

UN INDICATEUR DE DISPERSION POUR LES CARRIÈRES SALARIALES : APPLICATION À L'HÉTÉROGÉNÉITÉ DES CARRIÈRES DU SECTEUR PRIVÉ ET DU SECTEUR PUBLIC

Malik KOUBI¹()*,

() Insee, Département des études économiques*

Résumé

Cet article définit un indicateur de dispersion analogue à la variance mais adapté à l'objet intertemporel que sont les carrières salariales. Il propose une décomposition de cet indicateur faisant intervenir certaines caractéristiques des carrières individuelles comme leur niveau, leur progression ou leur instabilité. Il applique enfin cette décomposition aux carrières des salariés du secteur privé et des différents versants de la fonction publique pour comparer les sources d'inégalités spécifiques à chacun d'entre eux.

Abstract

This article defines a dispersion index similar to the variance but adapted to wage careers. It provides a breakdown of this indicator involving certain characteristics of individual careers as their level, their advance and their instability. This breakdown is finally applied to careers of employees in the public and private sector in order to compare sources of career dispersion specific to each of these sectors.

Mots-clés

Indicateurs d'inégalités, économétrie, analyse des données

¹ Malik Koubi, Insee
division Redistribution et politiques sociales
15 boulevard Gabriel Péri 92245 Malakoff
malik.koubi@insee.fr

Introduction

La notion de carrière salariale renvoie à l'évolution du salaire tout au long de la carrière professionnelle. Comme pour la théorie du cycle de vie, cette notion est sous-tendue par l'idée d'une certaine cohérence dans le temps de l'évolution du salaire au niveau individuel. Becker (1975) est sans doute le premier économiste à avoir tenté d'expliquer cette cohérence dans le temps en assimilant l'investissement en capital humain à un investissement matériel. Pour lui, le choix de la durée d'éducation au sein du système scolaire relève d'un arbitrage entre les coûts d'opportunité liés à l'acquisition d'une formation d'une part (frais de scolarité et manque à gagner induit par l'absence de rémunération) et la somme actualisée des gains futurs que permet de générer l'investissement d'autre part. Par ce modèle, Becker rend compte de la dispersion des salaires entre individus à l'aide de choix différents en matière d'éducation. Ces choix sont justifiés par des écarts dans les capacités individuelles et des accès inégaux aux sources de financement.

Dans le modèle de Mincer, qui s'inscrit dans la même perspective, le salaire est une fonction parabolique de l'âge. Un investissement donné réalisé par l'individu entraînerait un différentiel de salaire qui s'accroît avec le temps mais à un rythme de moins en moins rapide. La rémunération est liée à la productivité, elle-même dépendant du capital humain de l'individu, qui regroupe toutes les connaissances et savoir-faire accumulés par l'individu qui peuvent l'aider dans l'accomplissement de sa tâche. Les variables usuellement utilisées pour rendre compte du capital humain sont le diplôme, l'expérience (sur le marché du travail) et l'ancienneté (sur le poste ou dans l'établissement). Ces deux dernières, pas toujours observées dans les données, sont souvent approximées par l'âge.

L'approche de Mincer s'applique bien aux carrières salariales que l'on observe dans un environnement relativement stabilisé : entreprise de grande taille ou le secteur public, dans lesquels il existe une grille d'avancement en fonction du diplôme et de l'ancienneté et où le déroulement de carrière est encadré par un ensemble de règles constantes dans le temps. Les mobilités professionnelles s'opèrent souvent dans le cadre de marchés internes et assurent une progression régulière du salaire (théorie des contrats implicites). Elle décrit moins bien des carrières plus heurtées présentant des évolutions plus complexes à la suite d'évènements modifiant la trajectoire salariale : aléas économiques, restructurations, négociation salariale, changement d'employeur, etc. Ces évènements sont mieux pris en compte par les théories s'intéressant aux frictions du marché du travail mais au prix d'une schématisation importante : dans les modèles de recherche d'emploi du type Pissarides Mortensen (1994) par exemple, la dynamique d'emploi et des salaires est générée à travers un petit nombre de paramètres structurels.

Dans cette étude, nous généralisons l'approche de Mincer en proposant une décomposition des carrières plus souple que celle du modèle initial, captant des évolutions potentiellement plus complexes que la simple fonction parabolique proposée par Mincer. La première partie est consacrée à la présentation de la base de données qui permet de suivre un échantillon représentatif de salariés du public et du privé de 1988 à 2008. Cette première partie présente également des statistiques descriptives sur les carrières par sexe et par diplôme. Les salaires sont déflatés par le salaire médian de l'année car nous nous intéressons à la position relative de l'individu dans la hiérarchie salariale plutôt qu'à la valeur absolue de son revenu. Cette approche a de surcroît l'avantage empirique de rendre comparables des salaires observés à des dates éloignées, comme c'est le cas ici avec l'utilisation de données de panel

Dans la deuxième partie, chaque carrière individuelle est décomposée à l'aide d'un ensemble F de fonctions élémentaires dépendant de l'âge permettant de capter des dimensions particulières de la

carrière. Nous nous restreignons pour les applications essentiellement à des fonctions continues (polynômiales) mais la démarche peut être appliquée à des fonctions plus heurtées comme l'illustre un exemple en fin d'article.

Nous définissons dans la troisième partie un indicateur de dispersion adapté aux objets multidimensionnels que forment les carrières salariales et nous montrons que cet indicateur peut être décomposé en faisant intervenir les différentes caractéristiques de la carrière construites à partir des fonctions F de la partie I. Cette décomposition est appliquée aux carrières des salariés du secteur privé et des différents versants de la fonction publique pour comparer les sources d'inégalités spécifiques à chacun d'entre eux.

1. Les données : un panel sur 20 ans couvrant les salariés du privé et du public

Cette étude s'appuie sur des sources administratives, les déclarations annuelles de données sociales (DADS) et les fichiers de paie des agents de l'État, ainsi que sur les données l'échantillon démographique permanent.

1.1 Les Déclarations annuelles de données sociales

La Déclaration annuelle de données sociales (DADS) est un formulaire administratif que doivent remplir chaque année tous les employeurs, à destination des administrations sociales et fiscales. Les employeurs communiquent *via* ces déclarations le montant des rémunérations salariales versées à chaque salarié. Le champ des DADS exploitées par l'Insee recouvre l'essentiel du secteur privé et des entreprises publiques (hors salariés des particuliers employeurs), ainsi que les fonctions publiques territoriale et hospitalière.

1.2 Les fichiers de paie de la fonction publique de l'État

Jusqu'à l'exercice 2008 compris, les fichiers de paie de la FPE constituent la source de référence sur les rémunérations des agents de l'État (ministères civils et établissements publics administratifs).

1.3 L'échantillon démographique permanent

L'échantillon démographique permanent (EDP) est le premier panel sociodémographique de grande taille mis en place en France, pour étudier la fécondité, la mortalité, la mise en couple, les migrations géographiques au sein du territoire national, la mobilité sociale et la mobilité professionnelle, ainsi que les interactions possibles entre ces différents aspects. Conformément à l'arrêté du 24 juin 1998, il est possible d'utiliser l'information sur le diplôme issue de l'EDP dans le panel tous salariés.

Cette information sur le diplôme présente dans l'EDP provient des données recueillies lors des recensements de la population. Alors que les opérations de recensements étaient auparavant exhaustives et menées tous les 7 à 10 ans, à compter de 2004, elles deviennent annuelles mais sont menées chaque année sur une partie seulement du territoire. Pour cette raison, le diplôme n'est renseigné que dans l'EDP que pour une partie du panel. Pour les générations les plus anciennes, on dispose du diplôme grâce aux recensements exhaustifs antérieurs, mais ce n'est pas le cas pour les générations les plus jeunes, qui sont donc moins présentes dans le panel tous salariés après appariement avec l'EDP pour récupérer la variable diplôme. Afin d'en tenir compte, ces dernières sont surpondérées.

1.4. Le panel tous salariés

Le panel tous salariés cumule chaque année depuis 1988, pour chaque salarié présent dans les DADS ou dans les fichiers de paie de la fonction publique de l'État, l'ensemble de ses périodes d'emploi. L'arrivée du système d'information des agents des services publics en 2009 rend les comparaisons difficiles entre 2008 et 2009. C'est pourquoi, la présente étude s'intéresse aux emplois jusqu'en 2008 seulement.

1.5. Le concept de salaire retenu

L'étude porte sur le salaire net mensualisé. Ce concept permet de neutraliser l'effet du nombre de jours travaillés dans l'année (ce nombre joue sur le niveau du salaire annuel, mais pas sur le salaire mensualisé lorsqu'il est calculé sur la seule période travaillée). En revanche, il reflète la quotité de travail (nombre d'heures effectuées dans la journée : à salaire horaire égal, un salarié à mi-temps a un salaire mensualisé moindre qu'un salarié à plein temps). Le salaire net est le salaire perçu par le salarié. Dans le secteur privé, il comprend la partie de l'intéressement non placée sur un plan d'épargne entreprise (PEE), un plan d'épargne interentreprises (PEI) ou un plan d'épargne pour la retraite collectif (PERCO). Par ailleurs, le salaire considéré est relatif au salaire médian. En effet, l'analyse conduite sur longue période nécessite la définition d'un coefficient de correction temporel du salaire. La correction par le salaire médian est utilisée dans la majorité des cas. La raison est double. D'une part, le salaire rapporté au salaire médian permet de retracer l'évolution des salariés dans la

hiérarchie salariale. D'autre part, cette correction traite une bonne partie des fluctuations conjoncturelles.

Pour chaque salarié, les différents postes qu'il a occupés dans l'année ont été agrégés par secteur et versant de la fonction publique : enseignants, FPE hors enseignants, FPH, FPT, secteur privé.

1.6. Champ

Il est constitué des salariés à temps complet ou partiel du secteur privé (hors salariés des particuliers employeurs) et public.

Par ailleurs, seul le poste principal du salarié est retenu pour chaque année civile.

Entre 1988 et 2008, la frontière entre les secteurs public et privé a évolué. Le champ a été homogénéisé en conservant dans le secteur d'origine les salariés dont l'employeur a changé de catégorie juridique. Par exemple, quand la Poste est sortie de la FPE en 1991, le statut des salariés anciennement fonctionnaires est resté presque identique : grille de salaire et emploi à vie. En revanche, les contrats des nouveaux salariés relèvent du droit privé. Dans ce cas, pour notre analyse, les personnels présents avant 1991 sont conservés dans la FPE. Les nouveaux salariés de la Poste sont comptabilisés dans le secteur privé.

2. Statistiques descriptives

2.1. Des secteurs aux caractéristiques différenciées

Le panel tous salariés, apparié avec l'échantillon démographique permanent, permet de suivre individuellement sur longue période ces salariés, et d'analyser les carrières salariales par niveau de qualification.

Les caractéristiques démographiques des salariés (sexe, âge) ainsi que la structure de leurs qualifications ne sont pas uniformes dans tous les secteurs. Ainsi, sur la période de référence, 1988-2008, la moitié des salariés de la fonction publique de l'État (FPE) hors enseignants et de la fonction publique hospitalière (FPH) ont un diplôme supérieur au baccalauréat contre 35 % dans le privé et 29 % dans la fonction publique territoriale (FPT). Les enseignants ont un diplôme supérieur ou équivalent au baccalauréat dans 95 % des cas et plus des trois quarts sont titulaires d'un diplôme supérieur à bac + 2. Au cours du temps, la proportion de diplômés du supérieur a augmenté dans tous les secteurs suivant en cela la hausse de la qualification de l'ensemble de la population.

La notion de salaire utilisée dans cette étude est celle de salaire net mensualisé, y compris compléments de rémunération. Seul le salaire de l'activité principale est considéré. Le salaire progresse en moyenne avec l'âge, de manière différenciée selon le groupe de salariés considéré, notamment le niveau de diplôme. Dans la suite, les résultats sont systématiquement établis par groupes de salariés définis par le secteur (Privé, FPE, FPH, FPT), le diplôme et le sexe.

Par ailleurs, les salaires considérés sont déflatés par le salaire médian de l'année. En effet, les salaires et les trajectoires salariales étudiés couvrent une période longue, de 20 ans, au cours de laquelle les niveaux moyens de salaires, en francs puis en euros courants, ont globalement augmenté sous l'effet de l'inflation et des gains de productivité de l'ensemble de l'économie. Le salaire « relatif » ainsi défini, qui marque une position dans la hiérarchie salariale, a pour avantage de rendre comparables les salaires de deux individus observés à des dates éloignées.

Les hommes salariés les moins diplômés (BEP ou moins y compris les sans-diplôme) sont mieux rémunérés dans la FPE que dans le secteur privé et les autres versants de la fonction publique. En début de carrière, autour de 30 ans, ils gagnent 1,1 fois le salaire médian de l'ensemble des salariés, contre 1,0 dans la FPH, et 0,9 dans les grandes entreprises du secteur privé. Cet avantage de la FPE pour les moins diplômés subsiste pour les salariés plus âgés.

Les diplômes permettent d'obtenir une rémunération plus élevée, notamment dans le secteur privé. Ainsi, à 30 ans, les salariés masculins les plus diplômés (au moins bac + 3) qui travaillent dans une entreprise de plus de 100 salariés touchent 2 fois le salaire médian, et 1,7 fois pour ceux qui travaillent dans une entreprise de moins de 100 salariés. À ce niveau de diplôme et à cet âge, le salaire n'est que 1,2 fois le salaire médian dans la FPT et la FPE, la FPH occupant une position intermédiaire.

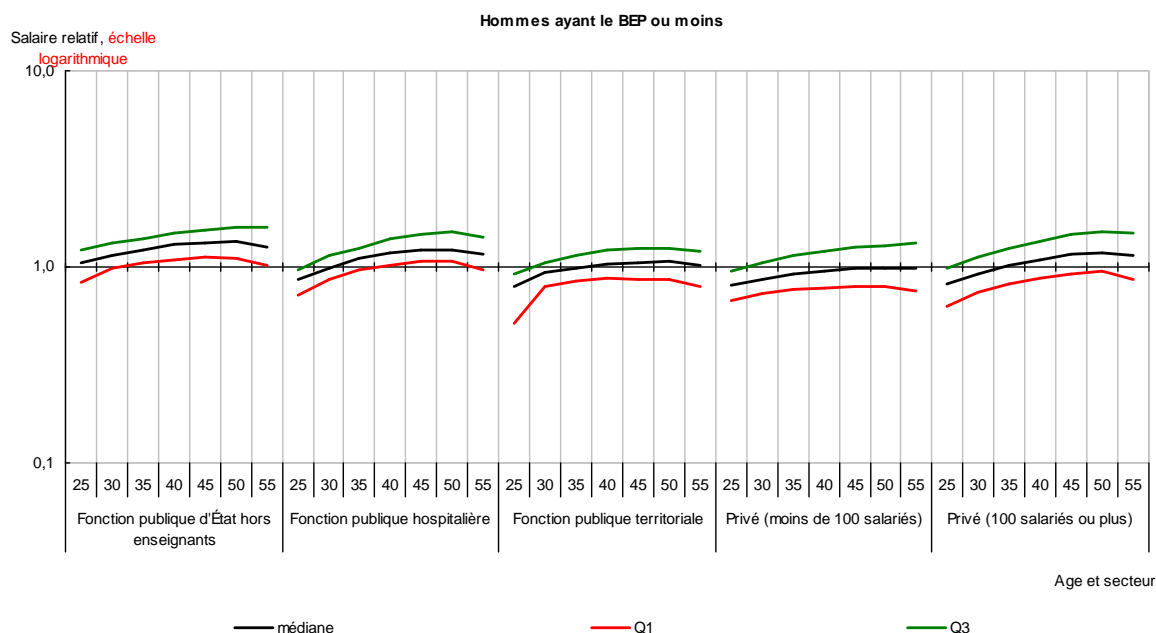
L'avantage procuré par le diplôme s'accroît avec l'âge et le dynamisme des carrières salariales des plus diplômés est plus marqué dans le secteur privé et la FPH.

Les écarts de salaire entre les hommes et les femmes sont plus grands dans le secteur privé que dans la fonction publique, en particulier pour les plus diplômés. À 30 ans, pour les moins diplômés, l'écart est de 1 % dans la FPH, 24 % dans la FPT et la FPE (hors enseignants), 23 % dans les grandes entreprises et 24 % dans les petites. Pour les plus diplômés, l'écart varie dans le public de 8 % pour les enseignants à 15 % pour la FPE hors enseignants ; dans le privé, il est d'au moins 26 %.

2.2. Les salaires progressent plus vite avec l'âge dans le secteur privé

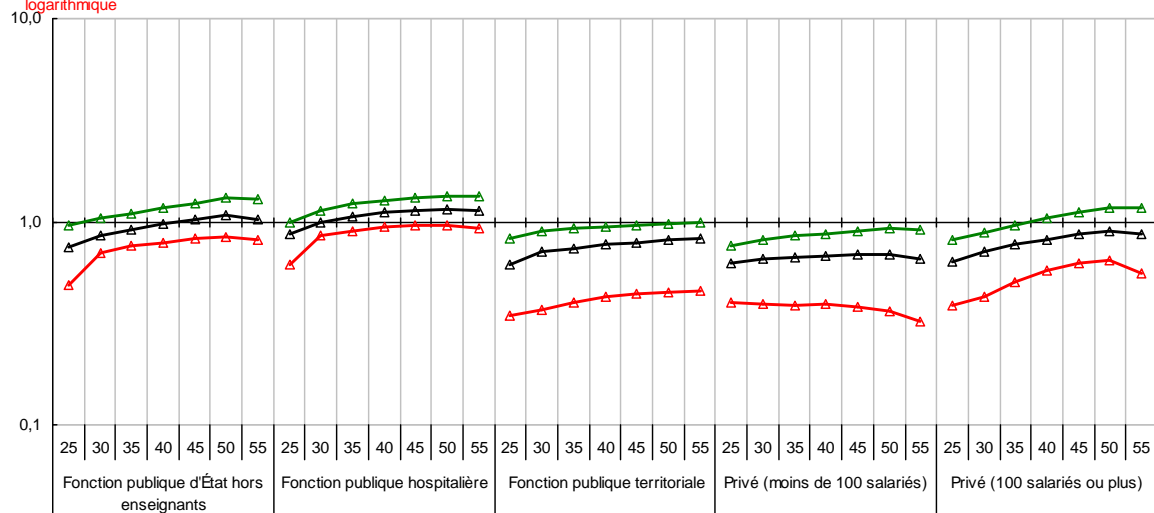
Plusieurs mécanismes peuvent expliquer que les carrières salariales diffèrent selon le type d'employeur et selon la qualification du salarié. Lorsqu'un employeur privé veut garder un salarié, qui a par exemple acquis des compétences qui le rendent difficile à remplacer, il peut mettre en place une carrière salariale ascendante, le salarié ayant alors intérêt à rester dans l'entreprise. À l'inverse, dans certains secteurs ne nécessitant pas de compétences spécifiques, les salaires augmenteront moins vite avec l'âge. Dans la fonction publique, beaucoup de métiers n'ont pas d'équivalent dans le privé ; il n'est alors pas nécessaire de fidéliser les salariés à travers des augmentations importantes de salaires au fur et à mesure de la carrière. En outre, la garantie de l'emploi représente un avantage non monétaire, qui compense en partie une carrière salariale moins dynamique ; à l'inverse, cette stabilité de l'emploi rend plus importante pour l'employeur l'enjeu de la motivation, qui est en partie alimentée par des hausses régulières de salaires.

De manière générale, les perspectives de progression entre 30 et 45 ans sont bien meilleures pour les plus diplômés que pour les moins diplômés. Ce constat vaut pour les trois fonctions publiques comme pour le secteur privé mais à des degrés divers. Dans la fonction publique, les salaires progressent avec l'âge plus régulièrement mais moins vite que dans le privé.



Femmes ayant le BEP ou moins

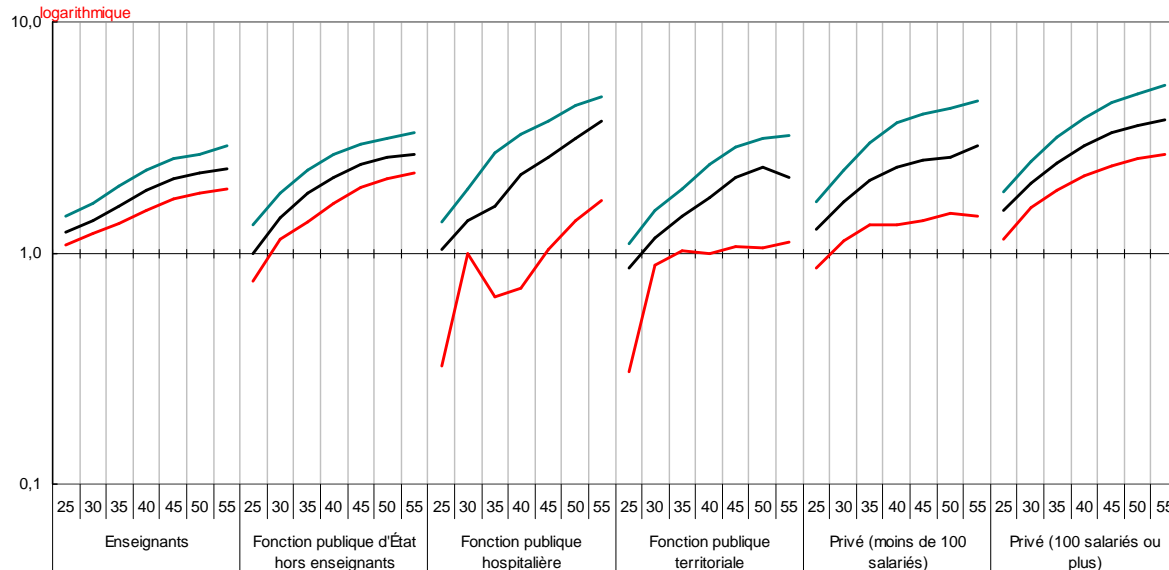
Salaires relatifs, échelle logarithmique



Age et secteur

Hommes ayant un diplôme au moins égal à bac+3

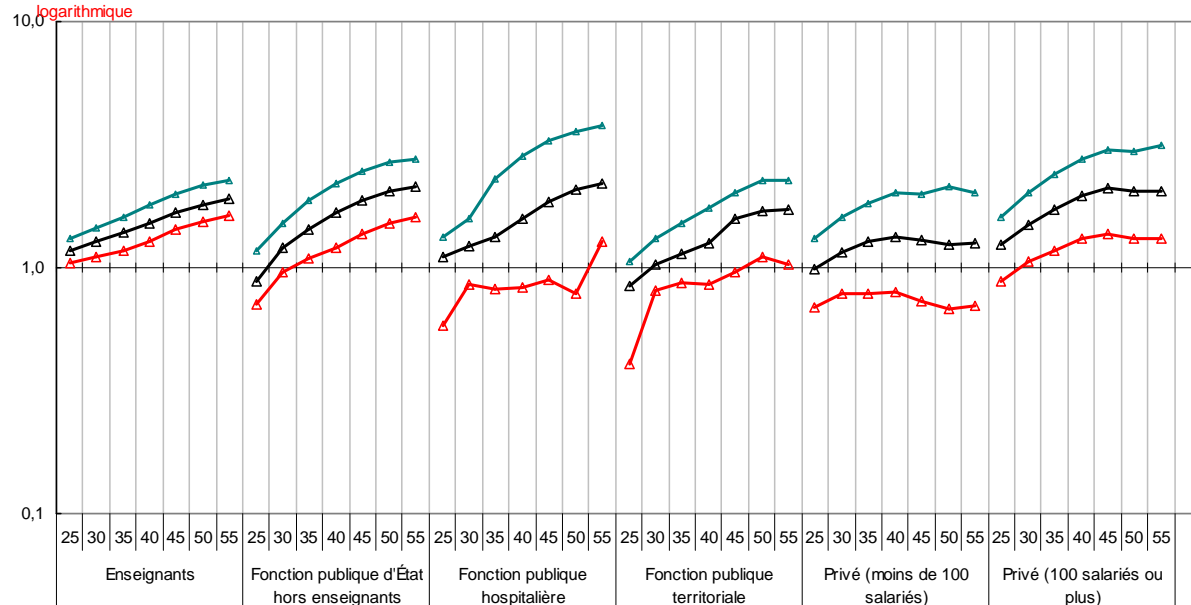
Salaires relatifs, échelle logarithmique



Age et secteur

Femmes ayant un diplôme au moins égal à bac+3

Salaires relatifs, échelle logarithmique



Age et secteur

3. Une modélisation des carrières à l'aide de fonctions élémentaires de l'âge

Les carrières agrégées considérées jusqu'ici peuvent être sujettes à différents biais. L'évolution globale du salaire avec l'âge peut ainsi être calculé sur des groupes de salariés différents à chaque âge (effet de sélection). De même, l'effet de l'âge mesuré en coupe peut inclure des effets de cohorte importants si certaines ont eu des carrières professionnelles particulièrement favorables. De ce fait, saisir la complexité des carrières nécessite d'aller au-delà d'une approche statique.

L'intérêt des données de panel dont nous disposons est précisément de pouvoir suivre les salariés individuellement sur une longue période. Dans cette partie, nous nous plaçons au niveau individuel et nous nous intéressons aux perspectives salariales et aux profils de carrière dans le secteur privé et les différents versants de la fonction publique. Nous décomposons chaque carrière individuelle à l'aide de fonctions élémentaires de l'âge. Les paramètres associés à cette décomposition captent chacun une caractéristique géométrique particulière du profil de carrière, tels que le niveau moyen de salaire, sa progression, etc.

3.1. Modélisation des carrières à l'aide d'un jeu de fonctions élémentaires

La carrière salariale du salarié i est définie comme la suite de ses salaires sur un intervalle d'âge IA donné. Pour simplifier, nous ne considérons ici que des carrières sans interruption et nous fixons une fois pour toutes l'intervalle d'âge $IA=[30,45]$. La carrière de l'individu i est donc simplement la suite des salaires $W_i = (w_{i,a})_{a \in IA}$.

Nous rappelons que les salaires sont déflatés par le salaire médian car nous nous intéressons à la position relative de l'individu dans la hiérarchie salariale plutôt qu'à la valeur absolue de son revenu. Par ailleurs, les variables de capital humain que nous utilisons sont le diplôme et l'âge, ce dernier étant un proxy de l'expérience sur le marché du travail. La dépendance par rapport à l'âge est le point important que nous précisons ici.

On se fixe un jeu de fonctions élémentaires de l'âge, définies sur le segment IA , qui serviront de base de décomposition du salaire.

$$F_K = (f_0, \dots, f_k, \dots, f_K) \text{ avec } f_0 \equiv 1$$

Par convention, la fonction f_0 est la fonction constante égale à 1.

Chaque carrière individuelle se décompose à l'aide de ce jeu de fonctions et d'un résidu qui représente les évolutions du salaire inexplicables par le jeu de fonctions élémentaires.

$$w_{i,a} = c_i^0 \cdot f_0(a) + c_i^1 \cdot f_1(a) + \dots + c_i^K \cdot f_K(a) + u_{i,a}$$

Ou, de manière vectorielle :

$$W_i = F.C_i + U_i \text{ avec } C_i = (c_i^0, c_i^1, \dots, c_i^K)'$$

Les termes de cette décomposition capturent certaines caractéristiques géométriques de la carrière. Par exemple, la valeur prédite c_0^i capte le niveau moyen de la carrière. Pour le jeu de fonctions principalement utilisé, les f_k sont des polynômes de degré k et $c_0^i + c_1^i.f_1(a)$ est la meilleure approximation linéaire de la carrière, $c_0^i + c_1^i.f_1(a) + c_2^i.f_2(a)$ la meilleure approximation parabolique, etc.

La décomposition obtenue dépend du jeu de fonctions que l'on s'est initialement fixé. Elle n'est par ailleurs pas « totale » car les fonctions élémentaires ne captent généralement pas l'ensemble de la variabilité de la carrière et il subsiste donc un résidu u_a^i non expliqué par les fonctions élémentaires F .

On peut supposer sans perte de généralité, quitte à renormaliser les fonctions élémentaires, que celles-ci sont non corrélées entre elles et de variance unitaire sur la plage d'âges IA (encadré 1). Pour un individu donné, le vecteur des résidus est par ailleurs également non corrélé aux f_k . Cela se traduit par les relations suivantes :

$$U_i' F = F' U_i = 0$$

$$F' F = F F' = I$$

Encadré 1

Les fonctions f_k servant à décomposer les carrières salariales individuelles sont a priori des fonctions quelconques définies sur le segment IA. Elles peuvent cependant être orthonormalisées à l'aide d'une procédure classique, de manière à ce qu'elles soient de variance unitaire et non corrélées entre elles. Le procédé d'orthonormalisation préserve l'espace engendré par ces fonctions. L'orthonormalisation n'a donc pas d'incidence sur la décomposition du salaire. Plus précisément, on peut définir un autre jeu de variables élémentaires \tilde{F} engendrant le même espace de fonctions que les fonctions F et vérifiant de plus les propriétés suivantes :

$$1) \forall k \neq 0, E(\tilde{f}_k) = 0$$

$$2) \forall (k, l), E(\tilde{f}_k \tilde{f}_l) = 1 \text{ si } k = l, 0 \text{ si } k \neq l$$

$$3) \forall k > 0, \langle \tilde{f}_0, \dots, \tilde{f}_k \rangle = \langle f_0, \dots, f_k \rangle \text{ et } \tilde{f}_k \perp \langle \tilde{f}_0, \dots, \tilde{f}_{k-1} \rangle$$

La preuve s'appuie sur le procédé d'orthonormalisation de Schmidt classique pour le produit scalaire défini par $\langle f, g \rangle = E(f.g)$. En pratique, l'algorithme itératif consiste à régresser f_k sur f_0, \dots, f_{k-1} et à définir \tilde{f}_k comme le résidu normalisé de cette régression. En particulier, de cette manière, les fonctions élémentaires sont bien orthonormales (propriété 1) et le processus conserve l'espace engendré par les fonctions élémentaires (propriété 3). La propriété 2 exprime simplement l'orthogonalité à f_0 des fonctions d'indice supérieur.

3.2. Application aux carrières du secteur public et privé avec des fonctions polynômiales

Dans cette application, les fonctions considérées sont de simples fonctions polynômiales de degrés étagés. Par ailleurs, l'analyse est systématiquement menée par groupe de salariés définis par le secteur, le diplôme et le sexe.

$$F_4 = (1, age, age^2, age^3, age^4)$$

L'ajustement est globalement assez bon. Nous ne pouvons décrire l'ajustement du modèle au niveau de chaque individu mais le tableau suivant donne la distribution des R^2 obtenus. L'ajustement est moins bon dans le secteur privé, ce qui traduit l'importance de la composante individuelle non expliquée par les fonctions polynômiales. Il est légèrement meilleur pour les plus diplômés que pour les moins diplômés. Il est également meilleur dans la FPT et la FPE.

Distribution des R^2 obtenus dans les régressions individuelles

Secteur	Diplôme	sexe	q1	q2	q3
Enseignants	Bac ou moins	Femmes	ns	ns	ns
		Hommes	ns	ns	ns
	Plus que Bac	Femmes	0,63	0,81	0,90
		Hommes	0,72	0,82	0,92
Etat (hors ens.)	Bac ou moins	Femmes	0,62	0,81	0,91
		Hommes	0,70	0,86	0,94
	Plus que Bac	Femmes	0,65	0,85	0,93
		Hommes	0,75	0,90	0,96
Fonction publique hospitalière	Bac ou moins	Femmes	0,60	0,80	0,90
		Hommes	0,64	0,81	0,92
	Plus que Bac	Femmes	0,63	0,83	0,92
		Hommes	0,67	0,90	0,94
Fonction publique territoriale	Bac ou moins	Femmes	0,70	0,87	0,94
		Hommes	0,63	0,88	0,94
	Plus que Bac	Femmes	0,59	0,86	0,94
		Hommes	0,71	0,90	0,96
Secteur privé	Bac ou moins	Femmes	0,47	0,73	0,88
		Hommes	0,45	0,70	0,87
	Plus que Bac	Femmes	0,50	0,73	0,88

	Hommes	0,52	0,80	0,92
--	--------	------	------	------

Champ : salariés du secteur privé et public âgés entre 30 et 45 ans.

Source : panel tous salariés 1988-2008.

Lecture : parmi les salariés enseignants ayant le bac ou plus, pour les 25% pour lesquels le modèle s'ajuste le moins bien, le R^2 obtenu est inférieur ou égal à 0,63. Pour les 25 % les mieux ajustés, le R^2 est supérieur ou égal à 0,90.

Les coefficients obtenus rejoignent en partie les statistiques descriptives de la partie I. Le coefficient c_0 , qui capture le niveau global de la carrière est croissant avec le niveau de diplôme au sein de chaque secteur. Pour les plus diplômés, il est plus élevé dans le secteur privé et la fonction publique hospitalière que dans les autres versants de la fonction publique, notamment la fonction publique territoriale. Le salaire « permanent » des moins diplômés est pour sa part plus élevé dans la FPE (hors enseignants) et la FPH. Il est moins élevé dans le secteur privé et la fonction publique territoriale.

La dispersion de ce premier coefficient est maximale pour les plus diplômés, particulièrement dans le secteur privé (les deux quartiles valent respectivement 1,31 et 2,54) et la FPH (1,31 et 2,35). La dispersion est moindre dans la FPT et la FPE.

Le second coefficient, assimilable à la progression salariale entre 30 et 45 ans, est également plus dispersé pour les diplômés, particulièrement ceux de la FPH (les quartiles extrêmes valent 0,12 et 0,71). Il est presque tout le temps positif sauf dans le secteur privé. Dans ce secteur, le premier quartile est négatif pour les femmes et les hommes les moins diplômés. Cela signifie qu'une partie non négligeables des salariés du secteur privé voient leur salaire progresser moins vite que le salaire médian entre 30 et 45 ans.

L'écart entre les carrières des hommes et celles des femmes est particulièrement marqué pour les plus diplômés. Cet écart concerne aussi bien le niveau moyen de salaire que sa progression entre 30 et 45 ans.

Éléments de distribution des coefficients c_0 , c_1 et c_2

			C0			C1			C2		
Secteur	Diplôme	sexe	q1	q2	q3	q1	q2	q3	q1	q2	q3
Enseignants	Bac ou moins	Femmes	ns	ns	ns	ns	ns	ns	ns	ns	ns
		Hommes	ns	ns	ns	ns	ns	ns	ns	ns	ns
	Plus que Bac	Femmes	1,167	1,320	1,521	0,063	0,122	0,188	-0,069	-0,026	0,017
		Hommes	1,319	1,518	1,826	0,097	0,153	0,261	-0,104	-0,034	0,013
Etat (hors ens.)	Bac ou moins	Femmes	0,824	0,963	1,115	0,026	0,060	0,114	-0,042	-0,015	0,011
		Hommes	1,028	1,233	1,426	0,044	0,084	0,141	-0,043	-0,020	0,002
	Plus que Bac	Femmes	1,025	1,200	1,445	0,036	0,094	0,187	-0,050	-0,015	0,023
		Hommes	1,286	1,537	2,009	0,090	0,172	0,300	-0,071	-0,033	0,005

Fonction publique hospitalière	Bac ou moins	Femmes	0,987	1,095	1,237	0,028	0,072	0,121	-0,066	-0,028	0,002
		Hommes	1,016	1,122	1,256	0,033	0,079	0,118	-0,050	-0,021	0,007
	Plus que Bac	Femmes	1,091	1,339	1,572	0,046	0,123	0,207	-0,061	-0,018	0,033
		Hommes	1,310	1,691	2,353	0,115	0,231	0,709	-0,076	-0,027	0,049
Fonction publique territoriale	Bac ou moins	Femmes	0,606	0,851	0,992	0,035	0,080	0,146	-0,048	-0,016	0,020
		Hommes	0,906	1,047	1,201	0,041	0,079	0,138	-0,047	-0,017	0,014
	Plus que Bac	Femmes	0,877	1,083	1,309	0,047	0,103	0,189	-0,063	-0,017	0,028
		Hommes	1,079	1,269	1,610	0,062	0,124	0,215	-0,066	-0,021	0,016
Secteur privé	Bac ou moins	Femmes	0,578	0,768	0,968	-0,017	0,043	0,114	-0,047	-0,003	0,047
		Hommes	0,840	1,018	1,258	-0,010	0,053	0,124	-0,050	-0,006	0,037
	Plus que Bac	Femmes	0,890	1,169	1,570	-0,003	0,092	0,206	-0,073	-0,008	0,059
		Hommes	1,307	1,801	2,544	0,031	0,163	0,361	-0,105	-0,021	0,055

Champ : salariés du secteur privé et public âgés entre 30 et 45 ans.

Source : panel tous salariés 1988-2008.

4. Un indicateur de dispersion pour les carrières salariales

4.1. Définition de l'indicateur de dispersion

Une partie de la dispersion des salaires tient à une évolution avec l'âge partagée peu ou prou par l'ensemble des salariés : en moyenne, on gagne tant de plus à 30 ans qu'à 20, tant de plus à 40 ans qu'à 30. Pour un salarié donné, un mauvais classement dans la hiérarchie à un âge jeune peut être compensé par une meilleure situation ultérieure. C'est pourquoi nous souhaitons autant que possible neutraliser ce phénomène de progression avec l'âge dans le calcul de l'indicateur de dispersion. L'indicateur que nous définissons mesure la dispersion des carrières du groupe de salariés considérés (définis par le secteur, le sexe et le niveau de diplôme) autour de la carrière moyenne de ce groupe. Celle-ci est simplement définie en prenant les moyennes à chaque âge.

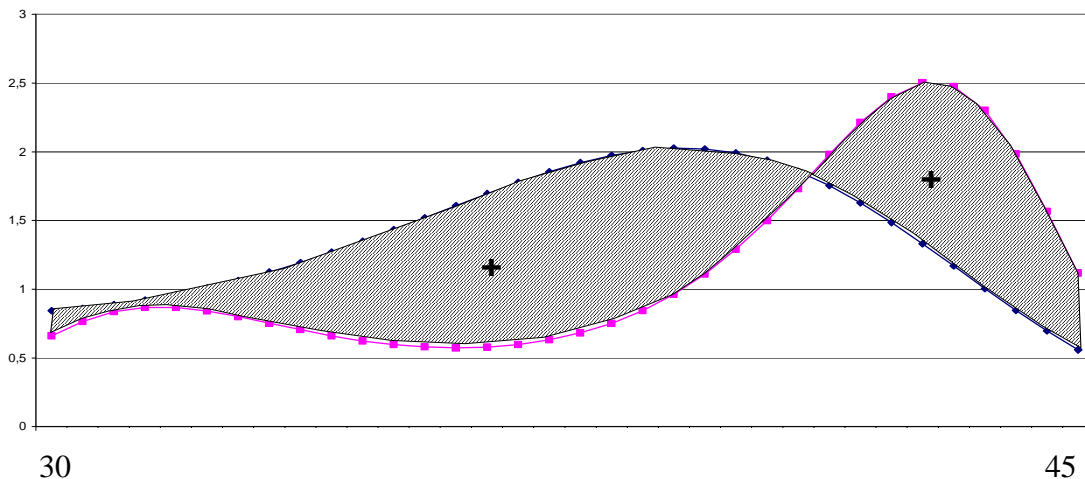
$$W_{\text{moy}} = (w_{\text{moy},a})_{a \in IA}$$

La « distance » de la carrière de l'individu i à la carrière moyenne du groupe est la distance euclidienne usuelle pour les vecteurs, normalisée par la longueur de la période considérée.

$$\|W_i - W_{\text{moy}}\|_{L2}^2 = \frac{1}{\#IA} \sum_{a \in IA} (w_{i,a} - w_{\text{moy},a})^2$$

Il s'agit simplement de l'écart quadratique moyen entre la carrière de l'individu i et la carrière moyenne du groupe auquel il appartient. Il peut être représenté géométriquement comme l'aire (quadratique) comprise entre les courbes définissant la carrière de l'individu i et la carrière moyenne du groupe. Une carrière est éloignée de la carrière moyenne dès qu'elle s'en éloigne à un instant donné.

Distance entre deux carrières : la contribution de chaque âge à la distance entre les deux carrières est toujours comptée positivement



L'indicateur d'inégalité que l'on étudie se contente de faire la moyenne de cette distance sur tous les individus (en nombre N).

$$I(W) = \frac{1}{N} \sum_i \|w_i - w_{\text{moy}}\|_{L^2}^2$$

Il s'agit donc de la « taille » du halo autour de la carrière moyenne.

4.2. Une formule de décomposition de l'indicateur

Une propriété intéressante de l'indicateur de dispersion ainsi défini est qu'il peut lui-même être décomposé de manière cohérente avec la décomposition en fonctions élémentaires. L'intérêt d'une telle décomposition est de faire apparaître les contributions à l'indicateur de dispersion des différentes caractéristiques de la carrière capturées par le jeu de fonctions $F_K = (f_0, \dots, f_k, \dots, f_K)$. En partant de la décomposition en fonctions élémentaires :

$$w_{i,a} = c_i^0 \cdot f_0(a) + c_i^1 \cdot f_1(a) + \dots + c_i^K \cdot f_K(a) + u_{i,a}$$

Sous la condition que le jeu de fonctions F_K est orthonormé, ce qu'on peut toujours supposer quitte à redéfinir légèrement le jeu de fonctions (encadré 1), la décomposition obtenue, assez simple, fait intervenir la variance interindividuelle des différents coefficients ainsi qu'un terme résiduel $I(U)$. La

contribution résiduelle $I(U)$ est simplement la valeur de l'indicateur de dispersion pour le résidu de l'équation de salaire.

$$I(W) = Var(c_0) + Var(c_1) + \dots + Var(c_k) + I(U) \quad (1)$$

Démonstration de la formule (1)

Pour la démonstration, nous adoptons une notation vectorielle plus concise.

$$W_i = F.C_i + U_i$$

$$W_{moy} = F.C_{moy} + U_{moy}$$

Avec cette notation :

$$\begin{aligned} \|W_i - W_{moy}\|_{L2}^2 &= \frac{1}{\#IA} (W_i' - W_{moy}') \cdot (W_i - W_{moy}) \quad (2) \\ &= \frac{1}{\#IA} [(C_i' - C_{moy}') \cdot F' + (U_i' - U_{moy}')] [F \cdot (C_i - C_{moy}) + (U_i - U_{moy})] \end{aligned}$$

Les termes croisés de ce produit sont nuls. En effet chaque vecteur U_i est orthogonal à l'ensemble des fonctions de F . il en va de même par linéarité de U_{moy} qui est la moyenne des U_i .

$$U_i' F = U_{moy}' F = F' U_i = F' U_{moy} = 0$$

Par conséquent :

$$(U_i' - U_{moy}') F (C_i - C_{moy}) = 0$$

$$(C_i' - C_{moy}') \cdot F' (U_i - U_{moy}) = 0$$

De sorte que :

$$\|W_i - W_{moy}\|_{L2}^2 = \frac{1}{\#IA} [(C_i' - C_{moy}') \cdot F' F \cdot (C_i - C_{moy}) + (U_i' - U_{moy}') (U_i - U_{moy})]$$

Comme par ailleurs les fonctions F sont supposées orthonormales (encadré 1) : $F' F = F F' = I$

L'équation (2) s'écrit finalement :

$$\begin{aligned} \|W_i - W_{moy}\|_{L2}^2 &= \frac{1}{\#IA} [(C_i' - C_{moy}') \cdot (C_i - C_{moy}) + (U_i' - U_{moy}') (U_i - U_{moy})] \\ &= \frac{1}{\#IA} \left[\sum_{k=1}^K (c_k^i - c_{moy}^k)^2 \right] + \|U_i - U_{moy}\|_{L2}^2 \end{aligned}$$

En prenant la moyenne sur les individus pour calculer l'indicateur d'inégalités :

$$I(W) = \sum_{k=1}^K Var(c_k) + I(U)$$

Fin de la démonstration

Ainsi, pour le jeu de fonctions polynômiales, cette décomposition permet de préciser quelle part de la dispersion provient des différences de niveaux moyens entre les individus (coefficient c_0), quelle part provient des différences de pentes (coefficient c_1), etc. Dans la suite, cette décomposition est appliquée pour chaque sous-population définie par le croisement du secteur, du diplôme et du sexe.

4.3. Décomposition de l'indicateur et comparaison public-privé avec des fonctions polynômiales

La décomposition de l'indicateur $I(W)$ a été appliquée aux sous groupes de salariés définis par le secteur, le diplôme et le sexe afin d'identifier les sources de dispersion des carrières salariales propres à chaque groupe de salariés. De manière générale, les carrières sont plus dispersées pour les hommes et pour les plus diplômés. Elles sont le plus dispersées dans le secteur privé et le moins dispersées parmi les enseignants. La dispersion de la première composante (niveau moyen) est toujours prépondérante. Cependant, pour certains groupes de salariés, les deuxième et troisième composante (pente et courbure) ne sont pas négligeables. C'est le cas dans le secteur privé, dans la FPH et pour les diplômés hommes de la FPT. Enfin, la composante purement individuelle est assez élevée pour l'ensemble des diplômés sauf pour les enseignants.

Décomposition de l'indicateur de dispersion obtenu avec le jeu de fonctions polynômiales, par sous groupe de salariés

Secteur	Diplôme	sexe	VarC0/W2	VarC1/W2	VarC2/W2	VarC3/W2	VarC4/W2	IU/W2
Enseignants	Bac ou moins	Femmes	ns	ns	ns	ns	ns	ns
		Hommes	ns	ns	ns	ns	ns	ns
	Plus que Bac	Femmes	0,064	0,011	0,007	0,005	0,003	0,019
		Hommes	0,089	0,011	0,007	0,004	0,003	0,043
Etat (hors ens.)	Bac ou moins	Femmes	0,065	0,012	0,008	0,007	0,004	0,052
		Hommes	0,048	0,008	0,009	0,010	0,006	0,042
	Plus que Bac	Femmes	0,143	0,016	0,008	0,006	0,004	0,083
		Hommes	0,134	0,021	0,010	0,006	0,004	0,063
Fonction publique hospitalière	Bac ou moins	Femmes	0,060	0,017	0,009	0,010	0,008	0,040
		Hommes	0,037	0,008	0,005	0,004	0,003	0,019
	Plus que Bac	Femmes	0,109	0,028	0,010	0,007	0,005	0,030
		Hommes	0,165	0,064	0,031	0,030	0,011	0,109
Fonction publique territoriale	Bac ou moins	Femmes	0,131	0,019	0,013	0,008	0,007	0,023
		Hommes	0,069	0,013	0,012	0,006	0,010	0,055

	Plus que Bac	Femmes	0,139	0,044	0,023	0,018	0,012	0,103
		Hommes	0,163	0,125	0,020	0,039	0,016	0,130
Secteur privé	Bac ou moins	Femmes	0,198	0,045	0,033	0,029	0,020	0,053
		Hommes	0,196	0,039	0,019	0,018	0,014	0,082
	Plus que Bac	Femmes	0,310	0,057	0,039	0,028	0,017	0,146
		Hommes	1,294	0,889	0,153	0,056	0,035	1,551

Champ : salariés du secteur privé et public âgés entre 30 et 45 ans.

Source : panel tous salariés 1988-2008.

La dispersion particulièrement élevée pour la dernière catégorie du tableau (salariés diplômés du secteur privé, particulièrement les hommes) est due à un petit nombre de salariés ayant des salaires particulièrement élevés. Lorsqu'on filtre de manière générale les 1% de très hauts salaires, les résultats s'en trouvent fortement atténués.

Décomposition de l'indicateur de dispersion obtenu avec le jeu de fonctions polynômiales, par sous groupe de salariés, avec filtrage des très hauts salaires

Secteur	Diplôme	sexe	VarP0/W2	VarP1/W2	VarP2/W2	VarP3/W2	VarP4/W2	IU/W2
Enseignants	Bac ou moins	Femmes	ns	ns	ns	ns	ns	ns
		Hommes	ns	ns	ns	ns	ns	ns
	Plus que Bac	Femmes	0,064	0,011	0,007	0,005	0,003	0,019
		Hommes	0,072	0,009	0,007	0,004	0,003	0,027
Etat (hors ens.)	Bac ou moins	Femmes	0,065	0,012	0,008	0,007	0,004	0,052
		Hommes	0,048	0,008	0,009	0,010	0,006	0,042
	Plus que Bac	Femmes	0,143	0,016	0,008	0,006	0,004	0,083
		Hommes	0,125	0,021	0,009	0,006	0,004	0,058
Fonction publique hospitalière	Bac ou moins	Femmes	0,060	0,017	0,009	0,010	0,008	0,040
		Hommes	0,037	0,008	0,005	0,004	0,003	0,019
	Plus que Bac	Femmes	0,109	0,028	0,010	0,007	0,005	0,030
		Hommes	0,165	0,064	0,031	0,030	0,011	0,109
Fonction publique territoriale	Bac ou moins	Femmes	0,131	0,019	0,013	0,008	0,007	0,023
		Hommes	0,069	0,013	0,012	0,006	0,010	0,055
	Plus que Bac	Femmes	0,139	0,044	0,023	0,018	0,012	0,103
		Hommes	0,163	0,125	0,020	0,039	0,016	0,130
Secteur privé	Bac ou moins	Femmes	0,198	0,045	0,033	0,029	0,020	0,053

	Hommes	0,160	0,031	0,019	0,016	0,012	0,080
Plus que Bac	Femmes	0,216	0,046	0,029	0,024	0,013	0,107
	Hommes	0,153	0,028	0,014	0,011	0,010	0,229

Champ : salariés du secteur privé et public âgés entre 30 et 45 ans.

Source : panel tous salariés 1988-2008.

4.4. Un exemple alternatif utilisant un jeu de variables non continues fonctions constantes ou affines par morceaux

Un autre jeu de variables intéressant consiste à prendre des fonctions présentant une ou plusieurs discontinuités. Dans le cas présent, nous avons défini un jeu de 4 fonctions en considérant deux polynômes (1, âge) sur chaque partie de carrière [30,37] et [38,45].

Plus précisément le jeu alternatif comprend 4 fonctions : $G_3 = (1, g_1, g_2, g_3)$ avec $g_1 = 1_{30 \leq \text{age} \leq 37}$, $g_2 = 1_{30 \leq \text{age} \leq 37} * \text{age}$, $g_3 = 1_{38 \leq \text{age} \leq 45} * \text{age}$. Ces fonctions permettent de générer toutes les carrières salariales formées d'un profil linéaire entre 30 et 37 ans et d'un autre profil linéaire, indépendant du premier, entre 38 et 45 ans.

Globalement, l'adéquation est légèrement moins bonne avec le jeu de fonctions G_3 qu'avec le jeu de fonctions F_4 . Toutefois, l'écart de distribution des R^2 est limité (moins de 0,06 pour tous les quartiles de R2) sauf pour les diplômés de la FPH où le premier quartile de R2 est sensiblement moins élevé (0,55 pour le jeu G contre 0,63 pour le jeu F_4).

Comme cela était prévisible, la décomposition obtenue est sensiblement différente de celle découlant de la décomposition polynômiale. Globalement, la composante individuelle est plus importante dans la deuxième décomposition, ce qui rejoint le fait que les fonctions choisies expliquent moins bien les évolutions salariales. Néanmoins, l'ajustement est assez bon pour le secteur privé et fait apparaître l'importance de la composante c_3 (croissance du salaire sur la deuxième partie de carrière, entre 38 et 45 ans) dans les évolutions salariales des diplômés du secteur privé. La divergence importante entre les trajectoires salariales de ce groupe se produirait donc en deuxième partie de carrière et tiendrait à des différences d'évolution salariale sur cette période.

Décomposition de l'indicateur de dispersion obtenu avec le jeu de fonctions affines par morceaux sur deux parties de carrière, par sous groupe de salariés

Secteur	Diplôme	sexe	VarC0/W2	VarC1/W2	VarC2/W2	VarC3/W2	IUW2
Enseignants	Bac ou moins	Femmes	ns	ns	ns	ns	ns
		Hommes	ns	ns	ns	ns	ns
	Plus que Bac	Femmes	0,064	0,009	0,006	0,007	0,021
		Hommes	0,089	0,012	0,007	0,004	0,043
Etat (hors ens.)	Bac ou moins	Femmes	0,065	0,015	0,011	0,003	0,050
		Hommes	0,048	0,007	0,009	0,006	0,052
	Plus que Bac	Femmes	0,143	0,014	0,013	0,005	0,080
		Hommes	0,134	0,016	0,011	0,007	0,068
Fonction publique hospitalière	Bac ou moins	Femmes	0,060	0,013	0,012	0,007	0,046
		Hommes	0,037	0,007	0,006	0,003	0,022
	Plus que Bac	Femmes	0,109	0,023	0,011	0,010	0,033
		Hommes	0,165	0,048	0,047	0,014	0,163
Fonction publique territoriale	Bac ou moins	Femmes	0,131	0,017	0,016	0,006	0,024
		Hommes	0,069	0,011	0,011	0,008	0,056
	Plus que Bac	Femmes	0,139	0,029	0,042	0,010	0,109
		Hommes	0,163	0,124	0,035	0,009	0,166
Secteur privé	Bac ou moins	Femmes	0,198	0,037	0,031	0,032	0,058
		Hommes	0,196	0,036	0,020	0,016	0,088
	Plus que Bac	Femmes	0,310	0,044	0,035	0,035	0,166
		Hommes	1,294	0,711	0,041	0,322	1,695

Bibliographie

Bessière S. et Pouget J., « Les carrières dans la fonction publique d'État - Premiers éléments de caractérisation », in *Les salaires en France*, coll. « Insee Références », édition 2007.

Koubi M., « Les carrières salariales par cohorte de 1967 à 2000 », *Économie et Statistique* n° 369-370, 2003.

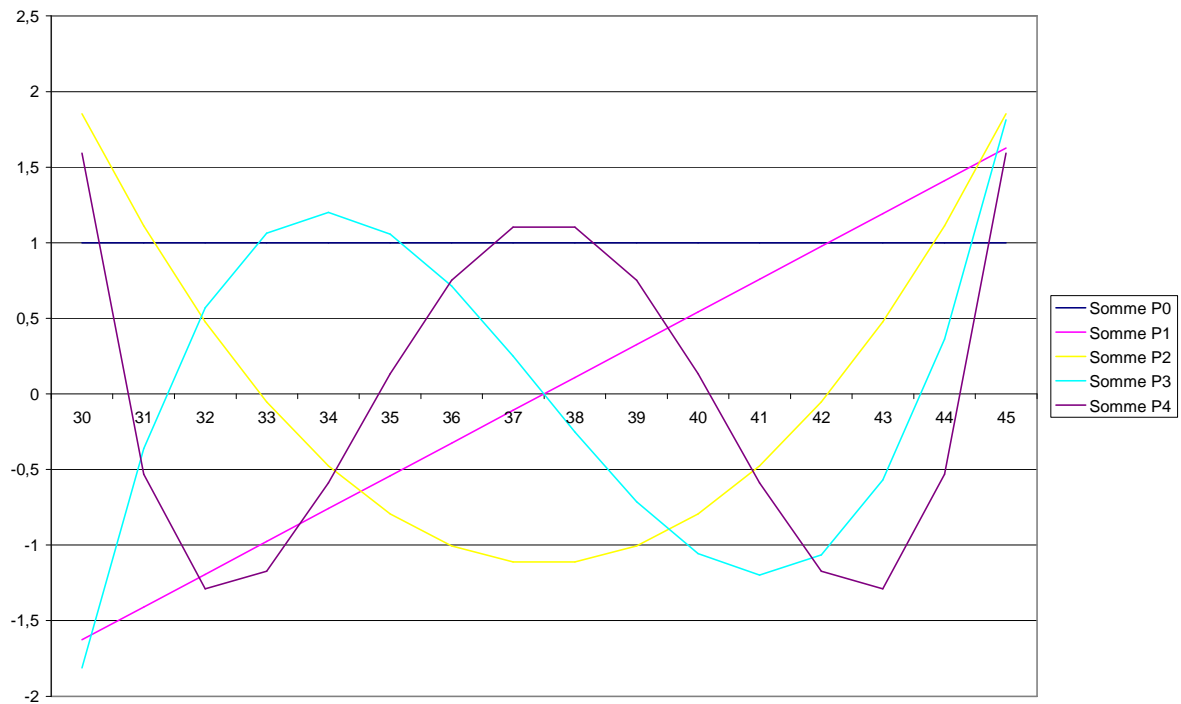
Lollivier S. et Payen J.F. (1990), « L'hétérogénéité des carrières individuelles mesurée sur données de panel », *Économie et Prévision*, n° 92-93, pp. 87-96.

Mincer J. (1958), « Investment in Human Capital and Personal Income Distribution », *Journal of Political Economy*, n° 4, pp. 282-302.

Mortensen, D. and C. Pissarides, "Job Creation and Job Destruction in the Theory of Unemployment", *Review of Economic Studies* 61 (1994), 397-415.

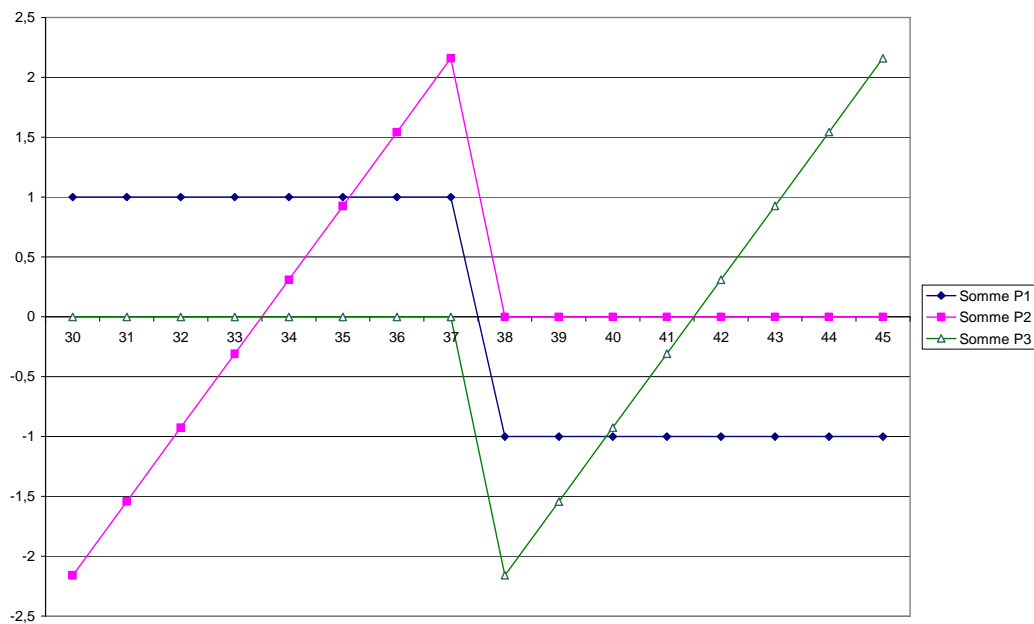
Annexe 1 : jeux de fonctions élémentaires de l'âge utilisées dans l'étude

1) Jeu de fonctions polynômiales principalement utilisé



Lecture : ce graphique représente le premier jeu de fonctions élémentaires, polynômiales entre 30 et 45 ans, une fois orthonormalisées, utilisées pour décomposer chaque carrière individuelle.

2) Jeu de fonctions affines sur chaque partie de carrière



Lecture : ce graphique représente le second jeu de fonctions élémentaires, affines sur chaque partie de carrière (30-37 ans et 38-45 ans), une fois orthonormalisées, utilisées pour décomposer chaque carrière individuelle.