

L'ADRESSE CONTRIBUE-T-ELLE A EXPLIQUER LES ECARTS DE SALAIRES? LE CAS DE JEUNES SORTANT DU SYSTEME SCOLAIRE

Emilia ENE JONES* et Florent SARI**

* Université Paris-Est Marne la Vallée, ERUDITE et TEPP (FR CNRS n° 3435), 5 boulevard Descartes, Champs sur Marne 77454 Marne la Vallée cedex 2, emilia.ene@univ-mlv.fr

** Université de Nantes, LEMNA et TEPP (FR CNRS n° 3435), Chemin de la Censive du Tertre 44322 Nantes Cedex 3, florent.sari@univ-nantes.fr

Version provisoire – Juillet 2014

Résumé

Le lieu de résidence peut avoir des effets spécifiques sur la situation sur le marché du travail des individus au travers différents mécanismes et peut donner naissance à des comportements discriminatoires. Ces préférences discriminatoires peuvent affecter les demandeurs d'emploi dans l'accès à l'embauche et/ou après le processus embauche. Dans ce travail, nous nous intéressons à un mécanisme particulier qui est la discrimination en raison du lieu de résidence, notamment à la discrimination après le processus d'embauche. L'objectif est de pousser la réflexion sur cette forme de discrimination en analysant l'impact potentiel du lieu de résidence sur la situation salariale des individus, au-delà des différences de caractéristiques individuelles. Plus précisément, nous cherchons à savoir si le quartier d'appartenance des individus permet ou non d'expliquer les inégalités salariales observées et, le cas échéant, nous cherchons à quantifier son importance.

Nous utilisons une extension des méthodes de décomposition des écarts de salaires proposée par Neuman et Oaxaca (2004). Notre analyse est confrontée à deux biais de sélection potentiels. D'un côté, il existe un probable biais de sélection lié au choix du lieu de résidence effectué par les parents ou par les jeunes concernés. D'un autre côté, il existe un biais de participation, lié au fait que l'on observe les salaires uniquement pour les jeunes qui sont effectivement en emploi. Pour prendre en compte ce double biais de sélection, nous mobilisons d'abord un probit bivarié afin d'estimer simultanément la probabilité d'habiter dans un quartier défavorisé et la probabilité d'être en emploi. Ensuite, nous calculons des termes de correction qui sont introduits dans les équations de salaires pour chacun des deux groupes de jeunes, c'est-à-dire ceux qui habitent dans des quartiers défavorisés et ceux qui habitent dans des quartiers non considérés comme tels.

Les résultats de nos décompositions des écarts de salaires montrent qu'une fois pris en compte ce double biais de sélection, les différences entre les deux groupes de jeunes ne sont plus forcément significatives. De plus, la partie de ces écarts qui peut être assimilée à de la discrimination n'est plus significative. Ces résultats suggèrent que sans prise en compte de ces biais de sélection, on risque de justifier les écarts de salaires observés entre les deux groupes de jeunes par l'existence d'une discrimination. Ce qui écarte ou minimise l'importance de l'autosélection ou du tri des individus dans l'espace.

Mots-clés : *discrimination, lieu de résidence, double biais de sélection, écarts de salaires.*

Introduction

Le lieu de résidence peut avoir des effets spécifiques sur la situation sur le marché du travail des individus. La théorie économique avance l'idée selon laquelle la localisation des individus et l'organisation spatiale des villes peuvent être sources d'une concentration spatiale du chômage, au travers différents mécanismes : la déconnexion physique entre les lieux de résidence et d'emplois (hypothèse de *Spatial Mismatch*), les effets de pairs ou de réseaux ou encore les préférences discriminatoires de la part des employeurs pour des localisations particulières (Kain, 1968 ; Gobillon *et al.*, 2007 ; Galster, 2012 ; Hellerstein et Neumark, 2012 etc.). Ces préférences discriminatoires peuvent affecter les demandeurs d'emploi dans l'accès à l'embauche et/ou après le processus embauche. Il s'agit, dans ce dernier cas, des différences de traitement de deux types de travailleurs qui manquent de justification objective et raisonnable à l'égard de l'accès aux promotions, à la formation et/ou à l'égard du salaire reçu.

L'analyse théorique des discriminations avance deux types d'explications pour ces comportements discriminatoires. La première hypothèse est de supposer qu'il existe une aversion des employeurs, des salariés ou des consommateurs pour certains groupes de la population (Becker, 1957). La deuxième hypothèse explique l'existence de la discrimination par un manque d'information des employeurs sur la productivité des employés. Ce fait va conduire l'employeur à une appréciation de la productivité des employés basée sur ses croyances et/ou préjugés relatifs à la productivité moyenne du groupe démographique auquel appartient le travailleur (Arrow, 1973 ; Phelps, 1972).

Essayer d'identifier et d'évaluer la discrimination dans l'accès à l'embauche ou après le processus d'embauche (cas de la discrimination salariale), suppose de choisir les méthodes appropriées. Dans le cas où l'on cherche à mesurer la discrimination dans l'accès à l'embauche, la méthode la plus adaptée semble être le *testing*, une méthode purement expérimentale (Petit, 2003) qui consiste en envoyer des candidatures équivalentes (à l'exception de la variable à tester) aux offres d'emploi et en comparer par la suite l'accès aux entretiens d'embauche ou à l'emploi des candidats. Dans les cas où l'on cherche à mesurer les discriminations salariales, les méthodes les plus répandues sont les méthodes de décomposition initialement développées par Oaxaca (1973) et Blinder (1973). Techniquement, ces méthodes ont pour principe de décomposer l'écart de salaire observé entre deux populations en une part qui s'explique par des différences de caractéristiques individuelles et une part inexplicée, liée à des différences dans le rendement des caractéristiques individuelles. La discrimination salariale correspond à cette part inexplicée, les différences dans le rendement des caractéristiques individuelles observées pouvant être le résultat d'un comportement discriminatoire de la part des employeurs.

Dans ce travail, nous nous intéressons à la discrimination après le processus d'embauche, notamment à la discrimination salariale liée au lieu de résidence des individus. L'objectif est de pousser la réflexion sur cette forme de discrimination en analysant l'impact potentiel du lieu de résidence sur la situation salariale des individus, au-delà des différences de caractéristiques individuelles. Plus précisément, nous cherchons à savoir si le quartier d'appartenance des individus permet ou non d'expliquer les inégalités salariales observées et, le cas échéant, nous cherchons à quantifier son importance.

Il existe une abondante littérature nationale et internationale consacrée à l'analyse des discriminations salariales, mais celle-ci visent surtout les différences hommes-femmes (Oaxaca, 1973; Miller, 1987 ; Kim et Polachek, 1994 ; Hellerstein *et al.*, 1999 ; Meurs et Ponthieux, 2000, 2006 etc.) ou les écarts entre les individus d'origines ethniques différentes (Boumahdi et Giret, 2005 ; Aeberhardt et Pouget, 2010 ; Belzil et Poinas, 2010 etc., pour la France ; voir notamment Altonji et Blank 1999 pour une revue de littérature des travaux dans les autres pays). Les études qui prennent en compte le lieu de résidence pour expliquer les écarts de salaires entre les individus sont relativement peu nombreuses et demeurent assez récentes. En France, Aeberhardt *et al.* (2007) introduisent pour la première fois des variables liées au lieu de résidence des individus (dans une ZUS ou en Île-de-France) dans la décomposition des écarts d'accès à l'emploi et des salaires observés entre les Français natifs et les Français ayant au moins un parent originaire d'Afrique. Utilisant des données de l'enquête Formation et Qualification Professionnelle (FQP) 2003 de l'INSEE, ils trouvent une part inexplicée des écarts plus prononcée dans l'accès à l'emploi qu'au niveau salarial (50% versus 5%). Par ailleurs, l'impact des variables liées à la résidence dans une ZUS ou en Île-de-France, qu'ils introduisent dans la décomposition des écarts, ne demeure pas notable. Rathelot (2010a, 2010b) confirme ces résultats sur des données de l'enquête Emploi en continu. L'auteur analyse les écarts d'emploi observés entre les descendants d'immigrés et les Français natifs, tout en

contrôlant de l'effet du lieu de résidence. Tandis que les écarts d'emploi observés s'expliquent à 25% par la qualité du quartier de résidence, les écarts de salaire s'expliquent juste dans une moindre mesure par la qualité du quartier de résidence. Aeberhardt *et al.* (2010) trouvent sur les mêmes données des écarts de salaires expliqués totalement par des différences de caractéristiques individuelles (comparaison individus originaires de Maghreb/Europe du sud versus françaises natifs). Le lieu de résidence (province/Ile-de-France, ZUS/hors-ZUS), semble affecter de la même manière les salaires des différents groupes analysés. En considérant séparément les hommes et les femmes, les résultats restent inchangés. Enfin, Couppié *et al.* (2010) étudient les écarts de salaires entre des jeunes habitant dans des zones urbaines sensibles (ZUS) et des jeunes résidant dans des unités urbaines qui comprennent des ZUS mais en dehors de la ZUS. Ils introduisent dans la décomposition des variables prenant en compte les barrières d'accès aux emplois qualifiés (notamment aux emplois de cadre) qui expliquent par la suite deux tiers de la différence salariale observée entre les deux populations (13%). Les écarts inexplicables pouvant être attribués à la discrimination salariale, sont par ailleurs faibles.

Par rapport à cette littérature, notre travail innove dans différentes directions. D'abord, pour mettre en évidence un lien entre le lieu de résidence et les écarts de salaires, nous comparons des individus qui résident dans des quartiers évalués comme « défavorisés » avec d'autres ne résidant pas dans ce type de quartier. Les quartiers considérés comme « défavorisés » sont déterminés à partir d'un indicateur composite qui reflète leur « qualité » en fonction d'un certain nombre de caractéristiques socio-économiques (la démographie, le niveau d'emploi, le niveau de diplôme, les qualifications etc.). Cette stratégie permet de s'affranchir des définitions légales, comme les zones urbaines sensibles, qui sont plus contraignantes et dont le découpage peut être sujet à critiques.

Par ailleurs, nous analysons les liens potentiels entre le lieu de résidence des individus et les écarts de salaires observés en faisant une distinction entre les hommes et les femmes. Nous partons du postulat que la discrimination en raison du lieu de résidence peut avoir un effet différencié selon le genre des individus. Cette hypothèse renvoie aux résultats d'un testing effectué par L'Horty *et al.* (2011) qui montre que les chances d'être embauché diminuent particulièrement pour les femmes issues des communes perçues comme étant les plus défavorisées.

Ensuite, d'un point de vue méthodologique, notre travail contrôle d'un double biais de sélection. Nous estimons un probit bivarié afin de prendre en compte deux sources de biais potentielles. En premier lieu, on peut suspecter un biais de sélection lié au choix du lieu de résidence effectué par les parents ou les jeunes. Le choix de ce lieu n'est pas effectué de manière aléatoire et est souvent volontaire ou contraint. Il est vraisemblable qu'il existe également des variables observables et/ou inobservables qui influencent le choix de localisation et donc faussent l'analyse. En second lieu, on considère le biais de sélection, dit « biais de participation » lié au fait que nous observons les salaires uniquement pour des personnes qui sont en emploi. Le recours au modèle probit bivarié permet d'expliquer simultanément la probabilité d'habiter dans un quartier défavorisé et la probabilité d'être en emploi. L'estimation de ce modèle nous permet d'obtenir des termes de correction qui sont introduits dans les équations de salaires, afin de contrôler de ces biais de sélection, selon la méthode de Neuman et Oaxaca (2004). Nous mobilisons finalement les méthodes de décomposition afin de décomposer l'écart de salaire observé entre les deux groupes en une part qui s'explique par des différences de caractéristiques individuelles, une part inexplicable liée à des différences dans le rendement de ces mêmes caractéristiques et une part liée à la sélection des jeunes.

La première section présente les bases de données utilisées. La deuxième section décrit les méthodologies utilisées. La troisième section présente de statistiques descriptives et les résultats obtenus. La dernière section conclut.

1. Présentation des données

1.1. L'enquête Génération 2004

Les données de l'enquête Génération 2004 réalisée par le Cereq (Centre d'Etudes et de Recherches sur les Qualifications) concernent un échantillon national de 33 655 jeunes sortis du système éducatif en 2004 de tous niveaux et spécialités de formation. Cet échantillon est représentatif des 750 000 sortants du système scolaire (Céreq, 2008). Le panel d'individus a été interrogé en 2007, soit trois ans après leur sortie du système éducatif.

L'enquête fournit trois bases de données complémentaires. Une première base correspond aux caractéristiques individuelles de jeunes sortants de formation initiale en 2003-2004. Une deuxième base de données correspond aux séquences en entreprise de l'individu. Enfin, une troisième base correspond aux séquences de non-emploi de l'individu. Différents thèmes sont abordés successivement dans ces bases : les caractéristiques individuelles et la situation familiale et de cohabitation ; le parcours scolaire et l'expérience du travail en cours de scolarité ; le calendrier professionnel et les perspectives professionnelles; le sentiment de discrimination sur le marché du travail. Organisée autour de ces trois bases de données, l'enquête Génération 2004 permet de suivre mois par mois la situation des jeunes à l'issue de leur formation initiale.

L'enquête Génération 2004 a été enrichie avec des informations sur la localisation résidentielle des individus suite à un partenariat avec le Secrétariat Général du Comité Interministériel des Villes¹. Ces données permettent de disposer du code IRIS du lieu de résidence des jeunes. Ce sigle signifie « Ilots Regroupés pour l'Information Statistique » et fait référence à un découpage du territoire en mailles de taille homogène (2 000 habitants) développé par l'INSEE. L'IRIS constitue « la brique de base en matière de diffusion de données infra-communales », conformément à l'INSEE. Disposer des codes IRIS permet l'appariement avec le recensement de la population réalisé par l'INSEE en 2007.

1.2. Le recensement de la population 2007 (INSEE)

Le recensement de la population représente une source exhaustive de données à l'IRIS sur les logements, les ménages et certaines caractéristiques socio-économiques des habitants tel que le sexe, l'âge, la catégorie socio-professionnelle, diplôme, situation professionnelle etc. En nous appuyant sur ces données, nous avons pu construire des indicateurs au niveau de l'IRIS, tels que : le taux de chômage de la population de 15-64 ans, le taux de scolarisation de la population de 16-24 ans, la part des logements HLM, la part des propriétaires, la part des étrangers, la part des ouvriers, la part des ménages de six personnes ou plus etc. Toutes ces variables vont nous permettre de caractériser le lieu de résidence des jeunes sortant du système scolaire.

2. Méthodologie

L'objectif de ce travail est d'analyser les effets de vivre dans un quartier considéré comme défavorisé sur la situation observée sur le marché du travail. Nous souhaitons en effet vérifier si des effets se manifestent sur la probabilité d'être en emploi, mais aussi sur les écarts de salaires observés entre individus. Au-delà, nous souhaitons également vérifier si le fait de vivre dans ce type de quartier peut-être une source de discrimination salariale, toutes choses égales par ailleurs.

2.1. Définir les quartiers défavorisés

Pour identifier ces quartiers considérés comme défavorisés, nous cherchons à construire un indicateur synthétique qui permet de mesurer la composition socio-économique du quartier. Cet indicateur est basé sur différentes caractéristiques qui reflètent chacune un aspect différent du quartier. Nous retenons des indicateurs relatifs à la qualification (la part des chefs de ménage sans diplôme ou diplômés du supérieur), à la catégorie socioprofessionnelle (la part des chefs de ménage « ouvriers » ou « cadres »), au statut d'emploi (le taux de chômage des 15-64 ans), à la composition démographique (la part des familles avec 3 enfants ou plus, la part des chefs de ménage étrangers) et au statut résidentiel (la part des logements HLM).

Nous avons recours à une analyse en composantes principales (ACP), avec l'objectif de trouver des facteurs non corrélés qui résument les indicateurs cités précédemment. Nous ne retenons que le premier facteur qui représente plus de 50 % de la variance de nos variables. Cet axe oppose les quartiers avec des caractéristiques favorables (part des chefs de ménage « cadres » et part des chefs de ménage diplômés du supérieur élevées) aux quartiers avec des caractéristiques plus défavorables (taux de chômage et part des chefs de ménages « ouvriers » importants et, dans une moindre mesure, la part des familles nombreuses). Afin d'évaluer la « qualité » d'un quartier, nous retenons la coordonnée de celui-ci sur le premier axe comme indicateur. En outre, les

¹ La Délégation interministérielle à la ville (DIV) a été renommée Secrétariat général du comité interministériel des villes (SG-CIV) depuis le 14 mai 2009, conformément à l'avenant numéro 2 à la convention constitutive du groupement d'intérêt public.

quartiers sont distribués selon une loi normale sur cet axe (voir l'annexe 1). Les plus désavantagés par leur composition sont ceux avec les valeurs les plus élevées, alors que les plus avantagés sont ceux avec les valeurs les plus faibles. Nous considérons comme « défavorisés » tous les quartiers appartenant aux 25% des quartiers qui ont la coordonnée la plus élevée sur ce premier axe (soit lorsqu'elle est supérieur à $x = 0.80$).

L'analyse en composantes principales est effectuée sur l'ensemble des IRIS pour lesquels nous avons identifié au moins un jeune sortant du système scolaire l'année d'interrogation de l'enquête (soit en 2004). Cette méthode permet de voir l'importance de chacune de ces caractéristiques dans notre évaluation de la « qualité » d'un quartier. En revanche, l'indicateur créé ne permet pas de voir quelles sont les caractéristiques, mobilisées pour définir le quartier, qui sont les plus pertinentes pour expliquer la situation d'emploi ou le niveau de salaire.

2.2. Evaluer le rôle du lieu de résidence dans l'explication des écarts de salaire

2.2.1. Méthodes économétriques

Dans ce travail, nous cherchons à comparer les effets d'habiter dans un quartier défavorisé sur les performances sur le marché du travail (et notamment sur les écarts de salaires observés) de jeunes sortant du système scolaire, toutes choses étant égales par ailleurs. Typiquement, l'idée est d'estimer des équations de salaires pour deux groupes d'individus distincts, l'un habitant dans un quartier considéré comme défavorisé tandis que l'autre non. Si on note $w_{1,i}$ et $w_{2,i}$ les salaires mensuels pour les individus habitant respectivement dans des quartiers défavorisés et non-défavorisés, les équations à estimer s'écrivent :

$$\ln w_{1,i} = X'_{1,i} \beta_1 + \varepsilon_{1,i}$$

$$\ln w_{2,i} = X'_{2,i} \beta_2 + \varepsilon_{2,i}$$

Où $\ln w$ représente le logarithme du salaire perçu par les individus, X_i est une matrice de variable explicative, β le vecteur de coefficients correspondants à estimer et ε le terme d'erreur. L'indice $j = 1, 2$ indique le fait de résider dans un quartier défavorisé ou non-défavorisé. Ces équations s'estiment traditionnellement par la méthode des MCO. Toutefois, dans notre cas, les estimations obtenues peuvent se révéler biaisées et inconsistantes. En effet, il importe de noter que l'analyse proposée souffre de deux biais potentiels.

D'abord, il est probable qu'il existe un biais de sélection, dit « biais de participation » lié au fait que nous observons les salaires uniquement pour des personnes qui sont en emploi. Ainsi, il est probable que les individus qui travaillent ont des caractéristiques observables et/ou inobservables différentes de ceux qui ne travaillent pas, ce qui biaiserait nos résultats. Ce problème est généralement pris en compte en incluant une variable explicative supplémentaire qui corrige du biais relatif à la décision de participation. Cette solution proposée par Heckman (1979) permet de corriger du fait que la variable dépendante n'est observée que pour une partie de l'échantillon.

Le second biais de sélection est relatif au lieu de résidence. Il est probable qu'il existe également des variables observables et/ou inobservables qui influencent le choix de localisation et donc faussent l'analyse. Ainsi, une potentielle auto-sélection des individus dans l'espace est susceptible d'influencer nos résultats (Jencks et Mayer, 1990 ; Galster, 2008 etc.) C'est le cas, par exemple, si l'on considère que les parents qui souhaitent optimiser les chances de réussite de leurs enfants, choisissent de résider dans des quartiers où l'environnement est plus favorable que d'autres. Ici le biais de sélection concerne notre variable de traitement qui est donc le lieu de résidence. Dans ce cas, la bonne identification du modèle est assurée par l'ajout de variables instrumentales dans le modèle (Maddala, 1983).

Afin de contrôler du biais de sélection relatif au lieu de résidence (pouvant influencer sur la probabilité d'être en emploi et le niveau de salaire) et du biais de participation (puisque l'on observe le salaire que pour les jeunes en emploi) nous estimons un probit bivarié. *Dans un premier temps*, on met en avant les déterminants du choix du lieu de résidence et, *dans un second temps*, on regarde son effet sur le statut d'emploi. Nous considérons donc le modèle probit bivarié suivant :

$$\begin{cases} L_i^* = \gamma Z_i' + u_i \\ E_i^* = \mu B_i' + v_i \end{cases}$$

Où L_i^* et E_i^* sont des variables latentes qui influencent la probabilité d'être localisé dans un quartier défavorisé et la probabilité d'être en emploi. Z et B représentent des vecteur des caractéristiques et u et v sont les termes d'erreurs de ces deux équations. Ces termes d'erreurs sont supposés suivre une loi normale bivariée de moyenne 0 et dont la matrice de covariance, après normalisation à 1 des éléments diagonaux, s'écrit :

$$\begin{pmatrix} u \\ v \end{pmatrix} \rightarrow N(0, \Sigma) \text{ où } \Sigma = \begin{pmatrix} 1 & \sigma \\ \sigma & 1 \end{pmatrix}$$

A noter que L_i^* et E_i^* étant des variables latentes, elles ne peuvent être observées. Seules les variables dichotomiques suivantes peuvent être représentées :

$$L_i = \begin{cases} 1 \text{ si } L_i^* > 0 \\ 0 \text{ sinon} \end{cases}$$

$$E_i = \begin{cases} 1 \text{ si } E_i^* > 0 \\ 0 \text{ sinon} \end{cases}$$

Où $L_i = 1$ et $L_i = 0$ indiquent que l'individu habite dans un quartier défavorisé ou non-défavorisé. $E_i = 1$ et $E_i = 0$ indiquent si l'individu est en emploi ou non. Notre modèle concerne quatre groupes de populations différents. Un groupe G_1 qui inclut les individus habitant un quartier défavorisé et qui sont en emploi (cas où $L_i = 1$ et $E_i = 1$), un groupe G_2 qui inclut les individus habitant dans un quartier non-défavorisé et qui sont en emploi (cas où $L_i = 0$ et $E_i = 1$), un groupe G_3 qui inclut les individus habitant dans un quartier défavorisé et qui ne sont pas en emploi (cas où $L_i = 0$ et $E_i = 0$) et un groupe G_4 qui inclut les individus habitant dans un quartier non-défavorisé et qui ne sont pas en emploi (cas où $L_i = 0$ et $E_i = 0$).

En résumé, nous considérons un modèle probit à deux équations dans lequel nous supposons que les termes d'erreurs sont corrélés et que la variable dépendante binaire de la première équation (L^*) est un régresseur endogène dans la deuxième équation (E^*). Le probit bivarié nous permet d'estimer les effets des variables explicatives et d'identifier si les termes d'erreurs des deux équations sont corrélés ou non. Si tel n'est pas le cas, alors le probit bivarié est inapproprié et il est préférable d'estimer séparément les deux équations.

L'estimation du modèle précédent par la méthode du maximum de vraisemblance permet d'obtenir les quatre termes de correction de la sélection suivants (également nommés inverses des ratios de Mills):

$$\lambda_{i,11} = \phi(\hat{\gamma}Z_i') \Phi \left(\frac{\hat{\mu}B_i' - \rho \hat{\gamma}Z_i'}{\sqrt{1-\rho^2}} \right) \cdot F(\hat{\mu}B_i', \hat{\gamma}Z_i', \rho)^{-1}$$

$$\lambda_{i,e1} = \phi(\hat{\mu}B_i') \Phi \left(\frac{\hat{\gamma}Z_i' - \rho \hat{\mu}B_i'}{\sqrt{1-\rho^2}} \right) \cdot F(\hat{\mu}B_i', \hat{\gamma}Z_i', \rho)^{-1}$$

$$\lambda_{i,12} = \phi(\hat{\gamma}Z_i') \Phi \left(-\frac{\hat{\mu}B_i' - \rho \hat{\gamma}Z_i'}{\sqrt{1-\rho^2}} \right) \cdot F(-\hat{\mu}B_i', \hat{\gamma}Z_i', -\rho)^{-1}$$

$$\lambda_{i,e2} = -\phi(\hat{\mu}B_i') \Phi \left(\frac{\hat{\gamma}Z_i' - \rho \hat{\mu}B_i'}{\sqrt{1-\rho^2}} \right) \cdot F(-\hat{\mu}B_i', \hat{\gamma}Z_i', -\rho)^{-1}$$

Où ϕ et Φ désignent respectivement la fonction de densité et la fonction de répartition de la loi normale, F représente la fonction de répartition d'une loi normale bivariée.

Dès lors, les équations de salaire précédentes peuvent être réécrites de la manière suivante :

$$\begin{aligned}
E(\ln w_{1,i} | X_i', L_i = 1, E_i = 1) &= X_{1,i}' \beta_1 + E(\varepsilon_{1,i} | L_i = 1, E_i = 1) \\
&= X_{1,i}' \beta + E(\varepsilon_{1,i} | u_i > -\gamma Z_i', v_i > -\mu B_i') \\
&= X_{1,i}' \beta + \sigma_1 \rho_{1u} \lambda_{i,l1} + \sigma_1 \rho_{1v} \lambda_{i,e1} \\
E(\ln w_{2,i} | X_i', L_i = 0, E_i = 1) &= X_{2,i}' \beta_1 + E(\varepsilon_{2,i} | L_i = 0, E_i = 1) \\
&= X_{2,i}' \beta + E(\varepsilon_{2,i} | u_i \leq -\gamma Z_i', v_i > -\mu B_i') \\
&= X_{2,i}' \beta + \sigma_2 \rho_{2u} \lambda_{i,l2} + \sigma_2 \rho_{2v} \lambda_{i,e2}
\end{aligned}$$

Où $\varepsilon_{1,i}$ et $\varepsilon_{2,i}$ correspondent aux termes d'erreurs avec $\varepsilon_{1,i} \approx N(0, \sigma_1^2)$ et $\varepsilon_{2,i} \approx N(0, \sigma_2^2)$. Ils sont supposés être corrélés avec les termes d'erreurs de nos deux équations de sélection (u et v). Ainsi, ρ_{1j} ($j = u, v$) correspond au coefficient de corrélation entre l'équation de sélection correspondante et le terme d'erreur $\varepsilon_{1,i}$ de l'équation de salaire pour les jeunes issus de quartiers défavorisés. Inversement, ρ_{2j} ($j = u, v$) correspond au coefficient de corrélation entre l'équation de sélection correspondante et le terme d'erreur $\varepsilon_{2,i}$ de l'équation de salaire pour les jeunes issus de quartiers défavorisés.

Nous adoptons une extension de la méthode en deux étapes proposée par Heckman (1979) pour l'estimation de ces équations. Dans la première étape, le modèle probit bivarié est estimé par la méthode du maximum de vraisemblance. Son estimation permet de calculer les termes de sélection $\lambda_{i,l1}$, $\lambda_{i,e1}$, $\lambda_{i,l2}$, $\lambda_{i,e2}$. Ensuite, dans la seconde étape, les équations de salaire sont estimées par la méthode des Moindres Carrés Ordinaires (MCO) en incluant les termes de correction comme variables additionnelles. Les termes $\lambda_{i,l1}$ et $\lambda_{i,e1}$ sont inclus dans l'équation de salaire des jeunes issus de quartiers défavorisés, alors que les termes $\lambda_{i,l2}$ et $\lambda_{i,e2}$ sont inclus dans l'équation de salaire des jeunes issus de quartiers non-défavorisés. Les coefficients de ces termes de sélection capturent l'effet de la corrélation des termes d'erreur dans l'estimation des processus de sélection avec celles des salaires. L'estimation de ces nouvelles équations de salaire permet alors d'obtenir des estimations non biaisées de nos coefficients.

2.2.2. Décomposition

Quand les estimations des équations de salaire obtenues sont sans biais et consistantes, les différences entre les deux types de quartier peuvent être décomposées en plusieurs parties. Nous nous appuyons sur une méthode de décomposition proposée par Neuman et Oaxaca (2004). Cette méthode permet de tenir compte du processus de sélection au sein de la décomposition par la combinaison de la méthode d'Oaxaca et Blinder (1973) et de la méthode de correction du biais de sélection développée par Heckman (1979). Selon cette méthode, les écarts de salaires peuvent être séparés en trois termes distincts. Formellement, la décomposition des écarts de salaire s'écrit de la manière suivante :

$$\begin{aligned}
\ln \bar{w}_1 - \ln \bar{w}_2 &= (\bar{X}_1 \hat{\beta}_1 + \hat{\sigma}_1 \hat{\rho}_{1u} \hat{\lambda}_{i,l1} + \hat{\sigma}_1 \hat{\rho}_{1v} \hat{\lambda}_{i,e1}) - (\bar{X}_2 \hat{\beta}_2 + \hat{\sigma}_2 \hat{\rho}_{2u} \hat{\lambda}_{i,l2} + \hat{\sigma}_2 \hat{\rho}_{2v} \hat{\lambda}_{i,e2}) \\
&= \underbrace{(\bar{X}_1 - \bar{X}_2) \hat{\beta}_1}_{\text{Partie expliquée}} + \underbrace{\bar{X}_2 (\hat{\beta}_1 - \hat{\beta}_2)}_{\text{Partie inexpliquée}} + \underbrace{[\hat{\sigma}_1 (\hat{\rho}_{1u} \hat{\lambda}_{i,l1} + \hat{\rho}_{1v} \hat{\lambda}_{i,e1}) - \hat{\sigma}_2 (\hat{\rho}_{2u} \hat{\lambda}_{i,l2} + \hat{\rho}_{2v} \hat{\lambda}_{i,e2})]}_{\text{Sélectivité}}
\end{aligned}$$

Où $\ln \bar{w}$ est le salaire moyen, exprimé en log et prédit par le modèle. \bar{X} correspond aux moyennes des caractéristiques, $\hat{\beta}$ représente les coefficients estimés au sein de l'équation de salaire et les $\hat{\lambda}$ correspondent à la moyenne des termes de correction estimés.

La décomposition de l'écart moyen des salaires comprend donc une « partie expliquée » représentant l'écart dû aux différences de caractéristiques entre les jeunes issus de quartiers défavorisés ou non, une composante reflétant la contribution de la sélection (dans les deux types de quartier et dans l'emploi) et enfin une « partie

inexpliquée » correspondant à de la discrimination. La discrimination est considérée comme l'écart de rendement des caractéristiques des individus selon leur quartier d'origine.

2.2.3. Choix des variables et identification des modèles

Dans ce travail, nous estimons différents modèles et spécifications afin d'estimer les probabilités conjointes de vivre dans un quartier considéré comme défavorisé et d'être en emploi, d'une part, et les équations de salaire, d'autre part. Toutefois, les estimations de ces différents modèles requièrent un certain nombre d'hypothèses d'identification.

En premier lieu, pour assurer l'identification du modèle probit bivarié, il est nécessaire de disposer d'au moins une variable supposée influencer la probabilité de vivre dans un quartier défavorisé mais pas la probabilité d'être en emploi. Dit autrement, il faut une variable comprise dans Z_i mais pas dans B_i ni dans le vecteur X_i de variables explicatives des équations de salaire. Concrètement, c'est le cas si l'on observe que $\rho_{uv} \neq 0$. Si $\rho_{uv} = 0$, le probit bivarié est inapproprié et il est plus judicieux d'estimer séparément les deux équations, avec L_i^* introduite simplement comme covariable dans la seconde équation E_i^* .

En second lieu, il est également nécessaire de disposer de variables supposées influencer le fait d'être en emploi mais pas les salaires. Ces variables sont donc incluses dans le vecteur B_i de l'équation E_i^* mais pas dans le vecteur X_i des équations de salaire ($\ln w_{i,1}$ et $\ln w_{i,2}$).

Afin d'expliquer la probabilité d'habiter dans un quartier défavorisé pour notre échantillon de jeunes sortant du système scolaire en 2004, nous mobilisons différentes variables qui renseignent sur les caractéristiques individuelles, sur les caractéristiques du ménage ou sur la localisation (voir l'annexe B pour l'ensemble des variables retenues et les statistiques qui leurs sont associées). Nous retenons les caractéristiques individuelles suivantes : le genre, l'âge, l'âge au carré², le niveau de diplôme le plus élevé obtenu à la sortie des études en 2004, l'expérience professionnelle antérieure³. Pour les caractéristiques du ménage, nous retenons : la nationalité des parents (français ou non), la situation d'emploi du chef de ménage à la fin des études du jeune. Pour les caractéristiques relatives à la localisation, nous retenons : la région de résidence afin de contrôler d'une potentielle hétérogénéité inobservée au niveau des régions (qui peut s'expliquer par des politiques régionales différentes) et la distance en kilomètre au plus grand centre d'emploi localisé dans un rayon de 50 kilomètres (évaluée en distance euclidienne).

Les variables instrumentales retenues pour expliquer la probabilité de vivre dans un quartier défavorisé, mais qui ne doivent pas influencer sur la probabilité d'être en emploi sont des variables d'aménités du quartier. Nous retenons le nombre de terrains multisports, le nombre d'aires de petits jeux, ainsi que le nombre de cinémas dans le quartier de résidence du jeune (qu'ils vivent encore chez leurs parents ou non en 2004)⁴. Ces variables sont susceptibles de déterminer le choix du lieu de résidence, car elles représentent des aménités que les parents peuvent souhaiter pour leurs enfants. Elles déterminent les possibilités de distraction et d'occupation des enfants et sont, par conséquent, des aménités demandées et recherchées. De plus, bien qu'elles soient susceptibles de rendre compte d'un certain niveau de développement économique et social du quartier de résidence des jeunes, elles n'ont pas d'influence directe sur la probabilité d'être en emploi.

Enfin, nous retenons également des variables d'exclusion qui déterminent la probabilité d'être en emploi, mais qui n'ont pas d'effet *a priori* sur les écarts de salaires observés entre les individus. Traditionnellement, les informations relatives au statut marital ou le nombre d'enfants sont des bons candidats (Hyslop, 1999 ; Buchinsky *et al.*, 2010 etc.). Elles permettent d'expliquer la décision de participer au marché du travail, mais n'ont pas d'effet direct sur le salaire à l'embauche. Concrètement, nous utilisons le nombre d'enfants pour le jeune au

² Cette variable permet de déceler une éventuelle relation non linéaire entre l'âge et notre principale variable à expliquer.

³ Cette catégorie renseigne sur d'éventuels emplois réguliers effectués au cours des études. Elle ne prend pas en compte les emplois saisonniers comme les jobs d'été, par exemple.

⁴ Il importe de noter que ces informations sont extraites de la Base permanente des équipements 2012, réalisée par l'INSEE. Les informations sur les équipements sont donc ultérieures à la date de l'enquête Génération. Toutefois, nous n'avons pas pu accéder à des versions plus anciennes de cette source de données. En outre, on peut raisonnablement supposer que ce type d'équipements ne varie pas trop dans le temps.

moment de l'enquête ainsi que sa situation familiale (le jeune vit seul, vit chez ses parents ou vit en concubinage). A noter que nous introduisons notre indicateur de quartier défavorisé dans l'équation de probabilité d'être en emploi. L'idée étant effectivement de mettre en évidence les effets de vivre dans ce type de quartier sur la situation d'emploi des jeunes.

Pour les équations de salaires, nous retenons un certain nombre de variables déjà utilisées dans les équations précédentes. En effet, à l'exception des variables sur les aménités, sur le nombre d'enfants et sur la situation familiale, les variables mobilisées sont celles du modèle probit bivarié. Toutefois, puisque ces équations de salaires ne concernent que des personnes déjà en emploi, nous ajoutons un certain nombre d'informations relatives à ces emplois et qui peuvent expliquer le salaire observé. On retient l'information relative au temps de travail (temps partiel ou temps plein) ainsi que la catégorie socio-professionnelle de l'emploi (de la catégorie ouvrier à celle d'artisan/chef d'entreprise)

3. Résultats

3.1. Statistiques descriptives

Avant de vérifier si oui ou non le lieu de résidence impacte sur les écarts de salaire observés entre les individus et de tenter de quantifier son effet, nous présentons quelques statistiques descriptives. Le Tableau 1 présente les différences de caractéristiques entre des individus qui résident dans les quartiers que l'on a identifié comme « défavorisés » et ceux qui ne résident pas dans ce type de quartier. Les variables présentées sont celles retenues dans le probit bivarié⁵.

D'emblée, il ressort de nos différents tests d'égalité de moyennes et de proportions que la très grande majorité des variables retenues affiche des écarts significatives entre les deux échantillons (très souvent au seuil de 1%). La moyenne d'âge est plus faible dans les quartiers « défavorisés » que les autres (20,9% contre 21,8%). La proportion de femmes y est également plus faible (47,2% contre 48,8%). Les jeunes sortant du système scolaire issus des quartiers « défavorisés » sont plus souvent non-diplômés (20,4% contre 11,2%) et atteignent moins les niveaux de diplômes les plus élevés. Par exemple, moins de 8% des jeunes de ces quartiers atteignent le niveau Bac+4 ou Bac+5, contre une proportion de plus de 13% dans les quartiers « non-défavorisés ».

On observe également des corrélations importantes entre le type de quartier de résidence et les caractéristiques des parents du jeune. Ainsi, dans ces quartiers, les parents sont plus souvent de nationalité étrangère : respectivement 8,8% et 7,6% pour le père et la mère dans quartiers « défavorisés » contre 6% et 5,6% dans quartiers non-défavorisés. Le chef de ménage est plus souvent chômeur ou inactif dans ce type de quartier que dans les autres (on observe 4,1% de chef de ménage au chômage contre 2,2%). Enfin, il importe de noter que les jeunes issus des quartiers « défavorisés » sont encore majoritairement chez leurs parents. La proportion de jeunes chez leurs parents est de plus de 51% contre à peine 39% pour les autres. Ces derniers vivent donc plus souvent seuls ou en couple.

Le fait d'avoir un enfant ou non ainsi que le nombre d'enfants ne révèlent pas de différences significatives entre les deux groupes de jeunes. Dans les deux groupes, la proportion de jeunes ayant au moins un enfant est proche de 11%. De même, le nombre moyen d'enfants (si on exclut les jeunes qui n'ont pas d'enfant) est proche de 1,2. Ces faibles valeurs s'expliquent assez logiquement par l'âge moyen des individus lorsqu'ils sortent du système scolaire. En ce qui concerne l'expérience professionnelle lors des études scolaires, peu d'écarts sont observés. Il ressort tout de même que les jeunes sortants issus de quartiers « défavorisés » exercent moins souvent un emploi régulier au cours de leurs études. Ce résultat peut sembler paradoxal, puisque l'on peut imaginer que ce sont ces jeunes qui ont davantage besoin de financer leurs études si leurs parents ont des revenus plus modestes que les parents dans les quartiers moins défavorisés.

Les écarts et différences observées pour la situation d'emploi actuelle témoignent d'importantes différences entre les deux groupes et justifient le recours à un probit bivarié pour contrôler notamment du biais de sélection lié au lieu de résidence. En effet, on observe que la proportion de jeunes en emploi issus des quartiers « défavorisés » est plus faible que celle dans les autres quartiers. Ces proportions sont respectivement de 75,3% et 82,4%. Il

⁵ Pour avoir la liste exhaustive des variables mobilisées dans nos différentes estimations, le lecteur peut se référer à l'Annexe B.

semble bien y avoir un lien entre les caractéristiques du lieu de résidence à la sortie des études (en 2004) et le fait de trouver un emploi dans les trois années qui suivent. De même, le salaire à l'embauche (du premier emploi) des jeunes de ces quartiers est plus faible que les autres⁶. Selon nos calculs, il est de près de 1 128 euros contre 1 212 euros dans les autres quartiers. Ces valeurs peuvent paraître faibles, mais elles regroupent des jeunes de tous niveaux scolaires, y compris des non-diplômés. Une fois que ce constat est posé, ce qui nous intéresse est de savoir si ces écarts de salaires s'expliquent par des différences de caractéristiques observables entre les jeunes des deux types de quartiers (puisque le Tableau 1 révèle que ces caractéristiques sont souvent très différentes) ou par des différences dans les rendements de ces caractéristiques et notamment du lieu de résidence. C'est notamment l'objet de la prochaine section.

Tableau 1 - Les différences de caractéristiques entre quartiers « défavorisés » et « non-défavorisés »

Variables	Quartier "défavorisé"		Quartier "non-défavorisé"	
	Moyenne	Ecart-type	Moyenne	Ecart-type
Situation en emploi***	75,3%	0,005	82,4%	0,003
Salaire à l'embauche*** (en ETP)	1 128,10	6,05	1 212,80	3,873
Etre une femme**	47,2%	0,006	48,8%	0,003
Age***	20,9	0,037	21,8	0,022
<i>Niveau de sortie des études</i>				
Non diplômé***	20,4%	0,005	11,2%	0,002
CAP-BEP-MC***	22,6%	0,005	18,7%	0,003
Bac**	24,7%	0,005	23,5%	0,003
Deug, BTS***	16,3%	0,004	21,0%	0,003
Licence L3***	5,3%	0,003	6,9%	0,002
Maîtrise, DEA***	7,9%	0,003	13,1%	0,002
Doctorat	2,8%	0,002	5,6%	0,002
<i>Expérience pro. Antérieure : emploi régulier</i>				
Oui, plusieurs	2,4%	0,002	2,4%	0,001
Oui, un seul	10,6%	0,004	11,2%	0,002
Non***	71,3%	0,005	68,9%	0,003
<i>Nationalité père : autre que française***</i>	8,8%	0,003	6,0%	0,002
<i>Nationalité mère : autre que française***</i>	7,6%	0,003	5,6%	0,002
<i>Situation du chef de ménage à la fin des études</i>				
Travaille***	75,4%	0,005	82,3%	0,003
Au chômage***	4,1%	0,002	2,2%	0,001
Inactif ***	11,7%	0,004	9,0%	0,002
<i>Caractéristiques du ménage</i>				
Présence d'au moins un enfant	11,5%	0,004	10,9%	0,002
Nombre d'enfants (si >0)	1,2	0,017	1,26	0,011
Vit chez ses parents***	51,1%	0,006	39,4%	0,003
Vit en couple***	32,3%	0,006	38,8%	0,003
Vit seul***	16,5%	0,004	21,7%	0,003
<i>Caractéristiques locales</i>				
Dist. au ctr. d'emploi le plus proche (en km)***	24,6	0,192	21,6	0,111
Terrains multisports***	49,3%	0,006	24,6%	0,003
Aires de petits jeux***	16,9%	0,004	9,3%	0,002
Cinéma***	3,0%	0,002	4,2%	0,001
Observations	6 976		20 392	

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : Le tableau présente les résultats de tests d'égalité de moyenne et de tests d'égalité de proportions.

* : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.

⁶ Les salaires sont ramenés en équivalent temps plein pour qu'ils soient directement comparables. Toutefois, nous ne disposons pas des heures exactes mensuelles. Nous avons l'information sur le fait que le jeune a travaillé à 80%, entre 50 et 80% ou moins de 50% d'un temps plein (soit 35 heures). Nous prenons le milieu de chaque classe pour déterminer un pourcentage de temps de travail et donc le salaire en équivalent temps plein.

Pour finir, nous nous intéressons aux différences de caractéristiques des quartiers de ces deux échantillons. Tout d'abord, nous constatons que la distance au centre d'emploi le plus proche est plus élevée pour les quartiers considérés comme « défavorisés ». Ce constat s'explique par le fait que ces quartiers sont plus généralement enclavés et à la périphérie des grandes villes et des centres d'emplois. En outre, ces quartiers abritent plus généralement des terrains multisports ou des aires de petits jeux que les autres quartiers (respectivement 49,3% et 16,9% contre 24,6% et 9,3%). Ce type d'installation est en effet plus courant dans les quartiers denses en population et a pour objectif notamment de divertir les plus jeunes. En revanche, les cinémas qui représentent un loisir plus coûteux pour les usagers sont moins développés dans ces quartiers que les autres (3% contre 4,2%).

3.2. Résultats des estimations

Les méthodes de décomposition de Neuman et Oaxaca (2004) supposent, en premier lieu, d'estimer des équations de salaire. Les résultats de cette estimation sont présentés dans le Tableau 3. Toutefois, comme précisé antérieurement, il importe de considérer des potentiels biais de sélection. L'un est lié au fait que les ménages ne choisissent pas aléatoirement leur lieu de résidence. L'autre est lié au fait que le salaire n'est observé que pour ceux déjà en emploi. Pour cette raison, nous estimons d'abord un modèle probit bivarié qui permet d'expliquer la probabilité d'être en emploi (Colonne 2 du Tableau 2) conditionnellement à la probabilité de vivre dans un quartier « défavorisé » (Colonne 1).

En premier lieu, nous observons que le fait d'être une femme augmente la probabilité de résider dans un quartier considéré comme « défavorisé ». A l'inverse, cette probabilité est décroissante avec l'âge de l'individu. Différentes hypothèses permettent d'expliquer ce constat. Il est probable que les plus jeunes habitent encore chez leurs parents et n'ont donc pas la possibilité de déménager pour vivre dans des quartiers de meilleure qualité (car la proportion de jeunes habitant encore chez leurs parents est plus élevée dans ce type de quartier). De même, l'âge étant souvent positivement corrélé avec le revenu, les plus âgés disposeraient de ressources plus importantes pour envisager de déménager vers d'autres quartiers.

Assez logiquement, nous retrouvons un effet négatif du diplôme. Plus l'individu est diplômé, plus sa probabilité de résider dans ce type de quartier diminue. Les effets les plus élevés se retrouvent effectivement pour les niveaux tels que « Bac +4 », « Bac +5 » et « Doctorat ». Le fait d'avoir des parents de nationalité étrangère contribue également à augmenter la probabilité de vivre dans un quartier « défavorisé ». Lorsque le chef de ménage est inactif ou chômeur, les risques sont plus importants d'habiter dans ces quartiers. Là encore, nous retrouvons un probable effet du revenu. Puisque l'on peut considérer que les personnes sans emploi ont moins la capacité de déménager. Il est aussi concevable qu'il y ait un problème de causalité inverse. C'est-à-dire que c'est parce que les individus habitent dans ce type de quartier qu'ils sont effectivement sans emploi. Si nous contrôlons de ce phénomène pour les jeunes de notre échantillon, notamment avec le probit bivarié, nous ne le faisons pas pour les parents.

Nous retrouvons un phénomène qui est déjà mis en avant avec les statistiques descriptives de la section précédente. En effet, il ressort que le fait de vivre chez ses parents augmente le risque d'habiter dans un quartier « défavorisé » par rapport à la référence qui est le fait de vivre en couple. Bien que les statistiques ne montrent pas d'écart significatif pour le fait d'avoir un enfant entre les deux groupes, les estimations montrent ici qu'avoir un enfant augmente encore la probabilité de vivre dans des quartiers de moindre qualité. Il est probable que les ménages qui ont un ou plusieurs enfants recherchent des logements avec davantage d'espace et ces logements sont certainement moins coûteux dans ces quartiers.

Les variables retenues comme caractéristiques locales démontrent que les quartiers « défavorisés » sont plus souvent éloignés des centres d'emplois, comme le suggère le coefficient positif et significatif pour la distance en kilomètres au centre d'emploi le plus proche. Ils sont aussi plus souvent équipés en aire de petits jeux et en terrains multisports, mais moins en cinéma. Ces variables d'aménités du quartier représentent nos instruments pour l'identification du modèle probit bivarié. Il est donc plutôt rassurant que l'on retrouve des effets significatifs sur la probabilité de vivre dans tel ou tel quartier⁷. Les indicatrices pour les vingt-deux régions montrent que les jeunes qui habitent en région Ile-de-France sont ceux qui ont la plus forte probabilité de vivre dans un quartier « défavorisé ».

⁷ En outre, il n'y a pas de raison a priori pour que ces caractéristiques jouent directement sur la probabilité d'être en emploi. Ce qui est une des conditions pour que nos instruments soient valides.

Enfin, le terme σ_{12} est positif et significatif. Il représente le coefficient de corrélation entre les résidus de chacune des deux équations. Il est statistiquement et significativement différent de 0 au seuil de 1%. Il confirme le fait que la localisation est potentiellement déterminée simultanément avec la probabilité d'être en emploi et, par la même occasion, confirme le recours aux méthodes qui contrôlent ce biais. Le signe est attendu car il révèle une corrélation positive. Ce qui revient à dire que les jeunes, qui ont une probabilité plus forte de d'être en emploi que celle expliquée par leurs caractéristiques observables, sont plus enclins à vivre dans un quartier « défavorisé ».

Tableau 2 – Résultats du probit bivarié

<i>Variable</i>	Probabilité d'habiter un quartier défavorisé		Probabilité d'être en emploi	
	Coefficient estimé	Ecart-type	Coefficient estimé	Ecart-type
Constante	-0,999**	0,389	-3,793***	0,371
Etre une femme	0,039***	0,019	-0,313***	0,019
Age	-0,142***	0,029	0,213***	0,033
(Age)²/100	0,250***	0,063	-0,364***	0,070
Niveau de sortie des études				
Non diplômé	0,275***	0,032	-0,176***	0,034
CAP-BEP-MC	0,085**	0,028	0,213***	0,028
Bac	<i>Réf.</i>		<i>Réf.</i>	
Deug, BTS	-0,149***	0,029	0,463***	0,032
Licence L3	-0,189***	0,042	0,088**	0,044
Maîtrise, DEA	-0,257***	0,037	0,130***	0,039
Doctorat	-0,370***	0,059	0,056	0,062
Expérience prof. antérieure : emploi régulier				
Oui, plusieurs	0,091	0,058	0,061	0,064
Oui, un seul	0,037	0,029	0,109***	0,031
Non	<i>Réf.</i>		<i>Réf.</i>	
Nationalité père : autre que française				
	0,240***	0,037	-0,028	0,039
Nationalité mère : autre que française				
	0,131***	0,039	-0,082**	0,039
Situation du chef de ménage à la fin des études				
Travaille	<i>Réf.</i>		<i>Réf.</i>	
Au chômage	0,354***	0,051	-0,105*	0,055
Inactif	0,251***	0,030	-0,083**	0,033
Caractéristiques du ménage				
A au moins un enfant	0,121***	0,031	-0,418***	0,033
Vit chez ses parents	0,096***	0,023	-0,467***	0,026
Vit en couple	<i>Réf.</i>		<i>Réf.</i>	
Vit seul	-0,011	0,026	-0,189***	0,028
Indicatrices de région				
	<i>Oui</i>		<i>Oui</i>	
Dist. au ctr. d'emploi le plus proche (en km)	0,011***	0,001	0,003***	0,001
Vit dans un quartier défavorisé			-0,446***	0,105
Aménités du quartier				
Terrains multisports	0,192***	0,012		
Aires de petits jeux	0,047**	0,021		
Cinémas	-0,191***	0,040		
σ_{12}	0,212***	0,065		
Log Likelihood	-25 535,884			
Observations	26 940			

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : * : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.

En ce qui concerne les variables expliquant la probabilité d'être en emploi, nous retrouvons des résultats assez classiques et attendus. On observe que le fait d'être une femme diminue la probabilité d'être en emploi, bien que l'on ne sache pas vraiment dire pourquoi. Cela peut s'expliquer par de la discrimination de la part des employeurs, par un choix délibéré de ne pas travailler pour s'occuper des enfants etc. L'âge joue un effet positif,

mais le signe négatif et significatif pour la variable au carré témoigne d'une relation convexe. L'âge jouerait positivement avec un maximum observé à l'âge de 29 ans environ.

Sans surprise également, le niveau de diplôme joue positivement sur les chances d'être en emploi. Néanmoins, la relation observée entre diplôme et statut d'emploi est loin d'être linéaire. En effet, si le fait d'être non diplômé diminue bien la probabilité d'être en emploi par rapport à la référence qui est d'avoir le Bac, l'effet le plus favorable n'est pas observé pour le niveau de diplôme le plus élevé. Avoir un Bac +3 ou au-delà augmente bien les chances de trouver un emploi, mais moins que le simple fait d'avoir un niveau Bac +2 (Deug ou BTS). Enfin, les résultats plus favorables de filières telles que les CAP ou BEP, par rapport au Bac, peut s'expliquer par le fait qu'elles correspondent à des filières professionnelles qui permettent souvent une bonne adéquation avec le marché du travail. En revanche, on n'observe pas ou peu d'effets significatifs concernant l'expérience antérieure des individus obtenue durant les études.

Les caractéristiques des parents, notamment leur nationalité et/ou leur statut d'emploi semblent jouer un rôle dans les chances d'accès à l'emploi. Par exemple, le fait d'avoir un père étranger ou le chef de ménage sans emploi affecte négativement la probabilité de trouver un emploi. On peut potentiellement imaginer qu'un chef de ménage sans emploi n'incite pas forcément le jeune à chercher lui-même un emploi, notamment si ce dernier suit explicitement ou implicitement le modèle que représentent ses parents.

Il ressort également de nos estimations que le fait de vivre dans un quartier considéré comme « défavorisé » diminue aussi les chances d'être en emploi. Ce résultat qui est conforme à la littérature en économie urbaine (voir notamment les revues de Gobillon *et al.*, 2007 ; Galster, 2012) regroupe une multitude d'effets difficiles sinon impossibles à départager. Cet effet négatif peut en effet s'expliquer par des effets de pairs (Crane, 1991 ; Benabou, 1993). C'est le cas si le comportement des voisins influe sur le propre comportement d'un individu. Dans ce cas, des voisins chômeurs n'inciteraient pas à chercher un emploi. Cela peut aussi s'expliquer par des effets de ségrégation ou de stigmatisation de la part des employeurs potentiels qui seraient réticents à embaucher des individus de quartiers stigmatisés (Holzer et Ihlanfeldt, 1998). Enfin, cela peut s'expliquer par un réseau social qui serait de moindre qualité (Holzer, 1988 ; Selod et Zenou, 2006). Or, le réseau est souvent un élément clef dans la recherche d'emploi, notamment pour les emplois les moins qualifiés. L'hypothèse d'un éloignement des quartiers « défavorisés » aux opportunités d'emplois (Kain, 1968) est à écarter, puisque nous contrôlons déjà de ce problème avec notre indicateur de la distance au centre d'emploi le plus proche.

Enfin, nos instruments mobilisés pour contrôler du biais de participation semblent bien affecter la probabilité d'être en emploi. Les coefficients pour les variables relatives à la présence d'enfants ou la situation familiale sont significatifs et vont dans le sens attendu. Les résultats montrent notamment que le nombre d'enfants joue négativement sur les chances d'être en emploi. Ce résultat s'explique évidemment par le fait qu'avoir des enfants impose de nombreuses contraintes en termes de disponibilité et de temps, ce qui est source d'empêchement dans l'accès à l'emploi. Pour finir, le fait de vivre seul ou avec ses parents constitue un désavantage dans la recherche d'emploi par rapport à un jeune qui serait déjà en couple. L'effet négatif s'explique assez bien pour les jeunes qui vivent chez leurs parents puisque, dans ce cas, l'incitation financière à avoir un emploi est moins importante. En revanche, il est difficile de trouver une explication pour les jeunes qui vivent seuls.

Le tableau 3 présente les résultats des équations de salaire. Ces équations sont estimées séparément pour les jeunes qui vivent dans des quartiers « défavorisés » et pour ceux qui vivent dans des quartiers « non-défavorisés ». L'Annexe C présente les résultats d'une estimation de l'équation de salaires pour l'échantillon global, tout en prenant en compte la correction des deux biais de sélection précédemment évoqués. Il importe de noter que, dans le Tableau 3, les tailles d'échantillons ne sont pas les mêmes. L'échantillon de jeunes issus des quartiers « défavorisés » est constitué de 5 480 individus, tandis que l'autre échantillon est constitué de plus de 16 000 observations⁸.

⁸ A noter que la taille de l'échantillon global est de 21 923 observations et diffère de l'échantillon utilisé pour le probit bivarié car nous raisonnons ici uniquement sur des individus déjà en emploi.

Tableau 3 – Résultats des équations de salaires selon le type de quartier

Variable	Quartier "défavorisé"		Quartier "non-défavorisé"	
	Coefficient estimé	Ecart-type	Coefficient estimé	Ecart-type
Constante	5,831***	0,204	6,338***	0,123
Etre une femme	-0,068***	0,01	-0,043***	0,006
Age	0,064***	0,015	0,053***	0,009
(Age)²/100	-0,102***	0,031	-0,078***	0,018
Niveau de sortie des études				
Non diplômé	0,075***	0,016	0,051***	0,01
CAP-BEP-MC	0,046***	0,013	0,048***	0,008
Bac	<i>Réf.</i>		<i>Réf.</i>	
Deug, BTS	0,115***	0,015	0,089***	0,009
Licence L3	-0,004	0,021	-0,016	0,011
Maîtrise, DEA	0,042**	0,019	0,048***	0,01
Doctorat	0,102**	0,033	0,080***	0,016
Expérience prof. antérieure : emploi régulier				
Oui, plusieurs	-0,054**	0,019	-0,046**	0,015
Oui, un seul	0,001	0,013	-0,020**	0,007
Non	<i>Réf.</i>		<i>Réf.</i>	
Nationalité père : autre que française	0,015	0,017	0,016	0,01
Nationalité mère : autre que française	0,023	0,018	0,018*	0,01
Situation du chef de ménage à la fin des études				
Travaille	<i>Réf.</i>		<i>Réf.</i>	
Au chômage	0,04*	0,021	-0,021	0,016
Inactif	0,001	0,013	-0,006	0,008
Caractéristiques de l'emploi				
Temps plein (<i>Réf.</i> Temps partiel)	0,046***	0,01	0,036***	0,006
Artisan, chef d'entreprise	0,111**	0,04	0,150***	0,025
Cadre	0,202***	0,021	0,251***	0,009
Profession intermédiaire	<i>Réf.</i>		<i>Réf.</i>	
Employé	-0,107***	0,012	-0,132***	0,007
Ouvrier	-0,062***	0,012	-0,079***	0,007
Indicatrices de région				
Dist. au ctr. d'emploi le plus proche (en km)	0,001	0,001	0,001	0,001
λ_l	0,482**	0,158	0,175*	0,105
λ_e	0,064	0,111	-0,219***	0,062
R²	0,208		0,301	
Observations	5 480		16 443	

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : La variable à expliquer représente les salaires mensuels du premier emploi obtenu. Ils sont exprimés en log et ont été transformés en équivalent temps plein.

* : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.

Quelles que soient l'équation de salaire, là encore, les résultats vont dans le sens attendu. Nous retrouvons des résultats fréquemment observés sur ces questions dans la littérature économique. Les femmes ont ainsi des salaires plus faibles que les hommes. Cela semble particulièrement vrai pour les jeunes de quartiers « défavorisés ». Dans ce cas, le fait d'être une femme diminue le salaire de 6,8% par rapport aux hommes contre 4,3% pour les autres. L'âge joue positivement et les résultats pour l'âge au carré montrent que la relation est convexe. On peut estimer l'âge pour lequel le niveau de salaire est le plus important à 31 ans environ pour le premier groupe de quartiers et 33 ans environ pour le second.

Le niveau de diplôme présente des effets quelque peu ambigus et difficiles à interpréter. En effet, on n'observe pas forcément une relation positive entre le niveau de diplôme et le salaire à l'embauche. Assez curieusement, il ressort que le fait de ne pas avoir de diplôme ou d'être faiblement diplômé est plus favorable que d'avoir le Bac. L'explication pour les filières professionnelles est très probablement similaire à celle qui justifie la probabilité élevée de trouver un emploi (dans le tableau précédent). Les niveaux de diplômes les plus favorables sont finalement les équivalents du Bac +2 (Deug et BTS) ou encore le Doctorat. Ces constats sont valables pour les

deux groupes de quartiers. On peut également observer que l'expérience acquise lors des études a un impact négatif sur les salaires ultérieurs, ce qui peut être perçu comme contre-intuitif. En effet, on considère souvent que des emplois réguliers lors des études sont valorisables dans le futur, car ils témoignent d'une certaine activité du jeune sortant du système scolaire. Les caractéristiques des parents n'ont plus d'influence ou affichent des effets faiblement significatifs (au seuil de 10%) si ce n'est le fait d'avoir un père inactif.

Sans surprise, nous retrouvons un lien étroit entre la catégorie socio-professionnelle de l'emploi obtenu et le niveau de salaire observé. L'effet le plus important est relevé pour la catégorie « cadre », puisqu'avoir un emploi de cadre augmente le salaire de 20% pour les jeunes de quartiers « défavorisés » et de plus de 25% pour les jeunes des autres quartiers (par rapport à la référence qui est d'avoir un emploi de profession intermédiaire). Inversement, l'effet le plus défavorable en termes de salaires n'est pas observé pour les ouvriers, mais les employés. Dans ce cas, le fait d'avoir un emploi de cette catégorie diminue le salaire de 10,7% pour les jeunes de quartiers « défavorisés » contre près de 13,2% pour les autres.

Enfin, les résultats montrent que le coefficient de corrélation entre les termes d'erreur de l'équation de sélection du lieu de résidence et de l'équation de salaire dans les deux groupes est significativement différent de zéro (aux seuils de 5% et 10% respectivement). De même, le coefficient de corrélation entre les termes d'erreur de l'équation participation et de l'équation de salaire dans les quartiers « non-défavorisés » est significativement différent de zéro (au seuil de 1%). Ces résultats confirment l'existence d'effets de sélection et justifie ainsi que l'on en tienne compte dans nos estimations.

3.3. Résultats des décompositions

Nous avons recours aux décompositions proposées par Neuman et Oaxaca (2004) afin de voir dans quelle mesure les écarts de salaires qui ne s'expliquent pas par des différences de caractéristiques observables sont attribuables au lieu de résidence et à une éventuelle discrimination qui lui serait associée. Les Tableaux 4 et 5 présentent la synthèse des résultats de ces décompositions. Les équations de salaire, avec correction du biais de sélection du lieu de résidence et du biais de participation, sont également estimées pour des sous-échantillons d'hommes et de femmes⁹. L'idée étant de voir si les effets sont les mêmes pour les hommes et les femmes pris séparément. Nous partons du postulat que la discrimination en raison du lieu de résidence peut avoir un effet différencié selon le genre des individus. Cette hypothèse renvoie aux résultats d'un testing effectué par L'Horty *et al.* (2011) qui montre que les chances d'être embauché diminuent particulièrement pour les femmes issues des communes perçues comme étant les plus défavorisées. Nous appliquons ensuite les méthodes de décomposition telles qu'elles ont été présentées précédemment. Le Tableau 4 présente les résultats de la décomposition des écarts de salaires sans prise en compte des biais de sélection, tandis que le Tableau 5 présente les résultats en tenant compte de ces biais.

Tableau 4 - Décomposition des écarts de salaires (sans correction des biais de sélection)

Groupe	Ecart de salaire	Ecart expliqué	Ecart non-expliqué	Observations
Population totale	0,050***	0,046***	92,0%	0,004 8,0% 21 946
Hommes	0,043***	0,045***	95,7%	-0,002 4,3% 11 645
Femmes	0,061***	0,050***	82,0%	0,011 18,0% 10 301

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : La variable relative au salaire est exprimée en log. Ce tableau synthétise les résultats de décompositions pour trois échantillons différents : la population totale, les hommes uniquement et les femmes uniquement.

* : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.

Les différences de salaires estimées entre les jeunes des deux groupes de quartiers peuvent être décomposées en une partie « expliquée » qui résulte des différences dans les caractéristiques des individus et une partie « inexpliquée » qui correspond à de la discrimination (la discrimination étant considérée comme l'écart de rendement des caractéristiques des individus selon leur quartier d'origine).

Si l'on ne tient pas compte des biais de sélection, nous trouvons que la partie « inexpliquée » des écarts attribuable à de la discrimination est faible et très inférieure à celle qui résulte des écarts dans les caractéristiques des jeunes. En effet, plus de 90% des écarts de salaires se justifient par des différences au

⁹ Les résultats ne sont pas détaillés ici afin de ne pas alourdir la présentation (voir l'Annexe D), mais ils restent globalement les mêmes pour les deux populations.

niveau des caractéristiques des individus ou de l'emploi occupé. Cette proportion approche les 96% pour les hommes et descend jusqu'à 82% pour les femmes. Inversement la part attribuable à de la discrimination varie entre 8% et 18% selon le groupe de référence. Outre des coefficients relativement faibles, les résultats montrent qu'ils ne sont pas non plus significatifs au seuil de 10%. Si les résultats des estimations précédentes laissent suggérer que le lieu de résidence impacte vraisemblablement sur la probabilité de trouver un emploi ou sur le salaire obtenu, les résultats des décompositions (sans correction) ne vont pas dans le sens d'une mise en évidence d'une discrimination liée au lieu de résidence.

Tableau 5 - Décomposition des écarts de salaires (avec correction des biais de sélection)

Groupe	Ecart de salaire	Ecart expliqué	Ecart non-expliqué	Sélection	Observations
Population totale	0,050***	0,036*** 4,3%	0,406** 48,7%	-0,392** 47%	21 946
Hommes	0,043***	0,032*** 4,5%	0,346* 48,5%	-0,335* 47%	11 645
Femmes	0,061***	0,041*** 4,5%	0,449** 48,9%	-0,429** 46,7%	10 301

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : La variable relative au salaire est exprimée en log. Ce tableau synthétise les résultats de décompositions pour trois échantillons différents : la population totale, les hommes uniquement et les femmes uniquement.

* : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.

Si l'on prend en compte les caractéristiques inobservables des jeunes sortant du système scolaire qui conduisent à une sélection du lieu de résidence et à la décision de participation, les résultats sont très différents (Tableau 5). Dans ce cas, il ressort que la part de l'écart « non-expliqué » est proche de 49%. Cette proportion reste similaire quel que soit le groupe de référence. En outre, les écarts de salaires se décomposent maintenant principalement en une part attribuable à de la discrimination salariale et une autre relative à la sélection des jeunes. La part de cette dernière est proche de 47%. Enfin, la part des écarts de salaires qui est due aux caractéristiques des jeunes et de l'emploi qu'ils occupent est d'environ 4,5%. Finalement, avec correction des biais de sélection, nous retrouvons un écart « non-expliqué » qui est significatif (aux seuils de 5% ou 10% selon le groupe de référence) et qui suggère donc une éventuelle discrimination relative au lieu de résidence. Ne pas tenir compte de cette double sélection conduit probablement à surestimer le rôle des caractéristiques individuelles et sous estimer le rôle de la discrimination dans les écarts de salaires. Les résultats obtenus pour chaque groupe n'attestent pas d'effet différencié selon le sexe. Cela invalide donc quelque peu l'hypothèse d'une discrimination selon le lieu de résidence qui serait plus pesante pour les femmes que pour les hommes.

Conclusion

Dans ce travail, nous avons cherché à mettre en évidence le rôle du lieu de résidence dans les écarts de salaires observés entre jeunes sortant du système scolaire. Nous nous sommes également intéressés à un mécanisme particulier qui pourrait rendre compte de ces écarts : la discrimination en fonction du lieu de résidence. Pour cela, nous avons mobilisé les méthodes de décomposition proposées par Neuman et Oaxaca (2004). L'idée est d'estimer des équations de salaires et de décomposer l'écart de salaire observé entre deux groupes de jeunes : (1) des jeunes vivant dans des quartiers considérés comme « défavorisés » et (2) d'autres ne vivant pas dans ce type de quartiers. Cette décomposition se partage en une part qui s'explique par des différences de caractéristiques individuelles et des emplois occupés, en une part inexpliquée liée à des différences dans le rendement de ces mêmes caractéristiques (la discrimination) et en une part liée à une double sélection. En effet, nos estimations doivent tenir compte d'un premier type de biais de sélection (dit « biais de participation ») qui est lié au fait que nous observons les salaires uniquement pour des personnes qui sont en emploi. Elles doivent également tenir compte d'un second biais de sélection qui est lié au fait que les individus ou leurs parents choisissent potentiellement leur lieu de résidence.

Globalement, les résultats de ce travail montrent que le lieu de résidence a une influence directe sur la probabilité de trouver un emploi à la sortie des études ainsi que le salaire associé à cet emploi. La méthode de décomposition permet d'aller plus loin et montre qu'il existe un probable problème de discrimination relatif au lieu de résidence des jeunes de notre échantillon. Les résultats montrent effectivement que même lorsque les différences de caractéristiques observables entre des jeunes de quartiers « défavorisés » et « non-défavorisés » sont prises en compte les écarts de salaires demeurent. Il convient également de noter que sans prise en compte

de ce double problème de sélection, on surestime le rôle de ces différences de caractéristiques et on sous-estime le rôle de la discrimination.

En termes de politiques publiques, nos résultats suggèrent que la localisation résidentielle est un déterminant important des performances observées sur le marché du travail. Dans ce contexte, des recommandations de politiques publiques pourraient être de développer et promouvoir la mixité sociale. Cela peut passer par une géographie repensée des logements sociaux, qui éviterait leur trop forte concentration dans certains quartiers. Dans la même logique, La rénovation des quartiers les plus défavorisés peut contribuer à attirer des populations différentes aux revenus plus élevés, qui n'auraient pas emménagé autrement. Indirectement, cette mesure peut favoriser la mixité sociale et par conséquent faire bénéficier aux populations les plus fragiles d'externalités positives. Enfin, en rendant un quartier plus agréable, on peut imaginer que cela atténue le signal négatif que celui-ci peut envoyer. Or, c'est ce signal qui est souvent source de discrimination territoriale. Cependant, tous ces mécanismes potentiellement favorables ne se mettent pas en place instantanément. L'amélioration de l'image d'un quartier et la perception que les individus en ont ne se développeront que sur le moyen/long terme.

Bibliographie

- Altonji J. Et Blank R. (1999), «Race and Gender in the Labor Market», In: Ashenfelter O, Card D (eds) Handbook of labor economics. Elsevier, Amsterdam, pp 3143-4259.
- Arrow K. J. (1973), «The theory of Job Discrimination», in Ashenfelter O. A. & Reeds A. (eds.) Discrimination in Labor Markets, Princeton University Press, p. 3-33.
- Aeberhardt R., Fougère D., Pouget J. et Rathelot R. (2010a), «Wages and Employment of French Workers with African Origin», *Journal of Population Economics*, vol. 23(3), pp. 881-905.
- Aeberhardt R., Fougère D., Pouget J. et Rathelot R. (2010b), « L'emploi et les salaires des enfants d'immigrés », *Économie et Statistique*, n° 433-434.
- Aeberhardt R. et Pouget J. (2010), «National origin differences in wages and hierarchical positions», *Annales d'Économie et de Statistique*, 99-100, pp. 117-139.
- Becker G. (1957), «The Economics of Discrimination», University of Chicago Press (Original edition, 1957).
- Belzil B. et Poinas F. (2010), «Education and Early Career Outcomes of Second-Generation Immigrants in France», *Labour Economics*, 17, pp. 101-110.
- Benabou R. (1993), « Working of a city: Location, Education and Production », *quarterly Journal of Economic*, 108, pp. 619-652.
- Blinder A.S. (1973), «Wage Discrimination: Reduced Form and Structural Estimates», *The Journal of Human Resources*, vol. 8(4), pp. 436-455.
- Boumahdi R. et Giret J.-F. (2005), « Une analyse économétrique des disparités d'accès à l'emploi et de rémunérations entre jeunes d'origine française et jeunes issus de l'immigration », *Revue Economique*, vol. 56, pp. 625-636.
- Buchinsky M., Fougère D., Kramarz F. et Tchernis R. (2010), «Interfirm Mobility, Wages and the Returns to Seniority and Experience in the U.S.», *The Review of Economic Studies*, vol. 77, pp. 972-1001.
- Couppié T., Giret J.-F. et Moullet S. (2010), « Lieu de résidence et discrimination salariale : le cas des jeunes habitant dans une zone urbaine sensible », *Economie et Statistique*, n° 433-434.
- Galster, G. (2008), «Quantifying the Effect of Neighbourhood on Individuals. Challenges, Alternative Approaches and Promising Directions», *Journal of Applied Social Sciences. Schmollers Jahrbuch* 128 (1) pp.7-48.
- Galster G. (2012), «The Mechanism(s) of Neighborhood Effects: Theory, Evidence, and Policy Implications», *Neighbourhood Effects Research: New Perspectives*, pp 23-56, Springer Science & Business.
- Gobillon L., Selod H. et Zenou Y. (2007), «The mechanisms of spatial Mismatch», *Urban Studies*, 44(12), pp. 2401-2427.
- Heckman J.J. (1979), «Sample Specification Bias as a Specification Error», *Econometrica*, 47, pp.153-161.
- Hellerstein J., Neumark D. et Troske K. (1999), «Wage, Productivity and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations», *Journal of Labor Economics*, vol. 17, n° 3, pp. 409-446.

- Hellerstein J. et Neumark D. (2012), «Employment in Black Urban Labor Markets: Problems and Solutions», The Oxford Handbook of the Economics of Poverty, Oxford University Press, chap. NA, pp. 164-202.
- Holzer H. (1988), «Search method use by unemployed youth», *American Economic Review*, 77 (3), pp. 446-52.
- Holzer H. et Ihlantfelt K. (1998), «Customer Discrimination and Employment Outcomes for Minority Workers», *The quarterly Journal of Economics*, 113 (3), pp. 835-867.
- Hyslop D. (1999), «State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women», *Econometrica*, vol. 67, pp. 1255–1294.
- Jencks C. et Mayer S. (1990) «The social consequences of growing up in a poor neighborhood», In: Lynn L., McGeary M. (Eds) Inner-city Poverty in the United States. Washington: National Academy Press.
- Kain J. F. (1968), «Housing segregation, Negro employment, and metropolitan decentralization», *Quarterly Journal of Economics*, 82(2), pp. 32–59.
- Kim M. et Polachek S. (1994), «Panel Estimates of Male-Female Earnings Functions», *Journal of Human Resources*, vol. 29, n° 2, pp. 406-428.
- L'Heritier J.-L. (1992), « Les déterminants du salaire », *Economie et Statistique*, n° 257
- L'Horty Y., Duguet E., du Parquet L., Petit P. et Sari F. (2011), « Les effets du lieu de résidence sur l'accès à l'emploi : un test de discrimination auprès de jeunes qualifiés », *Economie et Statistique*, n° 447, pp. 71-96.
- Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Economics*, New York: Cambridge University Press, pp. 257-91.
- Meurs D. et Ponthieux S. (2000), « Une mesure de la discrimination dans l'écart de salaire entre hommes et femmes », *Economie et Statistique*, n° 337-338, pp. 135-158.
- Meurs D. et Ponthieux S. (2006), « L'écart des salaires entre les femmes et les hommes peut-il encore baisser ? », *Economie et Statistique*, n° 398-399, pp. 99-129.
- Miller P. (1987), «The Wage Effect of the Occupational Segregation of Women in Britain», *Economic Journal*, vol. 97, n° 388, pp. 885-896.
- Neuman S. et Oaxaca R. (2004), «Wage Decompositions with Selectivity Corrected Wage Equations: A Methodological Note», *Journal of Economic Inequality*, vol.2, pp. 3-10.
- Oaxaca R. (1973), «Male-Female Wage Differentials in Urban Labor Markets», *International Economic Review*, vol. 14, pp. 693-709.
- Oaxaca R. et Ransom M. (1994), «On Discrimination and the Decomposition of Wage Differentials», *Journal of Econometrics*, vol. 65, pp. 5–21.
- Petit P. (2003), « Comment évaluer la discrimination à l'embauche ? », *Revue Française d'Economie*, 17 (3), pp. 55-87.
- Phelps E. S. (1972), «The Statistical Theory of Racism and Sexism», *American Economic Review*, vol. 62, n° 4, pp. 659-661
- Rathelot R. (2010), « Origine et quartier : expliquer le salaire et l'emploi des descendants d'immigrés », *Revue d'Economie Régionale et Urbaine*, n° 1, pp. 27-55.
- Selod H., Zenou Y. (2006), «City structure, Job Search and Labor Discrimination: Theory and Policy Implications ». *Economic Journal*, 116 (514), pp. 1057-1087.

Annexes

Annexe A – Construction de l'indicateur d'un quartier « défavorisé »

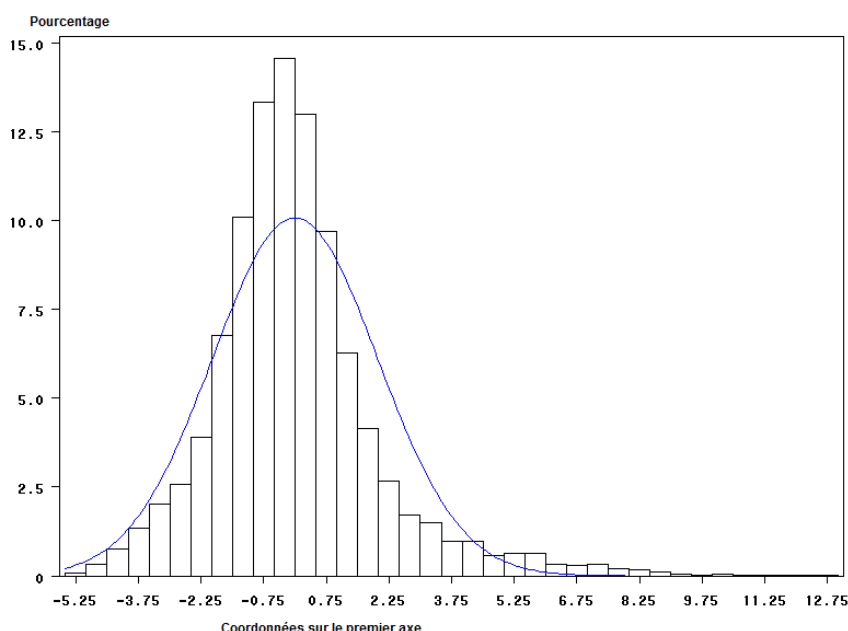
Tableau A.1 - Coordonnées et contributions des variables sur les axes

	Coordonnées		Contributions		Cosinus carrés	
	Axe 1	Axe 2	Axe 1	Axe 2	Axe 1	Axe 2
Taux de chômage des 15-64 ans	0,749	0,408	14,056	7,598	0,561	0,167
Part des chefs de ménage diplômés du supérieur	-0,646	0,703	10,449	22,554	0,417	0,494
Part des chefs de ménage sans diplôme	0,911	0,052	20,789	0,124	0,829	0,003
Part des chefs de ménage "cadres"	-0,704	0,649	12,415	19,193	0,495	0,421
Part des chefs de ménage "ouvriers"	0,749	-0,474	14,069	10,259	0,561	0,225
Part des chefs de ménage étrangers	0,464	0,716	5,390	23,383	0,215	0,512
Part des familles nombreuses (3 enfants ou plus)	0,668	0,431	11,199	8,464	0,447	0,185
Part des logements HLM	0,681	0,430	11,633	8,426	0,464	0,185

Champs : Analyse en composantes principales effectuée sur les 17 226 quartiers ou IRIS ayant un jeune sortant du système scolaire en 2004.

Sources : Recensement de la population 2007 (INSEE).

Graphique A.1 – Distribution des IRIS sur le 1^{er} axe de l'ACP



Champs : Analyse en composantes principales effectuée sur les 17 226 quartiers ou IRIS ayant un jeune sortant du système scolaire en 2004.

Sources : Recensement de la population 2007 (INSEE).

Tableau A.2 – Statistiques descriptives sur les variables retenues

Variables	Tous les quartiers		Quartiers défavorisés	
	Moyenne	Écart-type	Moyenne	Écart-type
Taux de chômage des 15-64 ans	10,6%	0,058	16,8%	0,073
Part des chefs de ménage diplômés du supérieur	11,0%	0,098	5,4%	0,033
Part des chefs de ménage sans diplôme	19,0%	0,087	29,6%	0,083
Part des chefs de ménage "cadres"	13,0%	0,106	6,0%	0,038
Part des chefs de ménage "ouvriers"	25,9%	0,115	36,4%	0,098
Part des chefs de ménage étrangers	5,0%	0,062	9,1%	0,089
Part des familles nombreuses (3 enfants ou plus)	2,3%	0,026	4,6%	0,037
Part des logements HLM	13,0%	0,194	31,7%	0,279
Observations	17 226		4 316	

Champs : Statistiques sur les 17 226 quartiers ou IRIS ayant un jeune sortant du système scolaire en 2004.

Sources : Recensement de la population 2007 (INSEE).

Annexe B – Caractéristiques de l'échantillon

Tableau B – Statistiques descriptives sur les variables retenues

Variables	Population totale		Hommes		Femmes	
	Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type
Situation : en emploi	0,806	0,002	82,2%	0,003	78,9%	0,003
Quartier défavorisé	25,5%	0,003	26,1%	0,004	24,9%	0,004
Etre une femme	48,4%	0,003	0		100,0%	
Age	21,8	0,022	21,3	0,031	22,3	0,032
<i>Niveau de sortie des études</i>						
Non diplômé	13,5%	0,002	17,8%	0,003	9,0%	0,002
CAP-BEP-MC	19,6%	0,002	24,0%	0,004	15,0%	0,003
Bac	23,8%	0,003	23,4%	0,004	24,2%	0,004
Deug	19,8%	0,002	12,8%	0,003	2,0%	0,001
Licence L3	6,5%	0,001	5,0%	0,002	8,2%	0,002
Maîtrise, DEA	11,7%	0,002	11,4%	0,003	12,1%	0,003
Doctorat	4,9%	0,001	5,6%	0,002	4,2%	0,002
<i>Exp. Prof. ant., emploi régulier</i>						
Oui, plusieurs	2,4%	0,001	1,9%	0,001	2,9%	0,001
Oui, un seul	11,1%	0,002	9,0%	0,002	13,2%	0,003
Non	69,5%	0,003	64,1%	0,004	75,2%	0,004
<i>Nationalité père : autre que française</i>	6,7%	0,002	6,4%	0,002	7,0%	0,002
<i>Nationalité mère : autre que française</i>	6,1%	0,001	6,0%	0,002	6,3%	0,002
<i>Situation du chef de ménage à la fin des études</i>						
Travaille	80,5%	0,002	81,1%	0,003	80,0%	0,003
Au chômage	2,6%	0,001	2,6%	0,001	2,7%	0,001
Inactif	9,7%	0,002	9,3%	0,002	9,9%	0,003
<i>Caractéristiques du ménage</i>						
Présence d'au moins un enfant	11,0%	0,002	6,8%	0,002	15,5%	0,003
Nombre d'enfants (si >0)	1,24	0,009	1,27	0,017	1,229	0,011
Vit chez ses parents	42,4%	0,003	53,9%	0,004	30,2%	0,004
Vit en couple	37,1%	0,003	25,8%	0,004	49,3%	0,004
Vit seul	20,4%	0,002	20,3%	0,003	20,5%	0,004
<i>Caractéristiques locales</i>						
Dist. au ctr. d'emploi le plus proche (en km)	22,1	0,107	22,4	0,146	21,8	0,158
Terrains multisports	19,6%	0,002	19,4%	0,003	19,8%	0,003
Aires de petits jeux	8,9%	0,002	8,7%	0,002	9,1%	0,002
Cinémas	3,0%	0,001	2,9%	0,001	3,2%	0,001
<i>Caractéristiques de l'emploi</i>						
Salaire à l'embauche	1 190,40	3,299	1 241,53	4,591	1 132,44	4,673
Temps plein	67,1%	0,003	73,3%	0,004	60,4%	0,004
Artisan, chef d'entreprise	0,8%	0,001	1,0%	0,001	0,5%	0,001
Cadre	11,1%	0,002	12,8%	0,003	9,3%	0,003
Profession intermédiaire	25,4%	0,003	18,1%	0,003	33,3%	0,004
Employé	21,0%	0,002	10,8%	0,003	31,9%	0,004
Ouvrier	26,4%	0,003	43,2%	0,004	8,1%	0,002
Observations	27 368		14 132		13 236	

Source : Enquête Génération 2004 du CEREQ – interrogation à 3 ans.

Annexe C – Equation de salaire sur la population globale

Tableau C – Résultats des estimations par MCO

<i>Variable</i>	Coefficient estimé	Ecart-type
Constante	6,095***	0,089
Etre une femme	-0,061***	0,004
Age	0,072***	0,007
(Age)²/100	-0,113***	0,015
Niveau de sortie des études		
Non diplômé	0,048***	0,008
CAP-BEP-MC	0,059***	0,006
Bac	<i>Réf.</i>	
Deug, BTS	0,117***	0,007
Licence L3	-0,008	0,009
Maîtrise, DEA	0,056**	0,009
Doctorat	0,089**	0,014
Expérience prof. antérieure : emploi régulier		
Oui, plusieurs	-0,044***	0,013
Oui, un seul	-0,008	0,006
Non	<i>Réf.</i>	
Nationalité père : autre que française	0,017*	0,009
Nationalité mère : autre que française	0,014	0,009
Situation du chef de ménage à la fin des études		
Travaille	<i>Réf.</i>	
Au chômage	-0,003	0,013
Inactif	-0,006	0,007
Caractéristiques de l'emploi		
Temps plein (<i>Réf.</i> Temps partiel)	0,039***	0,005
Artisan, chef d'entreprise	0,135***	0,021
Cadre	0,243***	0,009
Profession intermédiaire	<i>Réf.</i>	
Employé	-0,126***	0,006
Ouvrier	-0,074***	0,006
Indicatrices de région		
Dist. au ctr. d'emploi le plus proche (en km)	0,001*	0,001
Vivre dans un quartier « défavorisé »	-0,036***	0,008
λ	0,055***	0,012
R²	0,28	
Observations	21 946	

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : La variable à expliquer est le salaire obtenu à la première embauche. Il est exprimé en log et calculé en équivalent temps plein.

* : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.

Les résultats présentés sont ceux d'une estimation par MCO d'une équation de salaire sur l'ensemble de l'échantillon. Deux autres équations ont été estimées au préalable : (1) une équation de participation, à partir de laquelle les inverses des ratios de Mills (λ) ont été calculés. Ils sont ensuite introduits dans l'équation pour contrôler de ce biais de sélection. (2) une équation pour modéliser la probabilité de vivre dans un quartier « défavorisé ». Ici, on retient les valeurs prédites du modèle pour les inclure directement dans l'équation de salaire (dans une procédure équivalente aux 2MCO). Dans les deux cas, les instruments retenus sont ceux qui ont été présentés dans ce travail : le nombre d'enfant et la situation familiale pour l'équation (1) et les aménités du quartier pour l'équation (2).

Annexe D – Détails des estimations pour les hommes et femmes

Tableau D.1 – Résultats du probit bivarié (hommes)

Variable	Probabilité d'habiter un quartier défavorisé		Probabilité d'être en emploi	
	Coefficient estimé	Ecart-type	Coefficient estimé	Ecart-type
Constante	-1,039*	0,558	-1,126*	0,576
Age	0,014	0,047	0,206***	0,049
(Age)²/100	-0,061	0,101	-0,442***	0,105
Niveau de sortie des études				
Non diplômé	0,267***	0,043	-0,302***	0,047
CAP-BEP-MC	0,084**	0,038	0,125**	0,041
Bac	<i>Réf.</i>		<i>Réf.</i>	
Deug, BTS	-0,087***	0,044	0,181***	0,050
Licence L3	-0,085	0,065	0,020	0,071
Maîtrise, DEA	-0,244***	0,054	0,138**	0,059
Doctorat	-0,144*	0,085	0,062	0,092
Expérience prof. antérieure : emploi régulier				
Oui, plusieurs	0,007	0,093	0,091	0,104
Oui, un seul	0,051	0,044	0,072***	0,048
Non	<i>Réf.</i>		<i>Réf.</i>	
Nationalité père : autre que française	0,234***	0,052	0,043	0,058
Nationalité mère : autre que française	0,109**	0,055	-0,110*	0,057
Situation du chef de ménage à la fin des études				
Travaille	<i>Réf.</i>		<i>Réf.</i>	
Au chômage	0,325***	0,072	-0,187**	0,078
Inactif	0,302***	0,042	-0,138**	0,047
Caractéristiques du ménage				
A au moins un enfant	0,075*	0,040	0,123**	0,050
Vit chez ses parents	0,077**	0,034	-0,539***	0,040
Vit en couple	<i>Réf.</i>		<i>Réf.</i>	
Vit seul	-0,008	0,039	-0,251***	0,045
Indicatrices de région				
Dist. au ctr. d'emploi le plus proche (en km)	<i>Oui</i>		<i>Oui</i>	
	0,010***	0,001	0,004***	0,001
Vit dans un quartier défavorisé			-0,343**	0,149
Aménités du quartier				
Terrains multisports	0,217***	0,018		
Aires de petits jeux	0,060**	0,030		
Cinémas	-0,246***	0,061		
σ_{12}	0,135	0,090		
Log Likelihood	-13 099,044			
Observations	13 946			

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : * : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.

Tableau D.2 – Résultats des équations de salaires selon le type de quartier (hommes)

<i>Variable</i>	Quartier "défavorisé"		Quartier "non-défavorisé"	
	Coefficient estimé	Ecart-type	Coefficient estimé	Ecart-type
Constante	5,736***	0,281	6,512***	0,167
Age	0,062**	0,021	0,025**	0,012
(Age)²/100	-0,115**	0,044	-0,024	0,025
Niveau de sortie des études				
Non diplômé	0,054**	0,021	0,043***	0,014
CAP-BEP-MC	0,001	0,01	0,014	0,010
Bac	<i>Réf.</i>		<i>Réf.</i>	
Deug, BTS	0,073***	0,021	0,046***	0,012
Licence L3	-0,003	0,031	-0,015	0,016
Maîtrise, DEA	0,103**	0,028	0,059***	0,015
Doctorat	0,172***	0,045	0,063**	0,022
Expérience prof. antérieure : emploi régulier				
Oui, plusieurs	-0,096**	0,041	-0,035**	0,023
Oui, un seul	0,009	0,019	-0,025**	0,011
Non	<i>Réf.</i>		<i>Réf.</i>	
Nationalité père : autre que française	0,027	0,023	0,014	0,014
Nationalité mère : autre que française	0,003	0,025	0,021	0,014
Situation du chef de ménage à la fin des études				
Travaille	<i>Réf.</i>		<i>Réf.</i>	
Au chômage	0,006	0,028	-0,013	0,022
Inactif	0,019	0,018	0,010	0,012
Caractéristiques de l'emploi				
Temps plein (<i>Réf.</i> Temps partiel)	0,035**	0,015	0,011	0,009
Artisan, chef d'entreprise	0,157**	0,051	0,169***	0,030
Cadre	0,195***	0,028	0,264***	0,013
Profession intermédiaire	<i>Réf.</i>		<i>Réf.</i>	
Employé	-0,109***	0,019	-0,122***	0,011
Ouvrier	-0,038**	0,016	-0,055***	0,009
Indicatrices de région				
Dist. au ctr. d'emploi le plus proche (en km)	-0,000	0,000	0,000	0,000
λ_l	0,633**	0,226	0,410**	0,139
λ_e	0,054	0,282	-0,249**	0,085
R²	0,191		0,301	
Observations	2 999		8 646	

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : La variable à expliquer représente les salaires mensuels du premier emploi obtenu. Ils sont exprimés en log et ont été transformés en équivalent temps plein.

* : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.

Tableau D.3 – Résultats du probit bivarié (femmes)

<i>Variable</i>	Probabilité d'habiter un quartier défavorisé		Probabilité d'être en emploi	
	Coefficient estimé	Ecart-type	Coefficient estimé	Ecart-type
Constante	-1,031*	0,559	-2,418***	0,578
Age	0,008	0,046	0,230***	0,048
(Age)²/100	-0,004	0,095	-0,324***	0,101
Niveau de sortie des études				
Non diplômé	0,296***	0,051	-0,104**	0,052
CAP-BEP-MC	0,089**	0,041	0,239***	0,041
Bac	<i>Réf.</i>		<i>Réf.</i>	
Deug, BTS	-0,195***	0,038	0,595***	0,044
Licence L3	-0,273***	0,057	0,108*	0,058
Maîtrise, DEA	-0,268***	0,052	0,110**	0,052
Doctorat	-0,635***	0,092	0,032	0,091
Expérience prof. antérieure : emploi régulier				
Oui, plusieurs	0,150**	0,075	0,028	0,082
Oui, un seul	0,029	0,039	0,131***	0,042
Non	<i>Réf.</i>		<i>Réf.</i>	
Nationalité père : autre que française				
	0,242***	0,052	-0,096*	0,055
Nationalité mère : autre que française				
	0,149**	0,055	-0,043	0,057
Situation du chef de ménage à la fin des études				
Travaille	<i>Réf.</i>		<i>Réf.</i>	
Au chômage	0,381***	0,073	-0,005	0,079
Inactif	0,197***	0,042	-0,024	0,046
Caractéristiques du ménage				
A au moins un enfant	0,056*	0,030	-0,455***	0,031
Vit chez ses parents	0,095**	0,032	-0,401***	0,034
Vit en couple	<i>Réf.</i>		<i>Réf.</i>	
Vit seul	-0,039	0,035	-0,121***	0,037
Indicatrices de région				
	<i>Oui</i>		<i>Oui</i>	
Dist. au ctr. d'emploi le plus proche (en km)	0,011***	0,001	0,002*	0,001
Vit dans un quartier défavorisé			-0,586***	0,145
Aménités du quartier				
Terrains multisports	0,171***	0,018		
Aires de petits jeux	0,031**	0,031		
Cinémas	-0,159***	0,052		
σ_{12}	0,313***	0,093		
Log Likelihood	-12 279,744			
Observations	13 022			

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : * : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.

Tableau D.4 – Résultats des équations de salaires selon le type de quartier (femmes)

<i>Variable</i>	Quartier "défavorisé"		Quartier "non-défavorisé"	
	Coefficient estimé	Ecart-type	Coefficient estimé	Ecart-type
Constante	5,507***	0,313	6,044***	0,181
Age	0,090***	0,023	0,085**	0,014
(Age)²/100	-0,144***	0,045	-0,138***	0,027
Niveau de sortie des études				
Non diplômé	0,068**	0,025	0,045**	0,017
CAP-BEP-MC	0,095***	0,020	0,092***	0,012
Bac	<i>Réf.</i>		<i>Réf.</i>	
Deug, BTS	0,153***	0,022	0,135***	0,013
Licence L3	0,016	0,030	-0,013	0,016
Maîtrise, DEA	0,009	0,027	0,051***	0,015
Doctorat	0,070	0,054	0,115**	0,023
Expérience prof. antérieure : emploi régulier				
Oui, plusieurs	-0,021	0,033	-0,047**	0,019
Oui, un seul	0,003	0,018	-0,013**	0,010
Non	<i>Réf.</i>		<i>Réf.</i>	
Nationalité père : autre que française				
	0,010	0,025	0,014	0,015
Nationalité mère : autre que française				
	0,040	0,026	0,018	0,015
Situation du chef de ménage à la fin des études				
Travaille	<i>Réf.</i>		<i>Réf.</i>	
Au chômage	0,083**	0,032	-0,015	0,02.
Inactif	-0,018	0,019	-0,021*	0,012
Caractéristiques de l'emploi				
Temps plein (<i>Réf.</i> Temps partiel)	0,047***	0,014	0,048***	0,008
Artisan, chef d'entreprise	0,073	0,065	0,105**	0,046
Cadre	0,193***	0,031	0,236***	0,014
Profession intermédiaire	<i>Réf.</i>		<i>Réf.</i>	
Employé	-0,104***	0,016	-0,131***	0,009
Ouvrier	-0,104***	0,023	-0,121***	0,014
Indicatrices de région				
	<i>Oui</i>		<i>Oui</i>	
Dist. au ctr. d'emploi le plus proche (en km)	0,001	0,000	0,000	0,000
λ_l	0,161**	0,199	-0,144	0,129
λ_e	0,176	0,165	-0,212**	0,090
R²	0,219		0,302	
Observations	2 487		7 814	

Source : Enquête Génération 2004 du CEREQ – Interrogation à 3 ans.

Note : La variable à expliquer représente les salaires mensuels du premier emploi obtenu. Ils sont exprimés en log et ont été transformés en équivalent temps plein.

* : Significatif à 10% ; ** : Significatif à 5%. *** : Significatif à 1%.