

ENSAE
Projet de Statistiques Appliquées



Analyse des déterminants des promotions dans la fonction publique de l'État

Eve Samani, Manon Maugard, Sacha Nass, Claire du
Campe de Rosamel

Sous la direction de Benjamin Gilbert (Insee)

Mai 2025

Table des matières

Introduction	3
1 Cadre théorique et institutionnel	4
1.1 La Fonction publique française : un cadre formel d'égalité salariale	4
1.2 Revue de littérature : trajectoires standardisées, inégalités persistantes . . .	5
1.2.1 Un champ historiquement sous-étudié	5
1.2.2 Facteurs influençant les carrières dans la fonction publique	5
1.2.3 Genre et promotion : un enjeu central pour notre étude	5
1.2.4 Un cas d'étude pertinent : le ministère de la Justice	6
2 Les données et la description de l'échantillon	6
2.1 Les bases administratives	6
2.1.1 Fichier général de l'État (FGE)	7
2.1.2 Système d'Information sur les Agents des Services Publics (SIASP) .	7
2.1.3 Les enjeux de la définition des bases de données FGE et SIASP . . .	8
2.2 Quelques statistiques descriptives	9
3 Modélisation statistique des trajectoires professionnelles des fonctionnaires du Ministère de la Justice	11
3.1 Décomposition de la variance des salaires : la rémunération dans la fonction publique est-elle aussi transparente qu'annoncée ?	11
3.2 Des injustices au ministère de la Justice entre hommes et femmes ? Décomposition d'Oaxaca-Blinder	13
3.3 Les modèles de clustering : construction d'une typologie des trajectoires . . .	14
3.3.1 De la séquence au regroupement : les étapes de la méthode	14
3.3.2 Visualisation de résultats	15
3.4 Modéliser les temps d'accès à la promotion : approche par les modèles de durée	16
4 Les résultats : écarts de rémunération et trajectoires professionnelles	17
4.1 Une rémunération largement encadrée par les statuts administratifs, mais avec des marges résiduelles	17
4.2 Oaxaca-Blinder : dynamique d'explication de l'écart salarial entre hommes et femmes	18
4.3 Clustering : les trajectoires de salaires sur 25 ans de carrière	20
4.4 Les modèles de durée	23
4.4.1 Premiers modèles de Cox : limitations et tentatives de stratification .	23
4.4.2 Taux de promotion et <i>frailty effect</i> : comprendre l'effet structurel des départements de fonction sur les promotions au sein de la magistrature	25
Conclusion	29
Références	31

Appendix	32
A Théorie - Oaxaca-Blinder	32
A.1 Explication de la décomposition	32
A.2 Formalisation de la décomposition	32
B Robustness checks - Oaxaca-Blinder	33
B.1 Choix des variables et multicollinearité	33
B.2 Sensibilité aux filtres appliqués et bootstrap	34
B.3 Etude des années "instables" dans la décomposition Oaxaca-Blinder	34
C Modèles de durée	35
C.1 Fonctions de survie et de hasard	35
C.2 Estimations non-paramétriques	36
C.3 Modéliser la durée de vie : une première approche par le modèle de Cox simple	36
C.4 Les extensions du modèle de Cox - modèle de Cox stratifié et modèle Frailty	37
D Résultats complémentaires - modèles de survie	38
D.1 Résultats modèles de Cox avec stratification par genre	38
D.2 Statistiques descriptives complémentaires pour les modèles de survie	39

Disponibilité des scripts

Notre code R a été remis à l'encadrant via notre repo du Gitlab interne de l'Insee. Celui-ci requiert un accès aux serveurs de l'Insee.

Introduction

En apparence, les carrières dans la fonction publique sont encadrées par des règles claires et universelles : concours d'entrée, grilles indiciaires, progressions d'échelons selon l'ancienneté, avancements conditionnés à des critères statutaires. Toutes ces règles donnent donc une image de neutralité et d'égalité de traitement, dans un secteur fortement féminisé, où la transparence salariale est censée réduire les inégalités. Pourtant, de nombreux événements non attendus marquent les trajectoires de vie professionnelle des agents publics (changements de statut, changements de corps, promotions, détachements, sorties de la fonction publique, etc.). De nombreuses études mettent en évidence la persistance de disparités de carrière, notamment entre hommes et femmes, y compris dans la fonction publique. En particulier, le phénomène de plafond de verre semble encore à l'œuvre, en dépit des politiques d'égalité promues depuis plusieurs décennies.

C'est dans ce contexte que s'inscrit notre étude, consacrée aux trajectoires professionnelles dans la fonction publique de l'État et plus particulièrement aux déterminants des promotions dans celle-ci, avec une attention particulière portée aux inégalités de genre. Nous avons choisi d'analyser le cas du ministère de la Justice, qui présente un intérêt particulier car il a connu une croissance importante de ses effectifs, une féminisation marquée de plusieurs de ses corps, notamment dans la magistrature, et des réformes statutaires modifiant en profondeur l'organisation des carrières. Cela en fait un échantillon pertinent pour questionner les dynamiques de progression, les rigidités structurelles éventuelles, et les écarts de rémunération au sein d'un espace supposé égalitaire.

Notre analyse se déroule en deux temps. Dans un premier temps nous évaluons dans quelle mesure les écarts de salaire observés entre agents peuvent être expliqués par leurs caractéristiques observables (âge, corps, statut, quotité de travail, etc.). Il s'agit de mettre à l'épreuve l'hypothèse d'une transparence salariale effective dans la fonction publique, à travers des modèles de régression et des décompositions statistiques (variance, Oaxaca-Blinder). Dans un second temps, nous adoptons une approche longitudinale : les écarts de traitement ne se manifestent pas seulement à un instant donné, mais tout au long de parcours différenciés. En mobilisant des modèles de survie et des méthodes de classification des trajectoires (clustering), nous cherchons à identifier des profils types, à en décrire les régularités, et à observer leur répartition selon le sexe, le corps d'appartenance ou la catégorie hiérarchique.

Cette étude de cas porte donc sur les trajectoires de rémunération, en analysant à la fois les écarts instantanés entre individus et à la dynamique de carrière sur le long terme. Elle s'attache particulièrement à la dimension genrée de ces trajectoires : dans un contexte de forte féminisation de la fonction publique et d'apparente transparence salariale, il est crucial de comprendre si et comment des inégalités de traitement persistent. La question centrale devient alors : comment expliquer le maintien d'un écart salarial entre femmes et hommes dans un système censé garantir l'égalité ? Cette problématique est explorée en profondeur à travers l'exemple du ministère de la Justice.

Cette étude débute par une revue de littérature sur les trajectoires professionnelles, notamment dans le secteur public, et une présentation du cadre institutionnel de la Fonction publique. Nous décrivons ensuite les données mobilisées et les principaux défis liés à l'exploitation de bases administratives. La troisième partie détaille les méthodes statistiques utilisées : décomposition de variance, décomposition d'Oaxaca-Blinder, modèles de survie et méthodes

de clustering. Enfin, nous présentons les résultats de l'analyse, en identifiant les déterminants des rémunérations, les dynamiques d'explication de l'écart salarial entre femmes et hommes, ainsi que les trajectoires types de carrière et leur répartition.

1 Cadre théorique et institutionnel

1.1 La Fonction publique française : un cadre formel d'égalité salariale

La fonction publique française est fondée sur des principes de recrutement et d'évolution de carrière censés garantir l'égalité de traitement entre les agents. Elle comprend tous les agents appartenant à la fonction publique d'État (FPE), à la fonction publique hospitalière (FPH) et la fonction publique territoriale (FPT). Cependant tous les agents ne sont pas fonctionnaires car on comprend la fonction publique comme l'ensemble des personnes travaillant pour le secteur public. Ceux qui ne sont pas fonctionnaires (ou titulaires) sont contractuels (ou non titulaires, à l'exception des fonctionnaires stagiaires et élèves). Les fonctionnaires civils¹ sont répartis au sein de plusieurs niveaux hiérarchiques : les Catégories (A, B et C) elles-mêmes divisées en Corps qui constituent la base d'organisation de la fonction publique. Ces corps sont eux mêmes subdivisés en Grades (ou Classes) qui correspondent à des emplois, puis en échelons.

Le recrutement dans la fonction publique se fait par concours : concours externe pour les titulaires d'un diplôme, concours interne pour les agents pouvant justifier d'une ancienneté suffisante dans la fonction publique. Il existe trois formes d'évolution professionnelle : l'avancement d'échelon, la promotion de grade ou de classe ou le passage à un autre corps.

- L'avancement d'échelon se fait principalement à l'ancienneté. Dans certains corps ou cadres d'emplois peu nombreux, l'ancienneté requise peut être modulée en fonction de l'évaluation portée par la hiérarchie.
- L'avancement de classe ou de grade se fait au choix, c'est-à-dire en fonction de l'appréciation hiérarchique.
- Le changement de corps se fait en passant un concours interne, ou plus exceptionnellement, par promotion au choix. Ce type de promotion est souvent réservé aux fonctionnaires qui remplissent des conditions d'échelon, de temps passé dans cet échelon et d'âge déterminé et/ou d'ancienneté globale dans la fonction publique.

Ce cadre très normé laisse penser que les inégalités salariales devraient y être réduites, en particulier en comparaison avec le secteur privé. Pourtant, cette transparence *de jure* ne garantit pas l'absence d'inégalités *de facto*. Les écarts observés dans les trajectoires de carrière, notamment entre hommes et femmes, interrogent l'effectivité de cette égalité statutaire.

1. Les militaires ont des statuts spécifiques.

1.2 Revue de littérature : trajectoires standardisées, inégalités persistantes

1.2.1 Un champ historiquement sous-étudié

Les travaux de Cartier et Join-Lambert (2011) soulignent que les promotions et mobilités administratives ont longtemps été ignorées des analyses statistiques. Avant la création du compte des agents publics par l’Insee en 1946, le nombre exact d’agents publics était inconnu. À partir des années 1970, sociologues et statisticiens ont développé une réflexion critique sur les catégories utilisées dans le comptage des agents. Jusque-là, le ministère du Travail ne recensait que les salariés et indépendants, tandis que le ministère de la Fonction publique avait un rôle plus administratif que statistique. En 2000, la création de l’Observatoire de l’Emploi Public a marqué une structuration de l’analyse des carrières publiques.

Quelques travaux du tournant des années 1960, notamment Long et Blanc (1969) et Darbel et Schnapper (1969), traitent de la relation entre mobilité géographique et sociale, notamment en Île-de-France, au long des Trente Glorieuses. Depuis les années 1980, la recherche s’est concentrée sur les différences de trajectoires selon le genre et le niveau hiérarchique, mais avec pour limite le manque d’informations sur la catégorie socio-professionnelle des parents (Cartier et Join-Lambert, 2011).

1.2.2 Facteurs influençant les carrières dans la fonction publique

Les travaux de Bessière et Pouget (2007) indiquent que la décision de rester dans la fonction publique dépend principalement du statut à l’entrée (titulaire ou contractuel). Le taux de maintien diminue avec l’ancienneté, l’âge, le sexe, le niveau de diplôme et le grade. Ils soulignent également une accélération en 25 ans des promotions de catégorie B en A, vraisemblablement grâce à l’augmentation des qualifications. En revanche, les promotions de C en B sont plus rares. Ce phénomène s’est accompagné en parallèle de la titularisation des agents non titulaires restants via plusieurs plans de titularisation dans les années 1980 et 1990.

La fonction publique se distingue du secteur privé sur plusieurs points. Dubar *et al.* (2015) montre que les professions intermédiaires et les employés sont surreprésentés dans la fonction publique par rapport au secteur privé, tandis que les ouvriers y sont sous-représentés. Les femmes y sont également plus nombreuses que dans le privé. Gollac (2005) révèle qu’à diplôme égal, un enfant d’ouvrier a plus de chances de devenir fonctionnaire qu’un enfant de cadre, et que la mobilité ascendante y est plus fréquente que dans le privé. La fonction publique offre ainsi des opportunités d’ascension sociale pour les classes populaires, mais sous conditions : un bon niveau de diplôme est indispensable pour bénéficier de cette ascension sociale. Elle favorise particulièrement les femmes et les enfants de classes populaires, mais certaines filières, comme l’enseignement et l’administration, sont plus propices à cette ascension.

1.2.3 Genre et promotion : un enjeu central pour notre étude

Bessière et Pouget (2007) notent que les femmes sont plus enclines à rester dans la fonction publique, ce qui contribue à leur surreprésentation. En effet, Gollac (2005) confirme que les

femmes bénéficient d'un effet protecteur dans la fonction publique, où les inégalités de genre sont moins marquées que dans le privé.

Cependant, Marry *et al.* (2017) met en lumière les obstacles qui freinent la carrière des femmes dans l'administration publique, malgré des politiques en faveur de l'égalité professionnelle (quotas, actions de sensibilisation), notamment au sein des postes les plus prestigieux. Elle souligne une inégalité existant dès le recrutement, les concours reproduisant les inégalités de genre, et qui se perpétue ensuite via une culture du travail défavorable aux mères et une culture de la cooptation masculine. Malgré des politiques publiques bénéfiques, celles-ci restent limitées, contournées ou perçues comme contraignantes. Ces travaux révèlent l'existence d'un frein dans les administrations à l'accès des femmes aux postes de responsabilité.

1.2.4 Un cas d'étude pertinent : le ministère de la Justice

Le ministère de la Justice constitue un terrain particulièrement intéressant pour interroger ces dynamiques. Il représentait en 2023 plus de 91 000 agents en équivalent temps plein, dont environ 36 000 affectés à la justice judiciaire (magistrats, greffiers, personnels administratifs) et 43 400 à l'administration pénitentiaire (Ministère de la Justice, 2024). Ce ministère a aussi connu une forte féminisation : la proportion de femmes était de 70% parmi les juges et de 61% parmi les procureurs. Cette féminisation est plus marquée en première instance que dans les juridictions supérieures (cours d'appel, Cour de cassation, Conseil d'État).

Demoli et Willemez (2023) ont étudié les trajectoires des magistrats de l'ordre judiciaire. Le parquet, généralement plus masculin, est plus souvent investi par les hommes, tandis que les femmes sont davantage représentées au siège. Ces dernières accèdent moins fréquemment et plus tardivement aux fonctions de chef de juridiction, même à ancienneté et mode de recrutement équivalents. L'étude met en évidence un plafond de verre, renforcé par des écarts de capital social et d'ancienneté, identifiés via une Analyse de Correspondances Multiples (ACM).

L'étude met également en avant une double mobilité des magistrats : géographique et fonctionnelle, à la fois choisie et imposée, et qui joue un rôle central dans ces trajectoires. Bien que les magistrats soient globalement protégés par leur statut (inamovibilité, indépendance), cette mobilité est inégalement valorisée : la mobilité géographique, plus souvent masculine, favorise la promotion, tandis que la mobilité fonctionnelle, plus féminine, est moins récompensée.

Trois profils types ressortent de l'analyse d'une cohorte des années 1980 : une carrière au siège, une carrière au parquet, et une carrière administrative ou en détachement, avec des écarts marqués d'accès au hors hiérarchie. Ces résultats soulignent les spécificités du ministère de la Justice et justifient notre choix de le prendre comme terrain d'étude pour analyser les trajectoires salariales et leurs dynamiques genrées.

2 Les données et la description de l'échantillon

2.1 Les bases administratives

Pour analyser les trajectoires salariales dans la fonction publique, nous mobilisons deux bases administratives, le Fichier Général de l'État (FGE) et le Système d'Information sur les Agents des Services Publics (SIASP), couvrant plus de quarante années de carrière. Ces

données permettent de suivre l'évolution individuelle des agents, d'identifier les promotions, et d'observer les écarts de rémunération selon le sexe, le statut, le corps, etc.

Nous sommes soumis aux mêmes règles en matière de secret statistique qu'à l'Insee. Toute case doit ainsi comporter au moins cinq agents, et aucune personne ne doit représenter plus de 80% de la masse salariale de la case. De même, il nous est interdit de rapporter un résultat portant sur moins de trois établissements (SIRET) ou au sein duquel un seul établissement représente plus de 85% de l'effectif ou de la grandeur étudiée.

2.1.1 Fichier général de l'État (FGE)

Le Fichier Général de l'État (FGE) est une base de données historique permettant d'analyser les effectifs et les rémunérations des agents de la fonction publique de l'État en France. Il couvre une période allant de 1978 à 2009 (hors 1979, 1981, 1987 et 1993 à 1995 inclus), avant d'être remplacé par le Système d'Information sur les Agents des Services Publics (SIASP) à partir de 2010. FGE était notamment utilisé pour la production des estimations annuelles d'emploi et pour l'étude des rémunérations des agents de l'État.

La base FGE est constituée à partir de sources administratives, principalement les fichiers de paie de la Direction Générale des Finances Publiques (DGFIP), qui recensent l'ensemble du personnel payé sur crédits budgétaires de l'État. Pour compléter ces informations, les Déclarations Annuelles de Données Sociales (DADS) sont utilisées afin d'inclure les agents des établissements publics administratifs (EPA) et des établissements publics à caractère industriel et commercial (EPIC) qui ne sont pas directement rémunérés par la DGFIP, mais sur budget propre. Ainsi, certains agents rémunérés par des circuits spécifiques peuvent être hors champ. Toutefois, certaines situations particulières sont prises en compte, comme le personnel militaire, recensé grâce à des fichiers fournis par le ministère de la Défense. En définitive, la couverture de la base s'étend à l'ensemble des agents de la fonction publique d'État en France métropolitaine, dans les départements d'outre-mer (DOM) ainsi qu'à l'étranger.

Les principales variables de la base permettent d'étudier les effectifs des agents selon leur statut (titulaire, contractuel, militaire, etc.), les caractéristiques de leur emploi, leur catégorie statutaire, leur grade ou encore leur rémunération. Le suivi de cohorte est rendu possible grâce à la présence du NIR². Enfin, la base permet de distinguer l'emploi principal de l'emploi secondaire, ce qui évite les doubles comptes liés à la multi-activité des agents publics.

2.1.2 Système d'Information sur les Agents des Services Publics (SIASP)

Le Système d'Information sur les Agents des Services Publics (SIASP) est une base de données produite par l'INSEE, qui recense annuellement des informations sur l'emploi et les caractéristiques des agents de la fonction publique en France. Cette base couvre les trois versants de la fonction publique : État (FPE), territoriale (FPT) et hospitalière (FPH). Sa première année de référence est 2010, et les données sont mises à jour chaque année. SIASP est aujourd'hui la source centrale pour l'analyse de l'emploi public, notamment pour suivre les évolutions des effectifs, des statuts et des grades des agents des services publics.

2. A noter que nous n'avons pas accès directement aux NIRs mais à un identifiant non significatif qui nous permet néanmoins de suivre les individus dans le temps.

Tout comme FGE, SIASP est construite à partir de sources administratives. Les données proviennent principalement de quatre sources :

- les fichiers de paie de la DGFIP, qui couvrent les agents civils des ministères et certains établissements publics
- la Déclaration Annuelle de Données Sociales (DADS), qui fournit des informations sur les agents des fonctions publiques territoriale et hospitalière, ainsi que ceux de la fonction publique d’État non couverts par la DGFIP ;
- les données du ministère de la Défense, qui recensent les personnels militaires (y compris les gendarmes jusqu’en 2018) ;
- les fichiers de la Trésorerie Générale pour l’Étranger (TGE), qui incluent les agents en poste à l’étranger.

Depuis 2017, la Déclaration Sociale Nominative (DSN) a progressivement remplacé la Déclaration Annuelle de Données Sociales (DADS) comme source principale de données sociales transmises par les employeurs.³

SIASP recense tous les postes des agents publics sous contrat, qu’ils soient titulaires ou non, tant que leur employeur est une personne morale ou un organisme relevant du droit administratif. Un poste est défini par le cumul des périodes d’emploi d’un salarié dans un même établissement, quel que soit son employeur. La base fournit des informations détaillées sur les grades, la catégorie statutaire (A, B, C), le ministère de tutelle, la durée travaillée, etc. Enfin, tout comme FGE, le suivi de cohortes y est rendu possible par la présence du NIR (éventuellement, NIR haché). Comme toute base administrative, il existe également certaines imperfections et incohérences, notamment quant aux données sur les militaires, incomplètes sur certains millésimes en raison du secret Défense.

2.1.3 Les enjeux de la définition des bases de données FGE et SIASP

Le premier enjeu de ce projet a été de parvenir à fusionner deux systèmes de répertorisation des agents publics, FGE et SIASP, construits sur des nomenclatures différentes, sans table de conversion officielle permettant de faire correspondre les codes de grade et de corps d’une base à l’autre. Cette absence de référentiel commun a constitué la principale difficulté de traitement : comment sélectionner avec fiabilité les individus relevant du ministère de la Justice dans un univers administratif aussi vaste et hétérogène ?

Une variable d’identification du ministère de rattachement existe dans FGE, mais elle est connue pour être imparfaite et instable dans le temps. Nous l’avons néanmoins utilisée dans un premier temps comme filtre exploratoire pour extraire une majorité des libellés de grades et de corps potentiellement rattachés au ministère de la Justice. Ces libellés ont été vérifiés un par un, ce qui nous a permis de constituer une première liste de codes ANE (code de nomenclature FGE) plausibles.

Pour obtenir une correspondance avec SIASP, nous avons extrait les NIR (identifiants individuels) des agents présents dans cette sélection initiale, puis nous les avons suivis dans SIASP. Cela nous a permis de retrouver les codes NNE (nomenclature propre à SIASP) utilisés pour ces mêmes agents, ainsi que les libellés de grades associés. Grâce à quelques

3. En 2017, environ 75 % des entreprises privées utilisaient la DSN pour leurs déclarations sociales, et ce taux est passé à 99 % en 2018 . La DADS a été définitivement supprimée en 2023 . La DSN est devenue obligatoire pour les employeurs publics à partir de 2022.

correspondances ANE/NNE connues et à une comparaison systématique des libellés de grades dans les deux bases, nous avons pu établir une table de conversion entre les deux systèmes. À partir de cette table, nous avons défini des regroupements de grades en « corps » au sens large (des catégories plus englobantes que les corps statutaires de la fonction publique) afin de structurer les analyses, comme présenté dans la Figure 1c.

Une fois cette classification consolidée, nous avons pu réappliquer un filtrage plus fiable sur FGE et SIASP, cette fois fondé non plus sur la variable ministérielle, mais sur la liste de codes de grades validés.

Malgré ce travail de réconciliation, plusieurs incohérences subsistent. Certains grades classés dans les fonctions support présentent des effectifs très irréguliers au fil du temps, rendant leur inclusion problématique ; ils ont été écartés de l'étude. D'autres anomalies, comme la disparition totale des greffiers en chef à partir de 2016, sont probablement liées à des réformes statutaires intervenues à cette période, notamment celle de 2015 concernant les personnels de greffe. Enfin, une évolution brutale est observée à la fin de la période : à partir de 2018, les effectifs en catégorie A dépassent ceux de la catégorie B (cf. Figure 3b), ce qui correspond au reclassement du corps de commandement de l'administration pénitentiaire (lieutenants, capitaines et commandants) en catégorie A suite à une réforme statutaire (Ministère de la Justice, 2018).

Ces incertitudes sur la codification, la stabilité des nomenclatures et les ruptures de séries historiques constituent des limites importantes à avoir à l'esprit. Nous avons tenté de les contenir autant que possible, mais elles peuvent affecter certains résultats et appellent à une lecture prudente des analyses longitudinales.

Une fois ces traitements réalisés et les individus d'intérêt correctement identifiés, nous avons pu constituer un échantillon d'agents du ministère de la Justice cohérent sur le temps long. Ce socle de données permet d'examiner plusieurs caractéristiques clés des agents (notamment l'âge, le sexe, le statut, la catégorie), que nous présentons dans la section suivante à travers quelques statistiques descriptives.

2.2 Quelques statistiques descriptives

On observe sur la Figure 1a une augmentation quasi linéaire du nombre d'agents du Ministère de la Justice depuis 1978. Quand on décompose cette augmentation par corps (Figure 1c) on remarque qu'elle est principalement due à la forte augmentation des effectifs des services pénitentiaires. D'autres corps ont également connu une hausse de leurs effectifs mais dans une moindre mesure. À noter que le corps des greffiers en chef disparaît de nos bases de données à partir de 2016, comme précisé dans la partie 2.1.3. L'année 2010, signifiée par la droite en pointillés, marque le changement de base de données entre FGE et SIASP.

De plus, l'évolution des effectifs par sexe n'est pas la même (Figure 1b). Si en 1978 il y avait environ 10 000 femmes de moins que d'hommes, en 2021 cet écart s'est complètement résorbé. La part de femmes au sein des corps du Ministère de la Justice a donc considérablement augmenté, grâce à la forte féminisation du marché du travail, de la fonction publique et plus spécifiquement du Ministère de la Justice.

L'évolution des effectifs par corps est également marquée par sexe comme on peut le voir sur la Figure 1d. Ce sont les hommes qui constituent l'essentiel de l'augmentation des effectifs des services pénitentiaires tandis que leurs effectifs sont restés stables dans les autres corps.

Les femmes constituent donc la principale source d'augmentation des corps de magistrat, de greffier et des services de protection de la jeunesse, bien que leurs effectifs au sein des services judiciaires aient également augmenté.

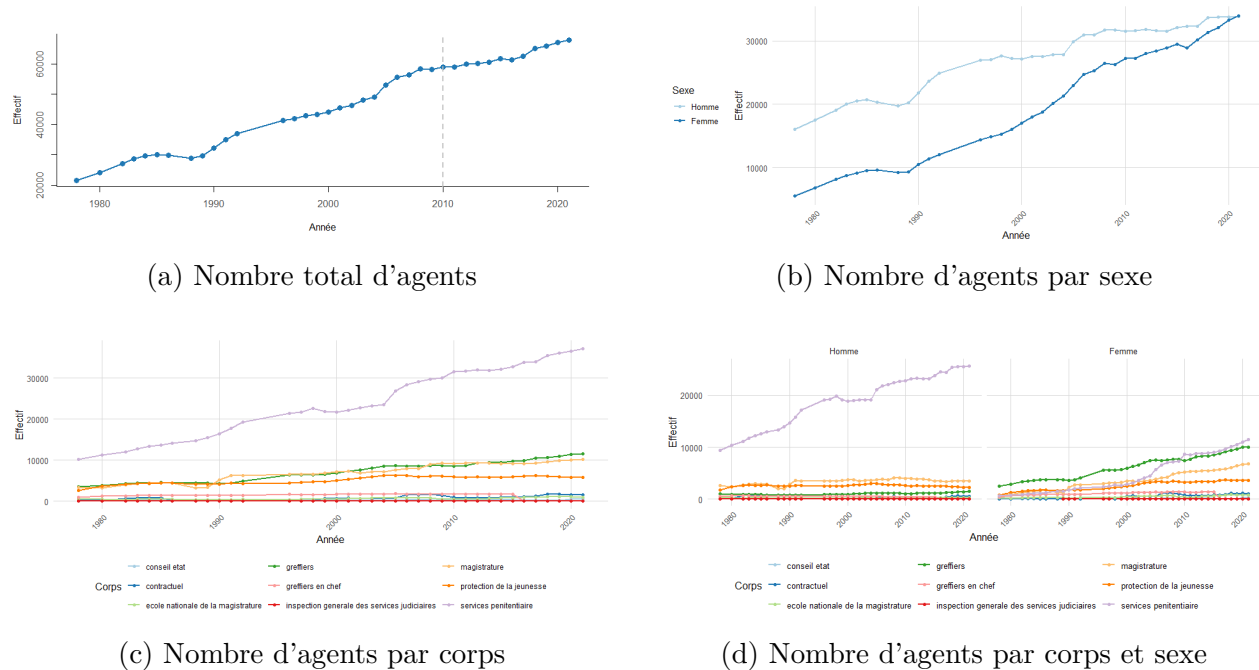


FIGURE 1 – Nombre d'agents du Ministère de la Justice

Par ailleurs, la répartition des âges des agents du Ministère de la Justice sur toute la période suit approximativement une loi normale, bornée par l'âge de fin d'études et l'âge de départ en retraite (Figure 2). Quand on différencie cette distribution par sexe on observe cependant que les femmes du Ministère de la Justice sont généralement plus jeunes que les hommes : il s'agit d'un effet de structure visible sur les stocks. On observe également que les écarts d'effectifs se réduisent voire s'annulent aux âges extrêmes.

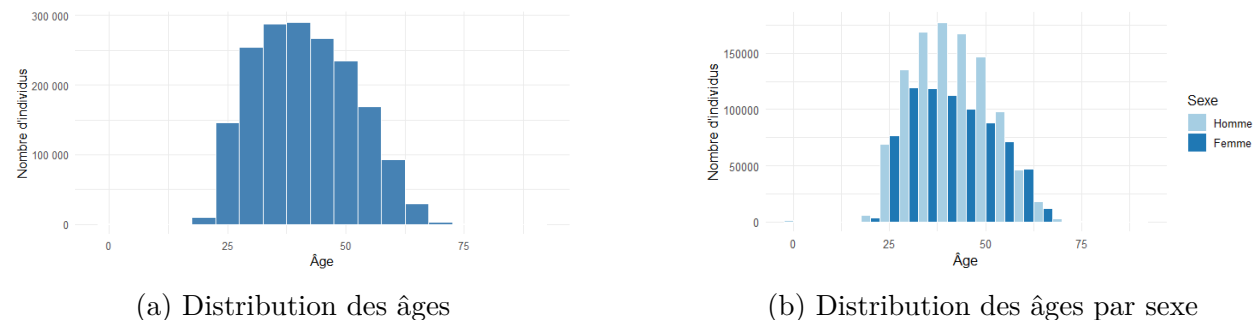


FIGURE 2 – Distribution des âges des agents du Ministère de la Justice

Enfin quand on s'intéresse aux effectifs par statut et catégorie (Figure 3) on obtient des résultats cohérents. En effet on observe un nombre d'agents titulaires bien supérieur au

nombre d’agents non titulaires. De plus on observe une évolution des catégories A, B et C similaire, bien que la catégorie C reste la plus importante. Cela est cohérent puisque c’est la catégorie la plus importante de la fonction publique d’État. Il est également cohérent de voir la catégorie D disparaître puisque les corps la composant ont été mis en extinction depuis le début des années 1990. On observe effectivement à la fin de la période le reclassement évoqué dans la partie 2.1.3.

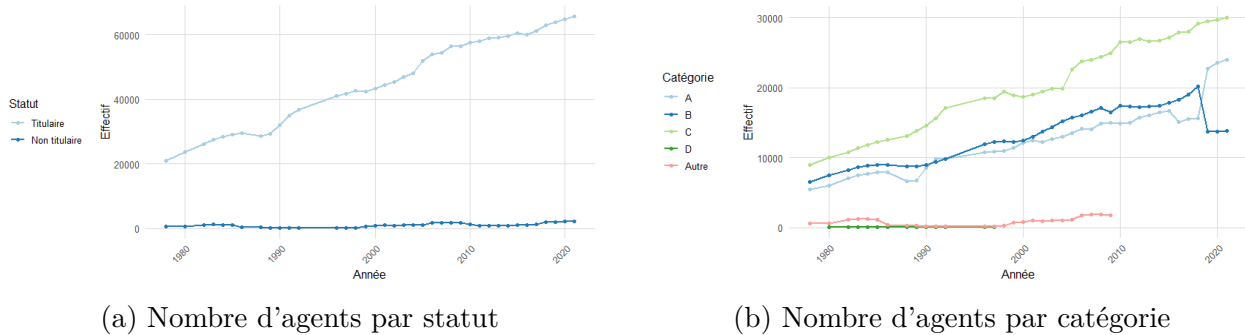


FIGURE 3 – Nombre d’agents du Ministère de la Justice par statut et catégorie

3 Modélisation statistique des trajectoires professionnelles des fonctionnaires du Ministère de la Justice

Cette section présente les outils statistiques mobilisés pour analyser les trajectoires salariales des agents du ministère de la Justice, avec une attention particulière portée aux inégalités de genre. Nous combinons des approches en coupe (cross-section) et longitudinales afin de cerner à la fois les écarts de rémunération à un instant donné, et leur évolution dans le temps.

Dans un premier temps, nous analysons les écarts de rémunération à un instant donné grâce à une décomposition de la variance et à une décomposition d’Oaxaca-Blinder. Dans un second temps, nous adoptons une perspective longitudinale, en identifiant des profils de carrière types grâce à des méthodes de classification, puis en modélisant les temps d’accès à la promotion à l’aide de modèles de durée.

3.1 Décomposition de la variance des salaires : la rémunération dans la fonction publique est-elle aussi transparente qu’annoncée ?

La fonction publique est souvent présentée comme un espace où la transparence salariale serait assurée par la codification stricte des carrières : les mêmes grilles indiciaires s’appliquent à tous, l’avancement est encadré par des règles d’ancienneté, et les rémunérations relativement prévisibles (Bessière et Pouget (2007)). Pourtant, au-delà de ce principe d’égalité théorique, il est pertinent d’interroger dans quelle mesure les déterminants à première vue purement

administratifs parviennent effectivement à expliquer les écarts de rémunération observés entre agents.

Dans cette perspective, nous proposons une première analyse exploratoire de la variance des salaires au sein du ministère de la Justice. L’objectif est double : mesurer quelle part de la variance du salaire net est effectivement expliquée par des facteurs administratifs et objectifs (corps d’appartenance, âge, statut, etc.), et comparer le pouvoir explicatif de deux regroupements hiérarchiques accessibles dans notre base de données : les catégories hiérarchique A, B, C d’une part⁴, et les corps détaillés d’autre part⁵.

Empiriquement, nous estimons plusieurs modèles linéaires (MCO) où la variable dépendante est le salaire net corrigé de l’inflation pour être exprimé en euros constants 2022, et où les variables explicatives sont l’âge, la catégorie hiérarchique ou le corps, le statut (titulaire/contractuel), le temps plein ou partiel⁶, la région de fonction, et le sexe⁷. Nous ne mettons qu’une variable de positionnement hiérarchique (catégorie hiérarchique ou corps) : une discussion de la multicollinéarité est proposée dans les *robustness checks* (Section B.1). Avec Y_i le salaire net de l’agent i , nous considérons les deux équations :

$$Y_i = \beta_0 + \beta_1 \text{AGE}_i + \beta_2 \text{CATEGORIE}_i + \beta_3 \text{TEMPSPLEIN}_i + \beta_4 \text{STATUT}_i + \beta_5 \text{SEXE}_i + \beta_6 \text{REGION}_i + \varepsilon_i \quad (1)$$

$$Y_i = \gamma_0 + \gamma_1 \text{AGE}_i + \gamma_2 \text{CORPS}_i + \gamma_3 \text{TEMPSPLEIN}_i + \gamma_4 \text{STATUT}_i + \gamma_5 \text{SEXE}_i + \gamma_6 \text{REGION}_i + u_i \quad (2)$$

Pour chaque modèle, nous calculons le coefficient de détermination R^2 afin de mesurer la part de variance expliquée, ainsi que l’importance relative de chaque régresseur selon la méthode de Lindeman *et al.* (1980), ou méthode LMG, à l’aide du package R *relaimpo*. Cette approche consiste à attribuer à chaque variable une contribution moyenne au R^2 , en tenant compte de l’ensemble des ordres possibles d’insertion des variables dans le modèle.

Cette méthode⁸ garantit que l’importance attribuée à chaque variable ne dépend pas de l’ordre arbitraire dans lequel les variables sont introduites dans le modèle, et traite équitablement les cas de multicollinéarité entre variables explicatives.

Enfin, afin de tenir compte d’éventuelles évolutions structurelles dans la fonction publique, cette analyse est menée séparément pour chaque année entre 1978 et 2022. Cela permet de documenter l’évolution de la transparence effective des rémunérations au fil du temps, dans un ministère marqué par des réformes successives et une forte féminisation des effectifs.

4. Cette variable est d’emblée présente dans FGE et SIASP, néanmoins cela n’exclut pas les ruptures de champ : la codification des contractuels diffère entre les deux bases, et la catégorie D a graduellement été menée à l’extinction.

5. Cette variable est un artefact statistique que nous avons réalisé. Il s’agit d’une sélection de corps relativement stables au cours de la période (1978-2021) et représentatifs du ministère de la Justice.

6. Notons néanmoins que notre nettoyage des deux bases a nécessité le retrait des temps de travail très courts, afin de ne conserver que les agents pour lesquels l’activité recensée n’est pas secondaire. Par conséquent, les temps partiels sont sous-représentés.

7. Cette dernière variable n’a pas d’effet *de jure*, mais nous l’incluons pour son apport résiduel.

8. Détaillée en Annexes A.1

L'analyse de la variance permet d'estimer la part des écarts de salaire attribuable à des facteurs observables, sans pour autant identifier la contribution spécifique de chaque variable ou groupe d'individus. Pour approfondir cette analyse, nous avons recours à la méthode de Oaxaca-Blinder, classiquement utilisée pour quantifier les écarts de salaire entre groupes à partir de régressions linéaires.

3.2 Des injustices au ministère de la Justice entre hommes et femmes ? Décomposition d'Oaxaca-Blinder

Au regard des faits stylisés de la littérature (Bessière et Pouget, 2007; Gollac, 2005), nous nous attendons avec la décomposition précédente à une forte explicabilité des différences de salaire en ne mobilisant que des variables relatives à l'emploi occupé et à l'âge. Pour autant, la persistance d'inégalités de genre dans la fonction publique (Cartier et Join-Lambert, 2011), notamment dans les promotions aux grades supérieurs ou les affectations prestigieuses, justifie que nous nous intéressions spécifiquement à l'explicabilité du *gender gap*. Or le ministère de la Justice s'est amplement féminisé, et le *gender gap* s'est réduit avec les années (Figure 4).

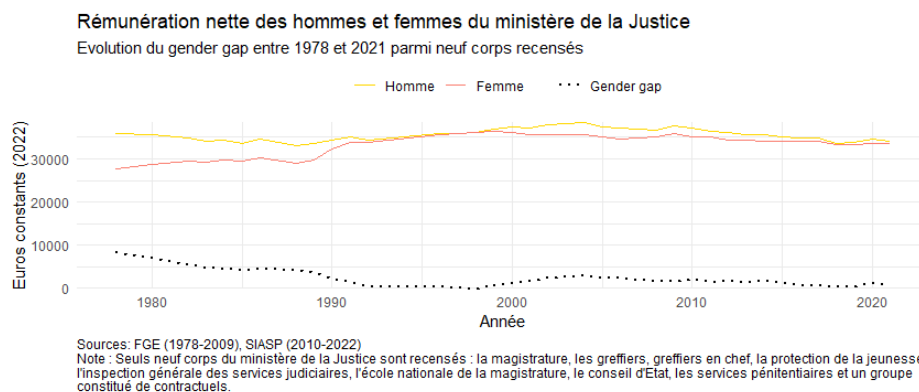


FIGURE 4 – Dynamique d'atténuation du *gender gap* au ministère de la Justice

Nous proposons ici d'évaluer si, à caractéristiques observables identiques, des écarts de rémunération subsistent entre hommes et femmes au ministère de la Justice. Pour ce faire, nous mobilisons une décomposition d'Oaxaca-Blinder (Blinder (1973); Oaxaca (1973))⁹, qui permet de séparer l'écart salarial moyen observé entre deux groupes (ici, hommes et femmes) en deux composantes :

- une part expliquée par des différences dans les caractéristiques observables (âge, corps, statut, région de fonction, temps de travail) ;
- une part inexpliquée, interprétée comme le résidu attribuable à des facteurs non observés, éventuellement de la discrimination.¹⁰

La décomposition est réalisée année par année sur la période 1978-2022 à l'aide du package R *oaxaca* (Hlavac, 2022), en utilisant les mêmes variables que dans l'analyse de variance précédente. Les erreurs standard sont estimées par bootstrap sur la bases constituées des

9. Il s'agit d'une décomposition *two-fold* d'Oaxaca-Blinder. Cette décomposition n'isole pas de terme d'interaction, contrairement à la décomposition *three-fold* (Hlavac (2022)).

10. Une formalisation est proposée en Annexe A.2

observations d’une année : cela a nécessité d’exclure les modalités pour lesquelles il n’y avait qu’un nombre très restreint d’observations¹¹. Ce design empirique permet de suivre l’évolution du *gender gap* dans le temps, et d’identifier si les écarts inexplicables tendent à se réduire, persister, ou s’inverser, au fur et à mesure des réformes et des évolutions du ministère de la Justice.

La décomposition d’Oaxaca-Blinder offre une lecture instantanée des écarts salariaux entre femmes et hommes. Toutefois, elle ne permet pas de rendre compte de la dynamique des parcours professionnels, ni des inégalités qui se construisent ou s’accumulent dans le temps. Pour explorer les dynamiques de carrière sur le temps long, nous avons recours à une méthode de classification des trajectoires, permettant d’identifier des profils types d’évolution salariale.

3.3 Les modèles de clustering : construction d’une typologie des trajectoires

Dans l’analyse biographique, qui s’intéresse aux parcours de vie des individus sur la durée, l’un des enjeux majeurs consiste à structurer la grande diversité des trajectoires observées. La typologie des trajectoires repose sur une démarche de clustering, c’est-à-dire de classification non supervisée visant à regrouper les trajectoires similaires en un nombre restreint de classes. Cette méthode permet de dégager des « trajectoires types » qui facilitent l’analyse, l’interprétation et, à terme, la mise en évidence de facteurs explicatifs ou de conséquences sociales associées à ces types. Ces méthodes sont présentées et expliquées par Robette dans Robette (2021) sur lequel nous nous sommes grandement appuyé.

3.3.1 De la séquence au regroupement : les étapes de la méthode

La première étape de cette démarche consiste à disposer de données biographiques codées sous forme de séquences d’états observés sur une période déterminée de la vie d’un individu. Par exemple, entre 14 et 50 ans, chaque année étant associée à un état professionnel comme « en études », « en emploi stable », « au chômage », etc. Pour pouvoir comparer ces trajectoires, il est crucial que la période d’observation soit commune à tous les individus et que le codage des états soit cohérent et restreint : trop d’états rendraient la typologie illisible, tandis qu’un codage trop grossier occulterait les nuances essentielles.

Une fois ces séquences établies, l’enjeu est de mesurer la dissimilarité entre chaque paire de trajectoires. La méthode la plus courante en sciences sociales, et celle que l’on choisira, est celle dite de l’Optimal Matching (OM). Elle repose sur une distance de type Levenshtein, qui quantifie le coût minimal pour transformer une séquence en une autre à l’aide d’opérations élémentaires : insertion, suppression, et substitution. Le choix des coûts associés à ces opérations est une étape cruciale de la démarche. Il peut être effectué de manière arbitraire, ou bien être fondé sur des considérations empiriques (par exemple, sur les probabilités observées de transition entre états). Les insertions/suppressions permettent de détecter des

11. Ce filtre n’a qu’un effet négligeable sur nos résultats : voir Annexe B.2 pour une discussion.

motifs communs même lorsqu'ils sont décalés dans le temps, tandis que les substitutions comparent strictement les états au même moment. On choisira un coût d'insertion / suppression (indel) fixé à 1. Pour les coûts de substitution on choisit une matrice de substitution probabiliste grâce à la méthode TRATE¹², ce qui permet d'ancrer les distances dans la structure empirique des trajectoires observées.

À partir de cette matrice de dissimilarité, on applique une méthode de classification, dont il existe deux grandes familles : les classifications hiérarchiques et les classifications par partitionnement. Dans une classification hiérarchique ascendante (comme celle fondée sur le critère de Ward), on commence par considérer chaque trajectoire comme une classe à part entière, puis on regroupe successivement les trajectoires ou groupes les plus proches afin de minimiser l'hétérogénéité interne. Cela produit un arbre de classification (dendrogramme), dans lequel chaque niveau correspond à une partition différente. À l'inverse, les méthodes de partitionnement, comme les k-means ou k-medoids, nécessitent de fixer a priori le nombre de groupes et procèdent à des itérations successives pour stabiliser les centres de gravité autour desquels sont agrégées les trajectoires. On choisira une méthode de classification hiérarchique ascendante avec critère de Ward, puisque cette classification vise à former des groupes aussi homogènes que possible (minimisant l'inertie intra-classe).

Le choix du nombre de classes est une étape importante. Trop de classes rendraient la typologie illisible et peu interprétable, tandis qu'un nombre trop restreint masquerait l'hétérogénéité réelle des trajectoires.

3.3.2 Visualisation de résultats

Une fois la typologie construite, on peut décrire chaque groupe en termes statistiques : durée moyenne passée dans chaque état, nombre moyen de transitions, degré de complexité (via l'entropie ou la turbulence), etc. On peut également identifier, pour chaque classe, un médoïde – c'est-à-dire l'individu dont la trajectoire est la plus représentative du groupe – et visualiser l'ensemble des trajectoires à l'aide de tapis de séquences ou de chronogrammes. Les chronogrammes représentent à chaque instant d'observation la proportion d'individus dans chaque état. Autrement dit le graphique présente à chaque instant les proportions cumulées dans chacune des situations, tandis que les tapis conservent la dimension longitudinale des trajectoires individuelles, au risque d'être difficilement lisible si le nombre d'individus représenté est élevé. Chaque individu est en effet une ligne horizontale où un changement de couleur indique un changement d'état.

Le clustering permet de repérer des formes récurrentes de carrières, mais ne renseigne pas directement sur la temporalité des événements, en particulier l'accès à la promotion. Pour cette raison, nous adoptons une approche longitudinale fondée sur des modèles de durée.

12. La méthode TRATE propose de calculer les coûts de substitution à partir des transitions observées dans les données. Autrement dit, deux états sont "proches" si on observe souvent des transitions entre eux dans les séquences. Elle reflète la dynamique réelle des trajectoires observées et est moins arbitraire qu'une matrice uniforme.

3.4 Modéliser les temps d'accès à la promotion : approche par les modèles de durée

L'étude des trajectoires professionnelles se prête bien à une modélisation statistique de survie. En effet, les modèles de survie permettent une approche pertinente des carrières par la caractérisation de la durée d'un état professionnel, l'estimation des risques de transition professionnelle, la prise en compte de données censurées (lorsqu'un individu ne connaît pas de changement de statut ou un événement lors de sa présence dans la base de données) et l'analyse de la pertinence de certains déterminants socio-démographiques dans les carrières.

A la différence des modèles de durée, les modèles de survie permettent la compréhension de la durée de survie qui représente la durée avant l'arrivée d'un événement spécifique (Quantin, 2018).

Les modèles de durée s'articulent autour de deux concepts fondamentaux : la fonction de survie et la fonction de risque (ou hazard). La fonction de survie représente la probabilité qu'un individu n'ait pas encore connu l'événement d'intérêt (promotion, changement de grade, etc.) à un instant donné. La fonction de risque, quant à elle, décrit le taux instantané de survenue de l'événement, en tenant compte de l'ancienneté et des caractéristiques individuelles.

Parmi les méthodes les plus couramment utilisées, on retrouve :

- L'approche non paramétrique avec les estimateurs de Kaplan-Meier et de Nelson-Aalen, qui permettent de décrire empiriquement la distribution des temps d'accès à une promotion. Ces méthodes, basées sur des fonctions en escalier, offrent une représentation claire des transitions observées.
- Les modèles semi-paramétriques de Cox, qui permettent d'analyser l'impact de variables explicatives (âge, grade, catégorie, etc.) sur le risque de promotion, tout en restant flexibles sur la forme de la fonction de risque de base.
- Les extensions du modèle de Cox, telles que le modèle stratifié et le modèle à effet aléatoire (frailty), permettant de capturer des hétérogénéités structurelles ou non observées dans les dynamiques de carrière.

Les détails techniques de l'estimation, ainsi que les démonstrations mathématiques associées, sont présentés en (Annexe C). Cette modélisation permet d'isoler les déterminants temporels et socio-professionnels influençant l'accès aux promotions, offrant ainsi un cadre robuste pour l'analyse des inégalités de carrière dans la fonction publique.

Les modèles de durée permettent de modéliser la probabilité d'un événement (comme une promotion) en fonction du temps. Ils fournissent une lecture analytique centrée sur un moment-clé de la carrière. En complément, nous mobilisons une approche plus exploratoire et descriptive : la classification des trajectoires, qui vise à regrouper les agents selon des profils de carrière similaires, sans hypothèse a priori sur leur évolution.

4 Les résultats : écarts de rémunération et trajectoires professionnelles

4.1 Une rémunération largement encadrée par les statuts administratifs, mais avec des marges résiduelles

L'analyse empirique confirme dans l'ensemble que la rémunération des agents du ministère de la Justice est fortement déterminée par des facteurs administratifs observables. Comme le montre la figure 5, le pouvoir explicatif global des modèles estimés reste élevé tout au long de la période 1978-2022. Lorsqu'on utilise la variable fine des corps (regroupant neuf grands corps de fonctionnaires représentatifs du ministère), les modèles expliquent chaque année entre 70% et 80% de la variance du salaire net. À titre de comparaison, en mobilisant uniquement la catégorie hiérarchique (A, B ou C), la part de variance expliquée reste significative mais inférieure, oscillant autour de 60%. Ces résultats confortent l'idée que, bien que les catégories administratives jouent un rôle important, la précision du classement par corps permet d'expliquer plus fidèlement les écarts de salaire.¹³

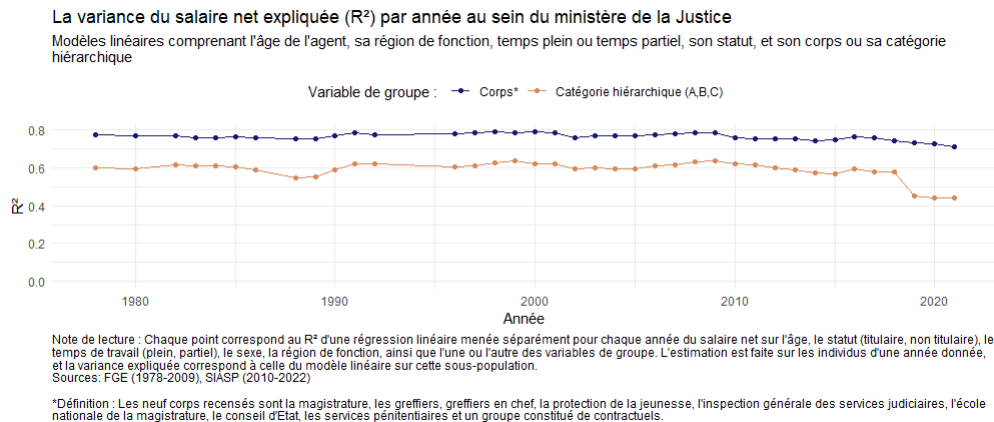


FIGURE 5 – Les disparités de salaires du ministère de la Justice sont en grande majorité explicables

La décomposition fine des contributions au R^2 (Figure 6) permet de préciser l'importance relative de chaque variable explicative. Dans les deux spécifications, la variable de groupe principale (catégorie ou corps) constitue de très loin le premier déterminant du salaire net. Lorsqu'on utilise la catégorie hiérarchique, celle-ci explique en moyenne 55% à 60% de la variance totale ; lorsqu'on utilise les corps, leur contribution grimpe jusqu'à 70% voire 75%.

L'âge de l'agent arrive en deuxième position : il explique entre 15% et 20% de la variance, reflétant l'effet de l'ancienneté sur les rémunérations. Les autres variables (temps plein ou partiel, statut titulaire/contractuel, région d'affectation, sexe) ont des contributions beaucoup plus modestes : chacune pèse moins de 5% individuellement. Le sexe, en particulier, ne

13. À ce titre, un niveau de précision plus important, par exemple via le grade ou l'échelon, permettraient vraisemblablement d'augmenter davantage le R^2 .

contribue que marginalement à la variance expliquée (moins de 1%), ce qui est cohérent avec l'absence d'effet direct prévu par les grilles salariales.

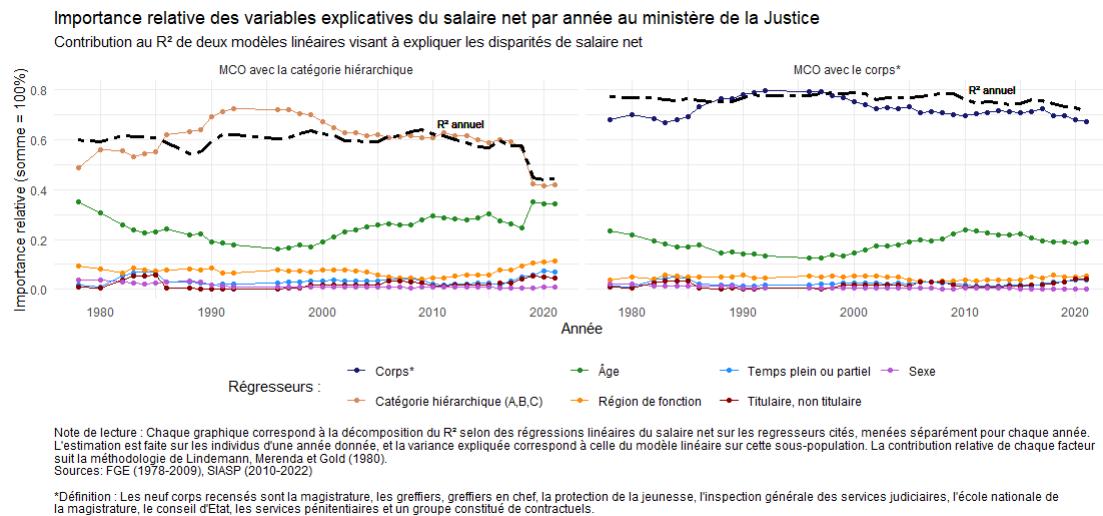


FIGURE 6 – Des salaires largement expliqués par le positionnement hiérarchique et l'ancienneté

Enfin, il est intéressant de noter la relative stabilité de ces résultats dans le temps. Malgré les réformes successives de la fonction publique, la part de variance expliquée par les corps et les catégories reste remarquablement constante, avec seulement un léger infléchissement dans les années récentes, vraisemblablement expliqué par un reclassement d'un corps de B à A. Cela suggère que, même face à des évolutions institutionnelles (fusions de corps, réforme des grilles, féminisation massive de certains métiers), la structure administrative demeure le principal déterminant des écarts salariaux au sein du ministère.

Si la rémunération est en grande partie déterminée par des variables administratives, cela ne signifie pas que le traitement salarial est identique entre femmes et hommes. Pour affiner cette lecture, nous décomposons l'écart de salaire selon le genre à l'aide de la méthode d'Oaxaca-Blinder, permettant de distinguer la part expliquée par les différences de caractéristiques de celle potentiellement liée à des discriminations indirectes.

4.2 Oaxaca-Blinder : dynamique d'explication de l'écart salarial entre hommes et femmes

La décomposition d'Oaxaca-Blinder réalisée année par année permet d'apporter plusieurs enseignements sur l'évolution du *gender gap* salarial au sein du ministère de la Justice entre 1978 et 2021.

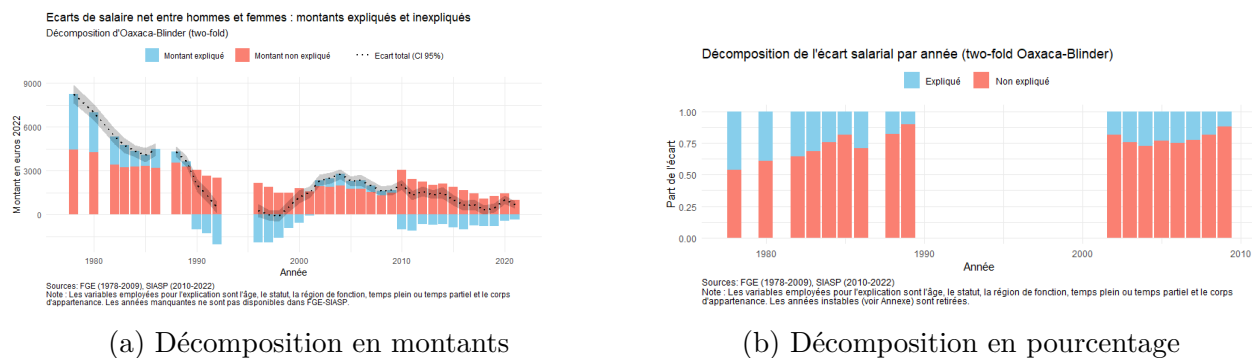


FIGURE 7 – Décomposition d'Oaxaca-Blinder

Tout d'abord, entre 1978 et la fin des années 1980, l'écart salarial entre les hommes et les femmes était positif et important au sein du ministère de la Justice (quelques milliers d'euros 2022 par an, cf. Figure 7a). Ainsi que l'illustre la figure 7b, environ 30 à 50% de cet écart étant attribuable à des différences de caractéristiques observables (âge, corps, région, etc.).

Ensuite, à partir du début des années 1990, la dynamique évolue sensiblement. L'écart salarial entre hommes et femmes se réduit très fortement, jusqu'à parfois s'inverser ponctuellement : certaines années, les femmes perçoivent en moyenne un salaire net légèrement supérieur à celui des hommes parmi les neuf corps étudiés. Dans ce contexte de faible différentiel de rémunération, la part inexpliquée de l'écart (c'est-à-dire la part non imputable aux caractéristiques observées) devient prédominante. Cela signifie que, même si les différences brutes de salaires se sont atténuées, celles qui subsistent échappent en grande partie à l'explication par les facteurs administratifs pris en compte dans l'analyse. Notons néanmoins que cette inversion rend la décomposition d'Oaxaca-Blinder moins stable, et pour plusieurs années (par exemple 1991, 1992, 1996), les résultats de la décomposition sont difficiles à interpréter (voir Section B.3, Robustness checks), avec des parts expliquées et inexpliquées très élevées ou négatives (cf. Figure 7a).

Dans les années 2000 et 2010, l'écart global reste faible, mais la décomposition semble plus stable, ainsi que l'illustre la Figure 7b. La composante inexpliquée représente alors la grande majorité de l'écart total, ce qui indique que les différences de caractéristiques observables expliquent peu des écarts résiduels de rémunération entre hommes et femmes. Ces années doivent néanmoins être interprétées avec prudence : elles traduisent vraisemblablement moins une montée brutale de l'inégalité de traitement que la difficulté technique d'appliquer la méthode de Oaxaca-Blinder dans un contexte de quasi-égalité salariale.

Dans l'ensemble, ces résultats suggèrent que si les différences objectives de carrières expliquaient initialement une part importante du *gender gap* salarial, la réduction de ces différences a mis au jour des écarts plus résiduels et inexpliqués, appelant à des analyses plus fines, notamment sur les trajectoires de promotion et les affectations. En définitive, ces deux sections montrent que les disparités salariales observées entre hommes et femmes s'expliquent d'abord par des différences de position hiérarchique et d'ancienneté, tandis que le sexe en tant que tel joue un rôle limité ; au fil du temps, le *gender gap* s'est progressivement resserré. Les écarts résiduels, majoritairement inexpliqués par des caractéristiques statiques, justifient de recourir à une approche dynamique : cette étude en *cross sections* est donc complétée par une étude des évolutions dynamiques dans la section suivante.

4.3 Clustering : les trajectoires de salaires sur 25 ans de carrière

On cherche à établir des trajectoires types de salaire au sein de la fonction publique du ministère de la Justice. Pour cela on met en place la méthodologie de Robette présentée dans Robette (2021) et dans la partie 3.3. L'idée est d'abord de définir des états dans lesquels les individus se trouvent au cours de leur carrière. On choisit de définir ces états comme des tranches de salaire net correspondant aux quintiles arrondis à la centaine. Autrement dit les états sont les suivants :

État	Borne sup	Borne inf
Tranche 1	0	26 300
Tranche 2	26 300	30 900
Tranche 3	30 900	35 600
Tranche 4	35 600	52 600
Tranche 5	52 600	Au delà

TABLE 1 – Tranches de salaire net

Cela nous assure une répartition des individus a priori relativement égale entre les états. On choisit également de se concentrer sur les individus présents dans la fonction publique pendant au moins 25 ans. Cela nous assure un nombre suffisant d'individus (19 894 NIR uniques) et une période suffisamment longue pour observer des trajectoires de salaire. Notre variable temporelle est ainsi non pas les années, mais le nombre d'années d'ancienneté dans la fonction publique. Après plusieurs essais on choisit de se concentrer sur 4 clusters. Plus de clusters aboutissait à des résultats redondants avec des catégories peu différentes les unes des autres tandis que moins ne donnait pas des résultats assez fins. En mettant en plus une méthode d'Optimal Matching avec une classification ascendante, on obtient les résultats qui suivent. La figure 8 présente ligne à ligne les trajectoires individuelles par cluster tandis que la figure 9 présente les proportions d'individus dans chaque état par cluster. Les tableaux 2 et 3 donnent les répartitions par sexe, catégories et corps de chaque cluster.

On observe que le cluster 1 est le cluster des trajectoires bloquées : il réunit les trajectoires les moins rémunérées qui stagnent dans les tranches 2 et 3. Il est fortement représenté par des agents de catégorie B et C, et par les greffiers et les agents des services pénitentiaires. Le cluster 2 est celui des carrières intermédiaires, avec une promotion graduelle et une arrivée tardive dans la tranche 4. Il est majoritairement représenté par des agents masculins, de catégorie C, travaillant également au sein des services pénitentiaires. Le cluster 3 est le cluster des trajectoires élitaires à progression rapide. C'est un groupe paritaire atteignant très rapidement les tranches hautes de la rémunération. Les magistrats y sont très concentrés ainsi que le personnel de l'École nationale de la magistrature. Enfin le cluster 4 est le cluster des carrières plafonnées. Leur progression est plutôt rapide, ils passent moins de temps dans les tranches 2 et 3 que les clusters 1 et 2 mais n'atteignent jamais la dernière strate.

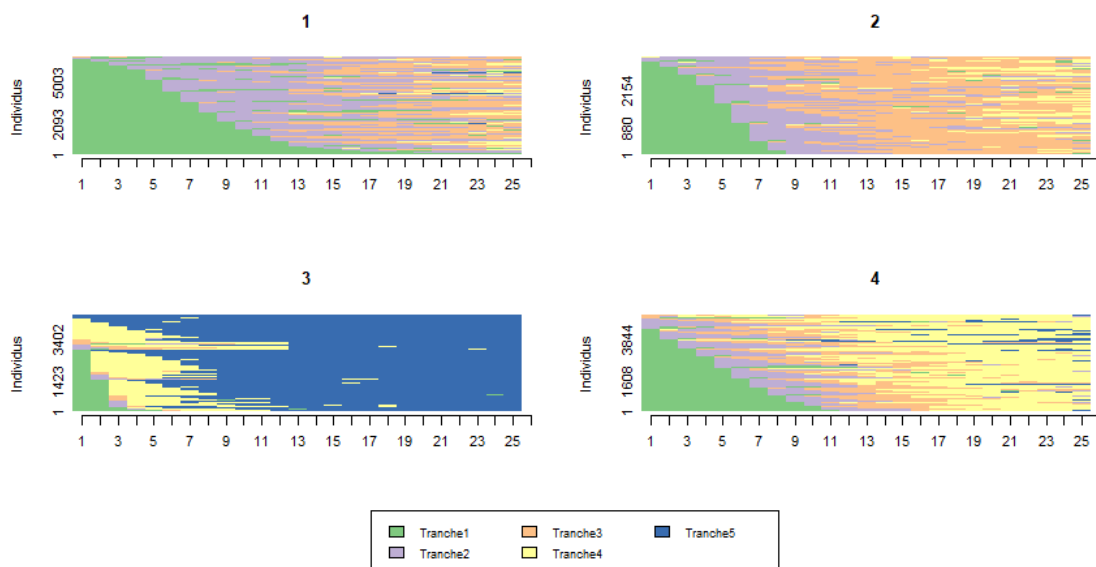


FIGURE 8 – Les trajectoires individuelles par cluster

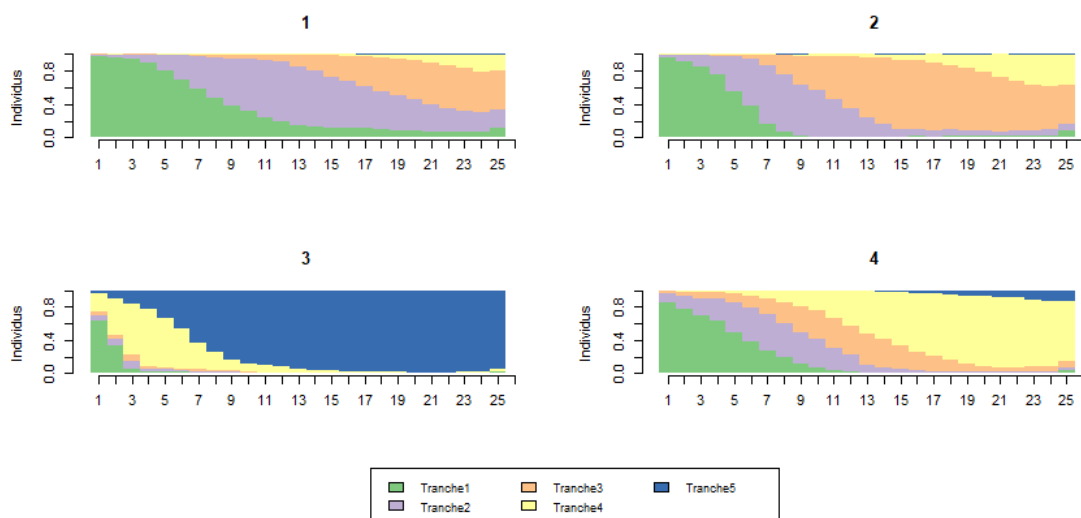


FIGURE 9 – La proportion d'états au fil du temps par cluster

Ainsi de manière logique, quand on regarde les tableaux 2 et 3, on observe que les catégories A sont très largement majoritaires dans le groupe 3, tandis qu'elles sont quasiment absentes des groupes 1 et 2. A l'inverse les catégories C sont majoritaires dans les clusters 1 et 2.

Similairement, la surreprésentation de la magistrature et de l'École nationale de la magistrature dans le cluster 3 traduit une certaine stabilité dans la rémunération tout au long de la carrière. Les corps des greffiers, des greffiers en chef et des services de protection de la jeunesse sont plus équitablement répartis dans les autres clusters. Les services pénitentiaires sont bien plus présents dans le cluster 2, traduisant une plus faible progression de salaire

Cluster	Effectifs	Sexe		Catégorie		
		Homme	Femme	A	B	C
1	6787	53.8	46.2	0.4	47.9	51.6
2	3278	78.2	21.8	0.6	30.6	68.7
3	4615	48.6	51.4	95.8	3.0	0.8
4	5214	65.1	34.9	10.5	43.3	46.0

TABLE 2 – Effectifs par cluster, par sexe et catégorie (en %)

des corps de ces services. La répartition par sexe est plus lisse, à l’exception du cluster 2, dominé par les hommes, ce qui peut s’expliquer par l’importante représentation d’hommes dans les services pénitentiaires. Les femmes sont majoritaires uniquement dans le cluster 3, probablement du fait de la forte féminisation de la magistrature. Ces observations rejoignent celles de la partie 3.2.

Cluster	ENM	Gref.	Gref. en chef	PJ	SP	Magistrature
1	0.1	35.0	0.3	11.6	53.0	NA
2	0.2	15.6	0.1	10.9	73.1	0.0
3	59.9	1.8	3.4	0.8	5.0	28.1
4	0.6	21.2	6.8	16.9	54.2	0.1

Note : les contractuels et le Conseil d’État ne sont pas représentés.
 ENM : Ecole nationale de la magistrature, Gref. : Greffiers, PJ : Protection de la jeunesse, SP : Services pénitentiaires

TABLE 3 – Effectifs par cluster et par corps (en %)

Une autre manière d’approcher ces trajectoires types est par le médoïde, autrement dit par l’individu central à chaque cluster, celui qui minimise la distance avec les autres individus de son groupe. C’est ce que présente la figure 10. On voit bien que le cluster 3 se distingue nettement des trois autres par la forte présence de la tranche 5 dans les trajectoires des individus. Les autres trajectoires se distinguent les unes des autres surtout par le temps passé dans chaque tranche de salaire. En moyenne, le cluster 1 n’accède que tardivement à la tranche 3, le cluster 2 n’atteint que très rarement la tranche 4, et le cluster 4 n’atteint jamais la tranche 5. Ainsi chacun de ces clusters passe la majorité de sa carrière dans la tranche de salaire juste inférieure.

La typologie des trajectoires met en évidence des profils différenciés selon le corps, la catégorie ou encore le sexe. Ces profils ne renseignent toutefois pas sur la dynamique précise d’accès aux promotions. Nous complétons donc cette approche par des modèles de durée, afin d’évaluer la probabilité d’atteindre un grade supérieur en fonction de l’ancienneté et d’autres caractéristiques.

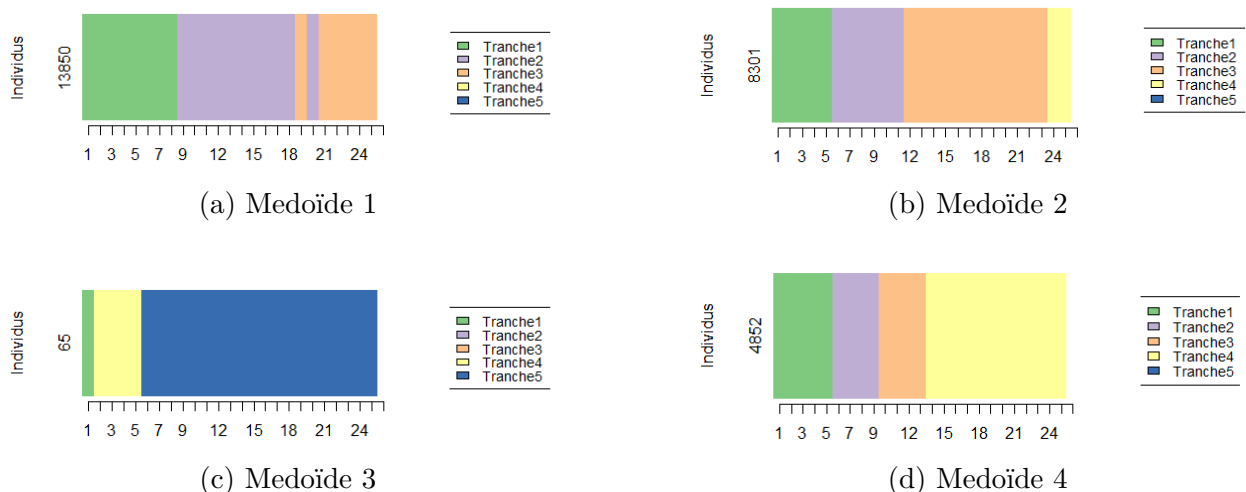


FIGURE 10 – La trajectoire typique (médoïde) par cluster

4.4 Les modèles de durée

Dans cette section, l'analyse porte uniquement sur les magistrats, afin d'étudier plus spécifiquement les dynamiques de promotion au sein de ce corps. Pour cela, une variable **event** a été construite pour capturer les promotions d'un grade à un autre. L'année 2010 a été exclue de la base de données en raison du changement de nomenclature des grades à cette date, empêchant l'identification fiable des promotions par changement de grade.

Les modèles sont spécifiés en termes d'intervalles (**start**, **stop**) plutôt qu'en temps unique, car les données sont de nature longitudinale avec des entrées différées dans la base de suivi. Cette approche permet de modéliser précisément les durées avant première promotion tout en tenant compte des variations des covariables dans le temps et des différents moments d'entrée des magistrats dans le suivi.

4.4.1 Premiers modèles de Cox : limitations et tentatives de stratification

Nous commençons l'analyse par un modèle de Cox standard (Modèle 1), intégrant les covariables **AGE**, **SEXE**, **STATUT** (fonctionnaire ou contractuel) et **TEMPS.PLEIN**. Ce modèle repose sur l'hypothèse de proportionnalité des risques et comme le montre la Table 4, cette hypothèse est massivement violée pour plusieurs covariables : (**AGE**) et (**SEXE**) ont des p -valeurs largement significatives (2×10^{-16}), signifiant que l'effet de ces variables sur le risque de promotion varie dans le temps, ce qui remet en cause la validité du modèle. Dans ces conditions, les coefficients estimés n'ont plus d'interprétation simple, et il devient nécessaire d'adapter la modélisation.

Une première tentative d'ajustement consiste à introduire une stratification sur le sexe (**strata(SEXE)**) dans le Modèle 2, permettant de relâcher l'hypothèse de proportionnalité sur la variable **SEXE**, en laissant le modèle estimer une fonction de base de risque distincte pour

TABLE 4 – Test de proportionnalité des risques de Schoenfeld pour les trois modèles

Variable	Modèle 1		Modèle 2		Modèle 2bis	
	χ^2	p-valeur	χ^2	p-valeur	χ^2	p-valeur
AGE	997.31	2e-16	718.69	2e-16	—	—
tranche_age3	—	—	—	—	818.00	2e-16
SEXE	287.91	2e-16	(strate)	—	(strate)	—
STATUT	0.57	0.45	0.01	0.916	0.001	0.972
TEMPS_PLEIN	18.43	9.9e-05	11.03	0.004	10.60	0.005
GLOBAL	1100.46	2e-16	739.31	2e-16	828.00	2e-16

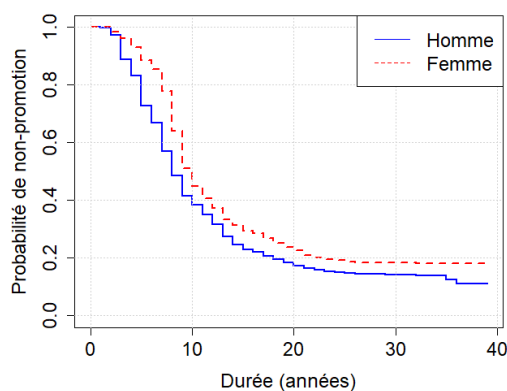
Note : Les modèles 2 et 2bis utilisent une stratification `strata(SEXE)`, ce qui empêche le test de proportionnalité sur cette variable. Le modèle 2bis remplace AGE par `tranche_age3`.

chaque modalité de sexe. Toutefois, le test de Schoenfeld¹⁴ montre que la non-proportionnalité persiste sur la variable AGE, avec une statistique toujours extrêmement significative. Le recours à la stratification sur le sexe améliore donc la situation, mais reste insuffisant.

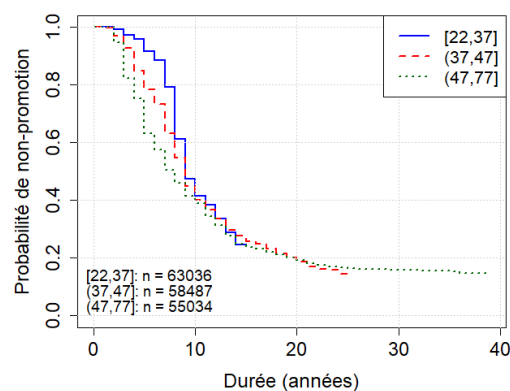
Nous construisons alors un Modèle 2bis dans lequel la variable continue AGE est remplacée par une variable catégorielle à trois tranches d'âge (`tranche_age3`), afin de mieux capturer d'éventuels effets non linéaires et de réduire la variabilité temporelle. La stratification sur le sexe est conservée. Néanmoins, l'hypothèse de proportionnalité des risques est toujours violée pour cette variable transformée (`tranche_age3`), avec un p -value toujours inférieur à 2×10^{-16} . Les résultats sont tout de même affichés en Annexe D.1 avec une proposition d'interprétation, afin de permettre la compréhension des sorties d'un modèle de Cox.

Pour mieux comprendre ces résultats, nous traçons les courbes de survie de Kaplan-Meier 11 (voir C.2) selon les tranches d'âge. Celles-ci révèlent que les courbes associées aux différentes tranches d'âge ont tendance à se croiser, ce qui suggère que l'effet de l'âge sur le risque de promotion est non seulement non proportionnel, mais également potentiellement non monotone. Cela explique en partie les difficultés d'estimation rencontrées avec les modèles de Cox classiques, et justifie le recours ultérieur à des modèles incluant des effets aléatoires, comme les modèles de Cox à frailty. Les courbes de survie illustrent la probabilité de ne pas encore avoir été promu au cours du temps selon le sexe et selon l'âge. On observe que les femmes accèdent en moyenne plus lentement aux promotions que leurs homologues masculins, ce qui peut traduire une inégalité d'accès à l'avancement ou des trajectoires de carrière différenciées. Concernant l'âge, les magistrats les plus jeunes progressent généralement plus rapidement en début de carrière, tandis que les courbes des tranches d'âge plus avancées montrent une plus grande inertie. Cependant, les modèles de survie tels que les modèles de Cox ne permettent d'appréhender que la première promotion.

14. Ce test vérifie si les résidus associés à chaque variable dépendent du temps. Une dépendance temporelle significative (p-valeur faible) indique une violation de l'hypothèse de proportionnalité des risques.



(a) selon le sexe



(b) selon les tranches d'âge

FIGURE 11 – Courbes de survie de Kaplan-Meier - magistrature

Source : FGE et SIASP — Traitement : estimation par *survfit*

4.4.2 Taux de promotion et *frailty effect* : comprendre l'effet structurel des départements de fonction sur les promotions au sein de la magistrature

Comme l'ont démontré Demoli et Willemez (2023), les trajectoires professionnelles dans ce corps d'État sont fortement influencées par des logiques spatiales qui méritent d'être explorées.

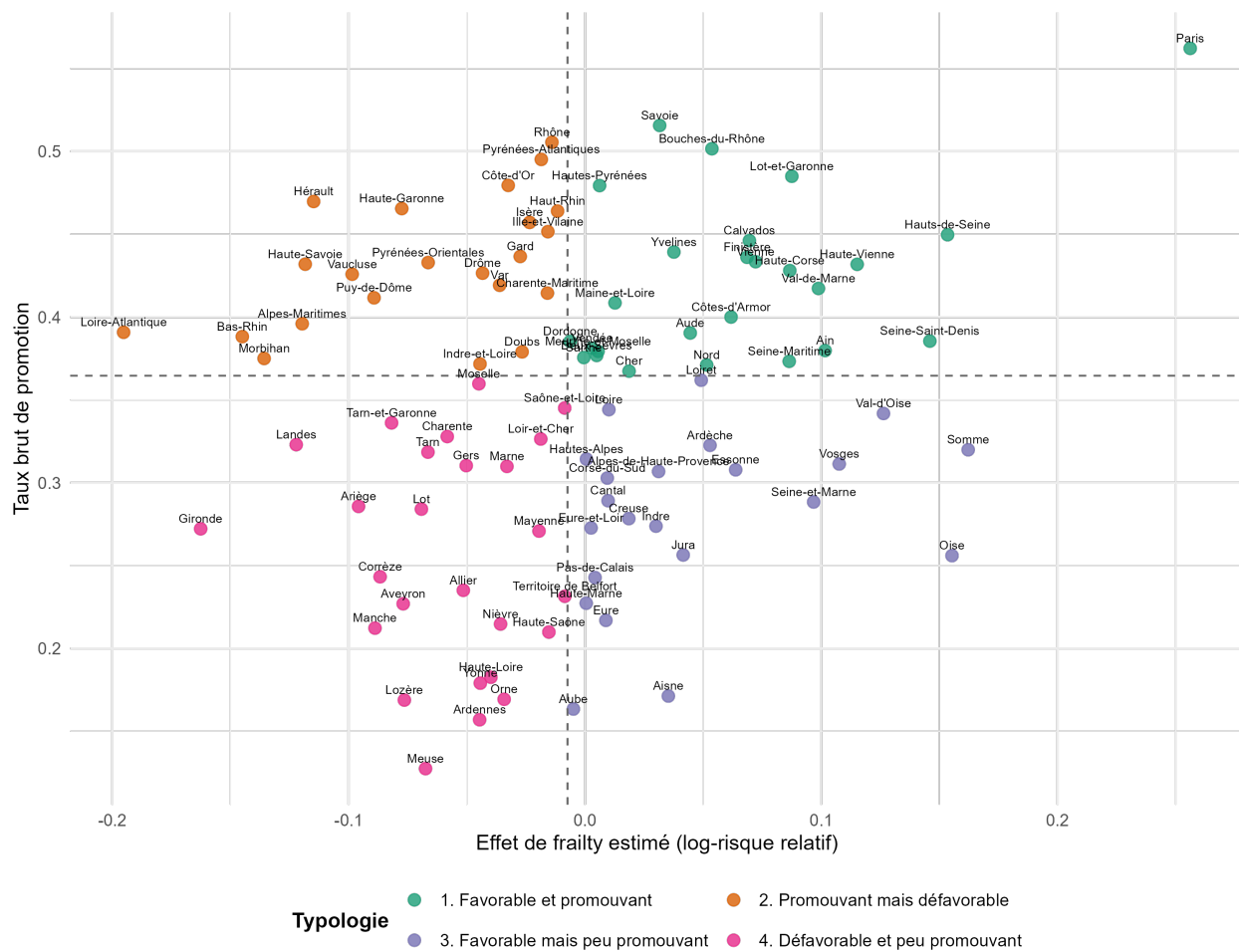
Le taux brut de promotion représente la proportion de magistrats promus dans chaque département (calculé comme le nombre de promotions par changement de grade sur le nombre total de magistrats suivis dans le département),¹⁵ sans tenir compte des caractéristiques individuelles. Il fournit une vision descriptive des disparités entre territoires, mais ne tient pas compte des spécificités locales. À l'inverse, l'effet de *frailty*, issu du modèle de Cox, mesure une spécificité territoriale ajustée, c'est-à-dire l'influence propre du département sur la probabilité de promotion, toutes choses égales par ailleurs. Un effet positif indique qu'à profil identique (SEXE, AGE, STATUT, TEMPS_PLEIN), un magistrat a une probabilité plus élevée d'être promu dans ce département par rapport à la moyenne nationale, tandis qu'un effet négatif traduit un désavantage structurel.¹⁶

L'analyse conjointe des deux indicateurs révèle des situations contrastées¹². Paris (75) se distingue par un fort taux de promotion (56 %) combiné à un effet de *frailty* positif (+0,245), soulignant un environnement particulièrement favorable, même après ajustement. En revanche, la Gironde (33) présente un effet de *frailty* négatif (-0,167) et un taux de promotion modeste (27 %), ce qui traduit des conditions locales moins propices. Le Rhône (69) illustre un cas de promouvant mais défavorable : malgré un taux de promotion élevé (51 %), son effet de *frailty* est légèrement négatif (-0,015), indiquant que les promotions dépendent davantage des profils des magistrats que du contexte départemental. À l'inverse, l'Aisne (02) affiche un effet de *frailty* positif (+0,037) mais un taux de promotion faible (17 %), suggérant que le

15. Une erreur d'encodage des départements d'Outre-mer nous ont contraint de les retirer de l'analyse.

16. Une carte des effectifs par département dans l'échantillon utilisé est disponible en annexe D.2

territoire est structurellement favorable, bien que les profils présents soient moins susceptibles d'être promus.

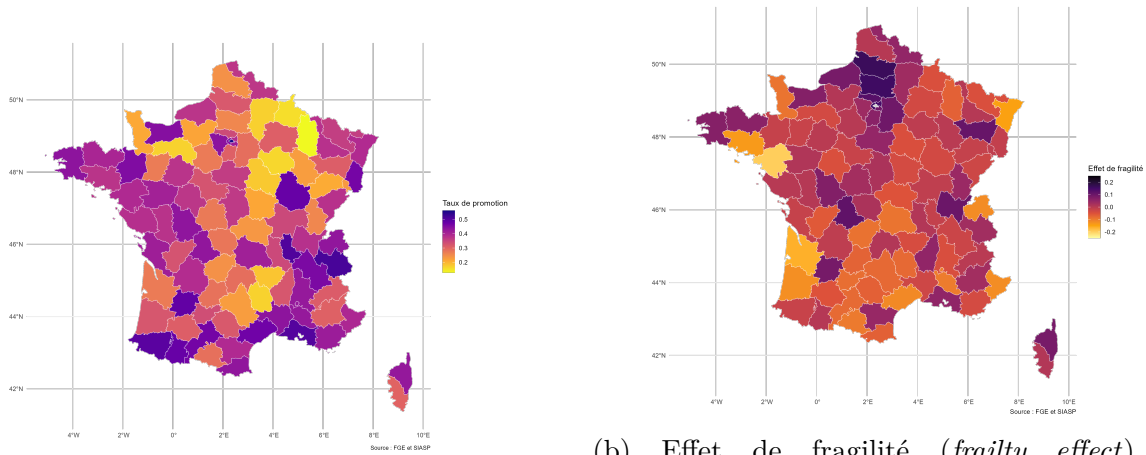


Source : FGE (1978–2009), SIASP (2010–2022)

FIGURE 12 – Typologie des départements en fonction du taux de promotion brut et de l'effet de fragilité en France.

L'analyse des cartes 13 montre que les départements colorés en jaune et orange sur la carte du taux de promotion brut présentent les pourcentages de promotion les plus faibles, particulièrement dans certaines régions du Nord-Est et de l'Ouest, tandis que les zones violettes affichent des taux plus élevés, souvent dans des régions littorales. Néanmoins on observe que ces régions sont également celles qui ont un frailty effect assez faible, suggérant que le taux brut observé est en partie dû aux profils spécifiques des magistrats plutôt qu'à un environnement local intrinsèquement avantageux. Cela s'explique par un effet de sélection : les magistrats en poste dans ces régions pourraient avoir des caractéristiques favorisant naturellement leur promotion (expérience, formation, réseaux), ce qui gonfle le taux brut sans pour autant refléter une structure départementale plus avantageuse. L'effet de frailty, en corrigeant cette dimension, permet de distinguer l'effet du territoire de celui des profils, révélant que l'attractivité des zones littorales n'implique pas nécessairement un contexte propice à

l'avancement.



(a) Taux de promotion brut par département

(b) Effet de fragilité (*frailty effect*) par département

FIGURE 13 – Comparaison entre le taux de promotion brut et l'effet de frailty par département en France.

Hiérarchie institutionnelle des juridictions et "désertification judiciaire" Les juridictions suprêmes parisiennes (Cour de cassation, juridictions spécialisées nationales) constituent le sommet hiérarchique, offrant une forte visibilité institutionnelle et des opportunités d'avancement accélérées. Les cours d'appel, situées dans les grandes métropoles régionales (Aix-en-Provence, Lyon, Bordeaux), représentent un niveau intermédiaire déterminant pour la progression de carrière. Les tribunaux judiciaires de taille modeste, principalement situés dans des départements ruraux ou moins densément peuplés (Meuse, Ardennes, Lozère), offrent structurellement moins d'opportunités d'avancement. Cette configuration explique en grande partie pourquoi Paris et les Hauts-de-Seine figurent parmi les départements les plus favorables, tandis que des départements ruraux comme la Meuse ou les Ardennes apparaissent comme défavorables.

Le rapport du Sénat sur « La justice de proximité : un service public au service des justiciables » (2021) confirme cette tendance en pointant l'aggravation des disparités territoriales dans l'accès à la justice, en particulier dans les départements ruraux correspondant à notre catégorie 4. Il serait donc intéressant, dans une perspective d'approfondissements de nos premiers travaux, d'étudier les mobilités géographiques au cours des carrières.

Facteurs sociologiques et opposition structurante Paris/province La sociologie de la magistrature a identifié plusieurs mécanismes explicatifs des disparités territoriales dont un "effet de centralité" professionnelle : la proximité avec les centres décisionnels (notamment parisiens) favorise la progression de carrière. L'importance des réseaux professionnels, plus denses et influents dans les grandes juridictions urbaines (Paris, Bouches-du-Rhône), joue également un rôle déterminant comme l'ont montré Dubar *et al.* (2015) dans leurs travaux sur les corps de la fonction publique. Le prestige différencié des affaires constitue un autre

facteur significatif : certaines spécialisations juridiques, plus fréquentes dans les grandes juridictions, sont davantage valorisées pour l’avancement. Enfin, des mécanismes d’auto-sélection sont à l’œuvre, les magistrats ayant des ambitions d’avancement rapide privilégiant naturellement les affectations dans les départements de catégories 1 et 2. Ces observations rejoignent les analyses de Bessière et Pouget (2007) sur les carrières dans la fonction publique d’État, qui démontrent comment les logiques spatiales recoupent des logiques sociales dans la structuration des trajectoires professionnelles.

Ces disparités territoriales soulèvent plusieurs enjeux pour l’organisation judiciaire française. La question de l’équité territoriale s’impose comme centrale et la gestion des carrières constitue un autre défi majeur : quels mécanismes incitatifs mettre en place pour rendre plus attractives les juridictions des départements identifiés comme défavorables ? Enfin, la formation des magistrats mérite d’être repensée : comment préparer les futurs magistrats à des carrières territoriales diversifiées plutôt qu’à un modèle unique centré sur les juridictions privilégiées ? Le rapport de la Commission d’enquête du Sénat sur “La situation de la justice en France” (2022) préconise justement la mise en place d’une “politique territoriale judiciaire ambitieuse” visant à réduire ces disparités. Les travaux de Demoli et Willemez (2023) proposent quant à eux une refonte des critères d’avancement prenant davantage en compte l’expérience acquise dans des juridictions périphériques pour valoriser ces parcours professionnels. En revanche, notre base de données actuelle, qui s’étend jusqu’en 2021, ne nous permet pas d’identifier les effets de ces réformes pour l’instant.

Cette première tentative de modélisation des promotions mérite davantage d’approfondissements et notamment l’étude des flux de mobilités géographiques. Enfin, un autre axe d’approfondissement pourrait être l’implémentation de modèles multi états markoviens en analyse de survie. Cela permettrait de comprendre les trajectoires non pas sur la première promotion seulement, mais sur l’ensemble des carrières, bien que cela nécessite un avancement poussé en statistiques.

Conclusion

L'analyse statistique menée au sein du ministère de la Justice a permis de caractériser les déterminants des promotions et des trajectoires salariales dans un environnement supposé transparent. À travers les modèles de régression, nous avons montré que les écarts de rémunération sont majoritairement expliqués par le positionnement hiérarchique et l'ancienneté, confirmant ainsi le poids des règles administratives. Cependant, la décomposition d'Oaxaca-Blinder a révélé que des différences résiduelles entre hommes et femmes persistent, même à caractéristiques observables identiques, suggérant l'existence de mécanismes non explicites freinant l'égalité salariale.

L'analyse longitudinale, appuyée sur des méthodes de clustering, a mis en évidence quatre trajectoires types de progression salariale. Ces trajectoires, fortement corrélées aux corps d'appartenance et aux catégories, illustrent des disparités marquées dans la dynamique de carrière. Si certaines fonctions, comme la magistrature, connaissent des progressions rapides et stables, d'autres corps, notamment les services pénitentiaires, se heurtent à une certaine stagnation.

Enfin, les modèles de survie ont permis d'explorer les transitions de grade et les changements de statut, révélant des inerties structurelles qui limitent les mobilités internes. Malgré l'apparente lisibilité des règles de promotion, les barrières à la progression restent importantes pour certaines catégories d'agents. Cependant, cette modélisation statistique, bien que pertinente, nécessite davantage de sophistication afin de permettre d'analyser significativement les déterminants sociodémographiques des carrières dans la magistrature.

Ces résultats invitent à s'interroger sur les mécanismes de promotion et d'accompagnement de carrière au sein de la fonction publique, en particulier pour mieux comprendre les freins à la mobilité et les écarts de progression observés. Dans ce contexte, le recours à une modélisation statistique poussée permet de dépasser les constats descriptifs pour quantifier précisément les déterminants de la mobilité professionnelle, les temporalités associées aux trajectoires de carrière, ainsi que les disparités entre groupes. Une telle approche vient compléter la littérature existante en objectivant les effets de facteurs structurels (corps, grade, âge, ancienneté, secteur d'affectation...) et individuels tout en contrôlant pour l'hétérogénéité non observée. L'intérêt serait d'autant plus grand si ces données étaient croisées avec d'autres sources d'information sur les individus, incluant par exemple l'origine sociale, le niveau de diplôme, le statut marital ou les responsabilités familiales. Ces analyses apportent des éléments solides pour orienter les réformes statutaires ou les politiques d'accompagnement. Elles contribueraient également à identifier les leviers concrets d'amélioration des parcours professionnels au sein de la fonction publique.

Références

- Sabine BESSIÈRE et Julien POUGET : Les carrières dans la fonction publique d'état. premiers éléments de caractérisation. *Les salaires en France*, pages 51–66, 2007.
- Alan S. BLINDER : Wage discrimination : Reduced form and structural estimates. *Journal of Human Resources*, 8(4):436–455, 1973.
- Marie CARTIER et Odile JOIN-LAMBERT : Promotions et migrations administratives : histoire, ethnographie, approches croisées. *Travail et emploi*, (127):5–14, 2011.
- Alain DARBEL et Dominique SCHNAPPER : *Morphologie de la haute administration française . 1, Les agents du système administratif / Alain Darbel, Dominique Schnapper*. Cahiers du Centre de sociologie européenne. Mouton, Paris La Haye, 1969.
- Yoann DEMOLI et Laurent WILLEMEZ : *Sociologie de la magistrature*. Armand Colin, 2023.
- Claude DUBAR, Pierre TRIPIER et Valérie BOUSSARD : Chapitre 7 - profession : fonctionnaire. les corps de la fonction publique. *Collection U*, pages 165–188, 2015.
- Sibylle GOLLAC : La fonction publique : une voie de promotion sociale pour les enfants des classes populaires ? une exploitation de l'enquête «emploi 2001». *Sociétés contemporaines*, (58):41–64, 2005.
- Ulrike GRÖMPING : Relative importance for linear regression in r : The package relaimpo. *Journal of Statistical Software*, 17(1):1–27, 2006. URL <https://www.jstatsoft.org/article/view/v017i01>.
- Marek HLAVAC : *oaxaca : Blinder-Oaxaca Decomposition in R*, 2022. URL <https://CRAN.R-project.org/package=oaxaca>. R package version 0.1.5.
- Richard H. LINDEMAN, Peter F. MERENDA et Ruth Z. GOLD : *Introduction to Bivariate and Multivariate Analysis*. Scott, Foresman, 1980.
- Marceau LONG et Laurent BLANC : *L'Economie de la fonction publique*. Collection SUP, 1969.
- Catherine MARRY, Laure BERENI, Alban JACQUEMART, Sophie POCHIC et Anne REVILLARD : *Le plafond de verre et l'État*. Armand Colin, 2017.
- MINISTÈRE DE LA JUSTICE : Les réformes des métiers de l'administration pénitentiaire entrent en vigueur, 2018. URL <https://www.justice.gouv.fr/actualites/actualite/reformes-metiers-ladministration-penitentiaire-entrent-vigueur>.
- MINISTÈRE DE LA JUSTICE : Références statistiques justice, 2024. URL <https://www.justice.gouv.fr/documentation/etudes-et-statistiques/references-statistiques-justice-2024>.
- Ronald OAXACA : Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709, 1973.

Simon QUANTIN : Modèles semi-paramétriques de survie en temps continu sous r. Rapport technique, Institut National de la Statistique et des Études Économiques (INSEE), 2018.

Nicolas ROBETTE : *L'analyse statistique des trajectoires*. Numéro 12. Ined Éditions, 2021.

A Théorie - Oaxaca-Blinder

A.1 Explication de la décomposition

Formellement, en adoptant les notations de Grömping (2006), soit S un sous-ensemble de l'ensemble des prédicteurs P :

$$P = (x_1, \dots, x_p) = (\text{AGE}, \text{CORPS}, \text{TEMPSPLEIN}, \text{STATUT}, \text{SEXE}, \text{REGION})$$

On note $R^2(S)$ le coefficient de détermination du modèle contenant les variables de S . Pour un ensemble S et une variable $x_j \notin S$, on définit la contribution séquentielle supplémentaire de x_j par :

$$\text{seq } R^2(x_j \mid S) = R^2(S \cup x_j) - R^2(S) \quad (3)$$

Si $r = (r_1, \dots, r_p)$ est une permutation des p prédicteurs¹⁷, on note $S(r, x_j)$ l'ensemble des prédicteurs apparaissant avant x_j dans l'ordre r . Alors, la contribution de x_j pour l'ordre r est :

$$\text{seq } R^2(x_j \mid S(r, x_j)) \quad (4)$$

La contribution globale de x_j selon la méthode LMG est obtenue en moyennant cette contribution séquentielle sur toutes les permutations possibles :

$$\text{LMG}(x_j) = \frac{1}{p!} \sum_{r \in \mathcal{P}} \text{seq } R^2(x_j \mid S(r, x_j)) \quad (5)$$

où \mathcal{P} désigne l'ensemble des $p!$ permutations de x_1, \dots, x_p .

Par construction, la somme des contributions LMG sur l'ensemble des prédicteurs restitue exactement le coefficient de détermination global du modèle :

$$R^2 = \sum_{j=1}^p \text{LMG}(x_j) \quad (6)$$

A.2 Formalisation de la décomposition

Formellement, soit Y_m et Y_f les salaires moyens des hommes et des femmes, et soit X_m et X_f les vecteurs moyens de caractéristiques observables pour chaque groupe. Les équations de salaires estimées séparément pour les hommes et les femmes sont :

$$E[Y_m] = X'_m \beta_m, \quad E[Y_f] = X'_f \beta_f \quad (7)$$

où β_m et β_f sont les vecteurs des coefficients de régression pour les hommes et pour les femmes. L'écart salarial moyen Δ est défini par :

$$\Delta = E[Y_m] - E[Y_f] \quad (8)$$

17. Par exemple, si l'on considère l'ordre $r = (\text{AGE}, \text{STATUT}, \text{CORPS}, \text{SEXE}, \text{REGION}, \text{TEMPSPLEIN})$, alors l'ensemble $S(r, \text{CORPS})$ est $\text{AGE}, \text{STATUT}$, car AGE et STATUT précèdent CORPS dans cet ordre.

La décomposition d'Oaxaca-Blinder repose alors sur l'identité suivante :

$$\Delta = (X_m - X_f)' \beta_f + X_m' (\beta_m - \beta_f) \quad (9)$$

où :

- la première composante $(X_m - X_f)' \beta_f$ mesure la part expliquée de l'écart, due aux différences de caractéristiques observables ;
- la seconde composante $X_m' (\beta_m - \beta_f)$ mesure la part inexpliquée.

B Robustness checks - Oaxaca-Blinder

B.1 Choix des variables et multicollinearité

Afin de vérifier l'absence de multicollinéarité excessive entre les variables explicatives du modèle linéaire utilisé dans la Section 3.1, nous avons calculé les facteurs d'inflation de la variance (GVIF) corrigés pour les degrés de liberté de chaque variable (Tableau 5).

Variable	GVIF	Df	GVIF ^{1/(2×Df)}
Âge	1,2	1	1,1
Corps	10,2	8	1,2
Sexe	1,4	1	1,2
Temps plein	1,5	2	1,1
Statut (titulaire, non titulaire)	5,0	1	2,2
Région de fonction	1,4	24	1,0

TABLE 5 – Facteurs d'inflation de la variance (GVIF) et GVIF corrigé pour les variables du modèle principal.

Les résultats montrent que l'ensemble des GVIF corrigés est inférieur à 2,5, c'est-à-dire bien en deçà des seuils d'alerte usuels (généralement fixés à 5 pour une multicollinéarité modérée, et à 10 pour une multicollinéarité problématique), ce qui suggère que la multicollinéarité ne compromet pas substantiellement l'interprétation des coefficients estimés. Parmi les variables catégorielles à nombreux niveaux, la variable Corps présente un GVIF brut relativement élevé (10,2), mais son GVIF corrigé¹⁸ est modéré (1,2), ce qui confirme que l'effet de la colinéarité reste contenu après ajustement.

Dans une spécification alternative intégrant simultanément les variables *Corps* et *Catégorie hiérarchique*, les facteurs d'inflation de la variance (GVIF) corrigés pour les degrés de liberté restent dans l'ensemble modérés (tableau 6), avec des valeurs corrigées inférieures à 2 pour la majorité des variables. Toutefois, deux observations importantes appellent à la prudence. Premièrement, les GVIF bruts atteignent des niveaux très élevés pour Corps (70,4) et Catégorie (12,5). Deuxièmement, il existe une forte redondance conceptuelle entre ces deux variables, car les corps sont un raffinement de la catégorie hiérarchique : notons, par exemple,

18. Une version normalisée du GVIF, appelée GVIF corrigé, permet de comparer les GVIF même lorsque les variables n'ont pas le même nombre de degrés de liberté.

que tous les magistrats sont de catégorie A. Par conséquent, malgré des GVIF corrigés relativement contenus, la structure même des données nous conduit à privilégier une spécification ne retenant qu’une seule de ces deux variables à la fois.

Variable	GVIF	Df	GVIF ^{1/(2×Df)}
Âge	1,2	1	1,1
Corps	70,4	8	1,3
Catégorie	12,5	3	1,5
Sexe	1,4	1	1,2
Temps plein	1,5	2	1,1
Statut (titulaire, non titulaire)	5,1	1	2,3
Région de fonction	1,4	24	1,0

TABLE 6 – Facteurs d’inflation de la variance (GVIF) et GVIF corrigé pour les variables de la spécification alternative intégrant catégorie et corps.

B.2 Sensibilité aux filtres appliqués et bootstrap

Afin d’assurer la stabilité du bootstrap utilisé pour l’estimation des erreurs standards dans la décomposition d’Oaxaca-Blinder (Section 3.2, nous avons filtré les modalités des variables catégorielles apparaissant trop rarement (moins de 10 observations dans l’année). Le tableau 7 présente, pour chaque année, le nombre d’observations avant et après ce filtrage.

Période	Pourcentage restant minimum	Pourcentage restant moyen
1978–1989	99,87	99,93
1990–1999	99,91	99,96
2000–2009	99,91	99,94
2010–2021	99,94	99,98

TABLE 7 – Taux d’observations conservées après filtrage des modalités rares, par période.

Les résultats montrent que l’impact du filtrage est extrêmement limité : dans toutes les années, plus de 99,9% des observations initiales sont conservées. La perte d’observations est donc négligeable.

B.3 Etude des années ”instables” dans la décomposition Oaxaca-Blinder

En complément de la section 4.2, l’analyse du tableau 8 met en évidence deux segments temporels au sein desquels l’écart salarial entre hommes et femmes représente moins de 2% du salaire moyen : d’une part, les années 1990 (1992, 1996 à 1999), et d’autre part, la période postérieure à 2015 (2016 à 2021).

Année	Gender gap (euros 2022)	Intervalle de confiance (95%)	Salaire moyen (euros 2022)	Gender gap sur salaire moyen (%)
1992	459	[70, 849]	34281	1.34
1996	281	[-148, 711]	35754	0.79
1997	-29	[-406, 348]	35818	0.08
1998	-84	[-476, 307]	36119	0.23
1999	574	[179, 969]	36788	1.56
2016	639	[304, 974]	34496	1.85
2017	686	[346, 1027]	34573	1.99
2018	326	[13, 640]	33518	0.97
2019	454	[137, 771]	33573	1.35
2021	675	[437, 912]	33842	1.99

TABLE 8 – Les années “instables” dans la décomposition d’Oaxaca-Blinder : gender gap de moins de 2% du salaire moyen

Dans les années 1990, l’interprétation est particulièrement délicate : non seulement l’écart est faible, mais il présente également une forte instabilité de signe, comme en 1997 et 1998 où l’écart salarial estimé devient négatif. Pendant 4 ans, 0 est présent dans l’intervalle de confiance à 95%. Ce phénomène est lié à la variabilité induite par le bootstrap, nécessaire pour avoir des erreurs standards sur les parts expliquées et inexpliquées. Dès lors, les décompositions Oaxaca-Blinder pour cette période ne sauraient être interprétées d’une manière usuelle.

À l’inverse, pour les années récentes (2016-2021), bien que l’écart salarial demeure modeste, il reste systématiquement positif et significativement différent de zéro. Cela suggère une situation où les différences salariales de genre existent toujours mais sont devenues plus limitées. Ces résultats autorisent ainsi une interprétation prudente de la décomposition d’Oaxaca-Blinder.

C Modèles de durée

C.1 Fonctions de survie et de hasard

Afin de décrire la durée de vie d’un événement, il est nécessaire d’analyser la distribution des temps de survie. Dans le cas des modèles de survie, on utilise les notions statistiques de fonction de survie et fonction de hasard (ou taux de hasard instantané).

La fonction de survie, notée $S(t)$, représente la probabilité qu’un individu n’ait pas encore connu l’évènement d’intérêt à l’instant t , par exemple une promotion de grade ou une augmentation significative de salaire. Elle est reliée à la fonction de répartition $F(t)$ par la relation $S(t) = 1 - F(t)$.

La fonction de hasard $h(t)$ exprime le taux instantané de survenue de l’évènement, condi-

tionnellement à la non-survenue jusque-là :

$$h(t) = \frac{f(t)}{S(t)},$$

où $f(t)$ est la densité de la durée de survie. L'étude conjointe de ces fonctions permet de mieux comprendre non seulement la probabilité de promotion au fil du temps, mais aussi l'évolution du risque de promotion selon l'ancienneté et les caractéristiques individuelles et socio-démographiques.

C.2 Estimations non-paramétriques

Pour décrire empiriquement la distribution des temps d'accès à une promotion ou une augmentation salariale, nous utilisons des méthodes non paramétriques. L'estimateur le plus fréquent est celui de Kaplan-Meier (ou estimateur produit limite). Il s'agit d'une fonction en escalier avec des sauts aux dates où un événement est observé. L'estimateur de Kaplan-Meier pour la fonction de survie est donné par :

$$\hat{S}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{n_j}\right),$$

où d_j est le nombre d'événements observés à t_j et n_j le nombre d'individus encore à risque juste avant t_j . Cet estimateur a pour intérêt d'être non-paramétrique (il n'y a donc pas d'hypothèses fortes sur la forme de la fonction), il prend en compte les données censurées, et il est consistant et sa variance est calculable.

De même, l'estimateur de Nelson-Aalen pour la fonction de risque cumulé permet une estimation de la fonction de hasard, qui correspond à la pente de la fonction de risque cumulé :

$$\hat{H}(t) = \sum_{j:t_j \leq t} \frac{d_j}{n_j}.$$

Ces estimateurs permettent une première exploration des dynamiques de carrière. Le lissage du risque cumulé par noyaux de convolution offre une estimation plus stable de la fonction de hasard instantané, essentielle pour identifier les périodes critiques de promotion.

C.3 Modéliser la durée de vie : une première approche par le modèle de Cox simple

Afin d'étudier l'influence des caractéristiques individuelles sur la probabilité de promotion, nous utilisons le modèle de Cox à hasards proportionnels (modèle semi-paramétrique). La fonction de hasard d'un individu i s'exprime comme :

$$h_i(t|X_i) = h_0(t) \exp(X_i' \beta),$$

où $h_0(t)$ est la fonction de base commune et β est le vecteur de coefficients à estimer.

L'interprétation repose sur le risque relatif : pour deux individus i et j ,

$$\frac{h_i(t|X_i)}{h_j(t|X_j)} = \exp((X_i - X_j)' \beta),$$

indiquant que le risque relatif est constant dans le temps.

Dans notre contexte, les évènements analysés sont le changement de grade : la promotion est effective si le grade supérieur est atteint. Cependant, cette variable n'est qu'une simplification d'un événement de promotion difficile à déterminer (évolution salariale sans changement de grade, augmentation des responsabilités, mutation dans l'objectif d'un avancement de carrière, etc). La variable dépendante est alors définie comme une indicatrice prenant la valeur 1 si l'évènement survient, 0 en cas de censure.

C.4 Les extensions du modèle de Cox - modèle de Cox stratifié et modèle Frailty

Lorsque l'on souhaite modéliser des temps de survenue d'évènements dans un cadre semi-paramétrique, le modèle de Cox proportionnel constitue une base robuste. Toutefois, certaines hypothèses de ce modèle peuvent s'avérer restrictives dans des contextes empiriques complexes, comme l'analyse des carrières dans la fonction publique. Deux extensions permettent d'en assouplir la structure : le modèle de Cox stratifié et le modèle à effet aléatoire, dit « frailty ».

Le modèle de Cox stratifié est pertinent lorsque l'on souhaite tenir compte d'une hétérogénéité structurelle non modélisée par les covariables explicites. Dans ce cadre, on suppose que l'effet des covariables explicatives \mathbf{X}_i sur le risque est constant à travers les strates, mais que la fonction de risque de base $h_0(t)$ peut varier selon une variable de stratification. La fonction de risque d'un individu i appartenant à la strate $s(i)$ s'écrit alors :

$$h_i(t|\mathbf{X}_i, s(i)) = h_{0,s(i)}(t) \exp(\mathbf{X}_i^\top \beta),$$

où $h_{0,s(i)}(t)$ est une fonction de base spécifique à chaque strate. Cette formulation conserve l'interprétation du modèle de Cox en termes de risques relatifs, tout en autorisant une flexibilité temporelle dans les intensités de base selon les groupes. Elle permet par exemple de tenir compte d'effets non paramétriques associés à des catégories comme les régions de fonction ou les corps d'appartenance, sans en estimer explicitement les coefficients.

Le modèle à effet aléatoire de type frailty, quant à lui, introduit une composante multiplicative non observée dans la fonction de risque, pour tenir compte d'une hétérogénéité inobservée entre groupes d'individus. Il s'écrit généralement :

$$h_i(t|\mathbf{X}_i, u_i) = h_0(t)u_i \exp(\mathbf{X}_i^\top \beta),$$

où u_i est une variable aléatoire non observable, propre à l'individu (ou à un groupe d'individus), souvent supposée suivre une loi Gamma de moyenne 1 et de variance θ . La présence de ce terme frailty u_i modifie la distribution du temps de survie, en capturant une vulnérabilité latente non expliquée par les covariables observées. Par exemple, deux agents ayant des caractéristiques observables identiques mais appartenant à des groupes aux dynamiques collectives différenciées (services, régions, fonctions, etc.) peuvent présenter des risques de promotion sensiblement distincts.

Ces deux extensions permettent ainsi d'améliorer la spécification du modèle en capturant des dimensions structurelles (stratification) ou aléatoires (frailty) qui échappent au cadre standard du modèle de Cox. Leur mise en œuvre repose sur l'estimation par vraisemblance partielle adaptée, et peut être réalisée dans la plupart des environnements statistiques, notamment sous R, via les fonctions `coxph()` du package `survival`.

D Résultats complémentaires - modèles de survie

D.1 Résultats modèles de Cox avec stratification par genre

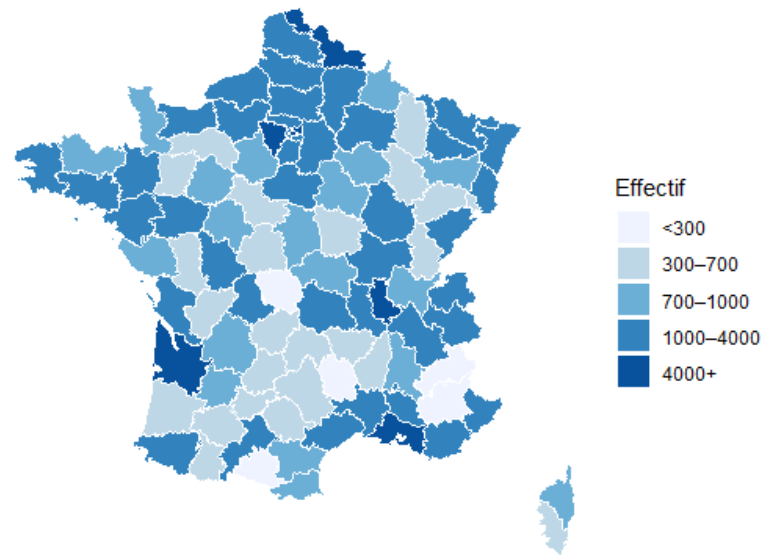
TABLE 9 – Résultats des modèles de Cox avec stratification par sexe

Variable	Modèle 2		Modèle 2bis	
	Coefficient	HR	Coefficient	HR
AGE (continu)	0.0070	1.007	—	—
tranche_age3 (37,47]	—	—	0.1246	1.133
tranche_age3 (47,77]	—	—	0.1169	1.124
STATUT : Non titulaire	1.5719	4.816	1.6932	5.437
TEMPS_PLEIN : Temps partiel	0.3050	1.357	0.2969	1.346
TEMPS_PLEIN : Autre	-0.7335	0.480	-0.7270	0.483

Note : HR = hazard ratio ($\exp(\text{coef})$), `strata(SEXE)`.

L'hypothèse des hasards proportionnels n'étant pas validée, les *hazard rates* ne peuvent être interprétés dans le contexte de notre étude. Cependant, afin de comprendre le sens des *hazard rates*, voici une proposition d'interprétation : «l'effet du statut non titulaire sur la probabilité de promotion est fortement significatif. Dans le Modèle 2bis, le coefficient estimé pour les non titulaires est de 1,693, ce qui correspond à un hazard ratio (HR) de 5,44. Cela signifie que, toutes choses égales par ailleurs, les non titulaires ont une probabilité de promotion instantanée plus de 5 fois supérieure à celle des titulaires.»

D.2 Statistiques descriptives complémentaires pour les modèles de survie



Source : FGE et SIASP

FIGURE 14 – Répartition des effectifs par département dans l'échantillon utilisé pour les modèles de Cox avec effet de frailty