrant de meilleures conditions de travail. Cet effet de sélection peut à terme entraîner une hausse du taux d'invalidité local.

3.3. TISSU SOCIAL LOCAL

Plusieurs études ont montré que l'environnement social dans lequel chacun évolue constitue un déterminant central de la santé et de la probabilité d'entrée en invalidité (Ahlburg, 1998 ; Anderson et al., 2003 ; Dahl et al., 2014 ; Sim et Yen, 1999 ; Trannoy, 2010). Nous détaillons ici plusieurs mécanismes par lesquels l'environnement social peut affecter le taux d'invalidité local. Premièrement, plusieurs études ont trouvé une corrélation intergénérationnelle du niveau de santé mentale et physique (Ahlburg, 1998 ; Trannoy, 2010). Cette corrélation résulte autant de facteurs biologiques d'ordre génétique ou liés à la santé des parents (Van den Berg et al., 2006) que de la manière dont l'environnement familial influence la situation socio-économique de chaque personne (Currie, 2009). Au-delà des liens familiaux, l'environnement social peut affecter la santé à travers de nombreux canaux tels que les caractéristiques socio-économiques du quartier d'habitation (Honold et al., 2012), l'existence de discriminations ethniques dans un quartier (Williams et al., 2003) ou encore, l'accès à des activités sociales et sportives. Enfin, au-delà de son effet sur la santé, l'environnement social peut également influencer la probabilité d'entrée en invalidité via des effets d'influence réciproque. En d'autres termes, un programme de soutien financier peut devenir plus attractif pour une

personne si des membres de son entourage viennent à en bénéficier. Dahl et al. (2014) ont par exemple montré pour la Norvège que la probabilité qu'un parent reçoive des allocations d'invalidité augmente de 6 (12) points de pourcentage la probabilité que l'enfant reçoive également des allocations d'invalidité après 6 (10) années. D'autres études ont également montré que des réseaux sociaux tels que les collègues, les voisins, ou encore le groupe ethnique peuvent exercer une influence sur la probabilité de bénéficier d'allocations d'invalidité. Ce résultat indique que la dynamique de l'invalidité est susceptible de s'auto-entretenir voire de s'accélérer au cours du temps en raison d'un effet multiplicateur (Dahl et al., 2014 ; Bratsberg et al., 2014 ; Markussen et Roed, 2015 ; Rege et al., 2012 ; Furtado et Theodoropoulos, 2012).

3.4. SYSTEME DE SANTE LOCAL

L'organisation du système de santé local peut influencer l'entrée en invalidité à travers un système de prévention effectif, la qualité des soins dispensés ou encore la façon dont les médecins évaluent l'inaptitude au travail. Plusieurs études ont en effet montré l'importance des services de prévention des risques relatifs à la santé au travail pour réduire les maladies chroniques et les troubles psychologiques (Loisel et al., 2005 ; Shaw et al., 2009). Cependant, à notre connaissance, aucune étude n'a encore tenté d'identifier les effets de la présence de services de prévention de santé au niveau local sur les inégalités spatiales de taux d'invalidité. Un autre aspect ayant reçu une attention particulière est la procédure d'évaluation de l'incapacité de travail. De manière assez naturelle, on peut s'attendre à ce qu'une procédure d'évaluation des applications plus stricte entraîne un niveau d'équilibre plus faible du nombre d'entrées en invalidité si une partie des entrants potentiels se voient soit refusés, soit dissuadés d'appliquer le programme d'invalidité. Pour estimer l'effet d'un contrôle plus strict des applications aux Pays-Bas, De Jong et al. (2011) ont effectué une expérience aléatoire dans laquelle une nouvelle méthode d'évaluation plus stricte[10] fut mise en place uniquement dans certaines régions du pays. Ceux-ci montrent que le nombre d'entrées en invalidité diminuèrent substantiellement dans les régions ou le programme fut mis en place. De plus, l'effet est observé principalement à travers une diminution du nombre d'applications et non

(10) Cette méthode consiste essentiellement à laisser plus de temps aux personnes en charge de l'évaluation de la demande d'allocations, en permettant notamment une plus grande proportion d'entretiens téléphoniques ou de rencontres personnelles avec les demandeurs d'allocations et leur employeur.

une augmentation du nombre de demandes rejetées, indiquant par là un effet dissuasif important. La restriction de l'accès à d'autres programmes de soutien financier peut également entraîner une hausse du nombre d'entrées en invalidité si le contrôle à l'entrée du programme d'invalidité est lui-même imparfait. Plusieurs études ont montré que des effets de substitution existent entre l'invalidité et d'autres programmes de soutien financier (Borghans et al., 2014 ; Stabuli, 2011 ; Petrongolo, 2008 ; Lammers et al., 2013), notamment suite à des changements législatifs restreignant l'accès à certains programmes. Etant donné que la procédure d'évaluation est définie en Belgique au niveau fédéral, les règles fixant l'accès aux allocations d'invalidité sont les mêmes dans tous les arrondissements. Cependant, le statut d'invalidité étant accordé au niveau local par les mutualités, une possibilité est que les médecins-conseils appliquent des critères d'évaluation différents, selon l'environnement de travail dans lequel ceux-ci évoluent. Bien qu'aucune étude n'ait approfondi cette hypothèse, des études en Norvège montrent que des caractéristiques telles que l'âge et le sexe des médecins influencent fortement leur propension à évaluer positivement l'incapacité de travail (Markussen et al., 2011 ; Markussen et al., 2013).

3.5. MOBILITE SPATIALE

La mobilité des personnes est susceptible d'exacerber les inégalités spatiales de taux d'invalidité si les personnes de même niveau de santé ont tendance à se regrouper dans les mêmes zones géographiques. Cela est dû au fait que la santé est en général corrélée avec un ensemble de caractéristiques socio-économiques (éducation, âge, revenu etc…) qui elle-même influencent la mobilité. Des effets de sélection peuvent donc apparaitre lorsque les personnes en bonne (mauvaise) santé quittent (arrivent dans) les zones dans lesquelles le niveau de santé moyen est faible. La mobilité est un facteur qui se distingue des quatre autres facteurs cités ci-dessus par le fait que celle-ci entraîne une interconnexion entre les différentes entités géographiques, ce qui n'est pas le cas des autres facteurs. En effet, les flux migratoires changent la composition de la population tant des zones géographiques de départ que d'arrivée et modifient l'état de santé moyen de chaque zone géographique selon que l'état de santé des arrivants/sortants est au-dessus ou au-dessous de celui-ci. Si l'effet

de court terme de la mobilité est très faible[11] , les effets de long terme peuvent en revanche jouer un rôle important.

3.6. EFFETS INCITATIFS VS SANTE OBJECTIVE

Une interrogation fréquente sur laquelle se sont penchées de nombreuses études est celle de distinguer les effets incitatifs des effets uniquement liés à la santé dite « objective » des bénéficiaires. Autrement dit, quelle part des fluctuations temporelles ou spatiales du taux d'invalidité peut être expliquée par l'évolution de la santé objective de la population ? La plupart des économistes sont divisés sur la question. En effet, certains chercheurs affirment que l'évolution du niveau de santé ne joue qu'un rôle marginal dans l'évolution du taux d'invalidité en raison du fait que l'on a observé aux Etats-Unis une amélioration de nombreux indicateurs de santé et une augmentation du taux d'invalidité entre 1990 et 2003 (Autor, 2011 ; Mc Vicar, 2007 ; Duggan et Imberman, 2008). En Belgique, l'enquête de santé rapporte une augmentation du pourcentage de la population (de 15 ans et plus) avec une ou plusieurs maladies de longue durée de 26,9 % en 2004 à 28,5 % en 2013, ainsi qu'une augmentation de la proportion de personnes souffrant de troubles dépressifs de 10 % (6 %) en 2004 à 18 % (11 %) en 2013 chez les femmes (hommes). Il est donc difficile d'établir la part des fluctuations du taux d'invalidité résultant des fluctuations du niveau de santé de la population dans le temps ou l'espace à partir d'indicateurs de santé agrégés.

Une manière plus précise de mesurer les effets de divers incitatifs consiste à exploiter des variations exogènes de ces incitatifs n'affectant que certains groupes d'individus, afin d'en isoler l'effet causal. Les incitatifs financiers (le montant des allocations d'invalidité ou leur écart par rapport à un emploi rémunéré) ont été la classe d'incitatifs sur laquelle les études se sont le plus penchées (Fevang et al., 2017 ; Kostol et Mogstad, 2014 ; Campolieti et Ridell, 2012 ; Borghans et al., 2014). Plusieurs études en Norvège ont ainsi montré un effet substantiellement négatif du montant des allocations d'invalidité sur le taux de retours à l'emploi (Fevang et al., 2017 ; Kostol et Mogstad, 2014),

(11) Dans nos données individuelles, nous observons qu'environ 3 % de la population change d'arrondissement d'une année à l'autre.

avec cependant des effets très hétérogènes selon l'âge et le niveau d'éducation[12] ainsi que des effets d'externalités sur les autres programmes de sécurité sociale (chômage et retraite anticipée). En revanche, Campolieti et Ridell (2012) ne trouvent aucun effet d'une diminution du montant des allocations d'invalidité au Canada sur le flux d'entrants en invalidité. En Belgique, la période caractérisée par l'augmentation la plus forte du taux d'invalidité (2007-2012) n'a pas connu de réforme importante ni du montant des allocations d'invalidité ni du montant des allocations de chômage ou d'autres programmes de soutien financier. En revanche, plusieurs réformes importantes mises en place entre 2006 et 2007 ont limité l'accès à plusieurs programmes de soutien financier. En particulier, le durcissement des conditions d'accès au régime de prépension instauré en 2007 ainsi que la mise en place de l'activation du comportement de recherche d'emploi pour les chômeurs de longue durée sont des explications potentielles, faisant l'objet d'actuelles recherches (De Brouwer et al., 2018).

3.7. CONCLUSIONS DE LA SECTION

En résumé, cette revue de la littérature a discuté le rôle respectif de cinq facteurs pouvant expliquer les variations géographiques du taux d'invalidité. L'utilité de cette revue de la littérature pour la présente étude est double. Premièrement, les résultats observés dans les sections suivantes seront systématiquement comparés avec les résultats obtenus dans la littérature scientifique. Deuxièmement, cette revue de la littérature nous permet d'identifier certaines variables à inclure dans les analyses économétriques menées ci-dessous. Notre étude se concentre principalement sur les facteurs d'ordre économiques et démographiques et laisse l'étude des autres facteurs (disparités de l'offre médicale, critères d'évaluation des médecins, effets de réseau) pour des recherches futures.

(12) Kostol et Mogstad (2014) montrent que l'effet d'incitatifs financiers pour encourager le retour à l'emploi est important uniquement pour les moins de 50 ans. En revanche, aucun effet n'est observé chez les plus de 50 ans.

4. ANALYSE MACROECONOMIQUE DES DETERMINANTS DU TAUX D'INVALIDITE

4.1. INTRODUCTION

Dans cette section, nous analysons le taux d'invalidité en prenant la commune comme unité d'analyse et relions celui-ci avec différentes variables d'ordre socio-économique. L'objectif ici est donc d'établir un modèle économétrique permettant d'identifier les variables capables d'expliquer les variations géographiques du taux d'invalidité et de déterminer la part de cette variation pouvant être expliquée par notre modèle. Notre base de données comprend un ensemble de variables annuelles au niveau communal pour la période 2009-2013. Cette période a été choisie principalement en raison de la disponibilité des données, mais également car elle coïncide avec l'augmentation la plus marquée du taux d'invalidité en Belgique. Nous avons récolté des données au niveau communal provenant de différentes sources (la Banque Carrefour de la sécurité sociale (BCSS), l'INAMI, l'Office national de sécurité sociale (ONSS) et l'Office belge de statistique (Statbel)). Les sources de chaque variable sont disponibles dans l'annexe 2. Le tableau ci-dessous montre la liste des variables utilisées ainsi que leur moyenne et écart type à travers les communes en 2011.

TABLEAU 1 : LISTE DES VARIABLES ET STATISTIQUES DESCRIPTIVES POUR L'ANNEE 2011

Variable	Moyenne	Ecart type
Taux d'invalidité	0,060	0,018
Taux de chômage	0,066	0,041
Revenu médian (ln)	9,714	0,076
Proportion de faillites par population active	0,002	0,001
Proportion d'ouvriers	0,526	0,118
Age moyen de la population active	43,287	0,789
Proportion de femmes actives	0,467	0,015
Densité de population (ln)	5,766	1,147
Proportion d'étrangers d'origine UE	0,052	0,070
Proportion d'étrangers d'origine non-UE	0,111	0,087
Taux de mortalité	0,009	0,002
Nombre total de communes	577	

Note : (1) les moyennes prennent la commune comme unité d'observation. (2) Les communes contenant moins de 1000 personnes actives ont été supprimées en raison de la forte variabilité des variables dans ces communes. Le nombre de communes passe donc de 589 à 577.

Le fait de disposer d'un échantillon longitudinal (contenant plusieurs années) nous permet d'utiliser ici les méthodes standards des données de panel. En d'autres termes, nous sommes en mesure d'établir un lien entre les variables explicatives et le taux d'invalidité en tenant compte d'un effet communal non observé et fixe dans le temps (modèle à effet fixe). Les sections suivantes détaillent le modèle économétrique utilisé et les résultats obtenus.

4.2. METHODE D'ESTIMATION

Formellement, notre modèle est décrit par l'équation suivante:

$$y_{it} = \gamma + \alpha_i + \delta_t + \sum_{k=1}^{K} x_{kit}\beta_k + \varepsilon_{it} \quad (1)$$

Où i, t représentent respectivement la commune et l'année, y est la variable dépendante (ici, le taux d'invalidité), α_i est un effet communal fixe dans le temps, δ_t est un effet temporel commun à toutes les communes, x_k (k=1,...,K) est un ensemble de variables explicatives et ε est le terme d'erreur.

Cette équation peut être estimée de plusieurs manières, en fonction des hypothèses faites sur la corrélation entre le terme d'erreur, l'effet fixe et les variables explicatives. Ici, nous estimons ce modèle de deux façons : premièrement nous estimons un simple modèle « pooled OLS », en régressant le taux d'invalidité communal sur la constante y, la variable temporelle δ_t et les variables explicatives en rajoutant un effet fixe d'arrondissement. Deuxièmement, nous estimons l'équation (1) dans un modèle à effet fixe[13], c'est-à-dire en régressant la variation annuelle du taux d'invalidité $\Delta_{(t+1,t)}\, y_i$ sur la variation annuelle des variables explicatives. L'avantage de ce modèle est qu'il supprime le biais potentiel des estimateurs, résultant de la corrélation entre l'effet fixe communal et le terme d'erreur. En revanche, il n'est dès lors plus

(13) Un modèle alternatif au modèle à effet fixe est le modèle à effet aléatoire (« random effect model »). Le test d'Hausman permet de déterminer si le modèle à effet fixe ou à effet aléatoire doit être utilisé. Ici, le test d'Hausman indique d'utiliser le modèle à effet fixe. Nous ne reportons donc pas les résultats du modèle à effets aléatoires.

possible d'estimer l'effet de l'arrondissement sur le taux d'invalidité en raison de la colinéarité entre l'effet fixe communal et l'effet fixe d'arrondissement. Les coefficients obtenus par le modèle à effet fixe seront moins biaisés par rapport à l'estimation par les moindres carrés si l'effet fixe communal est corrélé avec le terme d'erreur. Cependant, la comparaison entre les résultats des deux modèles est intéressante car elle nous permet de comparer la corrélation entre les différentes variables lorsque celles-ci sont prises en niveaux (pooled OLS) ou en différences premières (Effet fixe). Enfin, nous étudions comment la dispersion de l'effet d'arrondissement évolue avant et après inclusion des variables explicatives dans le modèle linéaire. Autrement dit, nous comparons pour chaque arrondissement le taux d'invalidité moyen avec le taux d'invalidité prédit sur base du modèle linéaire en contrôlant pour toutes les autres variables explicatives.[14] Cette façon de procéder permet de prédire quel serait le taux d'invalidité prédit pour une commune dans un certain arrondissement, toute autre chose égale par ailleurs. Notons enfin que l'invalidité étant une incapacité de travail de long terme il est donc probable que les variables affectent le taux d'invalidité avec un certain décalage. Pour résoudre ce problème, nous prenons des différences temporelles de deux années dans le modèle à effet fixe (2009-2011 et 2011-2013).

4.3. RESULTATS

Le tableau 2 indique les résultats de l'estimation de l'équation (1) selon le modèle « pooled OLS » et le modèle à effet fixe. Nous voyons plusieurs éléments intéressants apparaitre. Premièrement, le coefficient associé au taux de chômage est positif et significatif dans le modèle pooled OLS. En revanche, en estimant l'équation en différences premières, ce coefficient devient significativement négatif. Autrement dit, alors que le taux d'invalidité est élevé dans les communes à taux de chômage élevé, une augmentation d'une année à l'autre du taux de chômage est en général associée à une diminution du taux d'invalidité. Pour clarifier ce lien, la figure 8 représente graphiquement le lien entre le taux de chômage moyen et le taux d'invalidité moyen (à gauche) ainsi que leur relation en différences premières (à droite) pour chaque région. Nous voyons deux interprétations possibles pour ce résultat. Soit, la relation entre les deux variables est liée au fait que certains facteurs survenus entre 2009 et 2013 ont affecté positivement le taux d'invalidité et négativement le

(14) C'est-à-dire en assumant des niveaux équivalents des autres variables explicatives dans toutes les communes.

taux de chômage (ou inversement). Cette explication est peu plausible étant donné que la littérature scientifique a montré que la conjoncture économique affecte ces deux variables dans le même sens (Black et al., 2002 ; Charles et al., 2017). Soit, l'effet de l'évolution du taux de chômage sur le taux d'invalidité est direct, c'est-à-dire qu'un transfert s'est opéré entre les deux programmes. Nous reviendrons sur cette dernière hypothèse dans la section suivante en montrant que les demandeurs d'emploi ont des taux d'entrée en invalidité substantiellement plus élevés que les personnes en emploi.

Nous observons également que la conjoncture économique (mesurée par le revenu médian et le taux de faillites par personne active) semble avoir un effet faible, voire nul sur le taux d'invalidité. En effet, bien que les deux coefficients aient des signes attendus (négatif pour le revenu médian et positif pour le nombre de faillites par personne active), seul le coefficient associé à la variation temporelle du revenu médian est légèrement significatif. Ce résultat peut être interprété de deux manières ; soit le revenu médian n'est pas un indicateur reflétant correctement la conjoncture économique locale (par exemple en raison de la rigidité des salaires), soit la conjoncture économique affecte peu le taux d'invalidité à moyen terme.

Les autres résultats observés paraissent être en concordance avec les observations précédentes. En effet, on observe un lien positif entre la proportion d'ouvriers, l'âge moyen de la population active ainsi que la proportion de femmes actives (ayant des taux d'invalidité en moyenne plus élevés que les hommes) et le taux d'invalidité, tant dans le modèle « pooled OLS » que le modèle à effet fixe. Un autre résultat intéressant est que la densité de population est liée positivement au taux d'invalidité dans le modèle « pooled OLS » et négativement dans le modèle à effet fixe. Cela indique que, toute chose égale par ailleurs, les communes plus densément peuplées ont des taux d'invalidité en moyenne plus élevés. Le signe négatif obtenu avec le modèle à effet fixe est probablement un résultat mécanique, lié au fait qu'une augmentation de la densité de population augmente le dénominateur du taux d'invalidité (le nombre de personnes actives par commune), entraînant ainsi une relation négative entre l'évolution de ces deux variables.

Le coefficient associé au taux de mortalité n'est positif que dans le modèle « pooled OLS ». Autrement dit, l'évolution du taux de mortalité dans une commune ne semble pas affecter le taux d'invalidité. Ce résultat doit être interprété avec prudence. En effet, il est possible que le taux de mortalité soit un mauvais indicateur pour refléter la santé réelle des habitants d'une commune.

En effet, celui-ci n'est que faiblement influencé par l'existence de maladies chroniques ou de troubles psychologiques, dont la mortalité est faible. Enfin, nous n'observons aucun lien entre la proportion d'actifs d'origine étrangère et le taux d'invalidité, ce qui laisse supposer que ces groupes ont des taux d'invalidité similaires aux personnes d'origine Belge.

TABLEAU 2 : RESULTATS DE L'ESTIMATION DE L'EQUATION (1)

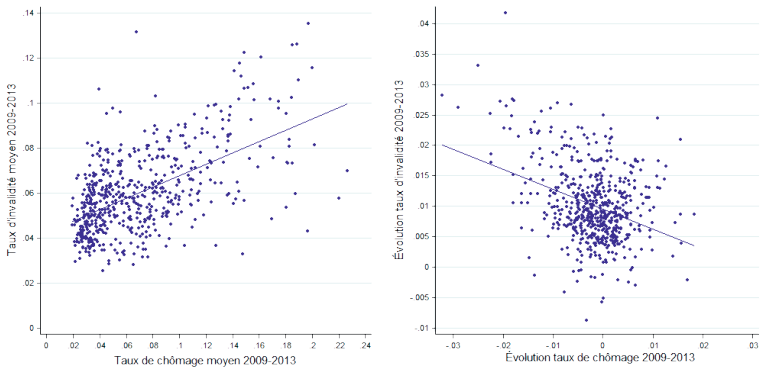
Variables	(1) **Pooled OLS**	(2) **Effet fixe**
Taux de chômage	0.0864** (0.0406)	-0.247*** (0.0331)
Revenu médian (ln)	-0.0189 (0.0129)	-0.0157* (0.00858)
Nombre de faillites par population active	0.140 (0.487)	0.0588 (0.197)
Proportion d'ouvriers	0.125*** (0.00861)	0.104*** (0.0158)
Age moyen population active	0.00279*** (0.000733)	0.00146 (0.000938)
Proportion de femmes actives	0.108** (0.0427)	0.109*** (0.0350)
Densité de population (ln)	0.00344*** (0.00102)	-0.0422*** (0.0153)
Proportion d'actifs d'origine étrangère non-UE	-0.0190 (0.0129)	0.00293 (0.0243)
Proportion d'actifs d'origine étrangère UE	-0.0128 (0.00946)	-0.0177 (0.0330)
Taux de mortalité	0.949*** (0.215)	-0.0818 (0.107)
Effet fixe d'arrondissement	Oui	Non
Effet fixe de commune	Non	Oui
Effet temporel	Oui	Oui
Cluster erreurs standard par arrondissement	Oui	Oui
Observations	1728	1728
R-carré	0.7928	0.7340
Nombre de communes	577	577

Erreurs standards entre parenthèses.

*** p<0.01, ** p<0.05, * p<0.1.

Note : ce tableau présente les résultats de l'estimation de l'équation (1) par les moindres carrés et en différences premières, pour la période 2009-2013.

FIGURE 8 : CORRELATION ENTRE LE TAUX DE CHOMAGE ET LE TAUX D'INVALIDITE EN NIVEAUX (A GAUCHE) ET EN DIFFERENCES PREMIERES (A DROITE))

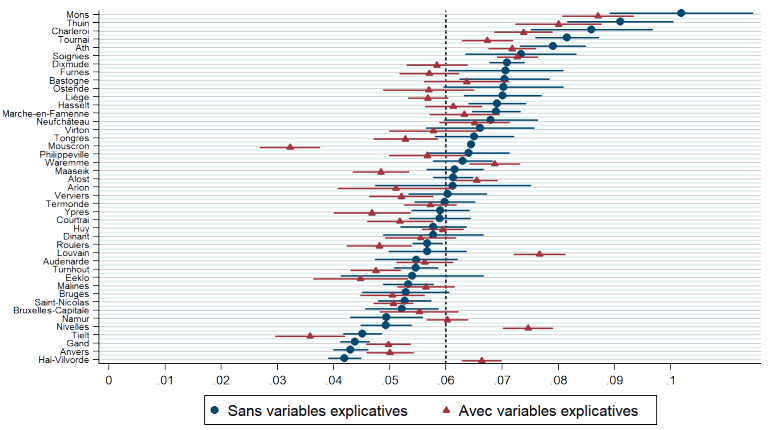


Note : cette figure indique à gauche le taux d'invalidité en fonction du taux de chômage et à droite le lien entre l'évolution temporelle de ces variables entre 2009 et 2013. Les droites de régression ont été ajoutées sur les graphiques.

Finalement, la figure 9 représente graphiquement la variation de l'effet d'arrondissement avec et sans inclusion des variables de contrôle dans le modèle « pooled OLS ». La ligne verticale en pointillés indique le taux d'invalidité moyen. En bleu (rouge) sont représentés les taux d'invalidité moyens par arrondissement sans (avec) inclusion des variables de contrôle dans le modèle. Ce graphique montre plusieurs éléments intéressants. Premièrement, comme observé précédemment, les arrondissements contenant les communes aux taux d'invalidité les plus élevés sont principalement rattachés à la province du Hainaut. En effet, sur la période étudiée, les taux d'invalidité communaux dans les cinq arrondissements les plus élevés variaient entre 8 % et 10 %, soit le double des niveaux observés dans les arrondissements où les taux d'invalidité sont les plus faibles. Deuxièmement, nous observons que même en incluant toutes les variables de contrôle dans notre modèle, l'hétérogénéité des effets d'arrondissements reste importante. En effet, les arrondissements dans le haut de la distribution restent significativement au-dessus de la moyenne nationale, tandis que de nombreux autres arrondissements sont significativement en-dessous de la moyenne nationale. Troisièmement, certains effets d'arrondissement varient fortement après inclusion des variables explicatives dans le modèle. Notamment, les arrondissements d'Hal-Vilvorde, Nivelles et Louvain ont des taux d'invalidité significativement au-dessus de la moyenne

uniquement en incluant les variables explicatives dans le modèle. Le cas inverse est observé pour les arrondissements de Tielt et Mouscron.

FIGURE 9 : TAUX D'INVALIDITE PAR ARRONDISSEMENT PREDITS SANS ET AVEC VARIABLES DE CONTROLE



4.4. CONCLUSIONS DE L'ANALYSE MACROECONOMIQUE

Dans cette section, nous avons étudié l'effet de différentes variables sur le taux d'invalidité en prenant la commune comme unité d'observation. L'inclusion d'un effet d'arrondissement dans le modèle nous a ensuite permis d'observer si ces variables peuvent expliquer les différences de taux d'invalidité observées entre arrondissements. La comparaison des résultats du modèle « pooled OLS » et à « Effet fixe » a été utile pour évaluer la robustesse des résultats, notamment lorsqu'on supprime le biais résultant de la corrélation entre l'effet fixe et le terme d'erreur. Les résultats ont ainsi montré que ce biais est particulièrement important si l'on veut évaluer l'effet du taux de chômage sur le taux d'invalidité. En effet, alors que le modèle OLS aurait indiqué un effet positif du chômage sur le taux d'invalidité, le modèle à effet fixe montre qu'une évolution positive du taux de chômage influence négativement le taux d'invalidité. Plus précisément, toute chose égale par ailleurs, une augmentation (diminution) d'un point de

pourcentage du taux de chômage entraîne une diminution (augmentation) de 0,247 points de pourcentage du taux d'invalidité. Une explication possible est qu'un transfert de personnes au chômage ait eu lieu sur la période étudiée. Une manière de vérifier cette hypothèse est d'analyser l'évolution des sorties du chômage vers le taux d'invalidité, ce que nous faisons dans la section suivante. Enfin, nous avons observé que l'effet d'arrondissement reste important même après avoir inclus nos variables explicatives dans le modèle, ce qui indique que d'autres variables influencent les disparités géographiques de taux d'invalidité. Toutefois, nous avons pu voir que l'effet d'arrondissement change substantiellement pour certains arrondissements après inclusion des variables dans le modèle (Hal-Vilvoorde, Nivelles, Louvain, Mouscron). Une étude plus attentive des raisons de ce résultat serait intéressante à l'avenir.

5. ANALYSE MICROECONOMIQUE

Dans cette section, nous menons une analyse microéconomique des déterminants de l'entrée en invalidité. Notre objectif est ici de déterminer quelle part de l'hétérogénéité spatiale de la probabilité d'entrée en invalidité peut être expliquée par des différences de composition de la population de chaque zone géographique. Nous estimons ensuite quels facteurs peuvent le mieux expliquer les différences de probabilités estimées en introduisant progressivement les variables de contrôle dans nos estimations.

5.1. DONNEES

Nous utilisons ici un échantillon représentatif de la population belge provenant de la BCSS. La BCSS est un organisme centralisant un large ensemble de données administratives provenant des différentes branches de la sécurité sociale pour chaque individu répertorié en Belgique.

Nous avons sélectionné un échantillon aléatoire de 6 % de l'ensemble de la population des 20-70 ans résidant en Belgique au cours de l'année 2007. L'échantillonnage a été stratifié par sexe et par arrondissement. A partir des individus sélectionnés, un ensemble de variables ont été extraites au dernier jour de chaque trimestre, entre 2003 et 2014. Au total, nous disposons d'un échantillon initial de 416 987 individus, suivis sur 48 trimestres. La quantité d'informations contenue par les données étant plus large que nécessaire pour réaliser nos estimations, nous réduisons donc le nombre de périodes et créons un certain nombre de variables qui synthétisent l'information contenue dans l'échantillon. Après le remaniement de nos données, nous observons donc le

statut[15] de chaque individu au début et à la fin de 4 périodes (voir méthode économétrique ci-après) à savoir : 2005q1-2006q4 ; 2007q1-2008q4 ; 2009q1-2010q4 ; 2011q1-2012q4 et 2013q1-2014q4.

Nous procédons également à la suppression de certains individus de notre échantillon. Etant donné que notre analyse porte sur les salariés et les chômeurs, nous ne conservons que les individus étant dans ce statut au premier trimestre de chaque période et supprimons les fonctionnaires (qui sont soumis à une législation distincte du reste des salariés), les indépendants et les inactifs. Nous supprimons aussi les personnes ayant été en invalidité au cours des 12 mois précédent le trimestre de base de chaque période ainsi que les personnes ayant connu plus de 5 trimestres d'incapacité primaire au cours des deux dernières années afin d'éviter les entrées liées à des rechutes ou à de graves problèmes de santé existant depuis longtemps. Enfin, pour éviter que les résultats soient guidés par des valeurs aberrantes, nous supprimons les individus dont le revenu est supérieur au 99ème pourcentile de la distribution. Nous disposons donc au total de 862 190 observations et 219 374 individus.

5.2. ANALYSE ECONOMETRIQUE

Notre analyse économétrique consiste à estimer la probabilité individuelle d'entrer en invalidité au cours d'une période donnée pour les salariés et les chômeurs. En raison du fait qu'une personne n'acquiert le statut d'invalide qu'après une année d'incapacité de travail, nous devons considérer une période relativement longue pour estimer cette probabilité. Pour cela, nous avons découpé notre échantillon en 4 périodes (2005q1-2006q4 ; 2007q1-2008q4 ; 2009q1-2010q4 ; 2011q1-2012q4 et 2013q1-2014q4) et modélisons la probabilité qu'un individu entre en invalidité au cours d'une période donnée comme une fonction d'un ensemble de variables explicatives et d'un terme d'erreur. L'entrée en invalidité est définie ici comme le fait de se trouver durant au moins un trimestre en invalidité au cours d'une période donnée.

(15) Le statut prend ici les modalités « En emploi », « Au chômage », « Indépendant », « en Invalidité » ou « Inactif ».

Formellement nous estimons le modèle de probabilité linéaire suivant :

$$Y_{ijt} = \tau + \delta_t + \gamma_j + x_{it_0}\beta + \varepsilon_{it} \quad (2)$$

Dans cette formule, i, j, t indiquent respectivement l'individu, l'arrondissement et la période Y_{ijt} est une variable égale à 1 si l'individu i dans l'arrondissement entre en invalidité entre le deuxième et le dernier trimestre de la période t (t= {2005q1-2006q4 ;2007q1-2008q4 ;2009q1-2010q4 ;2011q1-2012q4 et 2013q1-2014q4}), δ_t et $\gamma_j$représentent respectivement un effet temporel et un effet d'arrondissement, x_{it0} est un vecteur de caractéristiques individuelles observables au premier trimestre de chaque période et ε_{ijt} est un terme d'erreur. L'annexe 3 illustre graphiquement la méthode d'estimation de l'équation (2).

Notre objectif est ici d'expliquer la variation géographique de la probabilité d'entrée en invalidité par la variation géographique du vecteur de caractéristiques observables x_{it}. Dans un modèle linéaire classique, cet exercice peut être effectué simplement en observant le R^2 de la régression (qui représente le pourcentage de la variance de la variable dépendante expliquée par la variance des variables explicatives). Cependant, le fait d'opérer au niveau individuel complique l'analyse étant donné que nous avons ici un terme d'erreur ε_{it} et un effet d'arrondissement γ_j. Nous avons donc décidé de procéder comme dans la section précédente en analysant le changement de l'effet d'arrondissement avant et après inclusion des variables de contrôle dans le modèle. Afin de déterminer la capacité du modèle d'expliquer l'hétérogénéité géographique du taux d'invalidité, nous présentons différents indicateurs de dispersion des probabilités estimées, en incluant les variables explicatives progressivement. Une faible variation résiduelle indiquera que les différentiels peuvent être expliqués correctement par une inégale répartition géographique des caractéristiques individuelles, tandis qu'une forte variation résiduelle indiquera que d'autres éléments non capturés par le modèle influencent les disparités observées.

Notre analyse sera effectuée sur différents niveaux. Dans un premier temps, nous estimerons la probabilité d'entrer en invalidité par arrondissement pour les employés et chômeurs confondus. De cette manière, nous pourrons évaluer l'importance de la répartition géographique des catégories socio-économiques (employés et chômeurs) sur l'hétérogénéité géographique du taux d'entrée en invalidité. Nous procéderons ensuite à une analyse séparée pour les salariés et les demandeurs d'emploi.

Les variables explicatives introduites dans l'équation (2) ont été sélectionnées sur base de la littérature scientifique et de leur disponibilité dans la BCSS. En effet, nous avons discuté des facteurs pouvant potentiellement expliquer les variations géographiques du taux d'invalidité et regroupons à présent ceux-ci en 3 catégories : (1) les facteurs démographiques, (2) les facteurs relatifs à la situation sur le marché du travail et aux caractéristiques de l'emploi occupé, (3) des facteurs relatifs à des éventuels antécédents de santé des personnes. Malheureusement, nos données ne nous permettent pas d'identifier les effets de l'environnement social des personnes ou d'autres effets liés à l'offre de soins de santé dans les arrondissements. Le tableau 3 reprend l'ensemble des variables utilisées dans notre analyse.

TABLEAU 3 : VARIABLES EXPLICATIVES DANS L'ANALYSE MICROECONOMIQUE

Catégorie	**Variables**
vie privée	Sexe (2)
	Classe d'âge en 2007 (8)
	Type de ménage (8)
	Nationalité (3)
	Arrondissement (43)
Situation sur le marché du travail	Statut employé/chômeur (2)
	Nombre de trimestres au chômage au cours des 8 derniers trimestres
Caractéristiques de l'emploi	Salaire mensuel brut
	Ancienneté chez l'employeur actuel
	Secteur d'activité NACE 2digits (21)
	Secteur public/privé (2)
	Equivalent temps plein
	CPAS (2)
Historique santé	Nombre de trimestres en incapacité primaire au cours des 8 derniers trimestres
	Fond des Accidents de travail (2)
	Fond des maladies professionnelles (2)
	Fond des personnes handicapées (2)

Notes : (1) pour les variables catégorielles, le nombre de catégories est reporté entre parenthèses. (2) Les variables CPAS, Fond des accidents de travail, Fond des maladies professionnelles et Fond des personnes handicapées sont des variables 0/1 indiquant si la personne bénéficiait financièrement d'une aide d'un des fond au cours du trimestre de référence.

5.3. RESULTATS

Dans cette section, nous fournissons les résultats de l'estimation de l'équation (2) pour différentes catégories de personnes (salariés-chômeurs confondus, salariés et chômeurs) ainsi qu'en considérant séparément différentes catégories de maladies (troubles mentaux/maladies du système locomoteur et autres affections).

5.4. SALARIES ET CHOMEURS CONFONDUS

Les résultats de l'estimation de l'équation (2) sont fournis dans les tableaux 4, 5 et 6. Plus précisément, le tableau 4 présente l'effet de différentes caractéristiques de la vie privée, le tableau 5 se concentre sur l'effet de la situation sur le marché du travail et enfin le tableau 6 se concentre sur l'effet d'antécédents médicaux des individus sur la probabilité d'entrée en invalidité. Nous distinguons également l'entrée en invalidité selon deux groupes de maladies : (1) les troubles mentaux et maladies du système locomoteur[16] et (2) les « autres maladies »[17]. Cette distinction est importante étant donné la forte augmentation des troubles mentaux et maladies du système locomoteur ces dernières années. Dans les tableaux de résultats, pour les variables discrètes, nous montrons pour chaque modalité de référence la probabilité estimée pour un individu dans cette modalité et indiquons pour les autres modalités l'écart de probabilité par rapport à la modalité de référence.

Avant d'étudier comment l'estimation du modèle permet de comprendre les disparités géographiques du taux d'entrée en invalidité, nous présentons d'abord les résultats pour chaque variable incluse dans le modèle. Premièrement, nous observons que les personnes de nationalité non européenne ont des probabilités

(16) Nous n'avons pas voulu distinguer ces deux catégories de maladies afin de garder des résultats suffisamment précis.

(17) Le groupe des « autres maladies » est principalement constitué des maladies cardiovasculaires, des blessures accidentelles, des tumeurs et des maladies du système nerveux.

inférieures à entrer en invalidité pour les troubles mentaux et les maladies du système locomoteur. Toutes choses égales par ailleurs, une personne de nationalité hors UE a une probabilité de 0,21 points de pourcentage inférieure d'entrer en invalidité endéans les deux ans, soit une probabilité de 22 %[18] inférieure à une personne de nationalité belge. Deuxièmement, les ménages isolés avec et sans enfants ont des probabilités substantiellement plus élevées d'entrer en invalidité que les autres types de ménages. En particulier, un ménage isolé avec enfants a une probabilité d'invalidité de 0,5 points de pourcentage (67 %[19]) supérieure à une personne mariée avec enfants.

(18) 0,0021/0,0095=0,22.

(19) (0,0089+0.0043)/(0.0089-0.001)-1=0.67.

TABLEAU 4 : RESULTATS DE L'ESTIMATION DE L'EQUATION (2) : EFFETS DES CARACTERISTIQUES DE LA VIE PRIVEE SUR LA PROBABILITE D'ENTRER EN INVALIDITE

Variables	(1) **Total**	(2) **Troubles mentaux/malad. Du système locomoteur**	(3) **Autres affections**
Période			
2005q1-2006q4	0.0058 (réf.) (0.0002)	0.0034 (réf.) (0.0002)	0.0024 (réf.) (0.0001)
2007q1-2008q4	+0.0029*** (0.0003)	+0.0019*** (0.0002)	+0.0010*** (0.0002)
2009q1-2010q4	+0.0034*** (0.0003)	+0.0023*** (0.0002)	+0.0011*** (0.0002)
2011q1-2012q4	+0.0046*** (0.0003)	+0.0030*** (0.0003)	+0.0015*** (0.0002)
2013q1-2014q4	+0.0077*** (0.0003)	+0.0054*** (0.0003)	+0.0023*** (0.0002)
Classe d'âge en 2007			
20-24 ans	0.0018 (réf.) (0.0003)	0.0015 (réf.) (0.0002)	0.0003 (réf.) (0.0001)
25-29 ans	+0.0030*** (0.0003)	+0.0020*** (0.0002)	+0.0011*** (0.0002)
30-34 ans	+0.0059*** (0.0004)	+0.0038*** (0.0003)	+0.0020*** (0.0002)
35-39 ans	+0.0078*** (0.0004)	+0.0049*** (0.0003)	+0.0029*** (0.0002)
40-44 ans	+0.0101*** (0.0004)	+0.0062*** (0.0003)	+0.0039*** (0.0002)
45-49 ans	+0.0127*** (0.0004)	+0.0072*** (0.0004)	+0.0054*** (0.0002)
50-54 ans	+0.0121*** (0.0005)	+0.00601*** (0.0004)	+0.0061*** (0.0003)
55-59 ans	+0.0119*** (0.0006)	+0.0057*** (0.0005)	+0.0062*** (0.0004)

Nationalité			
Belgique	0.0095(réf.)	0.0060 (réf.)	0.0035 (réf.)
	(0.0001)	(0.00008)	(0.00007)
UE 28	-8.38e-05	-0.0003	+0.0002
	(0.0005)	(0.0004)	(0.0003)
Hors UE	-0.0021***	-0.0016**	-0.0006
	(0.0008)	(0.0007)	(0.0005)
Type de ménage			
Enfant	0.0089 (réf.)	0.0055 (réf.)	0.0034 (réf.)
	(0.0003)	(0.0002)	(0.0002)
Isolé sans enfants	+0.0030***	+0.0020***	+0.0010***
	(0.0004)	(0.0003)	(0.0003)
Isolé avec enfants	+0.0043***	+0.0036***	+0.0008**
	(0.0006)	(0.0005)	(0.0004)
Marié sans enfants	+0.0008*	+0.0002	+0.0005*
	(0.0005)	(0.0004)	(0.0003)
Marié avec enfants	-0.0010***	-0.0006**	-0.0004*
	(0.0004)	(0.0003)	(0.0002)
Cohabitant sans enfants	+0.0007*	+0.0003	+0.0004*
	(0.0004)	(0.0003)	(0.0002)
Cohabitant avec enfants	-0.0005	-0.0003	-0.0002
	(0.0004)	(0.0003)	(0.0002)
Autres	+0.0018***	+0.0019***	-2.51e-05
	(0.0007)	(0.0006)	(0.0004)
Observations	862 190	862 190	862 190

Erreurs standards robustes clustérisées au niveau individuel entre parenthèses.

*** p<0.01, ** p<0.05, * p<0.1.

Note : ce tableau présente les résultats de l'estimation de l'équation (2) par un modèle de probabilité linéaire. Pour les variables discrètes, nous avons indiqué pour chaque modalité de référence la probabilité estimée d'entrée en invalidité. Pour les autres modalités, le coefficient associé représente l'écart de probabilité par rapport à la modalité de référence. Chaque variable est observée au premier trimestre de chaque période (2005q1, 2007q1, 2009q1, 2011q1 et 2013q1).

Nous observons ensuite que les demandeurs d'emploi de courte et longue durée ont des probabilités d'entrée largement supérieures aux salariés. En effet, toutes choses égales par ailleurs, une femme au chômage depuis moins de 8 trimestres a une probabilité de 0,37 points de pourcentage (45 %) supérieure à une femme employée. On observe également une différence importante entre les employés et les ouvriers, ce qui n'est pas étonnant étant données les tendances observées dans les sections précédentes. Les écarts importants de taux d'entrée en invalidité entre les employés et les personnes au chômage, combinés avec les disparités importantes de taux de chômage entre les arrondissements constituent une piste intéressante pour expliquer les différences géographiques de taux d'invalidité. De manière plus générale, il apparait que les personnes plus éloignées du marché du travail ou ayant de faibles revenus sont les plus sujettes à entrer en invalidité.

TABLEAU 5 : RESULTATS DE L'ESTIMATION DE L'EQUATION (2) – SUITE : EFFETS DE LA SITUATION SUR LE MARCHE DE L'EMPLOI SUR LA PROBABILITE D'ENTRER EN INVALIDITE

Variables	(1) **Total**	(2) **Troubles mentaux/malad. Du système locomoteur**	(3) **Autres affections**
Sexe-type de contrat			
Employée femme	0.0082 (réf.) 0.0002	0.0051 (réf.) (0.0002)	0.0032 (réf.) (0.0001)
Employé homme	+9.18e-05 (0.0002)	-4.89e-06 (0.0002)	+9.69e-05 (0.0001)
Ouvrier femme	+0.0027*** (0.0004)	+0.0025*** (0.0004)	+0.0003 (0.0003)
Ouvrier homme	+0.0019*** (0.0002)	+0.0009*** (0.0002)	+0.0009*** (0.0002)
Chomeur <8 trimestres femme	+0.0037*** (0.0010)	+0.0045*** (0.0009)	-0.0007 (0.0006)
Chomeur <8 trimestres homme	+0.0021** (0.0010)	+0.0010 (0.0008)	+0.0010* (0.0006)
Chômeur >8 trimestres femme	+0.0042*** (0.0015)	+0.0038*** (0.0012)	+0.0005 (0.0008)
Chômeur >8 trimestres homme	+0.0031** (0.0015)	+0.0003 (0.0012)	+0.0028*** (0.0010)

Nombre de trimestres au chômage (au cours des deux dernières années)	+0.0005*** (0.0001)	0.0004*** (0.0001)	0.0001* (7.96e-05)
Revenu total annuel (ln)	-0.0075*** (0.0003)	-0.0046*** (0.0002)	-0.0029*** (0.0002)
CPAS			
Non	0.0094 (réf.) (0.0001)	0.0059 (réf.) (0.00008)	0.0035 (réf.) (0.00006)
Oui	+0.0046** (0.0020)	+0.0045** (0.0017)	+0.0002 (0.0010)
Observations	862 190	862 190	862 190

Erreurs standards robustes clustérisées au niveau individuel entre parenthèses.
*** p<0.01, ** p<0.05, * p<0.1.

Note : ce tableau présente les résultats de l'estimation de l'équation (2) par un modèle de probabilité linéaire. Pour les variables discrètes, nous avons indiqué pour chaque modalité de référence la probabilité estimée d'entrée en invalidité. Pour les autres modalités, le coefficient associé représente l'écart de probabilité par rapport à la modalité de référence. Chaque variable est observée au premier trimestre de chaque période (2005q1, 2007q1, 2009q1, 2011q1 et 2013q1). La variable CPAS est égale à 1 si la personne bénéficie d'une aide financière du CPAS lors du trimestre de référence.

Enfin, un résultat peu surprenant est que les personnes ayant connu des antécédents de santé (support d'un autre organisme de santé ou périodes d'incapacité primaire) ont des probabilités d'invalidité substantiellement plus importantes que la moyenne. Par exemple, le tableau 6 montre que le fait d'avoir été en incapacité primaire durant un trimestre au cours des deux années précédentes augmente la probabilité d'entrée en invalidité de 1 point de pourcentage (129 %). L'effet de bénéficier d'allocations du fond des accidents de travail augmente la probabilité d'entrée en invalidité de 4,2 points de pourcentage (446 %). Par contre, le fait de bénéficier d'allocations du fond des maladies professionnelles n'augmente pas significativement la probabilité d'entrée en invalidité.

TABLEAU 6 : RESULTATS DE L'ESTIMATION DE L'EQUATION (2) – SUITE : EFFET DES ANTECEDENTS DE SANTE SUR LA PROBABILITE D'ENTRER EN INVALIDITE

Variables	(1) **Total**	(2) **Troubles mentaux/malad. Du système locomoteur**	(3) **Autres affections**
Nombre de trimestres en incapacité (au cours des deux dernières années)			
0 trimestres	0.0077 (réf.)	0.0047 (réf.)	0.0030 (réf.)
	(0.0001)	(0.00007)	(0.00006)
1 trimestre	+0.0105***	+0.0074***	+0.0031***
	(0.0006)	(0.0005)	(0.0004)
2 trimestres	+0.0276***	+0.0200***	+0.0076***
	(0.0015)	(0.0013)	(0.0009)
3 trimestres	+0.0515***	+0.0398***	+0.0117***
	(0.0032)	(0.0027)	(0.0016)
4 trimestres	+0.0783***	+0.0532***	+0.0251***
	(0.0063)	(0.0052)	(0.0036)
5 trimestres	+0.123***	+0.0919***	+0.0309***
	(0.0205)	(0.0180)	(0.0112)
Allocation fond des personnes handicapées			
Non	0.0094 (réf.)	0.0059 (réf.)	0.0035 (réf.)
	(0.0001)	(0.00008)	(0.00006)
Oui	+0.0157***	+0.0104***	+0.0052**
	(0.0043)	(0.0035)	(0.0025)
Allocation fond des accidents de travail			
Non	0.0094 (réf)	0.0059 (réf.)	0.0035 (réf.)
	(0.0001)	(0.00008)	(0.00006)
Oui	+0.0422***	+0.0187***	+0.0235***
	(0.0065)	(0.00463)	(0.0047)
Allocation fond des maladies professionnelles			
Non	0.0094 (réf)	0.0059 (réf.)	0.0035 (réf.)
	(0.0001)	(0.00008)	(0.00007)
Oui	+0.0020	+0.0013	+0.0006
	(0.0023)	(0.0019)	(0.0015)
Observations	862 190	862 190	862 190

Erreurs standards robustes clustérisées au niveau individuel entre parenthèses.

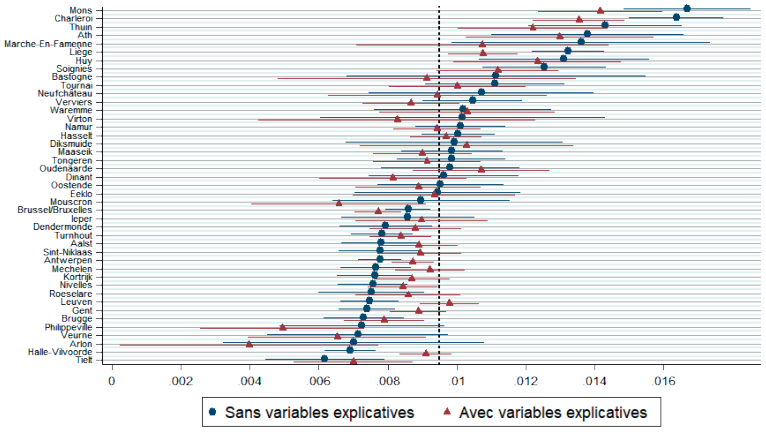
*** p<0.01, ** p<0.05, * p<0.1.

Note : ce tableau présente les résultats de l'estimation de l'équation (2) par un modèle de probabilité linéaire. Pour les variables discrètes, nous avons indiqué pour chaque modalité de référence la probabilité estimée d'entrée en invalidité. Pour les autres modalités, le coefficient associé représente l'écart de probabilité avec la modalité de référence. Chaque variable est observée au premier trimestre de la période de référence. Les variables 0/1 font ici référence au fait de bénéficier ou non d'un soutien financier de l'organisme concerné durant le trimestre de référence.

La figure 10 illustre graphiquement la probabilité d'entrée en invalidité par arrondissement (avec intervalles de confiance à 95 %) estimée par l'équation (2). Sont représentées en bleu, les probabilités estimées en introduisant uniquement l'effet d'arrondissement et l'effet temporel dans le modèle et en rouge, les probabilités estimées en introduisant l'ensemble des variables explicatives.[20] La ligne verticale en pointillés représente la moyenne des coefficients. Nous observons que l'inclusion des variables explicatives rapproche la grande majorité des coefficients de la moyenne nationale. En effet, l'inclusion des variables explicatives dans le modèle rend de nombreux coefficients non significativement différents de la moyenne nationale. Plus précisément, avant d'inclure les variables de contrôle dans le modèle, 23 arrondissements (sur 43) sont significativement au-dessus ou au-dessous de la moyenne nationale. Ce nombre est réduit à 13, lorsque nous incluons les variables de contrôle dans le modèle. Cependant, les arrondissements du haut et du bas de la distribution restent largement au-dessus et en dessous de la moyenne, ce qui indique que l'influence de facteurs non-inclus dans le modèle reste importante pour ces arrondissements. Une discussion approfondie sur ces facteurs sera présentée dans la section suivante. Tenons ici à l'esprit que ces facteurs non observés peuvent être soit des variables individuelles réparties inégalement à travers les arrondissements, soit des caractéristiques propres aux arrondissements eux-mêmes (tels que des caractéristiques liées au quartier, à l'offre de soins de santé, aux critères d'évaluation des médecins, etc…).

(20) Pour rappel, la liste des variables explicatives pour cette régression est la suivante : période, âge en 2007, nationalité, sexe-statut social, type de ménage, nombre de trimestres en incapacité (sur 2 années précédentes), nombre de trimestres au chômage (sur 2 années précédentes), revenu total annuel (logarithme), dummy CPAS, dummy fond des accidents de travail, dummy fond des maladies professionnelles, dummy fond des personnes handicapées.

FIGURE 10 : PREDICTION DES PROBABILITES D'ENTREE EN INVALIDITE PAR ARRONDISSEMENT SANS ET AVEC VARIABLES DE CONTROLEN



Afin de pouvoir identifier plus précisément les variables influençant le plus fortement l'hétérogénéité des effets d'arrondissement, nous ré-estimons l'équation (3), en incluant cette fois les variables explicatives progressivement et en observant l'évolution de la variabilité des coefficients. Pour chaque catégorie de maladie, nous menons quatre estimations. Premièrement, nous n'incluons que l'effet d'arrondissement. Deuxièmement, nous rajoutons les caractéristiques de la vie privée des individus (âge en 2007, nationalité, type de ménage). Troisièmement, nous ajoutons les variables liées à la position des personnes vis-à-vis du marché du travail (sexe-statut social, nombre de trimestres au chômage, revenu total annuel, dummy[21] CPAS). Finalement, nous ajoutons les variables liées à la santé des personnes (nombre de trimestres en incapacité, dummy FAT, dummy FMP, dummy personne handicapée). Le tableau 7 indique plusieurs indices de variabilité pour chaque régression et chaque catégorie de maladie et la figure 11 montre graphiquement les différents quantiles dans un graphique de type « boîte à moustache ». Nous observons que l'écart Max-Min reste quasiment inchangé à travers les spécifications du modèle, ce qui indique que les variables expliquent mal l'hétérogénéité des

(21) Le concept « dummy » fait référence à une variable binaire pouvant prendre les valeurs « vrai » ou « faux ».

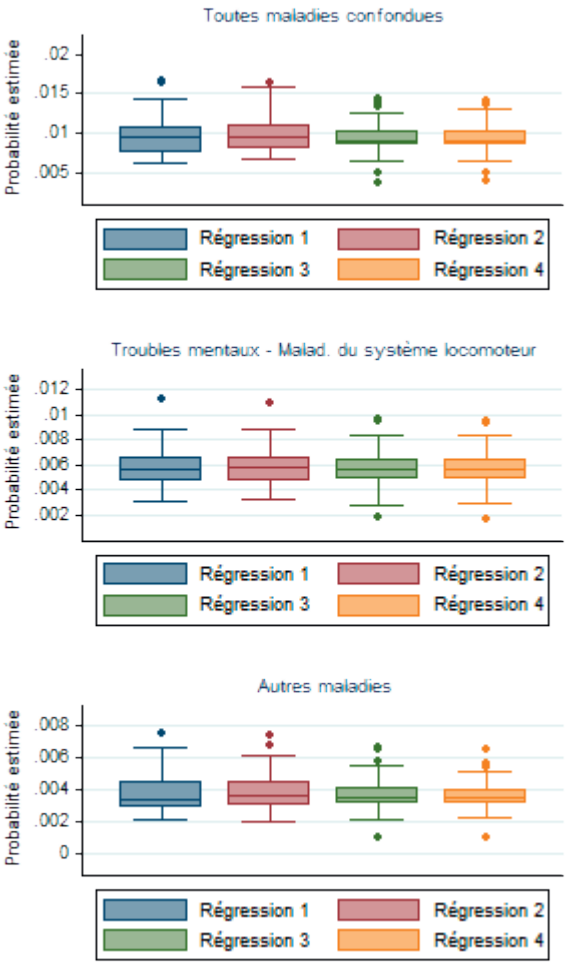
coefficients aux extrêmes de la distribution géographique. L'écart type lui non plus ne varie pas fortement, probablement en raison de l'influence des valeurs extrêmes. En revanche, l'écart entre le troisième et le premier quartile diminue fortement (de 0,32 pp. à 0,19 pp.) après inclusion des variables indiquant la position des personnes vis-à-vis du marché du travail. Nous concluons de ce résultat que parmi les facteurs étudiés, les inégalités d'accès à l'emploi sont les moteurs les plus importants des inégalités géographiques du taux d'invalidité. Une autre observation intéressante est que le modèle semble mieux expliquer l'hétérogénéité des coefficients pour la catégorie « autres maladies ». En effet, en observant l'écart interquartile, on voit que pour les troubles mentaux et maladies du système locomoteur, celui-ci passe de 0,19 à 0,16 pp. (diminution de 16 %) tandis que pour les autres pathologies, celui-ci passe de 0,15 à 0,09 pp. (diminution de 40 %). Il est donc plus difficile d'expliquer les différences géographiques du taux d'entrée en invalidité pour les troubles mentaux et les maladies du système locomoteur.

TABLEAU 7 : INDICES DE VARIABILITE DES COEFFICIENTS PAR TYPE DE MALADIE

Toutes maladies confondues	Rég. 1	Rég. 2	Rég. 3	Rég. 4
Max-Min	0,0105	0,0095	0,0105	0,0102
P75-P25	0,0031	0,0033	0,0019	0,0018
Ec. Type	0,0026	0,0025	0,0021	0,0020
Troubles mentaux/ malad. du système locomoteur	Rég. 1	Rég. 2	Rég. 3	Rég. 4
Max-Min	0,0082	0,0077	0,0078	0,0078
P75-P25	0,0019	0,0018	0,0017	0,0016
Ec. Type	0,0019	0,0018	0,0016	0,0015
Autres maladies	Rég. 1	Rég. 2	Rég. 3	Rég. 4
Max-Min	0,0053	0,0053	0,0056	0,0054
P75-P25	0,0015	0,0015	0,0010	0,0009
Ec. Type	0,0012	0,0012	0,0011	0,0010

Note : le tableau indique la dispersion des probabilités estimées d'entrée en invalidité de chaque arrondissement, en incluant progressivement les variables explicatives. Les variables de contrôle incluses dans chaque régression sont les suivantes: Rég. 1: aucune. Rég. 2 : période, âge, nationalité, type de ménage. Rég. 3 : période, âge, nationalité, type de ménage, sexe-statut social, nombre de trimestres au chômage, revenu médian (log), dummy CPAS. Rég 4 : période, âge, nationalité, type de ménage, sexe-statut social, nombre de trimestres au chômage, revenu médian (log), nombre de trimestres en incapacité, dummy fond des accidents de travail, dummy fond des maladies professionnelles, dummy fond des personnes handicapées.

FIGURE 11 : DISPERSION DES PROBABILITES ESTIMEES EN INCLUANT PROGRESSIVEMENT LES VARIABLES EXPLICATIVES

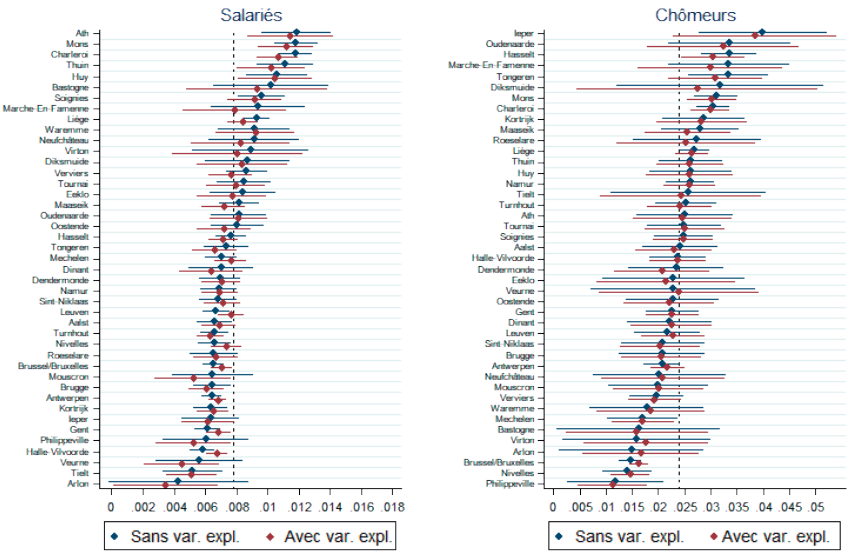


5.5. ANALYSE SEPAREE DES SALARIES ET CHOMEURS

Dans la section précédente nous avons estimé l'équation (2) pour les salariés et les chômeurs. L'avantage d'inclure ces deux catégories dans une même régression est que l'on peut ainsi expliquer l'hétérogénéité géographique de la probabilité d'entrée en invalidité par des différences géographiques du taux de chômage. Cependant, la limitation est qu'il n'est pas possible d'inclure des variables qui ne concernent qu'une seule catégorie, par exemple le salaire ou le secteur d'activité pour les salariés. Nous montrons donc dans cette section les résultats de l'estimation de l'équation (2) pour les salariés et chômeurs séparément. Nous pouvons ainsi ajouter plus de variables explicatives dans le modèle, en particulier pour les salariés, et observer comment celles-ci influencent l'hétérogénéité des coefficients observés. Un autre avantage est que nous pouvons observer si les effets d'arrondissement sont homogènes à travers ces deux catégories.

Les résultats sont illustrés graphiquement dans la figure 12. Nous observons ici que les coefficients avant et après inclusion des variables explicatives dans le modèle diffèrent très peu les uns des autres pour les salariés comme pour les chômeurs. De plus, pour les chômeurs, les intervalles de confiance sont particulièrement larges, ce qui rend toute interprétation compliquée. Ainsi, les caractéristiques des emplois occupés ne peuvent pas expliquer les différences géographiques du taux d'entrée en invalidité observées dans nos données. Une autre observation intéressante est que l'effet d'arrondissement semble hétérogène à travers les salariés et les chômeurs. En effet, chez les salariés, les arrondissements dans le haut de la distribution sont Ath, Mons, Charleroi et Thuin, tandis que chez les chômeurs, ce sont les arrondissements d'Ypres, Audenarde, Hasselt et Marche-en-Famenne pour lesquels les probabilités estimées sont les plus élevées. Ainsi, des facteurs différents influencent la variation résiduelle des coefficients observés pour les salariés et les chômeurs, ce qui montre l'importance de distinguer les risques d'entrée en invalidité selon la position occupée sur le marché du travail.

FIGURE 12 : PREDICTION DES PROBABILITES ESTIMEES D'ENTREE EN INVALIDITE – SALARIES ET CHOMEURS



5.6. DISCUSSION DES RESULTATS

En résumé, les deux sections précédentes ont montré un certain nombre de faits intéressants. Premièrement, la position des personnes vis-à-vis du marché du travail est un facteur particulièrement important pour expliquer les disparités géographiques de la probabilité d'entrée en invalidité. C'est en particulier le cas pour la catégorie « autres maladies ». Deuxièmement, alors que le modèle estimé semble relativement bien expliquer l'hétérogénéité des coefficients associés aux arrondissements dans le milieu de la distribution (entre le premier et le troisième quartile), celui-ci explique mal les valeurs extrêmes en haut et en bas de la distribution géographique. Troisièmement, l'effet d'arrondissement est différent chez les salariés et les chômeurs. En effet, l'ordre des arrondissements dans la distribution des coefficients estimés change substantiellement selon que l'on considère les salariés ou les chômeurs. Les facteurs non-observés à l'origine des différences géographiques du taux

d'entrée en invalidité sont donc différents pour les salariés et les chômeurs. Quatrièmement, l'inclusion des variables explicatives dans le modèle change très peu la dispersion des coefficients lorsqu'on considère les salariés et chômeurs séparément.

Ces résultats montrent ainsi qu'un pourcentage important de l'hétérogénéité géographique de la probabilité d'entrée en invalidité reste inexpliqué, en particulier pour les arrondissements dans le haut et le bas de la distribution et pour les troubles mentaux/maladies du système locomoteur. Plusieurs explications sont possibles. Premièrement, l'effet des variables incluses dans le modèle diffère à travers les arrondissements. Cela peut être le cas si des effets d'emballement existent (par ex. dans le cas d'effets de réseau) ou si différentes variables interagissent entre elles[22] (par ex. si l'effet du type de ménage diffère selon la position sur le marché du travail). Deuxièmement, des caractéristiques non-observées propres aux individus eux-mêmes peuvent influencer la dispersion des coefficients observés. En d'autres termes, il peut exister un effet de composition qui n'est pas pris en compte par notre modèle. Troisièmement, l'hétérogénéité des coefficients estimés peut être due à des caractéristiques propres aux arrondissements eux-mêmes. Malheureusement, il est impossible de distinguer ces deux dernières explications avec le modèle développé ici.[23] Les deux analyses économétriques tendent à montrer qu'un transfert de personnes au chômage vers l'invalidité a eu lieu durant la période étudiée. En effet, alors que l'analyse macroéconomique a montré une corrélation négative entre l'évolution du taux de chômage et du taux d'invalidité, l'analyse microéconomique a montré que les taux d'entrée en invalidité des personnes au chômage sont environ trois fois plus élevés que chez les salariés. Cette explication peut expliquer une part non-négligeable des disparités géographiques du taux d'entrée en invalidité. Cependant, le fait qu'on observe une importante hétérogénéité des coefficients estimés pour les salariés montre que d'autres facteurs sont à l'œuvre.

(22) Nous avons testé différents effets d'interaction, en incluant les interactions suivantes dans l'équation (3) : trimestres au chômage* sexe-statut social, type de ménage* sexe-statut social, revenu total* sexe-statut social, âge en 2007* sexe-statut social et âge en 2007* sexe-statut social* type de ménage. Les résultats restent cependant quasiment inchangés.

(23) Pour une méthode d'identification plus développée, voir Finkelstein et al. (2016).

6. CONCLUSIONS

Dans cet article, nous avons étudié les déterminants géographiques du taux d'invalidité en Belgique sur la période 2003-2013 en nous concentrant sur le régime des salariés. Nos analyses statistiques ont été divisées en deux parties principales : une partie macroéconomique et une partie microéconomique. Nous avons également fourni un ensemble de statistiques descriptives sur l'état de l'invalidité en Belgique ainsi qu'une discussion sur les causes possibles des disparités géographiques existantes. La partie descriptive a montré plusieurs éléments intéressants.

Premièrement, les arrondissements de la province du Hainaut (à l'exception de Mouscron) ont connu une forte augmentation de leur taux d'invalidité ces quinze dernières années et possèdent aujourd'hui avec Liège les taux d'invalidité les plus élevés en Belgique. Il serait donc intéressant que de futures études s'intéressent plus particulièrement à cette province. Deuxièmement, la méthode de standardisation a montré que ni les inégalités géographiques en termes de structure d'âge ou de ratios homme/femme et employé/ouvrier ne peuvent expliquer les fortes disparités géographiques de taux d'invalidité. Troisièmement, au-delà de son aspect géographique, l'invalidité et son augmentation est un phénomène touchant particulièrement les femmes et plus particulièrement encore les ouvrières et est en grande majorité lié à des troubles psychiques (dépression, burn out, fibromyalgie, SFC, …) et à des maladies du système locomoteur (arthrose, rhumatismes, lombalgies,…). En effet, l'augmentation du taux d'invalidité chez les femmes représente 80 % de l'augmentation du taux d'invalidité national alors que celles-ci ne représentent qu'environ 49 % de la population active. Enfin, alors que l'augmentation (faible) du taux d'invalidité chez les hommes peut être expliquée à 88 % par le vieillissement de la population active entre 2003 et 2014, seul 43 % de l'augmentation (forte) du taux d'invalidité chez les femmes peut être expliqué par le vieillissement de population. Dans la partie macroéconomique, nous avons tenté d'expliquer le taux d'invalidité local (au niveau communal) par une série de facteurs socio-économiques. Nous avons ainsi observé que celui-ci était influencé par une série de facteurs tels que l'âge moyen de la population active, le ratio employés/ouvriers, le taux de chômage, la densité de population ou encore le ratio hommes/femmes. Le résultat le plus intéressant est que la corrélation conditionnelle entre le taux de chômage et le taux d'invalidité devient significativement négative lorsque l'on effectue la régression en différences premières (modèle « à effet fixe »). Ce résultat semble indiquer qu'un transfert de personnes du chômage vers l'invalidité a eu lieu au cours de la période étudiée. La compréhension des causes d'un

tel phénomène dépasse les limites de la présente étude et des recherches sont actuellement menées dans cette direction.[24] En isolant l'effet d'arrondissement sur le taux d'invalidité communal, l'analyse macroéconomique a également montré que les arrondissements dans le haut de la distribution géographique restent largement au-dessus de la moyenne nationale même après inclusion des variables explicatives dans le modèle, ce qui indique que d'autres facteurs influencent l'hétérogénéité géographique des taux d'invalidité observés.

Afin d'explorer plus profondément les déterminants de l'entrée en invalidité, nous avons utilisé un échantillon représentatif de 6 % de la population résidant en Belgique en 2007. Dans cette analyse microéconomique, nous avons estimé la probabilité pour un individu d'entrer en invalidité sur un horizon de deux années, en fonction d'un ensemble de caractéristiques observables. Le modèle permet d'expliquer relativement bien les écarts de probabilités entre arrondissements pour les arrondissements dans le milieu de la distribution géographique (entre le 1cr et le 3ème quartile). En revanche, comme dans l'analyse macroéconomique, le modèle explique mal pourquoi certains arrondissements sont fortement éloignés de la moyenne nationale. Nous avons vu que parmi les caractéristiques observables, la situation vis-à-vis du marché du travail est le facteur le plus important pour expliquer l'hétérogénéité des coefficients estimés à travers les arrondissements. De plus, il semble que des facteurs différents influencent l'hétérogénéité des coefficients estimés pour les salariés et les chômeurs. Plus particulièrement, alors que chez les salariés, les taux d'entrée en invalidité sont les plus élevés dans les arrondissements du Hainaut, chez les chômeurs, ceux-ci sont particulièrement élevés dans les arrondissements d'Ypres, Audenarde, Hasselt ou encore Marche-en-Famenne. Nous pensons que cette observation est importante et devrait faire l'objet de recherches approfondies. Malheureusement, le modèle développé ici ne permet pas d'expliquer si l'hétérogénéité résiduelle est due à un effet de composition (caractéristiques individuelles non observées, réparties de façon hétérogène dans l'espace) ou à un effet propre aux arrondissements eux-mêmes. A la lumière de la littérature scientifique, nous estimons que ces deux explications sont également crédibles, voire difficilement différentiables. L'effet de composition comprendrait le tissu social local ainsi qu'un ensemble de pratiques (alimentaires, physiques, sociales) influencées par le milieu social. L'effet d'arrondissement englobe le système médical local, en particulier l'accès

(24) Voir De Brouwer et al. (2018).

à des soins préventifs et à un suivi médical de qualité. On peut également supposer des différences en termes de critères d'évaluation de l'invalidité, comme certaines études précédentes l'ont montré pour la Norvège (Markussen et al., 2011). L'exploration de ces diverses explications est laissée pour des recherches futures.

7. BIBLIOGRAPHIE

Ahlburg, D., Intergenerational transmission of health, *The American Economic Review*, 88(2), pp. 265-270, 1998.

Anderson, L. M., Scrimshaw, S. C., Fullilove, M. T. et Fielding, J. E., The Community Guide's model for linking the social environment to health, *American journal of preventive medicine*, 24(3), pp. 12-20, 2003.

Autor, D. H. et Duggan, M. G., The rise in the disability rolls and the decline in unemployment, *The Quarterly Journal of Economics*, 118.1, pp. 157-206, 2003.

Autor, D. H. et Duggan, M. G., The growth in the social security disability rolls: a fiscal crisis unfolding, *Journal of Economic perspectives*, 20.3, pp. 71-96, 2006.

Black, D., Daniel, K. et Sanders, S., The impact of economic conditions on participation in disability programs: evidence from the coal boom and bust, *American Economic Review*, 92(1), pp. 27-50, 2002.

Boone, J. et al., Recessions are bad for workplace safety, *Journal of health economics*, 30.4, pp. 764-773, 2011.

Borghans, L., Gielen, A. C. et Luttmer, E. FP., Social support substitution and the earnings rebound: evidence from a regression discontinuity in disability insurance reform, *American Economic Journal: Economic Policy*, 6.4, pp. 34-70, 2014.

Bratberg, Espen, Nilsen, Øivind Anti et Vaage, Kjell, Assessing the intergenerational correlation in disability pension recipiency, *Oxford Economic Papers*, 67.2, pp. 205-226, 2014.

Bratsberg, B., Fevang, E. et Røed, K., Job loss and disability insurance, *Labour Economics*, 24, pp. 137-150, 2014.

Campolieti, M. et Riddell, Ch., Disability policy and the labor market: evidence from a natural experiment in Canada, 1998-2006, *Journal of Public Economics*, 96.3-4, pp. 306-316, 2012.

Caroli, E. et Godard, M., Does job insecurity deteriorate health?, *Health economics*, 25(2), pp. 131-147, 2016.

Charles, K. K., Li, Y. et Stephens, Jr, M., Disability Benefit Take-Up and Local Labor Market Conditions, *Review of Economics and Statistics*, (0), 2017.

Colantone, I., Crinò, R. et Ogliari, L., *The hidden cost of globalization: Import competition and mental distress*, Working paper, 2015.

Currie, J., Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development, *Journal of Economic Literature*, 47(1), pp. 87-122, 2009.

Dahl, G. B., Kostøl, A. R. et Mogstad, M., Family welfare cultures, *The Quarterly Journal of Economics*, 129(4), pp. 1711-1752, 2014.

De Brouwer, O., Leduc, E. et Tojerow, I., The effect of changes in the unemployment benefit program on entry flows into the incapacity insurance: a focus on job search monitoring schemes of the older long-term unemployed, *DULBEA working paper*, 2018.

De Jong, Ph., Lindeboom, M. et Van der Klaauw, B., Screening disability insurance applications, *Journal of the European Economic Association*, 9.1, pp. 106-129, 2011.

Duggan, M. et Imberman, S. A., Why are the disability rolls skyrocketing? The contribution of population characteristics, economic conditions, and program generosity, in *Health at older ages: The causes and consequences of declining disability among the elderly*, University of Chicago Press, pp. 337-379, 2009.

Fevang, E., Markussen, S. et Røed, K., The sick pay trap, *Journal of Labor Economics*, 32(2), pp. 305-336, 2014.

Furtado, D. et Theodoropoulos, N., *Immigrant networks and the take-up of disability programs: evidence from US census data*, Working paper, 2012.

Honold, J., Beyer, R., Lakes, T. et van der Meer, E., Multiple environmental burdens and neighborhood-related health of city residents, *Journal of Environmental Psychology*, 32(4), pp. 305-317, 2012.

INAMI, *Facteurs explicatifs relatifs à l'augmentation du nombre d'invalides*, Service d'études de l'INAMI, 2014.

Jones, A. M. et Wildman, J., Health, income and relative deprivation: evidence from the BHPS, *Journal of health economics*, 27.2, pp. 308-324, 2008.

Kostol, A. R. et Mogstad, M., How financial incentives induce disability insurance recipients to return to work, *American Economic Review*, 104(2), pp. 624-55, 2014.

Lammers, M., Bloemen, H. et Hochguertel, S., Job search requirements for older unemployed: transitions to employment, early retirement and disability benefits, *European Economic Review*, 58, pp. 31-57, 2013.

Loisel, P., Buchbinder, R., Hazard, R., Keller, R., Scheel, I., Van Tulder, M. et Webster, B., Prevention of work disability due to musculoskeletal disorders: the challenge of implementing evidence, *Journal of occupational rehabilitation*, 15(4), pp. 507-524, 2005.

Maestas, N., Mullen, K. J. et Strand, A., Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt, *American Economic Review*, 103(5), pp. 1797-1829, 2013.

Markussen, S., Røed, K., Røgeberg, O. J. et Gaure, S., The anatomy of absenteeism, *Journal of health economics*, 30(2), pp. 277-292, 2011.

Markussen, S., Røed, K. et Røgeberg, O., The changing of the guards: can family doctors contain worker absenteeism?, *Journal of health economics*, 32(6), pp. 1230-1239, 2013.

Markussen, S. et Røed, K., Social insurance networks, *Journal of Human resources*, 50(4), pp. 1081-1113, 2015.

McManus, T. C. et Schaur, G., The effects of import competition on worker health, *Journal of International Economics*, 102, pp. 160-172, 2016.

McVicar, D., Why have UK disability benefit rolls grown so much?, *Journal of Economic Surveys*, 22(1), pp. 114-139, 2008.

Nordberg, M. et Røed, K., Economic Incentives, Business Cycles and Long-Term Sickness Absence, *Industrial Relations: A Journal of Economy and Society*, 48.2, pp. 203-230, 2009.

Organisation for Economic Cooperation and Development, Sickness, Disability and Work: Breaking the Barriers: A Synthesis of Findings across OECD Countries, OECD Publishing, 2010.

Petrongolo, B., The long-term effects of job search requirements: evidence from the UK JSA reform, Journal of Public Economics, 93(11-12), pp. 1234-1253, 2009.

Rege, M., Telle, K. et Votruba, M., Social interaction effects in disability pension participation: evidence from plant downsizing, The Scandinavian Journal of Economics, 114.4, pp. 1208-1239, 2012.

Saks, Y., Mieux comprendre l'évolution du nombre de bénéficiaires de l'assurance-invalidité, BNB Revue Economique, 2017.

Shaw, W. S., Van der Windt, D. A., Main, C. J., Loisel, P. et Linton, S. J., Early patient screening and intervention to address individual-level occupational factors ("blue flags") in back disability, Journal of occupational rehabilitation, 19(1), pp. 64-80, 2009.

Staubli, S., The impact of stricter criteria for disability insurance on labor force participation, Journal of Public Economics, 95.9-10, pp. 1223-1235, 2011.

Syme, S. L. et Yen, I. H., The social environment and health: a discussion of the epidemiologic literature, Annual review of public health, 20(1), pp. 287-308, 1999.

Trannoy, A., Tubeuf, S., Jusot, F. et Devaux, M., Inequality of opportunities in health in France: a first pass, *Health economics*, 19(8), pp. 921-938, 2010.

Van den Berg, G. J., Lindeboom, M. et Portrait, F., Economic conditions early in life and individual mortality, *American Economic Review*, 96(1), pp. 290-302, 2006.

Williams, D. R., Neighbors, H. W. et Jackson, J. S., Racial/ethnic discrimination and health: findings from community studies, *American journal of public health*, 93(2), pp. 200-208, 2003.

8. ANNEXES

8.1. ANNEXE 1 : DESCRIPTION DE LA METHODE DU « SHIFT SHARE »

Afin de calculer l'importance relative de chaque groupe de travailleurs (défini par sexe et statut social) dans l'augmentation totale du taux d'invalidité nous utilisons la méthode du « shift-share ». Cette méthode permet de séparer l'augmentation du taux d'invalidité national en deux effets :

Effet de composition : évolution du taux d'invalidité agrégé due à l'évolution de la part d'une catégorie de travailleurs dans l'emploi total.

Effet « interne » : évolution du taux d'invalidité agrégé due à l'évolution du taux d'invalidité à l'intérieur de chaque catégorie de travailleurs.

Formellement, cette formule repose sur des méthodes de log-linéarisation et est représentée comme suit :

$$\Delta Taux.inv = \sum_{i=1}^{I} \overline{\left(\frac{N_i}{N}\right)} * \Delta Taux.inv_i + \sum_{i=1}^{I} \Delta\left(\frac{N_i}{N}\right) * \overline{Taux.inv}_i$$

Dans cette formule, Δ représente la différence entre 2003 et 2014, N_i représente le nombre de titulaires indemnisables dans chaque catégorie (Ouvrier homme, Ouvrier femme,…), N est le nombre total de titulaires indemnisables et la barre indique la moyenne de la variable entre 2003 et 2014[25]. Le premier terme de la somme est l'effet intra-groupe tandis que le deuxième terme est l'effet de composition. Il est important de noter que chaque effet est le produit de deux éléments. L'effet intra-groupe est en effet composé de l'évolution du taux d'invalidité pour chaque catégorie, multiplié par la part relative de cette catégorie. Par exemple, on pourrait observer un effet intra-groupe important dû à une faible augmentation du taux d'invalidité dans un groupe fortement majoritaire et un effet intra-groupe peu important malgré une augmentation du taux d'invalidité dans un autre groupe fortement minoritaire.

(25) Nous utilisons l'année 2003 car il semble que cette date représente le début de l'accélération de l'augmentation du taux d'invalidité.

TABLEAU 8 : DECOMPOSITION PAR SEXE ET CATEGORIE SOCIALE (SALARIES)

Sexe-statut social	Nombre de TI-prép[26]		Part dans le nombre total de TI-prép salariés		Effet intra-groupe	Effet inter-groupe	Contribution dans l'évolution totale
	2003	2014	2003	2014	2003-2014	2003-2014	2003-2014
Ouvriers hommes	1.280.181	1.320.862	34,80 %	32,00 %	0,0062688	-0,00229	0,00398 (17,00 %)
Ouvriers femmes	718.312	832.919	19,50 %	20,20 %	0,01091	0,00072	0,01163 (49,70 %)
Employés hommes	681.181	774.057	18,50 %	18,80 %	0,00056	0,00007	0,00063 (2,70 %)
Employés femmes	998.573	1.197.253	27,10 %	29,00 %	0,00635	0,00082	0,00716 (30,60 %)
Total	3.678.247	4.125.091	100 %	100 %	0,02408	-0,0006849	0,0234 (100 %)

Sources: données INAMI et calculs propres..

Notes : les colonnes 2 et 3 (4 et 5) indiquent le nombre de TI-prép. dans chaque catégorie de travailleurs en 2003 et 2014. Les colonnes 6 et 7 indiquent l'effet intra et inter groupe pour chaque catégorie. Dans la colonne 8, le pourcentage entre parenthèses indique la part de chaque groupe dans l'augmentation totale du taux d'invalidité national.

(26) Pour rappel : TI-prép fait référence au nombre de titulaires indemnisables duquel on a retiré les prépensionnés.

8.2. ANNEXE 2 : ORIGINE DES VARIABLES EXPLICATIVES DANS L'ANALYSE MACRO-ECONOMIQUE

Variable	Source
Taux de chômage	BCSS
Population totale	BCSS
Proportion de la population active d'origine étrangère (UE et non UE)	BCSS
Proportion de femmes actives	BCSS
Âge moyen de la population active	BCSS
Taux d'invalidité	INAMI
Proportion d'emplois dans le secteur industriel	ONSS
Nombre de décès	Statbel - opendata
Revenu médian	Statbel - opendata
Nombre de faillites d'entreprise	Statbel - opendata

8.3. ANNEXE 3 : ILLUSTRATION GRAPHIQUE DE L'EQUATION (2)

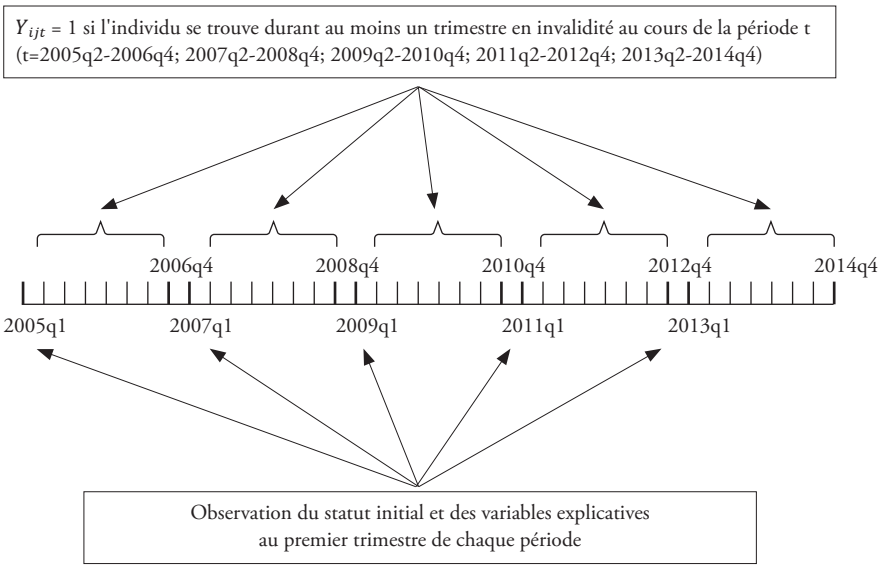


TABLE DES MATIERES

QUELS SONT LES DETERMINANTS DES DISPARITES GEOGRAPHIQUES DE TAUX D'INVALIDITE EN BELGIQUE ?

1. INTRODUCTION 5

2. ETAT DE L'INVALIDITE EN BELGIQUE 8

3. REVUE DE LA LITTERATURE 17
3.1. STRUCTURE DEMOGRAPHIQUE LOCALE . . . 17
3.2. MARCHE DU TRAVAIL LOCAL . . . 18
3.3. TISSU SOCIAL LOCAL . . . 19
3.4. SYSTEME DE SANTE LOCAL . . . 20
3.5. MOBILITE SPATIALE . . . 21
3.6. EFFETS INCITATIFS VS SANTE OBJECTIVE . . . 22
3.7. CONCLUSIONS DE LA SECTION . . . 23

4. ANALYSE MACROECONOMIQUE DES DETERMINANTS DU TAUX D'INVALIDITE 24
4.1. INTRODUCTION . . . 24
4.2. METHODE D'ESTIMATION . . . 25
4.3. RESULTATS . . . 26
4.4. CONCLUSIONS DE L'ANALYSE MACROECONOMIQUE . . . 31

5. ANALYSE MICROECONOMIQUE 32
5.1. DONNEES . . . 32
5.2. ANALYSE ECONOMETRIQUE . . . 33
5.3. RESULTATS . . . 36
5.4. SALARIES ET CHOMEURS CONFONDUS . . . 36
5.5. ANALYSE SEPAREE DES SALARIES ET CHOMEURS . . . 48
5.6. DISCUSSION DES RESULTATS . . . 49

6. CONCLUSIONS 51

7. BIBLIOGRAPHIE 54

8. ANNEXES 58
8.1. ANNEXE 1 : DESCRIPTION DE LA METHODE DU « SHIFT SHARE » 58
8.2. ANNEXE 2 : ORIGINE DES VARIABLES EXPLICATIVES DANS L'ANALYSE MACROECONOMIQUE . . . 60
8.3. ANNEXE 3 : ILLUSTRATION GRAPHIQUE DE L'EQUATION 61