



Bureau de l'égalité
entre les femmes
et les hommes - BEFH

Rue Caroline 11
1014 Lausanne

Etude

Réflexion empirique sur le seuil de tolérance
utilisé lors des contrôles de l'égalité salariale

Titre

Réflexion empirique sur le seuil de tolérance
utilisé lors des contrôles de l'égalité salariale

Editeur

Bureau de l'égalité entre les femmes et les hommes BEFH

Auteurs

Florian Chávez-Juárez, CORESO
Roman Graf, CORESO et R-G

Diffusion

Bureau de l'égalité entre les femmes et les hommes BEFH
Rue Caroline 11
1014 Lausanne
info.befh@vd.ch
www.vd.ch/egalite – www.vd.ch/befh

Les appréciations et interprétations de cette étude ne
reflètent pas forcément le point de vue du BEFH.

Lausanne, août 2021

Préface

Dans le canton de Vaud, la responsabilité des contrôles de l'égalité salariale est endossée par la Commission de contrôle des marchés publics et des subventions CoMPS. Depuis 2020, au moins dix contrôles sont réalisés chaque année auprès des entreprises qui ont obtenu des marchés publics dans le canton et des entités subventionnées par l'Etat de Vaud.

Ces contrôles sont réalisés à l'aide de l'outil d'analyse standard de la Confédération, Logib module 1. Ce dernier incorpore un seuil de tolérance de 5% pour l'effet lié au sexe à conditions égales. Depuis sa fixation en 2004, ce seuil a régulièrement fait l'objet de discussions du point de vue scientifique ou juridique.

Cette étude économétrique portant sur 2845 entreprises et organisations présente de manière détaillée les effets d'un abaissement théorique du seuil de tolérance sur les proportions d'entreprises et organisations respectant ou non l'égalité salariale. Elle montre aussi entre autres que celles qui emploient une part plus élevée de femmes, surtout parmi leurs cadres, respectent plus fréquemment l'égalité salariale.

Nous espérons que cette étude contribuera à éclairer la réflexion relative au seuil de tolérance de Logib module 1 afin d'encourager des pratiques salariales respectueuses de l'égalité salariale au sein des entreprises et organisations, notamment dans le cadre des marchés publics et des subventions.

Maribel Rodriguez

Cheffe du Bureau de l'égalité
entre les femmes et les hommes

Présidente de la Commission de contrôle
des marchés publics et des subventions



Réflexion empirique sur le seuil de tolérance utilisé lors du contrôle de l'égalité salariale dans les marchés publics du canton et les subventions de l'Etat de Vaud

FLORIAN CHÁVEZ-JUÁREZ, CORESO
ROMAN GRAF, CORESO et R-G

Août 2021

Résumé

Dans le canton de Vaud, la Commission de contrôle des marchés publics et des subventions (CoMPS) procède au contrôle du respect de l'égalité salariale entre les femmes et les hommes dans les marchés publics du canton et les subventions de l'Etat. Pour ce faire, elle utilise le modèle d'analyse standard de la Confédération (Logib). Au terme du contrôle, les entités employeuses présentant un écart salarial non expliqué significativement supérieur sur le plan statistique au seuil de tolérance de 5 % fixé par Logib sont considérées comme non conformes et peuvent être sanctionnées. Ce seuil de tolérance de 5 % ne possède aucune base scientifique ou juridique. Pourtant, il joue un rôle déterminant dans la définition du respect de l'égalité salariale au sens de l'outil Logib et donc dans la proportion d'entités considérées comme conformes lors des contrôles. La présente recherche examine la manière dont la proportion d'entités employeuses considérées comme respectant l'égalité salariale au sens de l'instrument Logib varie en fonction de plusieurs niveaux de seuil de tolérance. L'analyse empirique menée dans le cadre de cette recherche se fonde sur les données de 2'845 entités employeuses provenant de l'Enquête suisse sur la structure des salaires (ESS) de 2018. Les résultats montrent que la proportion d'entités employeuses considérées comme respectant l'égalité salariale au sens de l'instrument Logib est de 80,9 % en appliquant le seuil de tolérance de 5 %, de 65,0 % avec un seuil de 2 % et de 49,8 % sans seuil de tolérance. Des analyses complémentaires montrent que ces résultats varient en fonction du secteur (public vs. privé), de la taille de l'entité employeuse, de la branche économique ainsi que de la proportion de femmes et de femmes cadres au sein de l'entité. La principale conclusion du rapport est qu'une réduction, voire une suppression, du seuil de tolérance serait envisageable d'un point de vue économétrique. Un tel changement n'impliquerait pas nécessairement d'inclure des variables supplémentaires dans le modèle Logib, mais plutôt des tests statistiques complémentaires.

Synthèse (Executive summary)

Point de départ

Selon la loi fédérale sur les marchés publics (LMP), seuls les soumissionnaires qui respectent les dispositions relatives à l'égalité de traitement salarial entre les femmes et les hommes peuvent obtenir un marché public. Cette même loi prévoit que l'adjudicateur puisse contrôler le respect de cette exigence ou déléguer cette compétence à des tiers.

Au niveau du canton de Vaud, l'article 6 al. 1 let. f de la loi sur les marchés publics (LMP-VD) ainsi que l'article 3 al. 2 de la loi sur les subventions (LSubv) indiquent que seules les entités respectant l'égalité de traitement entre les femmes et les hommes peuvent respectivement obtenir des marchés publics dans le canton ou recevoir des subventions de l'Etat de Vaud. Selon l'art. 4b de la loi d'application dans le canton de Vaud de la loi fédérale sur l'égalité (LVLEg), la Commission de contrôle des marchés publics et des subventions (CoMPS) procède au contrôle du respect de l'égalité entre les femmes et les hommes en matière de salaires selon la procédure définie dans ses lignes directrices (CoMPS, 2021).

Les contrôles se fondent sur le modèle d'analyse standard de la Confédération (Logib) qui intègre un seuil de tolérance de 5 %. Si un effet lié au sexe est identifié et qu'il est supérieur à 5 % de façon statistiquement significative, l'entité employeuse visée par le contrôle est considérée comme non conforme sur le plan de l'égalité salariale et peut être sanctionnée.

Le seuil de tolérance de 5 % de Logib a été introduit au moment du développement de l'outil en 2004 et ne possède aucune base scientifique ou juridique. Le recours à un seuil est simplement justifié par le Bureau fédéral de l'égalité entre femmes et hommes (BFEG) par le fait que Logib est un modèle standard et qu'il n'est pas possible d'exclure qu'il existe d'autres variables objectives et non discriminatoires propres à une entité employeuse qui ne sont pas incluses dans l'outil. Sur la base des résultats des contrôles menés dans les marchés publics et les subventions, il est cependant observé que ce seuil de tolérance joue un rôle déterminant sur la conformité des entités contrôlées en matière d'égalité salariale. Sur les onze premiers contrôles réalisés par la CoMPS, quatre ont un résultat significativement supérieur à 0 % sur le plan statistique et zéro ont un résultat significativement supérieur au seuil de tolérance de 5 %. Au niveau fédéral, sur la base de 171 contrôles achevés par le Bureau fédéral de l'égalité entre femmes et hommes (BFEG, état au 23.06.2020), 77 entités possèdent un écart salariale non expliqué significativement supérieur à 0 % alors que 16 possèdent un écart salarial non expliqué significativement supérieur à 5 %.

Compte tenu des résultats des contrôles, il convient de s'interroger sur la pertinence du seuil de tolérance de 5 %. En effet, si l'objectif du seuil de tolérance est d'éviter de sanctionner à tort des entités qui respectent l'égalité salariale, celui-ci ne doit pas pour autant empêcher d'identifier à raison des entités qui ne respecteraient pas l'égalité salariale.

Objectifs du mandat

L'objectif de ce mandat est d'étudier la proportion d'entités employeuses qui respectent l'égalité salariale, au sens du modèle d'analyse standard de la Confédération (Logib), en fonction du niveau du seuil de tolérance utilisé. Ce mandat doit permettre à la CoMPS d'apprécier l'impact potentiel de différents seuils de tolérance sur le résultat de ses contrôles.

En particulier, il s'agit de répondre à trois questions :

1. Quelle proportion d'entités employeuses respectent l'égalité salariale au sens du modèle d'analyse standard de la Confédération lorsque le seuil de tolérance varie entre 0 % et 5 % ?
2. Quelles variables jouent un rôle significatif sur la proportion d'entités employeuses qui respectent l'égalité salariale et comment ce rôle change en fonction du seuil de tolérance utilisé ?
3. Quelles informations devraient être prises en considération par la CoMPS lors d'une réflexion sur le seuil de tolérance dans le cadre des contrôles menés dans les marchés publics et les subventions ?

Méthodologie

Pour répondre à ces questions, une analyse empirique est réalisée à partir des données de 2'845 entités employeuses issues de l'Enquête suisse sur la structure des salaires (ESS) de 2018. Cette analyse empirique se fonde sur le modèle d'analyse standard de la Confédération et consiste à le calculer pour toutes les entités employeuses de l'échantillon. Cette approche permet notamment de déterminer la proportion d'entités qui respectent l'égalité salariale entre les femmes et les hommes au sens du modèle d'analyse standard de la Confédération et d'apprécier la manière dont cette proportion varie en fonction de différents niveaux de seuil de tolérance (0 %, 1 %, 2 %, 3 %, 4 %, 5 %).

Cette analyse est également effectuée pour des groupes spécifiques en ventilant les résultats par secteur (privé vs. public), taille de l'entité, branche économique et proportion de femmes (cadres) au sein de l'entité. Ces analyses permettent de mieux comprendre quelles caractéristiques sont liées à un taux de conformité plus élevé et comment ces différents groupes seraient impactés par une modification du seuil de tolérance.

Sur la base d'une discussion théorique relative au seuil de tolérance, des analyses supplémentaires sont réalisées pour connaître la relation entre pouvoir explicatif du modèles et conformité des entités en matière d'égalité salariale, ainsi que l'effet de la prise en compte de variables supplémentaires sur la proportion de conformité. Enfin, ce rapport reprend une discussion de la littérature concernant une correction statistique qui est potentiellement nécessaire si une hypothèse concernant la distribution de la part non expliquée des salaires (résidus) n'est pas satisfaite¹. Une analyse empirique de cette discussion permet d'évaluer l'impact de cette correction sur les conclusions de Logib.

Résultats

Les principales réponses aux questions posées par le mandat sont les suivantes.

1. Quelle proportion d'entités employeuses respectent l'égalité salariale au sens du modèle d'analyse standard de la Confédération lorsque le seuil de tolérance varie entre 0 % et 5 % ?

Au niveau de l'ensemble de l'économie, la proportion d'entités employeuses respectant l'égalité salariale au sens du modèle d'analyse standard de la Confédération est de 80,9 % avec le seuil actuel de 5 %, de 65,0 % avec un seuil de 2 % et de 49,8 % sans seuil de tolérance. Autrement

1. Il s'agit de l'hypothèse de l'homoscédasticité. Voir l'annexe B.2 pour une explication détaillée.

dit, si le seuil actuel est maintenu, une entité sur cinq ne remplit pas les conditions requises, tandis qu'une sur deux ne les remplirait pas si aucun seuil n'était appliqué.

2. Quelles variables jouent un rôle significatif sur la proportion d'entités employeuses qui respectent l'égalité salariale et comment ce rôle change en fonction du seuil de tolérance utilisé ?

- Les entités employeuses du secteur public respectent significativement plus l'égalité salariale au sens du modèle d'analyse standard de la Confédération que les entités du secteur privé lorsque l'on tient compte du seuil de 5 %, mais cette différence est non significative lorsqu'il n'y a pas de seuil de tolérance.
- Les grandes entités sont légèrement plus susceptibles d'être en conformité que les petites avec un seuil de tolérance de 5 %. Lorsque le seuil se rapproche de zéro, l'inverse est vrai. En tenant compte des autres facteurs et sans seuil de tolérance, les entités employeuses les plus grandes ont jusqu'à 25 points de pourcentage de moins de probabilité de remplir les exigences en matière d'égalité salariale.
- De grandes différences dans la proportion d'entités conformes sont constatées entre les branches économiques. Les industries manufacturières ont la plus faible proportion de conformité. Ces différences ne changent pas de manière significative lorsque le seuil de tolérance est modifié.
- Les entités employeuses qui comptent une proportion plus élevée de femmes dans leur effectif et surtout parmi leurs cadres respectent davantage l'égalité salariale au sens de Logib. Les différences sont un peu plus faibles lorsque l'on tient compte d'autres facteurs, mais elles restent statistiquement significatives.

3. Quelles informations devraient être prises en considération par la CoMPS lors d'une réflexion sur le seuil de tolérance dans le cadre des contrôles menés dans les marchés publics et les subventions ?

- En 2015, une étude qualitative a révélé que des expert-e-s en matière d'égalité salariale estimaient que le seuil de tolérance avait prouvé son utilité. Ces personnes préconisaient de maintenir le modèle actuel avec un seuil de tolérance à 5 % plutôt que d'ajouter des variables supplémentaires et de réduire le seuil de tolérance. Dans une interview, un-e expert-e a toutefois proposé de baisser le seuil de tolérance.
- D'un point de vue statistique, l'utilisation d'un seuil de tolérance en plus du test de signification n'est pas justifiée pour éviter des conclusions erronées concernant le respect de l'égalité salariale au sein d'une entité employeuse. L'utilisation du seuil de tolérance actuel serait justifiable si des différences allant jusqu'à 5 % étaient explicitement tolérées par la loi, ce qui n'est pas le cas.
- L'analyse empirique du présent rapport montre que le lien entre le pouvoir explicatif du modèle Logib et l'écart salarial est faible. L'abaissement du seuil de tolérance ne nécessiterait donc pas automatiquement en contrepartie l'ajout de variables supplémentaires dans Logib. Les entités pour lesquelles le modèle Logib explique très bien leur pratique salariale sont en fait plus impactées par une variation du seuil de tolérance en termes de conformité en matière d'égalité salariale.
- L'analyse empirique a montré que près de 80 % de toutes les régressions Logib transgressent l'hypothèse sur la distribution des résidus de la régression (test d'homoscédasticité²). La

2. Voir annexe B.2 pour une explication

correction de ce problème n'a que peu d'impact sur les résultats des analyses Logib lorsque le seuil de tolérance de 5 % est pris en compte. Toutefois, la correction de ce problème modifie les conclusions relatives à la conformité des entités dans environ 5 % des cas lorsqu'aucun seuil de tolérance n'est pris en compte. Dès lors, il conviendrait de procéder à cette correction si le seuil de tolérance était diminué ou supprimé.

Conclusions

Sur la base de l'analyse empirique, de la discussion théorique et de la littérature, trois recommandations peuvent être formulées.

1. Supprimer le seuil de tolérance et se baser uniquement sur le test de signification statistique pour déterminer si une entité employeuse respecte ou non l'égalité salariale. La suppression du seuil pourrait se faire par étapes en passant par des seuils de tolérance moins élevés (par exemple 2 %) pendant une phase de transition.
2. En cas de suppression du seuil, faire systématiquement le test statistique pour examiner l'homogénéité de la distribution des résidus de la régression (homoscédasticité) et, en cas de rejet de l'hypothèse, utiliser des erreurs standard robustes.
3. En cas de suppression du seuil et de non conformité d'une entité employeuse en matière d'égalité salariale, envisager la possibilité pour l'entité de fournir des justifications objectives et conformes au droit afin de déterminer si elles pourraient avoir un impact sur la conclusion de l'analyse.

Ces recommandations sont basées sur des considérations statistiques mais cherchent aussi à prendre en compte les différents arguments en faveur du seuil de tolérance. La recommandation de supprimer le seuil de tolérance se base sur le constat que le système actuel permet à plus de 30 % des entités employeuses d'avoir un écart salarial significativement supérieur à zéro sur le plan statistique (donc contraire à la loi) tout en restant conformes selon les critères Logib.

La deuxième recommandation découle du fait que, sans seuil de tolérance, jusqu'à 5 % des conclusions par rapport à la conformité des entités en matière d'égalité salariale pourraient être changées en corrigeant le problème lié à la distribution des résidus de la régression. Cette correction n'a aucun coût et son introduction pourrait donc être envisagée.

Finalement, la troisième recommandation tient compte de l'argument de certaines parties selon lequel un modèle statistique ne peut jamais refléter complètement la réalité de chaque entité employeuse. Au terme d'une analyse sans seuil de tolérance débouchant sur un constat de non conformité en matière d'égalité salariale entre les femmes et les hommes, il serait judicieux de déterminer si l'entité possède des variables objectives et non discriminatoires susceptibles de faire changer la conclusion du contrôle. Il est toutefois à noter que l'appréciation du potentiel discriminatoire de nouvelles variables est une tâche très complexe (Pärli, 2019) et susceptible d'engendrer un coût disproportionné dans le cadre d'une procédure administrative. Actuellement, une telle démarche, bien que sensée sur le plan juridique, serait donc certainement difficile à mettre en œuvre. Le développement des connaissances et des outils en matière d'égalité salariale entre les femmes et les hommes pourraient peut-être permettre à l'avenir de concrétiser une telle démarche.

Table des matières

Résumé	2
Synthèse (Executive summary)	3
Point de départ	3
Objectifs du mandat	3
Méthodologie	4
Résultats	4
Conclusions	6
1 Introduction	10
1.1 Situation initiale et bases légales	10
1.2 Objectif du mandat	11
1.3 Méthodologie	11
1.4 Structure du rapport	12
2 Le modèle d'analyse standard de la Confédération (Logib)	13
2.1 Le modèle de régression	13
2.2 Identification d'un effet lié au sexe	14
3 Le rôle du seuil de tolérance	16
3.1 Littérature concernant le seuil de tolérance	16
3.1.1 Origine et justification du seuil de tolérance	16
3.1.2 Appréciation du seuil de tolérance par les expert·e·s	17
3.2 Importance du seuil de tolérance pour les contrôles d'égalité salariale	18
3.3 Lien entre statistique nationale et modèle d'analyse standard de la Confédération	19
3.4 Un regard statistique sur la discussion autour du seuil de tolérance	21
3.4.1 Risque d'une mauvaise conclusion	21
3.4.2 Lien entre pouvoir explicatif du modèle et l'écart estimé	23
3.4.3 Erreurs standard robustes	24
4 Données et méthodologie	26
4.1 Source de données	26
4.2 Choix des variables et traitement des données	26
4.3 Sélection de l'échantillon	28
4.4 Modèles économétriques	28
4.4.1 Approche générale	28
4.4.2 Modèle standard	29
4.4.3 Modèle simple	29
4.4.4 Modèle robuste	29
5 Résultats	30
5.1 Variation des résultats Logib par niveau de seuil de tolérance	30
5.2 Résultats Logib par niveau de seuil de tolérance ventilés par variable	33
5.2.1 Analyse par secteur	33
5.2.2 Analyse par taille d'entité employeuse	34
5.2.3 Analyse par branche économique	35

5.2.4	Analyse selon la proportion de femmes et de femmes cadres au sein de l'entité employeuse	36
5.2.5	Conformité par niveau de tolérance pour l'ensemble des facteurs	37
5.3	Lien entre le pouvoir explicatif du modèle et la différence salariale liée au sexe . .	41
5.4	Calcul de l'analyse Logib en tenant compte de l'hétéroscédasticité	44
6	Discussion et recommandations	46
6.1	Récapitulatif des principales constatations de l'étude	46
6.2	Conclusions	47
6.3	Limites de l'analyse	48
A	Statistiques descriptives	51
B	Explications techniques	54
B.1	Test de significativité	54
B.2	Hétéroscédasticité	55
B.3	Spécification d'un modèle économétrique	56
B.3.1	Diagramme causal et types d'effets	56

Table des figures

1	Illustration de la conclusion de Logib avec 3 cas de figure	15
2	Résultats Logib par niveau de seuil de tolérance	30
3	Distribution du coefficient estimé de la variable sexe (valeur absolue)	31
4	Conformité par niveau de tolérance pour les secteurs public et privé	33
5	Conformité par niveau de tolérance selon la taille de l'entité employeuse	34
6	Conformité par niveau de tolérance selon la branche économique (NOGA)	35
7	Conformité par niveau de tolérance selon la proportion de femmes au sein de l'entité employeuse	36
8	Conformité des entités employeuses par niveau de tolérance selon la proportion de femmes cadres au sein de l'entité employeuse	37
9	Distribution du coefficient de détermination (R^2)	41
10	Conformité par niveau de tolérance selon le niveau d'ajustement du modèle	42
11	Différence de l'effet lié au sexe entre le modèle standard et le modèle simple	43
12	Distribution de position professionnelle (par sexe)	52
13	Distribution de position professionnelle avec fonction de cadre (par sexe)	52
14	Distribution du niveau de compétences (par sexe)	53
15	Illustration de l'intervalle de confiance	54
16	Illustration de l'hétéroscédasticité	55
17	Diagramme causal	56
18	Diagramme causal théorique	57

Liste des tableaux

1	Définition actuelle de conformité	19
2	Variables pour l'estimation du modèle Logib	27
3	Variables supplémentaires utilisées	28
4	Modèle de probabilité linéaire : conformité avec l'égalité salariale	39
5	Comparaison des conclusions avec et sans erreurs standard robustes	44
6	Nombre d'entités par secteur et taille (non pondéré)	51
7	Statistiques descriptives de l'échantillon utilisé	51
8	Possibles effets à estimer avec un modèle économétrique	57

Remerciements

Les auteurs tiennent à remercier l'Office fédéral de la statistique (OFS), en particulier la Dre Julia Ignaczewska et Rahel Péter, pour la rapidité avec laquelle elles ont traité la demande d'obtention des données de l'ESS. Ils remercient aussi le Dr Steve Binggeli pour ses commentaires approfondis sur les premières versions de ce rapport.

1 Introduction

1.1 Situation initiale et bases légales

Depuis 1981, l'art. 8 al. 3 de la Constitution fédérale (Cst) intègre le principe selon lequel : *"L'homme et la femme ont droit à un salaire égal pour un travail de valeur égale"*. La loi fédérale sur l'égalité entre femmes et hommes (LEg) précise à l'art. 3 al. 2 que *"l'interdiction de toute discrimination s'applique notamment à l'embauche, à l'attribution des tâches, à l'aménagement des conditions de travail, à la rémunération, à la formation et à la formation continue, à la promotion et à la résiliation des rapports de travail."*

Selon l'art. 12 al. 1 de loi fédérale sur les marchés publics (LMP), seuls les soumissionnaires qui respectent les dispositions relatives à l'égalité de traitement salarial entre les femmes et les hommes peuvent obtenir un marché public. Cette même loi prévoit aussi que l'adjudicateur puisse contrôler le respect de cette exigence ou déléguer cette compétence de contrôle à des tiers.

Dans le canton de Vaud, l'article 6 al. 1 let. f de la loi sur les marchés publics (LMP-VD) ainsi que l'article 3 al. 2 de la loi sur les subventions (LSubv) indiquent que seules les entités respectant l'égalité de traitement entre les femmes et les hommes peuvent respectivement obtenir des marchés publics dans le canton ou recevoir des subventions de l'Etat de Vaud. Selon l'art. 4b de la loi d'application dans le canton de Vaud de la loi fédérale sur l'égalité (LVLEg), la Commission de contrôle des marchés publics et des subventions (CoMPS) procède au contrôle du respect de l'égalité salariale entre les femmes et les hommes auprès des entreprises qui ont obtenu des marchés publics dans le canton et des entités subventionnées par l'Etat. La procédure de contrôle est définie dans les lignes directrices de la CoMPS ([CoMPS, 2021](#)).

L'introduction dans la loi fédérale sur les marchés publics de pouvoir contrôler le respect de l'égalité salariale entre les femmes et les hommes au sein des entités employeuses est à l'origine du modèle d'analyse standard de la Confédération (Logib). Ce modèle, développé sur mandat du Bureau fédéral de l'égalité entre femmes et hommes (BFEG), est actuellement utilisé pour le contrôle du respect de l'égalité salariale non seulement au niveau fédéral, mais aussi au niveau cantonal et communal. Logib se fonde sur la méthode de l'analyse statistique des salaires appelée régression qui a été acceptée par les tribunaux dans des cas de discrimination salariale. Elle permet d'identifier l'écart salarial entre les femmes et les hommes qui ne peut pas être expliqué par les variables prises en compte dans le modèle. Dans le cas de Logib, ces variables sont la formation, l'ancienneté, l'expérience professionnelle potentielle, le niveau de compétence au sein de l'entité employeuse et la position professionnelle.

Si cet écart salarial non expliqué calculé par Logib est supérieur de façon statistiquement significative à un seuil de tolérance fixé à 5 %, alors la pratique salariale de l'entité est considérée comme non conforme sur le plan de l'égalité salariale entre les femmes et les hommes. Dans le cadre d'un contrôle de l'égalité salariale dans les marchés publics ou les subventions, un tel constat de non conformité de la part d'une entité peut déboucher sur des sanctions.

Le seuil de tolérance de 5 % de l'instrument Logib a été introduit au moment du développement de l'outil en 2004 et ne possède aucune base scientifique ou juridique. Le recours à un seuil est simplement justifié par le BFEG par le fait que Logib est un modèle standard et qu'il n'est pas possible d'exclure qu'il existe d'autres variables objectives et non discriminatoires propres à une entité employeuse qui ne sont pas incluses dans l'outil. Sur la base des résultats des contrôles menés dans les marchés publics et les subventions, il est cependant observé que ce seuil de tolérance joue

un rôle déterminant sur la conformité des entités contrôlées en matière d'égalité salariale. Sur les onze premiers contrôles lancés par la CoMPS en 2020, quatre ont un résultat significativement supérieur à 0 % sur le plan statistique et zéro ont un résultat significativement supérieur au seuil de tolérance de 5 %. Au niveau fédéral, sur la base de 171 contrôles achevés par le Bureau fédéral de l'égalité entre femmes et hommes (état au 23.06.2020), 77 entités possèdent un écart salariale non expliqué significativement supérieur à 0 % alors que 16 possèdent un écart salarial non expliqué significativement supérieur à 5 %.

Compte tenu des résultats des contrôles, il convient de s'interroger sur la pertinence du seuil de tolérance de 5 %. En effet, si l'objectif du seuil de tolérance est d'éviter de sanctionner à tort des entités qui respectent l'égalité salariale, celui-ci ne doit pas pour autant empêcher d'identifier à raison des entités qui ne respecteraient pas l'égalité salariale.

1.2 Objectif du mandat

L'objectif du mandat est d'étudier la proportion d'entités employeuses qui respectent l'égalité salariale, au sens du modèle d'analyse standard de la Confédération, en fonction du niveau du seuil de tolérance utilisé. Ce mandat doit permettre à la CoMPS d'apprécier l'impact potentiel du seuil de tolérance sur le résultat de ses contrôles.

En particulier, il s'agit de répondre à trois questions :

1. Quelle proportion d'entités employeuses respectent l'égalité salariale au sens du modèle d'analyse standard de la Confédération lorsque le seuil de tolérance varie entre 0 % et 5 % ?
2. Quelles variables jouent un rôle significatif sur la proportion d'entités employeuses qui respectent l'égalité salariale et comment ce rôle change-t-il en fonction du niveau du seuil de tolérance utilisé ?
3. Quelles informations devraient être prises en considération par la CoMPS lors d'une réflexion sur le seuil de tolérance dans le cadre des contrôles menés dans les marchés publics et les subventions ?

1.3 Méthodologie

Pour répondre à ces questions, une analyse empirique est réalisée à partir des données de 2'845 entités employeuses issues de l'Enquête suisse sur la structure des salaires (ESS) de 2018. Cette analyse empirique se fonde sur le modèle d'analyse standard de la Confédération et consiste à le calculer pour toutes les entités employeuses de l'échantillon. Cette approche permet notamment de déterminer la proportion d'entités qui respectent l'égalité salariale entre les femmes et les hommes au sens du modèle d'analyse standard de la Confédération et d'apprécier la manière dont cette proportion varie en fonction de différents niveaux de seuil de tolérance (0 %, 1 %, 2 %, 3 %, 4 %, 5 %).

Cette analyse est également effectuée pour des groupes spécifiques en ventilant les résultats par secteur (privé vs. public), taille de l'entité, branche économique et proportion de femmes (cadres) au sein de l'entité. Ces analyses permettent de mieux comprendre quelles caractéristiques sont liées à un taux de conformité plus élevé et comment ces différents groupes seraient impactés par une modification du seuil de tolérance.

Sur la base d’une discussion théorique relative au seuil de tolérance, des analyses supplémentaires sont réalisées pour connaître la relation entre pouvoir explicatif du modèles et conformité des entités en matière d’égalité salariale, ainsi que l’effet de la prise en compte de variables supplémentaires sur la proportion de conformité. Enfin, ce rapport reprend une discussion de la littérature concernant une correction statistique qui est potentiellement nécessaire si une hypothèse concernant la distribution de la part non expliquée des salaires (résidus) n’est pas satisfaite³. Une analyse empirique de cette discussion permet d’évaluer l’impact de cette correction sur les conclusions de Logib.

1.4 Structure du rapport

Ce rapport débute par la présentation du modèle d’analyse standard de la Confédération (Logib) dans la section 2. Cette présentation introduit d’abord le modèle Logib de façon générale, puis se concentre sur le rôle du seuil de tolérance de 5 %. La section 3 reprend les questions du mandat et les aborde de façon théorique. Quelques résultats clés de la littérature sont également présentés, ce qui permet de guider l’analyse empirique qui suit. La méthodologie de l’analyse empirique ainsi que les données utilisées sont décrites dans la section 4. La section 5 est dédiée aux résultats de l’analyse empirique en reprenant les étapes de recherche ainsi que la discussion théorique de la section 3. Sur la base de la discussion théorique et des résultats empiriques, la section 6 propose une conclusion et des recommandations. Les annexes fournissent des informations supplémentaires sur l’échantillon utilisé (annexe A) et sur certains aspects techniques (annexe B).

3. Il s’agit de l’hypothèse de l’homoscédasticité. Voir l’annexe B.2 pour une explication détaillée.

2 Le modèle d'analyse standard de la Confédération (Logib)

2.1 Le modèle de régression

Le module 1 de Logib a été développé par le Bureau fédéral de l'égalité entre femmes et hommes (BFEG) en collaboration avec des mandataires externes (Strub, 2004). Il s'agit d'un outil qui se base sur l'analyse de régression et qui cherche à identifier l'écart salarial qui ne peut pas être expliqué par des facteurs objectifs. Cette approche se distingue d'une simple comparaison des salaires en corrigeant l'écart salarial "brut" par la prise en compte de possibles différences entre les femmes et les hommes sur les facteurs objectifs formation, ancienneté, expérience professionnelle potentielle, niveau de compétence au sein de l'entité employeuse et position professionnelle. Autrement dit, la régression enlève la part de l'écart salarial entre les femmes et les hommes qui est due à ces facteurs et fournit une estimation de la part de l'écart salarial qui est attribuée directement au sexe. Il est à noter que le modèle est dit standard, car les mêmes facteurs sont utilisés quelle que soit l'entité au sein de laquelle l'analyse est effectuée.

Formellement, le modèle Logib s'écrit comme suit ⁴ :

$$\ln(\text{salaire}_i) = \beta_0 + \beta_{Train} \cdot Train_i + \beta_{Exp} \cdot Exp_i + \beta_{Exp^2} \cdot Exp_i^2 + \beta_{Tnr} \cdot Tnr_i + \beta_{Skl} \cdot Skl_i + \beta_{Pos} \cdot Pos_i + \beta_{Sexe} \cdot Sexe_i + \varepsilon_i \quad (1)$$

où

$\ln(\text{salaire}_i)$	logarithme naturel du salaire brut standardisé à plein temps pour l'employé-e i
$Train_i$	Nombre d'années de formation converties à partir du code de formation ⁵ .
Exp_i	Années d'expérience potentielle calculée à l'aide de l'âge et la durée moyenne de la formation acquise : $Exp_i = \max(Age_i - Train_i - 6, 0)$.
Tnr_i	Années de service au sein de l'entité employeuse actuelle.
Skl_i	Niveau de compétences au sein de l'entité employeuse. Cette variable catégorielle à 4 niveaux fait référence au niveau de qualification requis par le poste occupé. Elle intègre la régression sous forme d'un ensemble de 3 variables muettes selon les procédures standard en économétrie.
Pos_i	Position professionnelle du poste occupé par l'employé-e. Il s'agit d'une variable catégorielle à 5 niveaux (de <i>cadres supérieurs</i> à <i>employé-e-s sans fonction de cadre</i>) qui intègre la régression sous forme d'un ensemble de 4 variables muettes selon les procédures standard en économétrie.
$Sexe_i$	Variable muette du sexe (1 = femme, 0 = homme)

Les bêtas (β) dans l'équation (1) font référence aux coefficients estimés par le modèle économétrique et mesurent l'impact des différentes variables explicatives sur la variable expliquée (salaire). Une valeur bêta positive pour une variable explicative signifie qu'une valeur plus élevée de cette variable est associée à une valeur plus élevée de la variable expliquée et inversement pour une valeur négative.

4. Il s'agit d'une notation simplifiée du modèle comme il est également présenté par BFEG (2021b). Pour plus d'informations sur le modèle, veuillez consulter BFEG (2021b).

2.2 Identification d'un effet lié au sexe

Le coefficient d'intérêt pour les analyses d'égalité salariale est β_{sexe} . Une valeur négative indique un écart salarial en défaveur des femmes, tandis qu'une valeur positive fait référence à un écart salarial en défaveur des hommes.

La valeur réelle du coefficient est inconnue et le modèle statistique vise à estimer la valeur la plus probable. Cependant, cela reste une estimation et une certaine incertitude sur la valeur réelle persiste. Pour déterminer si la valeur estimée est significativement différente de zéro (ou de 5 % dans le cas de l'utilisation du seuil de tolérance), un test statistique de signification est utilisé⁶. Un intervalle de confiance autour de la valeur estimée est calculé et indique la plage des valeurs possibles de la valeur réelle pour un niveau de confiance donné. Si la valeur de zéro se trouve à l'extérieur de cet intervalle, la probabilité que la vraie valeur soit égale à zéro est inférieure à 5% et il est donc possible de conclure que la vraie valeur est effectivement différente de zéro. Pour Logib, un niveau de confiance de 95 % est utilisé, ce qui correspond à une pratique courante pour de tels modèles.

Tout d'abord, le modèle d'analyse standard de la Confédération vérifie si l'effet est différent de zéro. Si un effet différent de zéro est trouvé, un deuxième test est effectué pour évaluer si l'effet est significativement supérieur au seuil de tolérance de 5 %. Ce seuil de tolérance est utilisé depuis l'introduction de Logib. L'utilisation de deux tests consécutifs signifie que Logib peut arriver à trois conclusions différentes en ce qui concerne la conformité en matière d'égalité salariale.

La figure 1 illustre ces trois conclusions à titre d'exemple. Le cas indiqué en vert est l'exemple d'un coefficient estimé à -0,005, soit une différence salariale d'environ 0.5 % en défaveur des femmes dans le cas de Logib. La surface verte en dessous de la courbe indique l'intervalle de confiance. Étant donné que la valeur de zéro est comprise dans cet intervalle, Logib conclut qu'il n'y a pas d'effet lié au sexe.

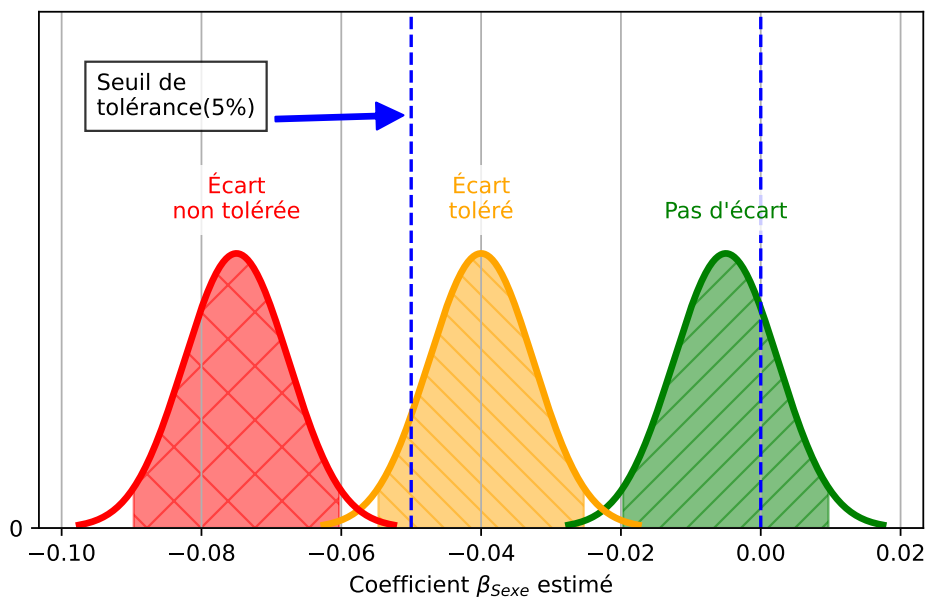
Le cas orange fait référence à un coefficient estimé de -0,04, soit une différence salariale non expliquée d'environ 4 % en défaveur des femmes selon Logib. Dans ce cas, la valeur de zéro ne fait pas partie de l'intervalle de confiance (surface orange). Cela indique qu'il existe un effet significatif lié au sexe (en défaveur des femmes). Ce constat amène à un deuxième test pour voir si la valeur est significativement inférieure à -0,05, soit le seuil de tolérance de 5 % de Logib, ce qui n'est pas le cas ici. La conclusion pour ce cas serait qu'il existe effectivement un effet lié au sexe, mais que cet écart salarial non expliqué ne dépasse pas le seuil de 5 %.

Enfin, l'exemple en rouge illustre un coefficient avec un intervalle de confiance allant de -0,09 à -0,06, ce qui veut dire que le coefficient est significativement différent de zéro et significativement plus grand que 5 %. Dans ce cas, la conclusion du modèle Logib est qu'il existe un important effet lié au sexe et que le seuil de tolérance est dépassé. Dans le cadre du contrôle de l'égalité salariale dans les marchés publics et les subventions, cette conclusion implique la non conformité de l'entité soumise au contrôle et donc de possibles sanctions.

A ce stade, il est important de souligner que Logib analyse la pratique salariale au niveau de l'entité employeuse en prenant en compte le personnel dans son ensemble. Le résultat obtenu avec Logib, soit la différence salariale non expliquée liée au sexe, ne fournit aucune indication concernant des discriminations salariales individuelles ou de groupe. C'est pourquoi, il n'est pas possible de conclure sur la base d'un résultat Logib non significativement différent de zéro ou

6. Une explication détaillée sur le fonctionnement de ces tests se trouve dans l'annexe B.1.

FIGURE 1 – Illustration de la conclusion de Logib avec 3 cas de figure



inférieur au seuil de tolérance de 5 % à l'absence de discrimination salariale au sein de l'entité. Par ailleurs, le résultat obtenu avec Logib n'offre aucune garantie quant à la probabilité que les autorités ou tribunaux compétents en la matière arrivent aux mêmes conclusions

Points clés 1 (Approche Logib).

- Logib calcule l'écart salarial entre les femmes et les hommes qui ne peut pas être expliqué par des facteurs objectifs tels que l'expérience professionnelle potentielle, l'ancienneté, la formation, le niveau de compétences requis pour le poste ou encore la position professionnelle.
- Une entité employeuse est seulement considérée comme non conforme en matière d'égalité salariale entre femmes et hommes, si l'écart salarial non expliqué par ces facteurs **dépasse le seuil de tolérance de 5 % de façon statistiquement significative**.

3 Le rôle du seuil de tolérance

Cette section est consacrée à une discussion théorique sur l'utilisation d'un seuil de tolérance pour le contrôle de l'égalité salariale entre les femmes et les hommes dans les marchés publics et les subventions.

Dans un premier temps, la section 3.1 présente une revue de la littérature sur le thème du seuil de tolérance, ce qui permet de résumer la discussion existante sur l'utilité de ce dernier.

Dans un deuxième temps, la section 3.2 traite de l'importance du seuil de tolérance sur le résultat des contrôles de l'égalité salariale.

Finalement, la section 3.4 traite certains aspects méthodologiques du seuil de tolérance à prendre en considération pour des valeurs inférieures à 5 % ou dans le cadre d'analyses sans seuil de tolérance.

3.1 Littérature concernant le seuil de tolérance

Depuis son introduction, Logib a fait l'objet de plusieurs rapports et études scientifiques. Le seuil de tolérance est un sujet traité dans la plupart de ces rapports.

3.1.1 Origine et justification du seuil de tolérance

L'origine du seuil de tolérance n'est pas très claire selon la littérature. L'appréciation générale est que le seuil de tolérance de 5 % n'a pas de base empirique, théorique ou juridique (Felfe *et al.*, 2015; Graf, 2016; Binggeli *et al.*, 2018). Plusieurs rapports mentionnent que le seuil de tolérance se justifie par le fait qu'il pourrait exister, au sein d'une entité employeuse, d'autres facteurs objectifs et non discriminatoires que ceux pris en compte par Logib et que ces facteurs pourraient permettre d'expliquer l'écart salarial non expliqué au delà de l'estimation calculée par Logib. Toutefois, ces rapports ne fournissent aucune explication sur la manière dont cet argument a conduit au choix de la valeur de 5 % (Felfe *et al.*, 2015; Binggeli *et al.*, 2018; Trageser *et al.*, 2011). Selon Strub (2004), la détermination du seuil de tolérance de 5 % se fonde sur les résultats des analyses de régression pour étudier le différentiel des salaires dans l'ensemble de l'économie pour l'année 1998. A cette époque, la différence salariale non expliquée entre les femmes et les hommes résultant de la prise en compte uniquement des facteurs formation, ancienneté et expérience professionnelle potentielle était de 20 %. Lorsque les variables niveau de qualifications requises pour le poste et position professionnelle étaient prise en compte dans le modèle, la différence salariale non expliquée était de 15 %. Autrement dit, le modèle complet réduisait l'écart salarial d'environ 5 points de pourcentage par rapport au modèle simple. En conséquence, Strub (2004) propose l'utilisation du modèle simple avec un seuil de tolérance de 5 %. Dans le cas de la version actuelle de Logib, cet argument ne peut toutefois pas être retenu pour justifier le seuil de tolérance fixé à 5 %, puisque l'outil utilise le modèle complet. Ainsi, l'articulation entre l'argumentation développée par Strub (2004) pour le modèle simple et l'utilisation du seuil de tolérance pour le modèle complet est peu claire. A cet égard, Trageser *et al.* (2011) soulignent encore que le seuil de tolérance avait été choisi de façon généreuse en 2006 pour donner aux entités employeuses le temps de s'adapter.

Sur le plan historique, il est aussi important de rappeler que les versions antérieures de Logib à celle de sa mise en ligne en juillet 2020 indiquaient expressément que le seuil de tolérance de 5 % s'appliquait uniquement dans le cadre du contrôle du respect de l'égalité salariale dans les marchés

publics de la Confédération. Ainsi, le seuil de tolérance était associé à un cadre d'application restreint, lié au fait qu'un résultat non conforme pouvait déboucher sur des sanctions. Avec la popularité croissante de l'outil, les entités employeuses réalisant l'analyse de façon volontaire avec Logib ont repris ce seuil de tolérance sans tenir compte qu'il s'appliquait dans les marchés publics, de même que des entités privées ayant développé d'autres modèles de l'analyse de l'égalité salariale comportant souvent davantage de variables explicatives que Logib. Ainsi, le seuil de tolérance de 5 % s'est généralisé avec le temps, de sorte qu'il est maintenant pris en compte dans toutes les analyses, notamment celles rendues obligatoires pour les entités d'au moins cent personnes dans le cadre de la modification de la LEg qui est entrée en vigueur en juillet 2020.

En résumé, l'origine du seuil de tolérance reste peu claire et aucune justification scientifique, empirique ou juridique n'a été identifiée dans la littérature. Toutefois, il a été considéré comme acquis à travers le temps même si la LEg ne prévoit aucun seuil de tolérance en matière de respect de l'égalité salariale.

3.1.2 Appréciation du seuil de tolérance par les expert-e-s

Malgré l'absence de justification relative à l'utilisation et au niveau du seuil de tolérance de 5 %, les rapports mentionnent qu'il a "fait ses preuves" depuis son introduction (Felfe *et al.*, 2015; Binggeli *et al.*, 2018; Stern *et al.*, 2015; Trageser *et al.*, 2011). Ces rapports basent ces affirmations sur des enquêtes auprès d'expert-e-s dans le domaine ainsi qu'auprès d'entités employeuses. L'opinion des expert-e-s semble être que la valeur du seuil est arbitraire, mais que le seuil a "fait ses preuves". Selon Felfe *et al.* (2015) qui ont interviewé des personnes en charge du contrôle de l'égalité salariale, cette évaluation se base sur la *structuration pragmatique du modèle autour de cinq variables de base et l'application d'un seuil de tolérance*. Trois spécialistes cité-e-s indiquent aussi que l'acceptation du modèle auprès des entités employeuses est plus grande grâce au seuil de tolérance. D'autre part, Felfe *et al.* (2015) citent également des expert-e-s qui sont favorables à une réduction du seuil de tolérance.

Felfe *et al.* (2015) indique que les entités employeuses sont favorables au maintien d'un seuil, tout en étant en léger désaccord sur sa valeur. Un tiers des entités employeuses considère le seuil de 5 % comme étant trop haut, un autre tiers comme étant trop bas et un dernier tiers comme étant à peu près juste. Un détail intéressant est que les entités employeuses qui ont déjà eu des expériences avec Logib sont légèrement plus favorables à un seuil plus bas.

Plusieurs rapports justifient le seuil de tolérance par l'impossibilité d'expliquer parfaitement les salaires et, par conséquent, le seuil ne peut être réduit que si des facteurs objectifs supplémentaires sont inclus (Felfe *et al.*, 2015; Trageser *et al.*, 2011; Stern *et al.*, 2015; Graf, 2016). Trageser *et al.* (2011) mentionne également l'appréciation d'expert-e-s selon laquelle les entités employeuses se montrent plus collaboratives s'il y a un seuil de tolérance. En même temps, Trageser *et al.* (2011) a également fait une interview avec un-e expert-e en statistique, qui est en désaccord avec le seuil de tolérance et suggère sa suppression. Cette étude propose de baisser le seuil de tolérance à 2 % ou 3 % si le modèle est légèrement adapté et la possibilité d'analyses approfondies est introduite.

En résumé, la littérature se base essentiellement sur l'avis d'expert-e-s et d'entités employeuses et conclue que le seuil de tolérance a fait ses preuves. Relativement peu d'arguments spécifiques pour ou contre le seuil ont été identifiés. Certain-e-s auteur-e-s concluent néanmoins qu'une réduction du seuil est envisageable. Une idée assez répandue est que le seuil de tolérance est directement lié au niveau explicatif du modèle et que, par conséquent, le seuil de tolérance ne

peut être abaissé que si des facteurs supplémentaires sont introduits dans le modèle. Toutefois, les expert·e·s semblent être plus favorables au modèle actuel avec un seuil de tolérance (*statu quo*), comparé à un modèle plus complet (inclusion d'autres variables) sans seuil de tolérance.

Il est à relever que l'enquête auprès des expert·e·s et des entités employeuses s'inscrivait notamment dans le cadre du traitement du postulat Noser (14.3388) "Egalité salariale. Améliorer la pertinence des statistiques" ⁷ qui proposait de prendre en compte d'autres variables comme l'expérience en matière de direction, les formations continues suivies ou encore les connaissances linguistiques. C'est face à ce débat sur la potentielle prise en compte de variables supplémentaires que les expert·e·s ont jugé qu'il valait mieux garder le modèle en l'état plutôt que d'intégrer de nouvelles variables et baisser le seuil de tolérance. C'est pourquoi, il est légitime de s'interroger sur la position des expert·e·s relative à une possible réduction du seuil de tolérance environ quinze ans après le lancement de Logib en 2006.

Points clés 2 (Révision de littérature).

- L'origine du seuil de tolérance ainsi que sa valeur de 5 % restent peu claires et il n'existe pas de justification théorique, empirique ou juridique relative au choix de cette valeur.
- Selon l'avis des expert·e·s, le seuil de tolérance a fait ses preuves. Ces expert·e·s proposent de considérer un seuil plus bas si d'autres facteurs objectifs et non discriminatoires sont pris en compte.

3.2 Importance du seuil de tolérance pour les contrôles d'égalité salariale

Le seuil de tolérance joue un rôle déterminant pour définir quelles sont les entités employeuses qui respectent l'égalité salariale dans le cadre des contrôles menés au sein des marchés publics et des subventions, car il introduit une catégorie d'entités employeuses, considérées comme respectant l'égalité salariale, tout en ayant un effet lié au sexe différent de zéro. En modifiant le seuil de tolérance, le nombre d'entités employeuses dans cette catégorie change et, par conséquent, le nombre d'entités considérées comme conformes ⁸.

Dans le canton de Vaud, sur les onze premiers contrôles lancés par la CoMPS en 2020, quatre ont un résultat significativement supérieur à 0 % sur le plan statistique et zéro ont un résultat significativement supérieur au seuil de tolérance de 5 %.

Au niveau fédéral, le tableau 1 présente les résultats de 171 contrôles du respect de l'égalité salariale menés dans les marchés publics de la Confédération (état au 23.06.2020) selon les trois conclusions possibles de l'analyse menée avec Logib.

On constate que la majorité des entités employeuses ont un écart significativement supérieur à zéro (56 %), mais que seulement 9 % dépassent le seuil de tolérance de façon statistiquement significative.

7. Voir <https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaefte?AffairId=20143388>

8. Il est important de rappeler encore une fois à ce stade du rapport que l'existence de discriminations salariales collectives ou individuelles au sens de la loi sur l'égalité (LEg) n'est pas l'objet du contrôle. Par conséquent, le résultat du contrôle ne peut pas être utilisé par comme preuve d'une absence globale de discrimination salariale au sens de l'art. 3 al. 2 LEg.

TABLEAU 1 – DÉFINITION ACTUELLE DE CONFORMITÉ

Condition	Conclusion	Conforme	Proportion
 $\beta_{\text{Sexe}} = 0$	Pas d'effet lié au sexe	Oui	44 %
 $0 < \beta_{\text{Sexe}} \leq 5\%$	Existence d'un effet lié au sexe	Oui	47 %
 $ \beta_{\text{Sexe}} > 5\%$	Important effet lié au sexe : seuil de tolérance dépassé	Non	9 %

Note : La colonne **Conforme** indique si une entité employeuse est considérée comme respectant l'égalité salariale. La colonne **Proportion** fait référence à la proportion de cas pour chaque catégorie lors des contrôles achevés dans les marchés publics de la Confédération (état au 23.06.2020).

Cette variation importante de la proportion d'entités employeuses considérées comme respectant l'égalité salariale en fonction de la prise en compte du seuil de tolérance montre à quel point ce seuil de tolérance est crucial dans la définition du respect de l'égalité salariale, notamment dans le cadre des contrôles menés dans les marchés publics de la Confédération. A ce jour, aucune étude ne s'est intéressée à la manière dont la proportion d'entités employeuses considérées comme respectant l'égalité salariale au sens de Logib pourrait varier selon le niveau de seuil de tolérance. L'examen de cette relation est l'objet de la première question de recherche de ce mandat.

Étape de recherche 1.

Comment la proportion des entités employeuses qui respectent l'égalité salariale entre femmes et hommes au sens de Logib change-t-elle en modifiant le niveau du seuil de tolérance ?

3.3 Lien entre statistique nationale et modèle d'analyse standard de la Confédération

L'étude spéciale de l'Enquête suisse sur la structure des salaires consacrée à l'égalité salariale entre femmes et hommes est mandatée tous les deux ans par l'Office fédéral de la Statistique (OFS). Cette étude propose des analyses détaillées relatives à l'écart salarial en le décomposant en part expliquée et non expliquée pour toute une série de facteurs. L'approche de la statistique nationale diffère de celle du modèle d'analyse standard à bien des égards, que cela soit au niveau de l'échantillon, de la méthode d'analyse ou des variables prises en considération. Toutefois, nous pouvons formuler l'hypothèse que les variations importantes de la part non expliquée de l'écart salarial entre les modalités d'une variable au niveau de la statistique nationale pourraient aussi être observées au niveau du modèle d'analyses standard entre les entités employeuses associées aux modalités de cette même variable. Par exemple, la statistique nationale montre que la part non expliquée de l'écart salarial entre femmes et hommes est de 6.2 % dans le secteur public et de 8.2 % dans le secteur privé. Dès lors, on peut formuler l'hypothèse que davantage d'entités employeuses du secteur public que du secteur privé respectent l'égalité salariale au sens du modèle d'analyse standard de la Confédération.

Selon la LVLEg, la CoMPS désigne, en principe par tirage au sort, les entités adjudicatrices de marchés publics ou subventionnées qui seront contrôlées. Au niveau fédéral, les directives des contrôles du respect de l'égalité salariale dans les marchés publics de la Confédération indiquent que les propositions des entités à contrôler se font en fonction des risques ou de manière aléatoire (cf. BFEG (2021a, chapitre 4)). Le BFEG tient compte des informations relatives au résultat de l'analyse de l'égalité salariale conformément à la déclaration du soumissionnaire concernant

le respect de l'égalité salariale entre les femmes et les hommes dans les marchés publics de la Confédération; de la part de l'écart salarial par branche selon l'enquête sur la structure des salaires de l'OFS, de la part non expliquée des différences salariales entre les sexes par branche, région et taille d'entreprise selon l'analyse de décomposition des données ESS réalisée sur mandat de l'OFS; et enfin selon la fréquence, la durée et le volume des marchés publics de la Confédération.

Sur la base de ce qui précède, on constate qu'il peut exister un lien entre les risques de non respect de l'égalité salariale dans certains secteurs de l'économie selon l'étude spéciale de l'OFS, la sélection des entités dans le cadre des contrôles dans les marchés publics et les subventions, ainsi que potentiellement les résultats de ces contrôles. Il est donc intéressant de s'interroger sur l'impact de la variation du niveau de seuil de tolérance de Logib sur la proportion d'entités respectant l'égalité salariale en fonction des modalités de certaines variables de l'ESS. Cette interrogation fait écho à la seconde question de recherche du présent mandat.

L'étude [Kaiser et Möhr \(2021\)](#), mandatée par l'OFS pour analyser les différences salariales entre femmes et hommes sur la base de l'ESS 2018, montre qu'il existe des écarts salariaux non expliqués importants, notamment au sein des variables suivantes :

- **Secteurs privé et public** : Comme indiqué ci-dessus, l'écart salarial non expliqué est de 6.2 % dans le secteur public et de 8.2 % dans le secteur privé.
- **Taille des entités employeuses** : Dans l'ensemble de l'économie, les écarts salariaux non expliqués se situent entre 6.9 % dans les entités d'au moins 1000 personnes et 9.7 % dans les entités de moins de 20 personnes.
- **Branche économiques** : Dans l'ensemble de l'économie, l'écart salarial non expliqué varie de 3 % dans la branche "hébergement et restauration" à 16 % dans la branche "fabrication de textiles, industrie du cuir".

En plus des variables mentionnées ci-dessus, il semble approprié d'analyser si la proportion de femmes dans les entités employeuses joue un rôle sur leur conformité en matière d'égalité salariale selon le modèle d'analyse standard de la Confédération. D'une part, l'estimation de l'effet du sexe sur le salaire peut être plus précise pour les entités comptant approximativement autant de femmes que d'hommes, ce qui peut conduire à un effet plus fort d'une variation du seuil de tolérance pour ces entités employeuses. D'autre part, les entités employeuses avec une grande majorité de personnes du même sexe sont plus susceptibles d'opérer sur un marché typiquement associé à ce sexe, ce qui peut modifier les évaluations de l'importance de l'égalité salariale. Dans ce même ordre d'idée, l'examen de la relation entre la proportion de femmes parmi les cadres et le respect de l'égalité salariale au sens de l'outil Logib est intéressant à examiner.

Étape de recherche 2.

Comment le secteur, la taille des entités, la branche économique et la proportion de femmes (cadres) influencent la proportion d'entités conformes en matière d'égalité salariale selon Logib pour différents niveaux du seuil de tolérance ?

3.4 Un regard statistique sur la discussion autour du seuil de tolérance

La revue de la littérature présentée dans la section 3.1 a déjà permis d'aborder quelques arguments sur l'origine et le maintien du seuil de tolérance. Cette section introduit quelques réflexions statistiques qui permettent d'enrichir la réflexion concernant l'impact possible d'une variable du seuil de tolérance sur la conformité des entités employeuses au sens de Logib. Ces réflexions statistiques se basent sur trois points de consensus identifiés dans la littérature au sujet du contrôle du respect de l'égalité salariale entre femmes et hommes :

1. Proposer une approche **efficace** et **efficiente** qui réduit la charge administrative pour les entités employeuses tout en permettant de trouver un résultat correct.
2. **Éviter d'accuser une entité employeuse à tort** de non conformité en matière d'égalité salariale.
3. **Éviter de déclarer à tort une entité employeuse comme conforme** en matière d'égalité salariale.

Le premier point fait essentiellement référence à la volonté de réduire au maximum la charge administrative pour les entités employeuses lors des contrôles du respect de l'égalité salariale. Le fait d'utiliser les variables que les entités employeuses doivent fournir de toute façon à l'Enquête suisse sur la structure des salaires (ESS) y contribue fortement. Beaucoup d'arguments en faveur du maintien du modèle actuel plutôt que d'ajouter plus de complexité et de variables vont dans ce même sens. En ce qui concerne le seuil de tolérance, ce premier point n'est pas particulièrement important, car l'utilisation d'un seuil n'a pas d'effet direct sur la charge administrative pour faire l'analyse.

Les points 2 et 3 font référence à un dilemme bien connu en statistique inférentielle et présentent un lien direct avec le seuil de tolérance. Ce dilemme est expliqué de façon plus approfondie dans la sous-section suivante.

3.4.1 Risque d'une mauvaise conclusion

Le dilemme, introduit dans la section précédente se réfère au fait d'éviter de tirer des conclusions erronées par rapport à la valeur réelle du coefficient⁹. A ce propos, deux types d'erreur se distinguent :

1. Erreur type 1 : conclure qu'il existe un effet lié au sexe alors qu'il n'y en a pas en réalité.
2. Erreur type 2 : conclure qu'il n'existe pas d'effet lié au sexe alors qu'il y en a un en réalité.

En statistique inférentielle, la manière d'aborder ce dilemme se fonde sur l'utilisation de tests statistiques avec un niveau de confiance. La probabilité de commettre une erreur de type 1 est directement définie par le niveau de confiance. Logib utilise un niveau de confiance de 95 %, ce qui veut dire que la probabilité de faire une erreur de type 1 est de 5 %. En augmentant le niveau de confiance à 99 %, on pourrait réduire ce risque, mais cela entraînerait une augmentation des chances de commettre une erreur de type 2. La probabilité des deux types d'erreur est inversement

9. La valeur réelle du coefficient fait référence à la valeur réelle de ce qu'on cherche à estimer. L'estimation peut être imprécise et être légèrement différente de la valeur réelle. Un exemple illustratif est l'indicateur de vitesse dans une voiture qui indique une estimation de la vitesse, mais qui peut être légèrement différent à la vitesse effective.

proportionnelle. Les tests statistiques permettent alors de calculer avec précision les probabilités de faire une mauvaise conclusion dans l'une ou dans l'autre direction ¹⁰.

L'utilisation d'un seuil de tolérance est une autre manière de réduire le risque de conclure qu'il existe un effet lié au sexe alors qu'il n'y en a pas en réalité. Cette approche est utilisée quand on dispose d'une seule mesure et que cette dernière peut contenir une erreur de mesure. Un exemple d'un tel cas de figure donné dans la littérature est le contrôle de vitesse sur les routes à l'aide de radars ¹¹. Dans le cadre d'un contrôle de l'égalité salariale au sein d'une entité employeuse, les données analysées se fondent sur des dizaines, des centaines voire des milliers de personnes. En ce sens, la correction requise dans le cadre d'une mesure unique n'est pas nécessairement pertinente dans la situation où de nombreuses mesures sont disponibles.

En somme, l'objectif d'un intervalle de confiance et d'un seuil de tolérance est le même. Les deux définissent une plage de valeurs plausibles autour de la mesure. Au lieu de baser la conclusion uniquement sur une seule valeur de la mesure (ou estimation), on se base sur l'ensemble de cette plage de valeurs. La différence entre les deux concepts vient de leur définition. Alors qu'un intervalle de confiance est calculé à partir de la distribution des données utilisées lors d'une analyse statistique, un seuil de tolérance est souvent défini à partir des spécifications techniques de l'appareil de mesure (par exemple un radar vitesse). Utiliser les deux approches simultanément revient, par conséquent, à tenter de résoudre le même problème à deux reprises.

Dans le cadre de Logib, on utilise effectivement les deux approches, ce qui peut avoir des conséquences potentiellement problématiques, car cette solution introduit un grand risque de déclarer une entité employeuse comme conforme alors qu'en réalité elle ne l'est pas. Pour illustrer ce cas, imaginons qu'une entité employeuse décide de rémunérer tous ses employés de sexe masculin, toute chose étant égale par ailleurs, 6 % de plus que ses employées de sexe féminin. La valeur réelle de l'effet lié au sexe est alors de 6 %. Dans ce cas, le modèle Logib estime l'effet lié au sexe à 6 % et indique, par exemple, un intervalle de confiance de 4 % à 8 %. La conclusion d'un point de vue statistique serait que la valeur réelle se situe quelque part entre 4 % et 8 %, ce qui est clairement différent de zéro. Cette situation serait donc contraire à la loi, pourtant l'entité ne serait pas considérée comme non conforme sur le plan de l'égalité entre femmes et hommes par Logib puisque la différence salariale non expliquée ne serait pas significativement supérieure au seuil de tolérance de 5 % (la valeur de 5 % se situe à l'intérieur de l'intervalle de confiance).

S'il s'agit effectivement de vouloir tolérer des discrimination allant jusqu'à 5 %, l'utilisation des deux approches de façon simultanée (test statistique et seuil de tolérance) est justifiable. En revanche, si le seuil de tolérance est utilisé pour éviter des conclusions erronées comme décrit ci-dessus, il semble plus pertinent d'utiliser uniquement les méthodes statistiques bien établies et de supprimer le seuil de tolérance.

10. Une explication plus détaillée sur le fonctionnement des tests statistiques se trouve dans l'annexe B.1

11. Un exemple de ce cas de figure sont les contrôles de vitesse à l'aide d'un radar, également cité par [Trageser et al. \(2011\)](#). La mesure de vitesse à l'aide d'un radar n'est pas précise car la vitesse réelle peut être légèrement inférieure ou supérieure à la valeur mesurée. Par exemple, si le radar a une marge d'erreur de +/- 3 km/h, la mesure peut indiquer 82 km/h alors que la voiture roulait à 80 km/h et il serait faux de conclure que la voiture roulait trop vite. En ajoutant un seuil de tolérance de 5 km/h, ce risque est évité, car une mesure affichée de plus de 85 km/h indique un excès de vitesse même si la mesure a surestimé la vitesse réelle de 3 km/h.

3.4.2 Lien entre pouvoir explicatif du modèle et l'écart estimé

La littérature sur le seuil de tolérance (section 3.1) montre que plusieurs expert·e-s et certaines entités employeuses considèrent qu'il existe un lien entre le pouvoir explicatif du modèle et l'écart salarial estimé. Cet argument amène à l'appréciation qu'une possible réduction du seuil de tolérance doit être accompagnée par l'inclusion de variables supplémentaires. D'un point de vue économétrique, cette relation présumée entre le pouvoir explicatif du modèle et l'écart estimé, n'est pas nécessairement vraie pour les deux raisons suivantes :

1. La part non expliquée du modèle (résidus) n'est pas automatiquement attribuée à un effet lié au sexe. De ce fait, réduire la part non expliquée n'influence pas nécessairement l'effet estimé.
2. L'objectif de Logib n'est pas de maximiser le pouvoir explicatif du modèle, mais d'identifier de la meilleure façon possible l'effet lié au sexe selon la définition conceptuelle ou légale.

Contrairement à une conception assez répandue, l'objectif en micro-économétrie n'est normalement pas de trouver un modèle qui explique le mieux possible la variable dépendante, mais plutôt d'estimer l'effet d'intérêt avec la meilleure précision possible et sans biais. Un modèle avec un coefficient de détermination (R^2) plus élevé n'est pas nécessairement préférable à un modèle plus simple qui, avec moins de variables, n'atteint pas le même R^2 , mais dont chacune des variables contribue proportionnellement d'une manière importante et d'un point de vue théorique d'une manière pertinente à l'explication de la variable dépendante¹². L'analyse de l'égalité des salaires entre les femmes et les hommes n'y fait pas exception. Voir l'annexe B.3 pour une explication plus détaillée sur le choix de variables et comment ce choix peut modifier la nature de l'effet estimé par un modèle.

A titre illustratif, il est souvent relevé que le modèle d'analyse standard de la Confédération ne tient pas compte de la performance des employé·e-s. Cette variable est souvent mise en lien avec la part variable du salaire, élément intégré à l'analyse Logib. Pour autant que l'on puisse imaginer une mesure de la performance applicable à tout le tissu économique et toutes les fonctions, il est vraisemblable que son intégration dans Logib permette de mieux comprendre la rémunération. Statistiquement, cela se traduira par un coefficient de détermination plus élevé. Toutefois, cette variable de performance n'aura un impact sur le résultat de l'analyse que si elle est, toutes choses étant égales par ailleurs, liée au sexe des employé·e-s. Ainsi, l'ajout de cette variable peut ne pas avoir d'effet sur le résultat de l'analyse, réduire l'écart salarial non expliqué ou l'augmenter. Or, si la variable de performance devait être corrélée au sexe lorsque l'on tient compte des différences existantes entre les femmes et les hommes sur toutes les autres variables du modèle, alors se poserait la question juridique de savoir pourquoi les femmes seraient systématiquement plus performantes ou moins performantes que les hommes au sein de l'entité. On comprend ici que l'ajout de variables supplémentaires est relativement complexe et ne saurait être un argument déterminant dans la justification du seuil de tolérance. Pour une réflexion détaillée relative à l'analyse juridique du potentiel de discrimination des variables explicatives servant à mesurer l'égalité salariale entre les femmes et les hommes, se référer à [Pärli \(2019\)](#).

Afin d'apporter des résultats empiriques à cette discussion, le lien entre le pouvoir explicatif du

12. Le coefficient de détermination (R^2) est une mesure qui indique la proportion de la variation de la variable dépendante qui est expliquée par le modèle. Ce coefficient varie entre 0 et 1 et peut s'exprimer sous la forme d'un pourcentage. Par exemple, une valeur de 0.7, dans le cas d'une analyse Logib, indique que 70 % de la variation du salaire est expliqué par les variables incluses dans le modèle. Une réaction assez naturelle - mais pas nécessairement correcte - est de croire qu'une valeur plus élevée est toujours préférable.

modèle Logib et la proportion d'entités employeuses respectant l'égalité salariale sera examiné dans ce rapport. Par ailleurs, nous étudierons un exemple lié à l'impact de l'ajout de variables supplémentaires à un modèle d'analyse de régression sur le paramètre de la variable sexe. Cette analyse, basée sur les données de l'EES, permettra de fournir des informations utiles sur la nécessité de considérer l'inclusion de variables supplémentaires en cas de réduction du seuil de tolérance.

Étape de recherche 3a.

Quel est le lien entre le pouvoir explicatif (R^2) du modèle Logib et la proportion d'entités employeuses qui respectent l'égalité salariale ? Comment ce lien change-t-il en fonction du niveau du seuil de tolérance ?

3.4.3 Erreurs standard robustes

Pour conclure ce regard statistique sur le seuil de tolérance afin d'identifier des informations qui devraient être prises en compte par la CoMPS lors d'une réflexion sur le seuil de tolérance, il convient de reprendre un sujet traité dans le rapport de [Trageser *et al.* \(2011\)](#) : l'utilisation d'erreurs standard robustes dans l'analyse de régression. Le modèle d'analyse standard de la Confédération se fonde sur une régression linéaire semi-logarithmique calculée avec l'estimateur des moindres carrés ordinaires (MCO). L'analyse de régression se base sur cinq hypothèses, aussi connues comme les conditions de Gauss-Markov ([Wooldridge, 2006](#)). Une des conditions est l'homoscédasticité des résidus, c'est-à-dire que les résidus ont la même distribution tout au long de la ligne de régression (voir l'annexe B.2 pour plus de détails). Si cette condition n'est pas satisfaite - ce qui est souvent le cas quand on utilise des données au niveau des individus - les erreurs standard sont erronées et les conclusions sur la significativité des écarts salariaux non-valables ([Cameron et Trivedi, 2005](#); [Wooldridge, 2006](#); [Greene, 2005](#)). Si les erreurs standards ne sont pas valables, l'intervalle de confiance qui en résulte peut être trop large ou trop petit, ce qui peut amener à une conclusion erronée par rapport à la significativité de l'effet.

Il existe des tests statistiques simples pour identifier si la condition est satisfaite ou pas, ainsi qu'une version robuste des erreurs standards qui peut se calculer très facilement afin de corriger un problème d'hétéroscédasticité ([White, 1980](#); [Wooldridge, 2006](#)). L'utilisation de ces erreurs standard robustes pour des analyses économétriques avec des données individuelles est une pratique tout à fait ordinaire et tous les logiciels statistiques peuvent les calculer sans problèmes.

Les données de l'ESS fournissent des informations détaillées sur la nécessité de disposer d'erreurs standard robustes, ce qui permet de définir la quatrième étape de recherche comme suit :

Étape de recherche 3b.

Est-il nécessaire de considérer l'utilisation d'erreurs standard robustes lors du contrôle du respect de l'égalité salariale avec Logib ? En particulier :

1. Dans quel pourcentage d'analyses, l'hypothèse d'homoscédasticité n'est pas satisfaite ?
2. Combien de conclusions concernant la conformité avec l'égalité salariale dépendent du choix des erreurs standards ?
3. Est-ce que le nombre de conclusions qui dépendent du choix des erreurs standards change en fonction du seuil de tolérance ?

Une analyse avec l'échantillon complet de la ESS permettra d'obtenir une appréciation générale du besoin d'utilisation d'erreurs standard robustes et d'analyser le pourcentage de conclusions qui peuvent changer en fonction du modèle choisi.

Points clés 3 (Regard statistique).

- D'un point de vue statistique, l'utilisation d'un seuil de tolérance n'est pas nécessaire si on se base sur les tests statistiques de significativité des effets. Ces tests existent pour éviter de conclure à tort à l'existence d'un effet quand il n'y a pas réellement.
- La part non expliquée d'un modèle n'est pas automatiquement attribuée à l'écart salarial. Il n'est donc pas nécessaire de chercher à maximiser le pouvoir explicatif du modèle.
- L'objectif de Logib n'est pas d'expliquer le mieux possible les salaires au sein d'une entité employeuse, mais de fournir une estimation précise et non biaisée de l'écart salarial lié au sexe.
- Il serait pertinent de considérer l'utilisation de tests d'homoscédasticité et, le cas échéant, d'erreurs standards robustes lors de contrôles.

4 Données et méthodologie

4.1 Source de données

Pour effectuer les étapes de recherche définies dans la section 3, l'analyse empirique se base sur les données issues de l'Enquête suisse sur la structure des salaires 2018 (ESS)¹³. L'ESS se présente sous la forme d'un questionnaire, envoyé par l'OFS tous les deux ans depuis 1994 à quelques 35'000 entreprises actives en Suisse. Elle permet de décrire régulièrement, à partir de données représentatives, la structure des salaires dans l'ensemble des branches économiques des secteurs secondaire et tertiaire. La pertinence du recours aux données de l'ESS pour le présent mandat découle du fait que les variables utilisées dans le modèle d'analyse de la Confédération proviennent à l'origine de l'ESS, même si certaines différences sont apparues avec le temps. Ainsi, il est possible de calculer le modèle Logib pour les entités employeuses participant à l'ESS.

4.2 Choix des variables et traitement des données

Du fait de la similarité entre les données de l'ESS et celles requises pour l'analyse Logib, la préparation des données se limite à de simples recodages et ajustements. Le tableau 2 présente les variables utilisées dans l'étude, ainsi que leur équivalent dans la notation de Logib et le traitement nécessaire pour leur utilisation pour l'analyse.

13. Office fédéral de la statistique, Enquête suisse sur la structure des salaires 2018. <https://www.bfs.admin.ch/bfs/fr/home/statistiques/travail-remuneration/enquetes/ess.html>

TABLEAU 2 – VARIABLES POUR L'ESTIMATION DU MODÈLE LOGIB

Variable Logib	Variable ESS	Traitement
$\ln(\text{salair}_i)$	mbls	Application du logarithme naturel. Il s'agit du salaire brut standardisé à 4 1/3 semaines à 40 heures, avec les indemnités légales, ainsi qu'un 1/12 du 13ème salaire et des paiement spéciaux. Cette définition correspond à la définition utilisée pour Logib.
Train_i	ausbild	Conversion du niveau de formation en années de formation selon la définition de Logib (BFEG, 2021b, 2.2).
Exp_i	alter	Calcul à partir de l'âge (variable alter) et la variable Train_i : $\text{Exp}_i = \max(\text{alter} - \text{Train}_i - 6, 0)$.
Tnr_i	dienstja	Aucun ajustement nécessaire.
Pos_i	berufst	Aucun ajustement nécessaire. Pour l'estimation du modèle, cette variable catégorielle a été convertie en un ensemble de 4 variables muettes.
Skl_i	va_ps07	Aucun ajustement nécessaire. Pour l'estimation du modèle, cette variable catégorielle a été convertie en un ensemble de 3 variables muettes. La variable va_ps07 qui correspond au niveau de compétences dans Logib est automatiquement déduite par l'OFS du groupe de professions selon la nomenclature CIP-08 (<i>Classification internationale type des professions</i> , angl. ISCO-08). Le groupe de professions est quant à lui déterminé à l'aide des informations sur la profession exercée (colonne 10 du questionnaire ESS). Un code CIP (ISCO_2, 2 digit) est donné à la profession exercée sur la base de la banque de données des professions de l'OFS. De là, on attribue quatre catégories (niveaux de compétences). Dédutions : ISCO_2 = 01 : va_ps07 = 4 ; ISCO_2 = 02 : va_ps07 = 2 ; ISCO_2 = 03 : va_ps07 = 1 ; ISCO_2 = 10-26 : va_ps07 = 4 ; ISCO_2 = 30-35 : va_ps07 = 3 ; ISCO_2 = 40-83 : va_ps07 = 2 ; ISCO_2 = 90-96 : va_ps07 = 1 (Source : <i>Codebook, Enquête suisse sur la structure des salaires 2012-2018</i>). Cependant, dans le cadre des contrôles, ce sont les entreprises qui recodent leur système de fonction selon les indications de Logib. Ce classement peut se faire <i>sur la base du profil du poste et du cahier des charges (tâches) ainsi que du profil des compétences (aptitudes requises)</i> . L'homogénéité (à l'intérieur de l'entreprise) et la pertinence de ce recodage sont ensuite examinées par les expert·e·s chargé·e·s du contrôle.
Sexe_i	geschle	Recodage selon les spécifications du BFEG (2021b) : (1 = femme, 0 = homme)

D'autres variables ont été utilisées pour identifier les caractéristiques des entités employeuses ainsi que pour obtenir les estimations pondérées. Le tableau 3 présente ces variables avec une description de leur utilisation.

TABLEAU 3 – VARIABLES SUPPLÉMENTAIRES UTILISÉES

Variable ESS	Définition	Utilisation
entid_n	Identifiant (anonymisé) de chaque entité employeuse	Possibilité d'associer chaque employé-e à une entité employeuse.
privoef	Indicateur concernant le secteur privé et public	Division de l'échantillon pour faire des analyses par groupe d'entités employeuses selon secteur privé et public.
k_untgroe	Taille de l'entité employeuse (variable catégorielle)	D'une part pour limiter l'échantillon et d'autre part pour effectuer l'analyse par taille de l'entité employeuse.
noga08_ent_2	Code NOGA à deux chiffres	Effectuer l'analyse par branche économique.
gewicht	Facteur d'expansion	Ce facteur est utilisé pour obtenir les estimations pondérées lorsqu'il s'agit de présenter des résultats au niveau de toutes les entités employeuses.

4.3 Sélection de l'échantillon

Le module 1 de Logib qui s'utilise lors du contrôle de l'égalité salariale peut être appliqué pour des entités employeuses à partir de 50 employé-e-s. Pour cette raison, l'échantillon a été limité aux entités employeuses pour lesquelles au moins 50 registres complets¹⁴ sont observés dans la base de données. Afin d'éviter une estimation de l'effet lié au sexe qui se base sur un nombre très limité de personnes, seulement les entités employeuses avec au moins 20 employé-e-s de chaque sexe sont considérées dans l'échantillon.

Au total, un échantillon de 2'845 entités employeuses a été retenu¹⁵. L'annexe A présente quelques statistiques descriptives de cet échantillon.

4.4 Modèles économétriques

4.4.1 Approche générale

La logique générale de l'approche consiste à estimer le modèle Logib pour chaque entité employeuse de l'échantillon et d'extraire le coefficient lié au sexe (β_{sexe}), l'erreur standard associée et le coefficient de détermination (R^2) de la régression. Le coefficient et l'erreur standard permettent par la suite de calculer la proportion d'entités employeuses respectant l'égalité salariale

14. Un registre est considéré comme complet lorsqu'il fournit l'information de toutes les variables pertinentes pour l'estimation de Logib.

15. La réduction de l'échantillon, par rapport à l'échantillon complet de l'ESS de 36125 entités employeuses, s'explique par la condition d'avoir au moins 50 employé-e-s (87 % de la réduction), la condition d'avoir au moins 20 femmes et 20 hommes (5 % de la réduction) et par des valeurs manquantes (8 % de la réduction).

selon différent niveau du seuil de tolérance et pour différents groupes d'entités employeuses. Ces proportions sont toujours calculées en utilisant les poids d'échantillonnage (variable *gewicht*) afin d'obtenir des résultats représentatifs pour l'ensemble de l'économie.

4.4.2 Modèle standard

L'équation (1) présentée dans la section 2 est utilisée pour estimer le modèle Logib standard. Afin d'éviter des erreurs de programmation du modèle dans le logiciel statistique, cette analyse a été effectuée avec deux codes indépendants et deux logiciels indépendants.

4.4.3 Modèle simple

Pour fournir des informations par rapport au lien entre le pouvoir explicatif du modèle et l'effet lié au sexe (voir étape de recherche 3a), un modèle plus simple de Logib a aussi été estimé. Ce modèle n'inclut pas la position hiérarchique ni le niveau de compétence, et il s'écrit de la manière suivante :

$$\ln(\text{salaire}_i) = \beta_0 + \beta_{Train} \cdot Train_i + \beta_{Exp} \cdot Exp_i + \beta_{Exp^2} \cdot Exp_i^2 + \beta_{Tnr} \cdot Tnr_i + \beta_{Sexe} \cdot Sexe_i + \varepsilon_i \quad (2)$$

Ce modèle simple était inclus dans Logib depuis sa création et jusqu'à sa mise en ligne en juillet 2020. Il était aussi appelé régression QP pour qualifications personnelles. L'idée de ce modèle était notamment de pouvoir comparer son résultat avec celui du modèle intégrant les variables niveau de compétence et position professionnelle (modèle standard). Une diminution importante du coefficient lié au sexe entre le premier et le second modèle s'expliquait par le fait que les variables niveau de compétence et position professionnelle étaient corrélées au sexe. La différence de résultat entre les deux modèles pouvait être interprétée comme une forme de ségrégation au sein de l'entité, liée potentiellement à une discrimination à l'embauche ou à la promotion.

4.4.4 Modèle robuste

Dans le but d'apporter des informations concernant l'utilisation des erreurs standard robustes (voir étape de recherche 3b), une version robuste du modèle standard a été estimée. Pour chaque entité employeuse, un test d'homoscédasticité a été effectué et, en cas de rejet de l'hypothèse d'homoscédasticité, les erreurs standard robustes ont été calculées en utilisant l'estimateur de White (1980).

5 Résultats

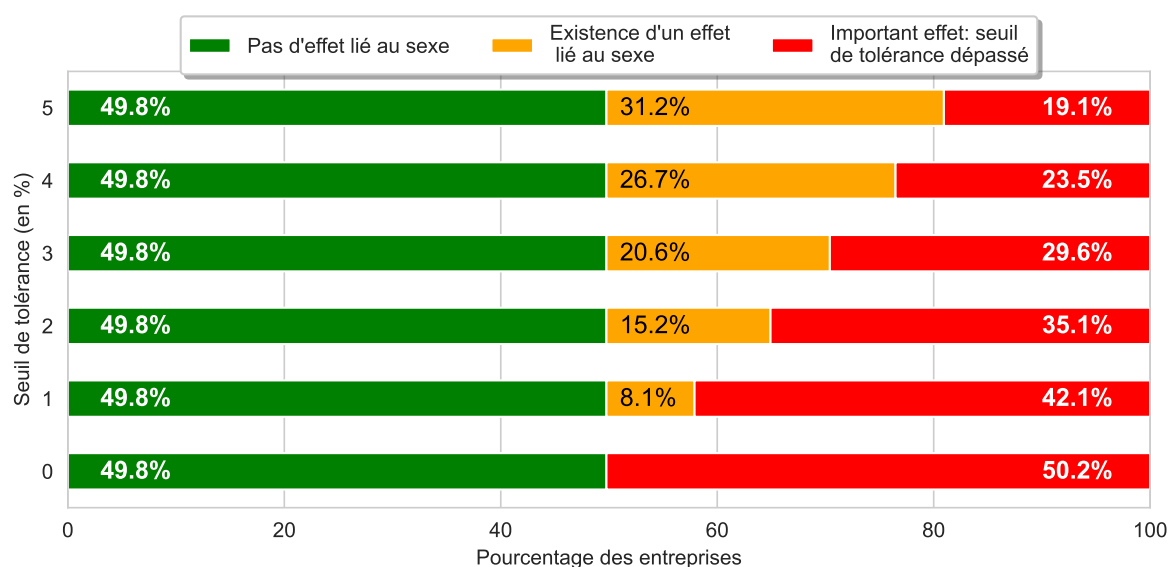
Cette section est dédiée aux résultats de l'analyse empirique en suivant l'ordre des étapes de recherche définies dans la section 3. Dans un premier temps, la section 5.1 révèle la proportion d'entités employeuses qui respectent l'égalité salariale en fonction de différents niveaux du seuil de tolérance (étape de recherche 1). Ces résultats sont présentés pour l'ensemble de l'économie. La section 5.2 expose une analyse similaire, mais en fonction de différentes caractéristiques des entités employeuses selon l'étape de recherche 2.

La section 5.3 est consacrée à la discussion du lien entre le pouvoir explicatif du modèle et l'effet lié au sexe (étape de recherche 3a). Enfin, la section 5.4 montre une analyse concernant l'utilisation des erreurs standard robustes (étape de recherche 3b).

5.1 Variation des résultats Logib par niveau de seuil de tolérance

La figure 2 montre la proportion d'entités employeuses respectant l'égalité salariale au sens du modèle Logib en fonction de niveaux de seuil de tolérance variant de 0 % à 5 %. Chaque barre horizontale est divisée en trois couleurs représentant les trois différentes conclusions que Logib fournit. Ces résultats permettent de répondre à la première question de recherche de ce mandat.

FIGURE 2 – Résultats Logib par niveau de seuil de tolérance



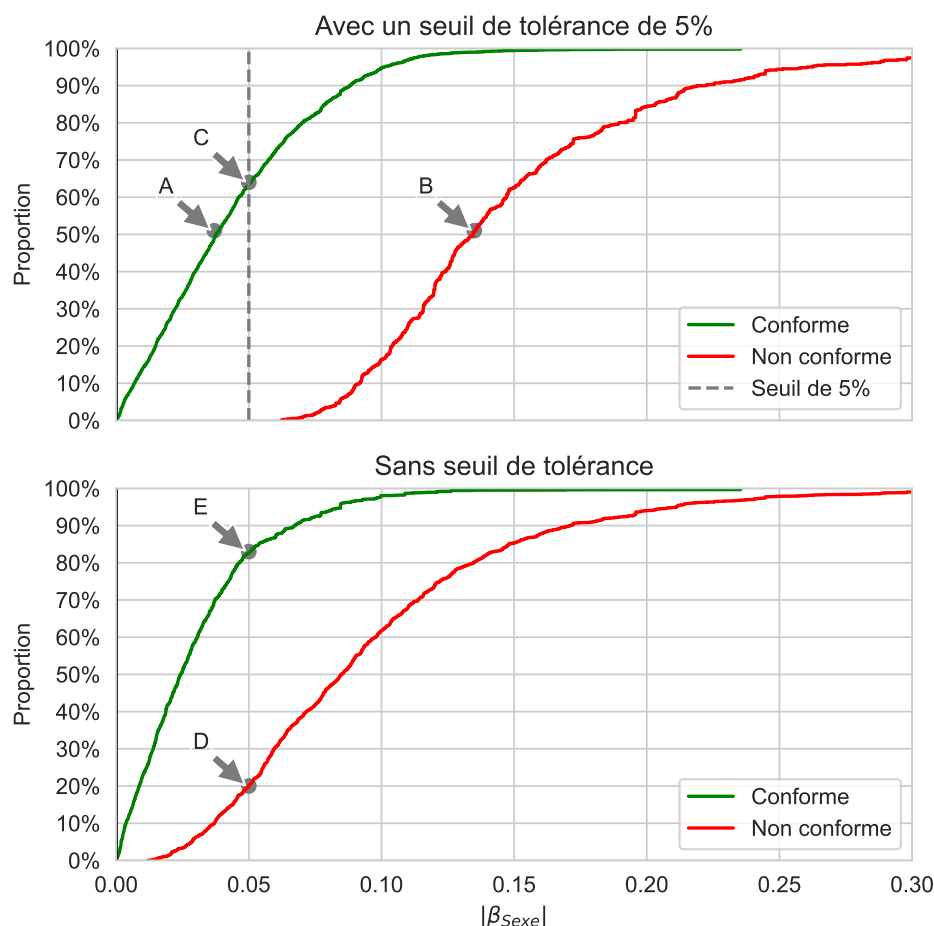
En prenant un seuil de tolérance de 5 %, 80,9 % (49,8 % + 31,2 %) des entités employeuses respectent l'égalité salariale entre femmes et hommes au sens de l'instrument Logib de sorte que seules 19,1 % sont considérées comme non conformes sur le plan de l'égalité salariale. On constate que 31,2 % des entités bénéficient directement du seuil, car ces dernières présentent une différence salariale non expliquée significativement supérieure à zéro mais non significativement supérieure à 5 %.

Grâce à ce graphique, on constate que la proportion d'entités employeuses bénéficiant du seuil de tolérance diminue à mesure que le seuil se réduit. Pour un seuil de 2 %, cette proportion

d'entités dans le résultat orange se situe à 15,2 % et celle dans le résultat rouge à 35,1 %. Si le seuil de tolérance était supprimé, 50,2 % des entités employeuses ne respecteraient pas l'égalité salariale au sens de Logib.

En complément à ces résultats, la figure 3 montre la distribution de la valeur absolue du coefficient de la variable sexe estimé pour le seuil de tolérance de 5 % (en haut) et sans seuil de tolérance (en bas) en fonction de la conclusion de l'analyse Logib (conforme et non conforme).

FIGURE 3 – Distribution du coefficient estimé de la variable sexe (valeur absolue)



La ligne verte décrit la distribution des entités employeuses conformes (vert + orange) et la ligne rouge représente la distribution des entités employeuses considérées non conformes par Logib.

L'axe vertical indique la proportion cumulée d'entités employeuses et l'axe horizontal fait référence à la valeur absolue du coefficient β_{sexe} . A titre d'exemple, le point A se situe au niveau de la médiane des entités employeuses conformes. Reporté sur l'axe horizontal, ce point A est associé à un coefficient d'environ 0.037. Cela signifie que, lorsque l'on tient compte d'un seuil de tolérance de 5 %, 50 % des entités conformes ont une différence salariale non expliquée inférieure à environ 3.7 % tandis que l'autre moitié des entités conformes a une différence salariale supérieure à 3.7 %. Pour les entités considérées comme non conformes, cette médiane (point B)

se situe à 0,135. Cela signifie que lorsque l'on tient compte d'un seuil de tolérance de 5 %, 50 % des entités non conformes ont une différence salariale non expliquée inférieure à environ 13.5 % alors que l'autre moitié a une différence salariale supérieure à 13.5 %. Enfin, le point C montre que, parmi les entités conformes, près de 40 % ont effectivement une différence salariale non expliquée supérieure à 5 %. Ce résultat vient du fait que la différence salariale non expliquée doit être significativement supérieure sur le plan statistique au seuil de 5 % pour qu'une entité employeuse soit considérée non conforme.

Sur la partie inférieure de la figure sont présentées les distributions des entités conformes et non conformes lorsqu'aucun seuil de tolérance n'est pris en compte. On observe que les deux courbes sont considérablement plus vers la gauche en comparaison aux courbes de la partie supérieure de la figure. Le point D sur la courbe rouge montre que seules 20 % des entités non conformes ont une différence salariale non expliquée inférieure à 5 %. Autrement dit, même si aucun seuil de tolérance n'est appliqué, 80 % des entités employeuses considérées comme non conformes affichent un effet lié au sexe supérieur à 5 %. La valeur la plus faible de cette courbe rouge est environ 0.015, soit un écart salarial non expliqué de 1.5 %. C'est le point à partir duquel l'écart salarial non expliqué devient significativement différent de zéro sur le plan statistique au sein de l'échantillon examiné. Pour terminer, le point E montre que, même sans seuil de tolérance, un effet supérieur à 5 % n'est pas toujours synonyme d'une non conformité. Ici, autour de 20 % des entités ont un coefficient supérieur à 0,05, mais ce dernier n'est pas significativement différent de zéro.

Points clés 4 (Résultat de l'étape de recherche 1).

La proportion des entités employeuses qui respectent l'égalité salariale diminue de 80,9 % pour un seuil de tolérance de 5 %, à 65,0 % pour un seuil de 2 % et 49,8 % sans seuil de tolérance.

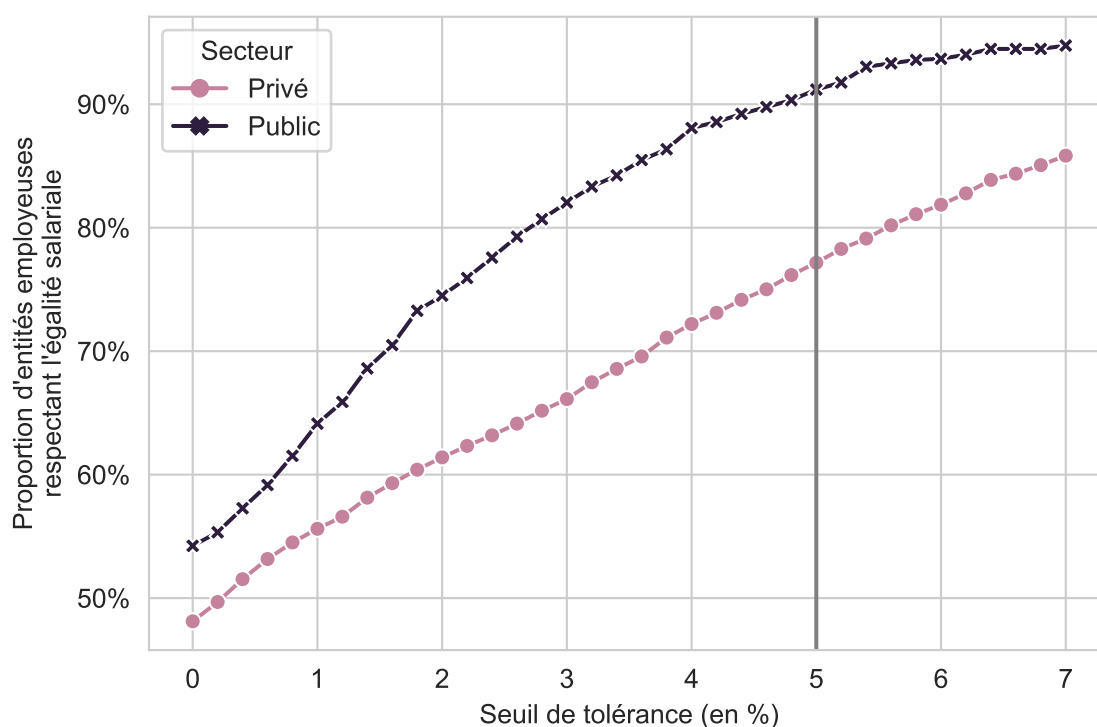
5.2 Résultats Logib par niveau de seuil de tolérance ventilés par variable

Cette section présente une série de figures qui montrent comment la proportion d'entités employeuses respectant l'égalité salariale au sens de Logib change par niveau de seuil de tolérance pour les variables secteur (privé et public), taille de l'entité employeuse, branche économique et proportion de femmes au sein de l'entité.

5.2.1 Analyse par secteur

La figure 4 présente la proportion d'entités conformes par niveau de seuil de tolérance en fonction du secteur (privé/public).

FIGURE 4 – Conformité par niveau de tolérance pour les secteurs public et privé



Pour un seuil de tolérance de (5 %), une grande différence entre les deux secteurs est observée. Dans le secteur public, le taux de conformité au sens de Logib est de 91,2 %, alors qu'il atteint un niveau de 77,2 % dans le secteur privé. Ce résultat fait écho à ceux des études de [Kaiser et Lobsiger \(2019\)](#) et [Kaiser et Möhr \(2021\)](#) qui ont identifié avec une autre méthodologie des écarts non expliqués plus importants dans le secteur privé que dans le secteur public. Il est aussi cohérent avec les résultats d'une récente étude de l'Université de Bâle ([Felder et Wunsch, 2021](#)), basée sur les données de l'ESS 2018, qui montre que 10 % des entités du secteur public et 24 % des entités du secteur privé ont une différence salariale non expliquée significativement supérieure au seuil de tolérance de 5 %. A noter que l'échantillon de l'université diffère légèrement du nôtre puisque seules les entités d'au moins 60 personnes sont prises en compte.

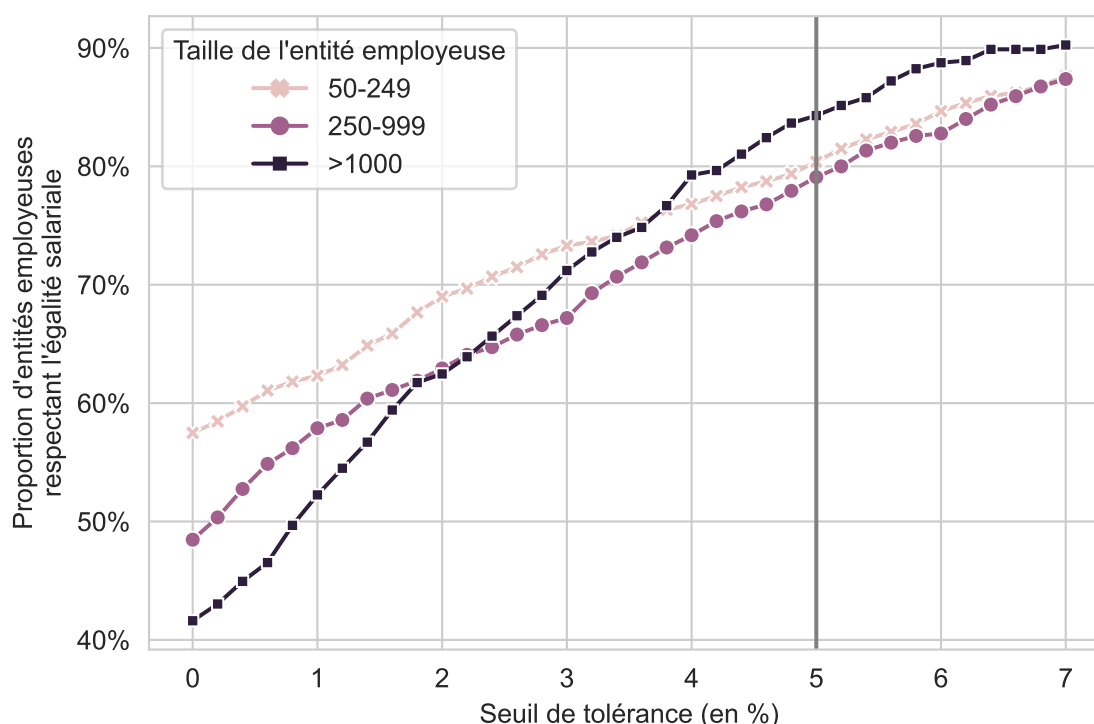
Il est intéressant de noter que cette différence entre les secteurs diminue considérablement à

mesure que le seuil de tolérance est réduit. En éliminant le seuil, les proportions de conformité sont respectivement de 54,2 % et 48,1 % environ dans le secteur public et le secteur privé. La différence entre ces deux courbes se réduit d'environ 15 points de pourcentage avec un seuil de tolérance de 5 % à environ 7 points de pourcentage sans seuil de tolérance.

5.2.2 Analyse par taille d'entité employeuse

La figure 5 présente la proportion d'entités conformes par niveau de seuil de tolérance en fonction de la taille des entités employeuses.

FIGURE 5 – Conformité par niveau de tolérance selon la taille de l'entité employeuse

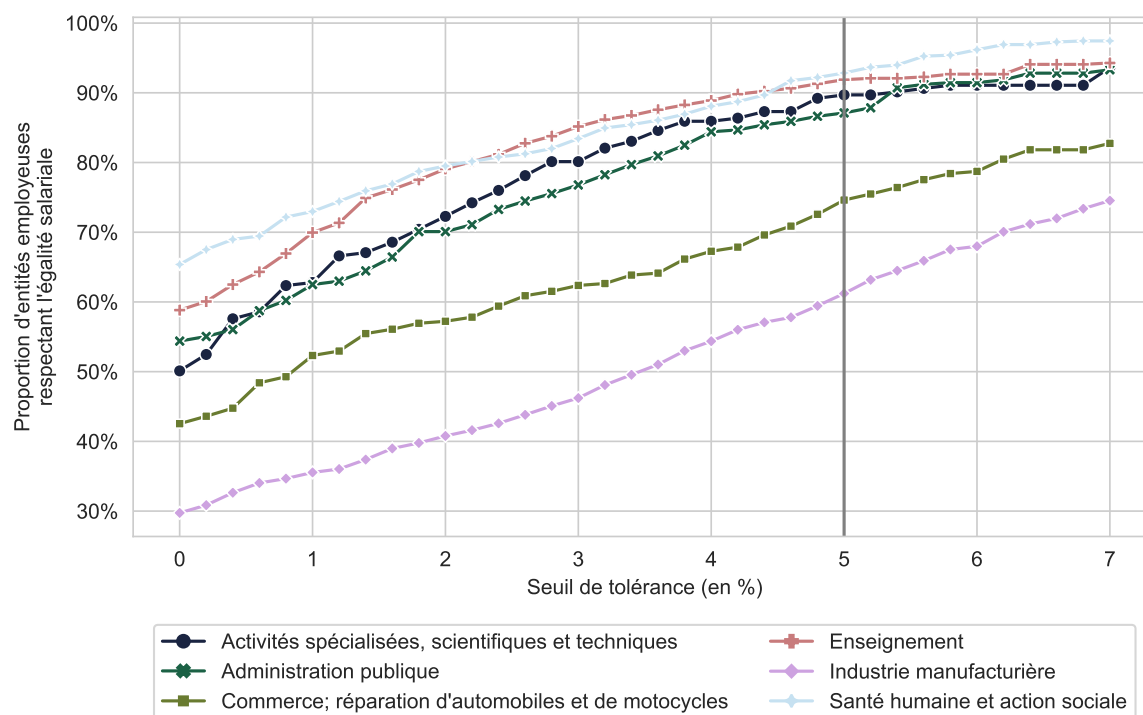


Pour un seuil de tolérance de 5 %, les différences de proportion d'entités employeuses conformes selon Logib sont faibles sur la base de leur taille, même si la proportion de conformité est légèrement plus élevée pour celles employant plus de 1000 personnes. Ce résultat est en accord avec les conclusions de [Kaiser et Lobsiger \(2019\)](#) et [Kaiser et Möhr \(2021\)](#). En l'absence de seuil de tolérance, la proportion d'entités conformes selon Logib est la plus élevée parmi les plus petites entités et la plus faible parmi les plus grandes. Une explication de ce résultat est la meilleure précision des estimations pour les entités les plus grandes, ce qui permet d'identifier des effets plus petits comme étant statistiquement significatifs. En revanche, l'imprécision des estimations pour les petites entités leur permet d'avoir des effets légèrement plus élevés tout en restant statistiquement non significatifs.

5.2.3 Analyse par branche économique

La figure 6 présente la proportion d'entités conformes par niveau de seuil de tolérance en fonction de la branche économique. Seules les branches ayant au moins 150 entités dans notre base de données ont été prises en compte pour ce graphique.

FIGURE 6 – Conformité par niveau de tolérance selon la branche économique (NOGA)

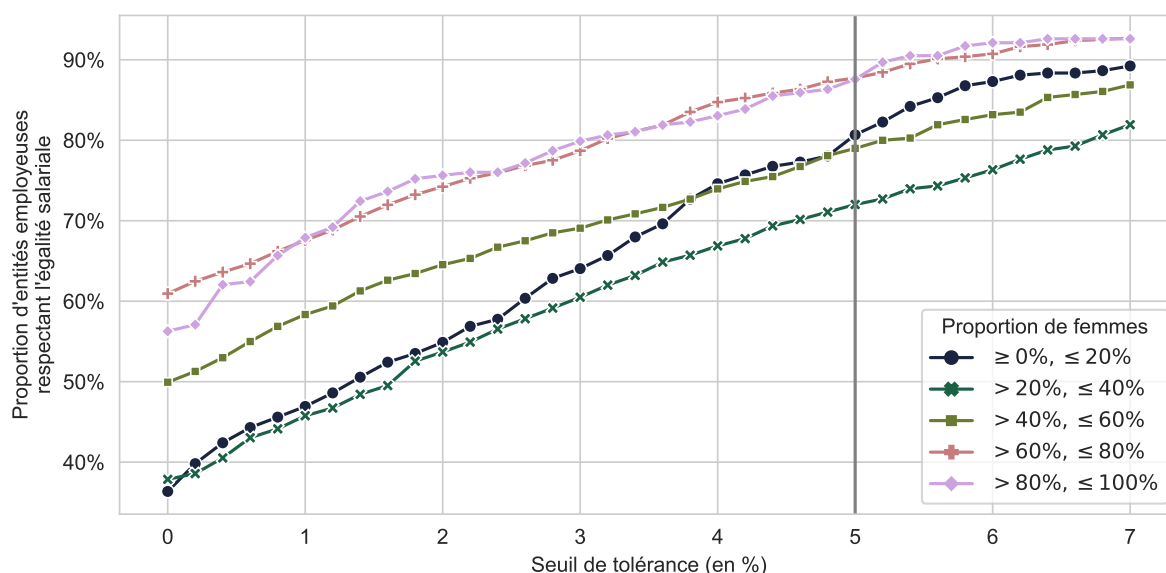


Compte tenu du nombre de branches, il serait laborieux de décrire tous les résultats obtenus. Toutefois, deux résultats clés ressortent de ce graphique. Premièrement, la modification du seuil de tolérance a un effet plutôt homogène sur les branches, le niveau de conformité est réduit d'environ 25 à 30 points de pourcentage lorsque le seuil passe de 5 % à 0 %. Deuxièmement, une forte séparation verticale s'observe entre les différentes branches, confirmant les conclusions de [Kaiser et Lobsiger \(2019\)](#) et [Kaiser et Möhr \(2021\)](#). En outre, la proportion de conformité la plus basse est affichée pour la branche "industries manufacturières", comme c'est le cas dans des études de [Kaiser et Lobsiger \(2019\)](#) et [Kaiser et Möhr \(2021\)](#), alors que la proportion d'entités conformes la plus élevée selon Logib concerne la branche "santé humaine et action sociale".

5.2.4 Analyse selon la proportion de femmes et de femmes cadres au sein de l'entité employeuse

La figure 7 présente la proportion d'entités conformes par niveau de seuil de tolérance selon la proportion de femmes employées.

FIGURE 7 – Conformité par niveau de tolérance selon la proportion de femmes au sein de l'entité employeuse

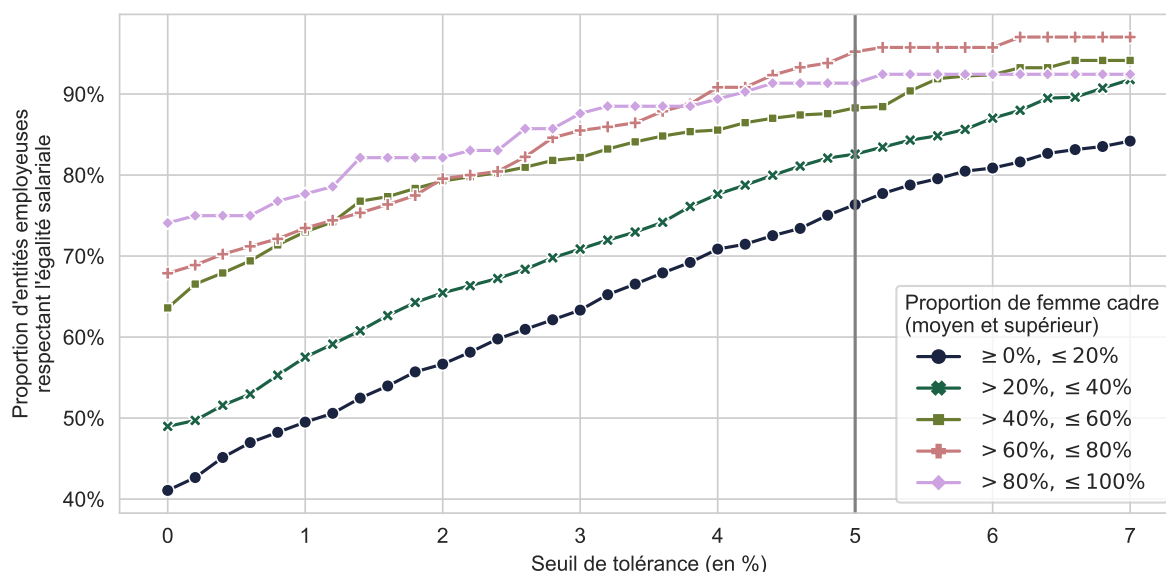


A l'exception des entités ayant moins de 20 % de femmes, les différentes courbes sont relativement parallèles quelque soit le niveau de seuil de tolérance. La proportion d'entités employeuses respectant l'égalité salariale au sens de Logib tend à augmenter avec la proportion de femmes employées au sein de l'entité. Les plus hauts niveaux de conformité se retrouvent pour les entités ayant plus de 60 % de femmes, indépendamment du niveau du seuil de tolérance.

La figure 8 se base sur la proportion de femmes parmi les employé·e·s exerçant une fonction de cadre moyen ou supérieur.

Les résultats montrent que les entités employeuses ayant un faible pourcentage de femmes cadres (<40 %) sont moins conformes en matière d'égalité salariale quelque soit le niveau de seuil de tolérance. Pour les entités comptant 40 % ou plus de femmes cadres, aucune différence n'est constatée. La réduction du seuil de tolérance ferait augmenter cette différence dans la conformité des entités entre celles employant moins de 40 % de femmes cadres et celles employant au moins 40 % de femmes cadres.

FIGURE 8 – Conformité des entités employeuses par niveau de tolérance selon la proportion de femmes cadres au sein de l'entité employeuse



5.2.5 Conformité par niveau de tolérance pour l'ensemble des facteurs

Le tableau 4 présente un modèle de probabilité linéaire qui intègre simultanément toutes les variables traitées dans les sections précédentes. La variable dépendante est le fait de respecter l'égalité salariale entre les femmes et les hommes selon Logib. L'avantage d'analyser toutes les variables simultanément est de pouvoir identifier si les effets précédemment identifiés sont liés. Par exemple, on peut imaginer que les entités publiques sont, en moyenne, plus grandes que les entités privées. Par conséquent, la plus grande proportion d'entités employeuses conformes dans le secteur public pourrait au moins en partie être expliquée par la taille de ces entités. L'analyse simultanée de toutes les variables permet d'identifier l'effet de chaque variable en maintenant les autres à une valeur constante. Dès lors, il est possible de trouver des résultats légèrement différents à ceux présentés dans les précédentes sections.

Selon les résultats présentés dans le tableau 4, la différence entre le secteur public et le secteur privé est d'environ 8,5 points de pourcentage, toute chose étant égale par ailleurs. Cette différence est hautement significative pour un seuil de tolérance de 5 %, mais devient non significative une fois que le seuil est réduit à zéro. Ce résultat correspond à ce que la figure 4 montre, mais la différence est un peu moins importante. Selon la figure 4, la différence entre les deux secteurs est de 14 points de pourcentage, tandis que l'analyse simultanée de toutes les variables indique une différence de 8,5 points de pourcentage. Autrement dit, une partie de la différence entre le secteur public et le secteur privé s'explique par d'autres variables comme la taille de l'entité ou encore la branche économique.

Pour la taille des entités employeuses, les différences ne sont pas statistiquement significatives pour le seuil de 5 %, mais deviennent significatives lorsque le seuil est réduit. La différence dans la proportion de conformité entre les entités les plus petites (50-250) et les plus grandes (>1000) atteint un niveau en dessus de 25 points de pourcentage ($0,125 + 0,128 = 0,253$).

Il faut préciser que l'interprétation des paramètres par branche s'avère compliquée. Cependant, dans l'ensemble, aucun changement significatif d'effet positif à négatif et inversement n'est observé lorsque le niveau du seuil de tolérance change. L'effet reste donc à peu près similaire pour la plupart des branches. Ceci est en accord avec les résultats présentés dans la figure 6. L'analyse de régression confirme également les grandes différences du pourcentage de conformité entre branches économiques. A titre d'exemple, on constate une différence de 43,8 points de pourcentage entre la branche *Arts, spectacle et activités récréatives* et *construction* ($0,251 - (-0,187) = 0,438$).

Il est raisonnable d'examiner la proportion de femmes pour les deux indicateurs ensemble, proportion de femmes et proportion de femmes cadres, car ils sont probablement corrélés. Pour le seuil de 5 %, et en maintenant tout le reste constant, les entités qui ne comptent presque pas de femmes parmi leurs effectifs ont tendance à respecter davantage l'égalité salariale au sens de Logib. Ce constat découle de l'observation des coefficients négatifs pour toutes les catégories de la variable proportion femmes. En comparaison à la figure 7, ce résultat peut surprendre, mais peut être expliqué justement par les autres variables considérées dans l'analyse, notamment la branche économique.

En même temps, il y a un effet clairement positif et croissant de la proportion de femmes cadres sur la conformité des entités employeuses. Par exemple, la valeur de 0,74 pour une proportion de femmes cadres de 0.4 à 0.6 signifie que la proportion de conformité est 7.4 % plus élevée pour les entités employeuses avec 40 % à 60 % de femmes cadres en comparaison avec le groupe de référence (entités avec moins de 20 % de femmes cadres).

Lorsque le seuil de tolérance est plus bas, les coefficients estimés pour la variable *proportion femmes* deviennent non significatifs, à l'exception de la catégorie allant de 20 % à 40 %. Ce résultat indique que la proportion de femmes dans l'ensemble de l'effectif n'est plus une variable déterminante pour expliquer la proportion de conformité.

En revanche, un résultat intéressant est observé pour la proportion de femmes cadres. Les catégories 40 % à 60 % ainsi que 80 % à 100 % affichent des taux de conformité significativement plus élevés comparés aux autres catégories. Pour la catégorie allant de 80 % à 100 % le taux de conformité est 14 % plus élevé que pour la catégorie de référence (< 20 %).

TABLEAU 4 – MODÈLE DE PROBABILITÉ LINÉAIRE : CONFORMITÉ AVEC L'ÉGALITÉ SALARIALE

Facteur	Seuil de 0 %			Seuil de 2 %			Seuil de 5 %		
	Coef.	ES	Sig.	Coef.	ES	Sig.	Coef.	ES	Sig.
Constante	0.414	0.052	***	0.571	0.054	***	0.878	0.044	***
Secteur privé	(Référence)			(Référence)			(Référence)		
Secteur public	0.027	0.034		0.100	0.032	***	0.085	0.024	***
Taille 250-999	(Référence)			(Référence)			(Référence)		
Taille : 50-249	0.125	0.022	***	0.080	0.021	***	0.022	0.018	
Taille : 1000	-0.128	0.027	***	-0.097	0.029	***	-0.008	0.025	
Activités de services administratifs et de soutien	(Référence)			(Référence)			(Référence)		
Activités financières et d'assurance	-0.116	0.056	**	-0.078	0.059		-0.122	0.053	**
Activités immobilières	-0.033	0.112		0.045	0.110		-0.040	0.097	
Activités spécialisées, scientifiques et techniques	0.053	0.058		0.159	0.056	***	0.098	0.044	**
Administration publique	0.167	0.059	***	0.093	0.057		0.007	0.042	
Arts, spectacles et activités récréatives	0.251	0.085	***	0.231	0.073	***	0.100	0.056	*
Autres activités de services	0.122	0.088		0.137	0.084		-0.026	0.076	
Commerce; réparation d'automobiles et de motocycles	-0.049	0.052		-0.040	0.053		-0.058	0.045	
Construction	-0.187	0.074	**	-0.058	0.088		-0.121	0.076	
Enseignement	0.200	0.056	***	0.214	0.053	***	0.100	0.039	**
Hébergement et restauration	0.234	0.059	***	0.252	0.054	***	0.152	0.041	***
Industrie manufacturière	-0.164	0.048	***	-0.168	0.049	***	-0.199	0.042	***
Industries extractives	-0.039	0.358		-0.151	0.358		0.099	0.044	**
Information et communication	-0.070	0.072		-0.092	0.074		-0.085	0.065	
Production et distribution d'eau; assainissement, gestion des déchets et dépollution	-0.051	0.224		-0.010	0.233		-0.063	0.164	
Production et distribution d'électricité, de gaz, de vapeur et d'air conditionné	-0.136	0.106		-0.258	0.114	**	-0.260	0.118	**
Santé humaine et action sociale	0.148	0.051	***	0.152	0.049	***	0.102	0.038	***
Transports et entreposage	0.132	0.067	*	0.130	0.065	**	-0.002	0.054	
Proportion femmes : <0.2	(Référence)			(Référence)			(Référence)		
Proportion femmes : (0.2, 0.4]	-0.077	0.034	**	-0.104	0.036	***	-0.148	0.031	***
Proportion femmes : (0.4, 0.6]	-0.040	0.037		-0.059	0.038		-0.145	0.032	***
Proportion femmes : (0.6, 0.8]	0.013	0.040		-0.029	0.040		-0.124	0.033	***
Proportion femmes : (0.8, 1.0]	-0.035	0.052		-0.022	0.049		-0.148	0.039	***
Proportion femmes cadres : <0.2	(Référence)			(Référence)			(Référence)		
Proportion femmes cadres : (0.2, 0.4]	-0.009	0.023		0.020	0.022		0.041	0.019	**
Proportion femmes cadres : (0.4, 0.6]	0.102	0.030	***	0.120	0.026	***	0.074	0.020	***
Proportion femmes cadres : (0.6, 0.8]	0.045	0.043		0.031	0.039		0.085	0.023	***
Proportion femmes cadres : (0.8, 1.0]	0.140	0.058	**	0.078	0.050		0.082	0.030	***

Notes : Modèle de probabilité linéaire. La variable dépendante est un indicateur de la conformité en matière d'égalité salariale selon Logib. Coef : coefficient, ES : erreur standard, et Sig. : niveau de significativité (*** = 1%, ** = 5%, * = 10%). Pour les intervalles la notation courte est utilisée, par exemple (0.4, 0.6] signifie $0.4 < \leq 0.6$

Points clés 5 (Résultats de l'étape de recherche 2).

L'analyse concernant le rôle de différents facteurs sur la proportion d'entités employeuses conformes par niveau de seuil de tolérance a révélé plusieurs points intéressants :

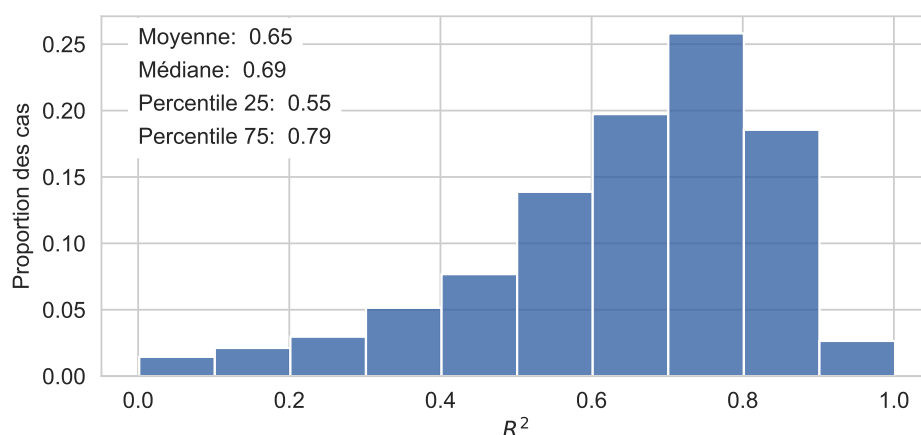
- En utilisant un seuil de tolérance de 5 %, la proportion d'entités conformes est nettement supérieure dans le secteur public par rapport au secteur privé. Si le seuil de tolérance est supprimé, la différence se réduit considérablement.
- Les entités plus grandes ont tendance à être plus conformes en matière d'égalité salariale selon Logib que les plus petites lorsqu'un seuil de tolérance de 5 % est appliqué. C'est l'inverse qui se produit pour les seuils de tolérance inférieurs, car les entités employeuses les plus grandes sont celles qui affichent les plus faibles niveaux de conformité. En ce sens, la réduction du seuil aurait un impact plus fort parmi les grandes entités employeuses.
- De fortes différences dans la proportion de conformité entre les branches économiques ont été constatées. Ces différences demeurent indépendamment du niveau du seuil de tolérance.
- Une plus grande proportion de femmes, en particulier de femmes cadres, est associée à une plus grande proportion de conformité en matière d'égalité salariale selon Logib.

5.3 Lien entre le pouvoir explicatif du modèle et la différence salariale liée au sexe

Pour analyser l'influence du pouvoir explicatif du modèle Logib sur la proportion d'entités conformes, ainsi que la manière dont cette relation change par niveau de seuil de tolérance, une analyse similaire à celles présentées précédemment peut être réalisée.

Toutefois, avant de passer aux résultats de cette analyse, il est pertinent de présenter la distribution du coefficient de détermination (R^2) pour avoir une meilleure appréciation du pouvoir explicatif du modèle Logib ainsi que de l'importance relative des groupes d'entités définies par la valeur du R^2 . La figure 9 montre la distribution du coefficient de détermination pour l'ensemble des régressions réalisées lors de l'analyse empirique.

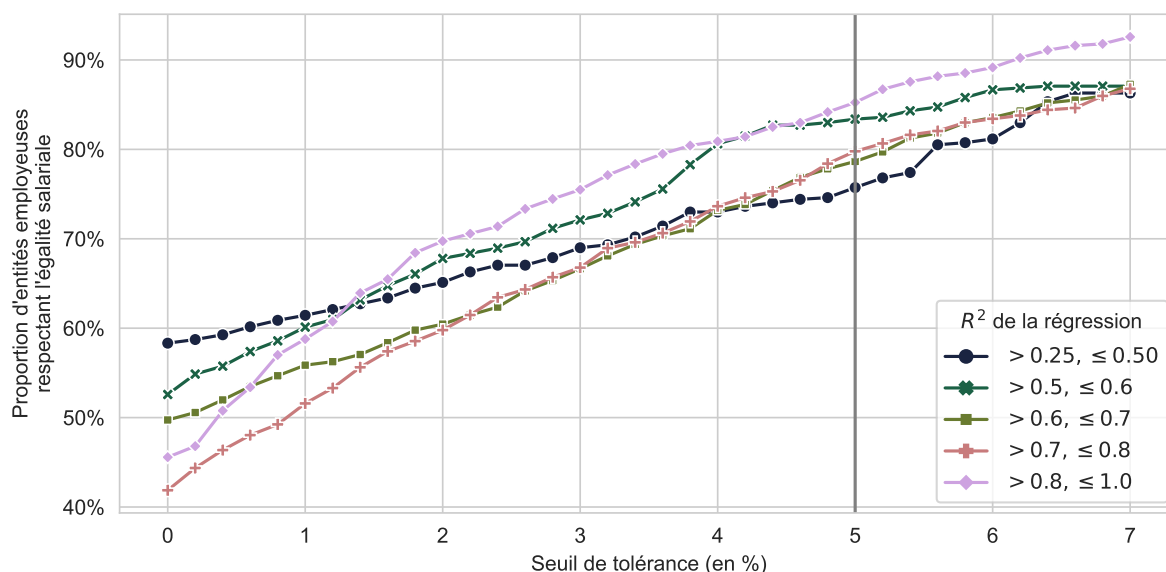
FIGURE 9 – Distribution du coefficient de détermination (R^2)



On constate que pour la grande majorité des régressions, le coefficient de détermination (R^2) se situe entre 0,6 et 0,9, ce qui correspond à une proportion entre 60 % et 90 % de la variation des salaires au sein des entités employeuses qui est expliquée par le modèle. La moyenne du coefficient de détermination est à 0,65 et la médiane à 0,69. [Binggeli et al. \(2018\)](#) rapportent une valeur moyenne de 84 % pour le coefficient de détermination dans le cadre des contrôles menés par le BFEG dans les marchés publics de la Confédération. Les différences entre les résultats de notre étude et ceux des contrôles pourraient notamment venir du codage de la variable niveau de compétences. Dans le cas de l'ESS, cette variable est codée de manière automatique (cf. Tableau 2), alors qu'elle correspond à la réalité interne de l'entité dans le cadre des contrôles.

La figure 10 montre la proportion d'entités conformes par niveau de seuil de tolérance selon la statistique R^2 du modèle Logib.

FIGURE 10 – Conformité par niveau de tolérance selon le niveau d'ajustement du modèle



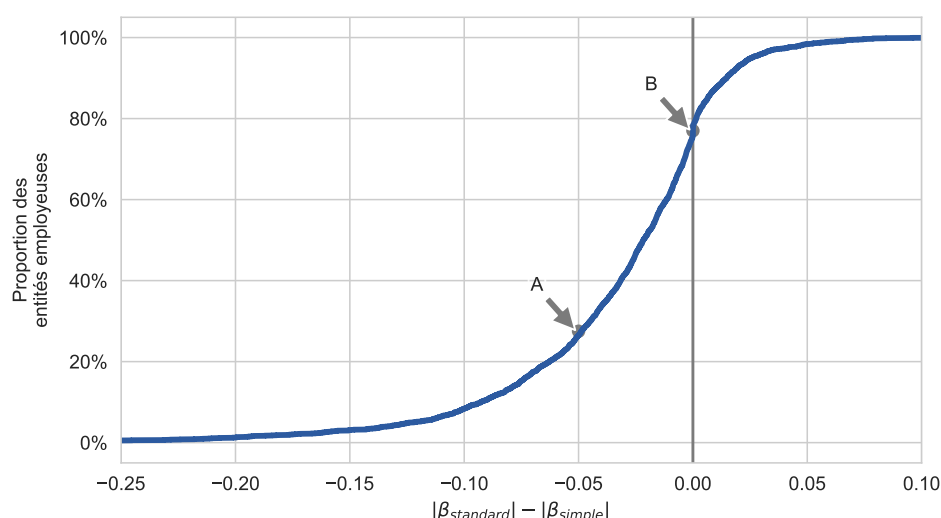
Les résultats de cette analyse sont intéressants à plusieurs égards. Premièrement, les courbes sont relativement proches les unes des autres, ce qui suggère que le R^2 n'est pas pertinent pour expliquer la proportion d'entités conformes en matière d'égalité salariale. Ce résultat est en accord avec l'argument avancé dans la section 3.4.2.

Deuxièmement, on peut voir que l'ordre des courbes pour un seuil de tolérance de 5% n'est pas lié au niveau du R^2 , alors que l'ordre devient au moins partiellement lié lorsque le seuil de tolérance est égal à zéro. Dans ce cas, les entités pour lesquelles le modèle s'adapte le mieux aux données ont tendance à avoir un niveau de conformité plus faible. Par conséquent, la réduction du seuil affecterait davantage les entités dont le R^2 est élevé en comparaison à celles dont le pouvoir explicatif est plus faible. D'un point de vue statistique, ce résultat n'est pas particulièrement surprenant, car la précision de l'estimation est plus faible pour les entités employeuses ayant un faible niveau de R^2 et il est donc plus difficile d'identifier les effets significatifs.

Les résultats de la figure 10 sont basés sur les différences d'ajustement du modèle entre les entités employeuses utilisant toutes le même modèle. La figure 11 examine le lien entre l'ajustement du modèle et l'effet lié au sexe d'un point de vue différent. Elle compare l'effet lié au sexe du modèle standard avec celui du modèle simple dans lequel les variables niveau de compétences et position professionnelle ne sont pas prises en compte. La soustraction des coefficients estimés du modèle standard avec ceux du modèle simple révèle l'effet de l'inclusion desdites variables sur la valeur absolue de l'effet lié au sexe.

La figure montre la distribution de cette différence. Sur l'axe horizontal se trouve le résultat de la soustraction des résultats des deux modèles (standard - simple). Une valeur négative indique que l'inclusion des deux variables supplémentaires (niveau de compétences et position professionnelle) réduit l'écart salarial non expliqué. Une valeur positive, en revanche, indique que l'inclusion des deux variables augmente l'écart salarial non expliqué. L'axe vertical montre la proportion des entités employeuses.

FIGURE 11 – Différence de l'effet lié au sexe entre le modèle standard et le modèle simple



Le graphique est intéressant puisqu'il montre que l'inclusion du niveau de compétences et de la position professionnelle peut avoir des conséquences très différentes selon les entités. Le point A indique que pour 27 % des entités employeuses, l'effet lié au sexe se réduit d'au moins 5 % lorsque les variables niveau de compétences et position professionnelle sont ajoutées au modèle. Le point B montre que pour 77 % des entités employeuses l'inclusion de ces deux variables réduit l'effet lié au sexe, tandis que pour 23 % des entités employeuses cette inclusion fait augmenter l'estimation de l'effet lié au sexe. Enfin, on constate aussi que la différence de l'écart non expliqué se situe entre $\pm 5\%$ pour une large majorité d'environ 70 % de entités employeuses.

En considérant les deux principales analyses de ce chapitre, il apparaît que l'ajustement du modèle, que cela soit par l'ajout ou non de variables supplémentaires, et l'effet lié au sexe ne sont pas aussi étroitement liés que ce que laisse penser la littérature. L'argument selon lequel une réduction du seuil de tolérance devrait s'accompagner de la prise en compte de variables supplémentaires dans le modèle ne paraît donc pas si évident sur le plan statistique.

Points clés 6 (Résultats de l'étape de recherche 3a).

- Le lien entre le R^2 du modèle Logib et l'effet lié au sexe est très faible, voire inexistant si on utilise un seuil de tolérance de 5 %. Par contre, les entités avec un R^2 plus élevé ont un taux de conformité légèrement inférieur à celles avec un R^2 inférieur sans seuil de tolérance.
- L'inclusion du niveau de compétences et de la position professionnelle dans un modèle basé sur les qualifications personnelles des employé·e·s réduit en moyenne la différence salariale non expliquée, mais pas pour toutes les entités employeuses. Pour 27 % des entités employeuses la réduction est de plus de 5 %, tandis que pour 23 %, l'inclusion fait même augmenter l'effet lié au sexe. Pour les 50 % restantes, la réduction est de moins de 5 %.

5.4 Calcul de l'analyse Logib en tenant compte de l'hétéroscédasticité

L'étape de recherche 3b vise à évaluer la pertinence d'envisager l'utilisation d'erreurs standard robustes dans le cadre de l'analyse Logib. A ce sujet, trois questions se posent :

1. Combien de modèles Logib satisfont l'hypothèse d'homoscédasticité permettant l'utilisation des erreurs standard normales ?
2. Combien de conclusions Logib par rapport à la conformité en matière d'égalité salariale changeraient, si les erreurs standard robustes étaient utilisés ?
3. Est-ce que le nombre de conclusions qui dépendent du choix des erreurs standards change en fonction du seuil de tolérance ?

Pour répondre à la première question, il suffit d'utiliser le test statistique qui permet de voir si l'hypothèse est satisfaite ou pas. Sur les 2'845 régressions, seules 606 satisfont l'hypothèse d'homoscédasticité, tandis que pour 2'239 (soit 79 %), il est nécessaire de calculer des erreurs standard robustes. Autrement dit, pour quatre entités sur cinq, l'hypothèse sur l'homoscédasticité est transgressée dans le cadre de l'analyse Logib et par conséquent le calcul des erreurs standard n'est pas correct.

Pour répondre à la deuxième question concernant la proportion de conclusions qui changeraient, il suffit de comparer la conclusion faite par Logib avec et sans prise en compte des erreurs standard robustes. Le tableau 5 montre les résultats de conformité avec un seuil tolérance de 5 % et sans seuil de tolérance pour la version actuelle ainsi que pour la version robuste de Logib. La comparaison entre les proportions avec un seuil tolérance de 5 % et sans seuil de tolérance permet de répondre à la troisième question.

TABLEAU 5 – COMPARAISON DES CONCLUSIONS AVEC ET SANS ERREURS STANDARD ROBUSTES

	Logib	Logib robuste	Seuil de 5 %	Seuil de 0%
1	Conforme	Conforme	79.9 %	46.6 %
2	Non conforme	Non conforme	18.2 %	48.7 %
3	Conforme	Non conforme	1.0 %	3.2 %
4	Non conforme	Conforme	0.9 %	1.6 %
	Total		100 %	100 %

La première ligne (1) montre la proportion des entités employeuses qui seraient considérées conformes au sens de Logib, avec ou sans erreurs standard robustes. En appliquant un seuil de tolérance de 5 %, 79.9 % des entités employeuses sont considérées conformes au sens de Logib, indépendamment de l'utilisation des erreurs standard robustes. Sans seuil de tolérance, cette même proportion est de 46.6 %. La deuxième ligne (2) montre la proportion des entités qui sont considérées non conformes par le modèle Logib actuel et la version robuste de Logib. Ces proportions sont respectivement de 18.2 % et 48.7 % pour le seuil de tolérance de 5 % et 0 %.

Ces deux premières lignes indiquent la proportion d'entités employeuses pour lesquelles la conclusion de Logib ne change pas si les erreurs standard robustes sont introduites.

En revanche, les lignes 3 et 4 présentent les proportions d'entités employeuses pour lesquelles le choix d'erreurs standard est critique, puisque la conclusion par rapport à la conformité au sens de Logib change en fonction du type d'erreur standard utilisé.

Sans seuil de tolérance, 3.2 % des entités seraient erronément déclarées conformes en matière d'égalité salariale alors que cela n'est pas le cas si l'on utilise les erreurs standard robustes. Dans l'autre sens, 1.6 % des entités sont actuellement considérées comme non-conformes alors qu'elles seraient conformes si les erreurs standard robustes étaient utilisées. En faisant la somme des lignes 3 et 4, on constate que dans 4.8 % des cas, la conclusion de Logib serait erronée si on n'utilisait pas de seuil de tolérance et si on ignorait la violation de l'hypothèse d'homoscédasticité. Cette proportion de conclusions erronées est inférieure (1.9 %) si on utilise Logib avec un seuil de tolérance de 5%. La proportion de conclusions erronées augmente donc si le seuil de tolérance est réduit.

Il convient de rappeler que la correction avec les erreurs standard robustes est une procédure tout à fait ordinaire dans la littérature et ne nécessite pas de données supplémentaires.

Points clés 7 (Résultats de l'étape de recherche 3b).

- L'hypothèse d'homoscédasticité n'est pas satisfaite dans environ 80 % des analyses Logib.
- Si les erreurs standard ne sont pas corrigées pour l'hétéroscédasticité, en absence du seuil de tolérance, 4.8 % des conclusions par rapport à la conformité des entités seraient erronées.

6 Discussion et recommandations

6.1 Récapitulatif des principales constatations de l'étude

Pour la discussion et l'élaboration des recommandations, il convient de reprendre les questions générales auxquelles le mandat cherche à répondre et de résumer les éléments que l'analyse empirique et la révision de la littérature ont permis d'obtenir.

1. Quelle proportion d'entités employeuses respectent l'égalité salariale au sens du modèle d'analyse standard de la Confédération lorsque le seuil de tolérance varie entre 0 % et 5 % ?

Au niveau de l'ensemble de l'économie, la proportion d'entités employeuses respectant l'égalité salariale au sens du modèle d'analyse standard de la Confédération est de 80,9 % avec le seuil actuel de 5 %, de 65,0 % avec un seuil de 2 % et de 49,8 % sans seuil de tolérance. Autrement dit, si le seuil actuel se maintient, une entité sur cinq ne remplit pas les conditions requises, tandis qu'une sur deux ne les remplirait pas si aucun seuil n'était appliqué.

2. Quelles variables jouent un rôle significatif sur la proportion d'entités employeuses qui respectent l'égalité salariale et comment ce rôle change en fonction du seuil de tolérance utilisé ?

- Les entités employeuses du secteur public respectent significativement plus l'égalité salariale au sens du modèle d'analyse standard de la Confédération que les entités du secteur privé lorsque l'on tient compte du seuil de 5 %, mais cette différence est non significative lorsqu'il n'y a pas de seuil de tolérance.
- Les grandes entités sont légèrement plus susceptibles d'être en conformité que les petites avec un seuil de tolérance de 5 %. Lorsque le seuil se rapproche de zéro, l'inverse est vrai. En tenant compte des autres facteurs et sans seuil de tolérance, les entités employeuses les plus grandes ont jusqu'à 25 points de pourcentage de moins de probabilité de remplir les exigences en matière d'égalité salariale.
- De grandes différences dans la proportion d'entités conformes sont constatées entre les branches économiques. Les industries manufacturières ont la plus faible proportion de conformité. Ces différences ne changent pas de manière significative lorsque le seuil de tolérance est modifié.
- Les entités employeuses qui comptent une proportion plus élevée de femmes dans leur effectif et surtout parmi leurs cadres respectent davantage l'égalité salariale au sens de Logib. Les différences sont un peu plus faibles lorsque l'on tient compte d'autres facteurs, mais elles restent statistiquement significatives.

3. Quelles informations devraient être prises en considération par la CoMPS lors d'une réflexion sur le seuil de tolérance dans le cadre des contrôles menés dans les marchés publics et les subventions ?

- En 2015, une étude qualitative a révélé que des expert·e·s en matière d'égalité salariale estimaient que le seuil de tolérance avait prouvé son utilité. Ces personnes préconisaient de maintenir le modèle actuel avec un seuil de tolérance à 5 % plutôt que d'ajouter des variables supplémentaires et de réduire le seuil de tolérance. Dans une interview, un·e expert·e a toutefois proposé de baisser le seuil de tolérance.
- D'un point de vue statistique, l'utilisation d'un seuil de tolérance en plus du test d'hypothèse n'est pas justifiée pour éviter des conclusions erronées concernant le respect de

l'égalité salariale au sein d'une entité employeuse. L'utilisation du seuil de tolérance actuel serait justifiable si des différences allant jusqu'à 5 % étaient explicitement tolérées par la loi, ce qui n'est pas le cas.

- L'analyse empirique a montré que le lien entre le pouvoir explicatif du modèle Logib et l'écart salarial est faible. L'abaissement du seuil ne nécessiterait pas automatiquement l'ajout de variables supplémentaires dans Logib. Les entités présentant les R^2 la plus élevée dans le modèle Logib seraient en fait plus impactées par le changement de seuil de tolérance en termes de proportion de conformité.
- L'analyse empirique a montré que près de 80 % de toutes les régressions Logib transgressent l'hypothèse sur la distribution des résidus de la régression (test d'homoscédasticité¹⁶). La correction de ce problème n'a que peu d'impact sur les résultats des analyses Logib lorsque le seuil de tolérance de 5 % est pris en compte. Toutefois, la correction de ce problème modifie les conclusions relatives à la conformité des entités dans environ 5 % des cas lorsqu'aucun seuil de tolérance n'est pris en compte. Dès lors, il conviendrait de procéder à cette correction si le seuil de tolérance était diminué ou supprimé.

6.2 Conclusions

Au vu des résultats précédents et de la discussion avec une perspective statistique abordée dans la section 3.4, une adaptation du seuil de tolérance semble pertinente pour deux raisons principales. Premièrement, l'utilisation d'un tel seuil n'est pas scientifiquement justifiée si l'objectif est de réduire à zéro les écarts salariaux. Deuxièmement, les entités employeuses ont maintenant eu plusieurs années pour ajuster leurs politiques et réduire les écarts salariaux liés au sexe. En laissant le seuil à 5 %, on court le risque que les entités n'essaient pas de réduire complètement les écarts en choisissant de se conformer aux exigences d'une méthodologie utilisée pour analyser l'égalité salariale avec un seuil de tolérance de 5 % plutôt que de se conformer à la loi qui ne prévoit pas de seuil de tolérance.

Les résultats de cette étude ont également montré que l'utilisation d'un modèle ayant un plus grand pouvoir explicatif n'est pas une condition nécessaire pour abaisser le seuil de tolérance. Par conséquent, sans ajout de variables, la réduction du seuil ne devrait pas entraîner une charge administrative plus importante pour les entités employeuses.

De ce fait, nos recommandations générales peuvent se résumer comme suit :

- Supprimer le seuil de tolérance et se baser uniquement sur le test de signification statistique pour déterminer si une entité employeuse respecte ou non l'égalité salariale. La suppression du seuil pourrait se faire par étapes en passant par des seuils de tolérance moins élevés (par exemple 2 %) pendant une phase de transition.
- En cas de suppression du seuil, faire systématiquement le test statistique pour examiner l'homogénéité de la distribution des résidus de la régression (homoscédasticité) et, en cas de rejet de l'hypothèse, utiliser des erreurs standard robustes.
- En cas de suppression du seuil et de non conformité d'une entité employeuse en matière d'égalité salariale, envisager la possibilité pour l'entité de fournir des justifications objectives et conformes au droit afin de déterminer si elles pourraient avoir un impact sur la conclusion de l'analyse.

16. Voir annexe B.2 pour une explication

Ces recommandations sont basées sur des considérations statistiques et elles cherchent à prendre en compte les différents arguments en faveur du seuil de tolérance. La première recommandation qui adresse la suppression du seuil de tolérance se base sur le constat que le système actuel permet à plus de 30 % des entités employeuses d'avoir un écart salarial significatif (donc contraire à la loi) tout en restant conformes en matière d'égalité salariale selon les critères Logib. L'objectif n'est pas de pénaliser davantage d'entités employeuses, mais plutôt de ne pas désavantager les entités qui respectent l'égalité salariale vis-à-vis de celles qui ne la respectent pas. Les lignes directrices de la CoMPS soulignent ce point en indiquant que *"les employeuses qui respectent les prescriptions légales ne doivent pas être défavorisées par rapport à celles qui ne les respectent pas."* (CoMPS, 2021).

La deuxième recommandation découle du fait que jusqu'à 5 % des conclusions par rapport à la conformité des entités en matière d'égalité salariale pourraient être changées en corrigeant le problème lié à la distribution des résidus de la régression. Cette correction n'a aucun coût et son introduction pourrait donc être envisagée. D'un point de vue économétrique, l'idéal serait de tester systématiquement l'hétéroscédasticité et d'appliquer des erreurs standard robustes le cas échéant. Alternativement, l'application de ces tests uniquement en cas de découverte d'un effet significatif lors d'un contrôle pourrait être une solution pratique.

La troisième recommandation tient compte de l'argument de certaines parties selon lequel un modèle statistique ne peut jamais refléter complètement la réalité de chaque entité employeuse. Au terme d'une analyse sans seuil de tolérance débouchant sur un constat de non conformité en matière d'égalité salariale entre les femmes et les hommes, il serait judicieux de déterminer si l'entité possède des variables objectives et non discriminatoires susceptibles de faire changer la conclusion du contrôle. Il est toutefois à noter que l'appréciation du potentiel discriminatoire de nouvelles variables est une tâche très complexe (Pärli, 2019) et susceptible d'engendrer un coût disproportionné dans le cadre d'une procédure administrative. Actuellement, une telle démarche, bien que sensée sur le plan juridique, serait donc certainement difficile à mettre en œuvre. Le développement des connaissances et des outils en matière d'égalité salariale entre les femmes et les hommes pourraient peut-être permettre à l'avenir de concrétiser une telle démarche.

6.3 Limites de l'analyse

Bien que l'analyse présentée dans cette étude ait été réalisée avec le plus grand soin, quelques limites sont à prendre en compte.

La base de données utilisée ne permet pas d'identifier les entités qui participent activement aux marchés publics. Il n'est pas à exclure que ces entités soient différentes et que les proportions estimées changent en conséquence.

De plus, lors d'un contrôle, les expert·e·s sont en contact avec l'entité à évaluer et effectuent une plausibilisation des données. En cas de doute, des explications ou des ajustements des données sont demandés par ces expert·e·s. Ce processus n'a pas été reflété ni fait pour cette analyse.

Une deuxième limite est l'année de référence à partir de laquelle toutes les données sont tirées, à savoir 2018. Depuis lors, certains progrès ont pu avoir lieu en matière d'égalité salariale.

Néanmoins, dans l'ensemble, l'analyse devrait fournir une image précise de la situation.

Références

- BFEG : Directive - Contrôles du respect de l'égalité salariale entre femmes et hommes dans les marchés publics de la Confédération. Bureau fédéral de l'égalité entre femmes et hommes BFEG, Version : janvier 2021, Disponible en ligne : https://www.ebg.admin.ch/dam/ebg/fr/dokumente/lohngleichheit/infos-zu-analysen/richtlinie_kontrolle_lohngleichheit_beschaffungswesen_bund.pdf.download.pdf/richtlinie_kontrolle_lohngleichheit_beschaffungswesen_bund.pdf, 2021a.
- BFEG : Modèle d'analyse standard de la Confédération - Descriptif méthodologique Logib Modules 1 et 2, 2021b. Version 2021.1. Disponible en ligne : <https://www.ebg.admin.ch/ebg/fr/home/prestations/logib-triage/logib-modul-1/standardanalysemodell-bund.html> (Téléchargé le 14.7.2021).
- Steve BINGGELI, Oliver SCHRÖTER et Jennifer BIERI : Analyse de l'égalité salariale. *Jusletter*, 2018.
- A. CAMERON et Pravin TRIVEDI : *Microeconometrics*. Cambridge University Press, 2005. URL <https://EconPapers.repec.org/RePEc:cup:cbooks:9780521848053>.
- CoMPS : Lignes directrices - Contrôle du respect de l'égalité salariale entre les femmes et les hommes au sein des marchés publics dans le canton et dans le cadre des subventions octroyées par l'État de Vaud. Commission de contrôle des marchés publics et des subventions, Version : 02.2021, 2021.
- Rahel FELDER et Conny WUNSCH : Evaluation des Analyse-Tools Logib Modul 1 für den Lohnvergleich zwischen Frauen und Männern. WWZinsights No 3, Wirtschaftswissenschaftliche Fakultät, Universität Basel. Disponible en ligne : https://wwz.unibas.ch/fileadmin/user_upload/wwz/99_WWZ_Forum/FINAL_WWZInsights_No3.pdf, 2021.
- Christina FELFE, Judith TRAGESER et Rolf ITEN : Etude des analyses appliquées par la Confédération pour évaluer l'égalité des salaires entre femmes et hommes - Rapport final, 2015. Disponible en ligne : https://www.ebg.admin.ch/dam/ebg/fr/dokumente/lohngleichheit/berichte_studien_gutachten/studie_zu_den_statistischenanalysendereidgenossenschaftbetreffen.pdf.download.pdf/etude_des_analysesappliqueesparlaconfederationpourevaluerlegalit.pdf (20.07.2021).
- Roman GRAF : L'égalité femme-homme : une nouvelle exigence. in : *Jean-Baptiste Zufferey (Éd.), Marchés publics 2016, Schulthess, Zürich*, 32:pp. 371–391, 2016.
- William GREENE : *Econometric Analysis*. Sixth edition édition, 2005.
- Boris KAISER et Michael LOBSIGER : Lohnunterschiede zwischen Frauen und Männern : Evaluation alternativer statistischer Methoden und der Integration weiterer Faktoren aus administrativen Registern der Schweiz. Analyse der Lohnunterschiede anhand der Schweizerischen Lohnstrukturerhebung (LSE). B,S,S. Volkswirtschaftliche Beratung. Studie im Auftrag des Bundesamts für Statistik (BFS), 2019.
- Boris KAISER et Thomas MÖHR : Analyse des différences salariales entre femmes et hommes sur la base de l'enquête suisse sur la structure des salaires (ESS) 2018. B,S,S. Volkswirtschaftliche Beratung. Étude sur mandat de l'Office fédéral de la statistique (OFS), 2021.

- Stephen L. MORGAN et Christopher WINSHIP : *Counterfactuals and Causal Inference - Methods and Principles for Social Research*. Cambridge University Press, New York, USA, 2007.
- Sophie PASCHOUD : Égalité salariale : la manipulation d'un juste principe - Quelques remarques critiques au regard d'un sondage réalisé auprès des entreprises romandes. *Etudes & Enquêtes*, 2015. Publication du Centre Patronal.
- Kurt PÄRLI : Diskussionspapier zum Diskriminierungspotenzial einzelner Erklärungsfaktoren zur Messung der Lohnungleichheit zwischen Frau und Mann aus juristischer Sicht. Étude sur mandat du Bureau fédéral de l'égalité entre femmes et hommes BFEG. Disponible en ligne : <https://www.ebg.admin.ch/dam/ebg/fr/dokumente/lohnungleichheit/Diskussionspapier.pdf.download.pdf/Document%20de%20r%C3%A9flexion.pdf>, 2019.
- Susanne STERN, Judith TRAGESER, Bettina RÜEGGE, Andrea SCHULTHEISS et Rolf ITEN : Regulierungsfolgenabschätzung zu den geplanten Massnahmen zur Durchsetzung der Lohnungleichheit - Schlussbericht, 2015. Disponible en ligne : https://www.infras.ch/media/filer_public/b8/39/b839cf56-f5f0-400f-af84-03e9a2e2483e/schlussbericht_rfa_lohnungleichheit_15_09_07.pdf (20.007.2021).
- Silvia STRUB : Contrôle du respect de l'égalité de salaire entre femmes et hommes dans les marchés publics de la confédération, 2004.
- Judith TRAGESER, Susanne STERN et Rolf ITEN : Evaluation der Kontrollen Im Beschaffungswesen - Schlussbericht, 2011. Disponible en ligne : https://ub.unibas.ch/digi/a125/sachdok/2012/BAU_1_6014049.pdf (20.07.2021).
- Halbert WHITE : A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica*, 48(4):817-838, 1980.
- Jeffrey M. WOOLDRIDGE : *Introductory Econometrics - A modern approach*. Thomson South-Western, 3rd édition, 2006.

A Statistiques descriptives

Cette annexe fournit quelques statistiques descriptives de l'échantillon utilisé dans le cadre de l'analyse empirique. Sauf indication contraire, toutes les statistiques sont pondérées par la variable *gewicht*.

Le tableau 6 présente le nombre¹⁷ d'entités employeuses par secteur et taille pour lesquelles le modèle Logib a pu être estimé.

TABLEAU 6 – NOMBRE D'ENTITÉS PAR SECTEUR ET TAILLE (NON PONDÉRÉ)

Taille\Secteur	Privé	Public	Total
50-249	988	75	1'063
250-1'000	767	114	881
>1'000	213	688	901
Total	1'968	877	2'845

Au total, près de 3'000 modèles Logib ont été estimés, dont près de 2'000 concernent des entités employeuses du secteur privé. La distribution selon la taille des entités est assez uniforme avec près de 1'000 observations par catégorie.

Le tableau 7 présente quelques statistiques descriptives des personnes travaillant dans ces 2845 entités.

TABLEAU 7 – STATISTIQUES DESCRIPTIVES DE L'ÉCHANTILLON UTILISÉ

Variable	Population	Moyenne	Écart type	P25	Médiane	P75
Salaire standardisé	Femmes	6905.23	3000.7	4857.0	6292.0	8135.0
	Hommes	8422.98	4843.7	5667.0	7348.0	9835.0
	Total	7658.39	4093.2	5193.0	6793.0	8945.0
Années de formation	Femmes	12.41	2.9	12.0	12.0	14.0
	Hommes	12.82	2.8	12.0	12.0	15.0
	Total	12.61	2.9	12.0	12.0	15.0
Expérience professionnelle	Femmes	23.31	12.4	13.0	23.0	34.0
	Hommes	24.01	11.9	14.0	24.0	34.0
	Total	23.66	12.2	13.0	24.0	34.0
Années de service	Femmes	7.93	8.4	1.0	5.0	12.0
	Hommes	9.33	9.7	2.0	6.0	14.0
	Total	8.63	9.1	2.0	6.0	13.0

Notes : P25/P75 font référence aux percentiles 25 et 75. Par souci de cohérence avec l'analyse empirique, toutes les statistiques ont été pondérées avec la variable *gewicht*. Source de données : ESS 2018

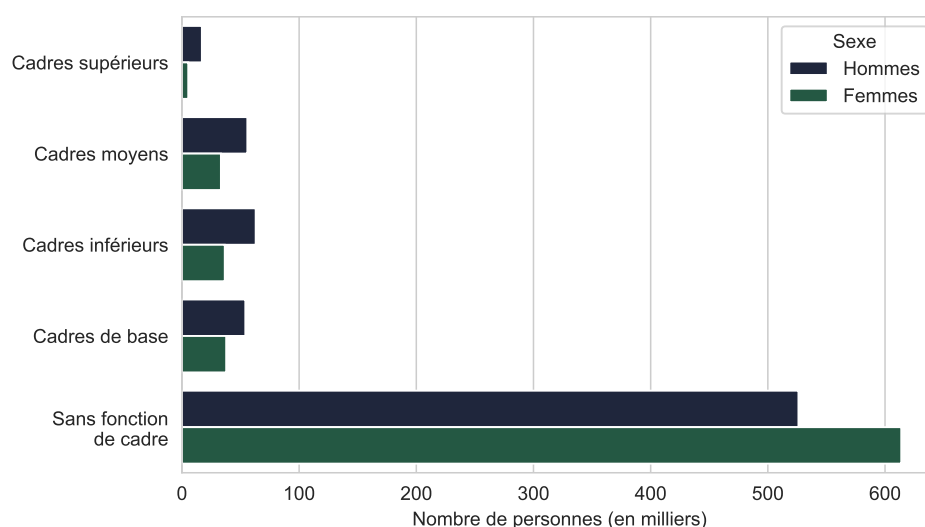
Le salaire moyen standardisé est de 8'422 CHF pour les hommes contre 6'905 CHF pour les femmes. En ce qui concerne les années de formation, les années d'expérience et les années de

17. Il s'agit du nombre non pondéré d'unités employeuses pour montrer combien de modèles (c'est-à-dire de coefficients) ont été estimés.

service dans l'entité employeuse, les hommes ont également des valeurs légèrement supérieures à celles des femmes.

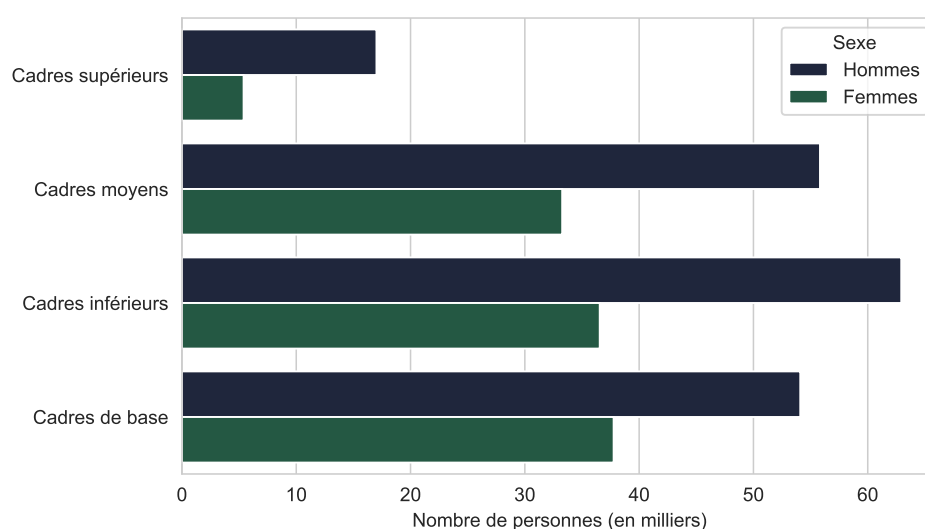
La figure 12 montre la répartition de la position professionnelle par sexe. La grande majorité des employé·e·s n'ont pas de poste de cadre et le nombre de femmes dans cette catégorie est plus élevé que celui des hommes.

FIGURE 12 – Distribution de position professionnelle (par sexe)



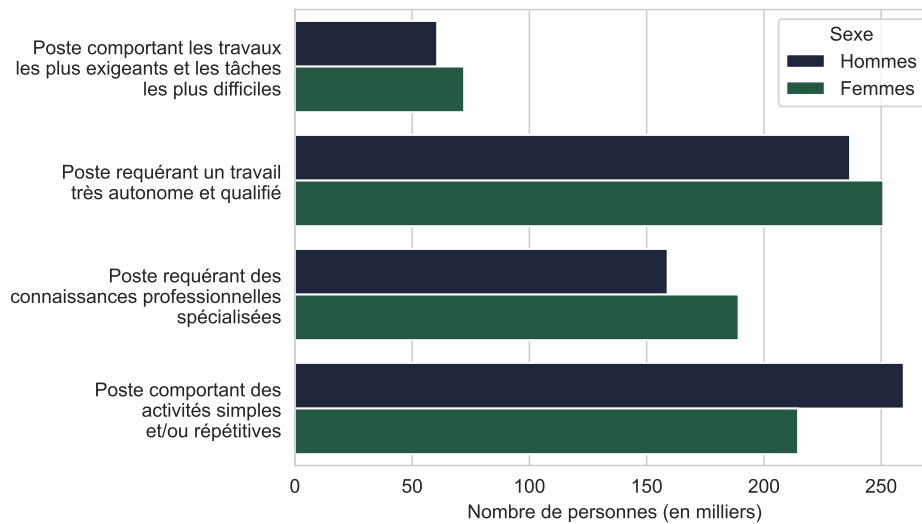
Le nombre de personnes occupant des postes de cadre diminue lorsque la position professionnelle augmente. Pour des raisons de lisibilité, la figure 13 se concentre uniquement sur les employé·e·s ayant des fonctions de cadre. Pour tous les niveaux de cadre, il y a plus d'hommes que de femmes et la proportion de femmes diminue avec le niveau de responsabilité.

FIGURE 13 – Distribution de position professionnelle avec fonction de cadre (par sexe)



La figure 14 présente la répartition du niveau de compétences par sexe. On retrouve une image similaire à celle de la position professionnelle. Parmi les niveaux de qualification les plus bas, les femmes sont plus nombreuses que les hommes, tandis que l'inverse est vrai pour les autres catégories. Néanmoins, les différences sont moins frappantes que pour la position professionnelle.

FIGURE 14 – Distribution du niveau de compétences (par sexe)



B Explications techniques

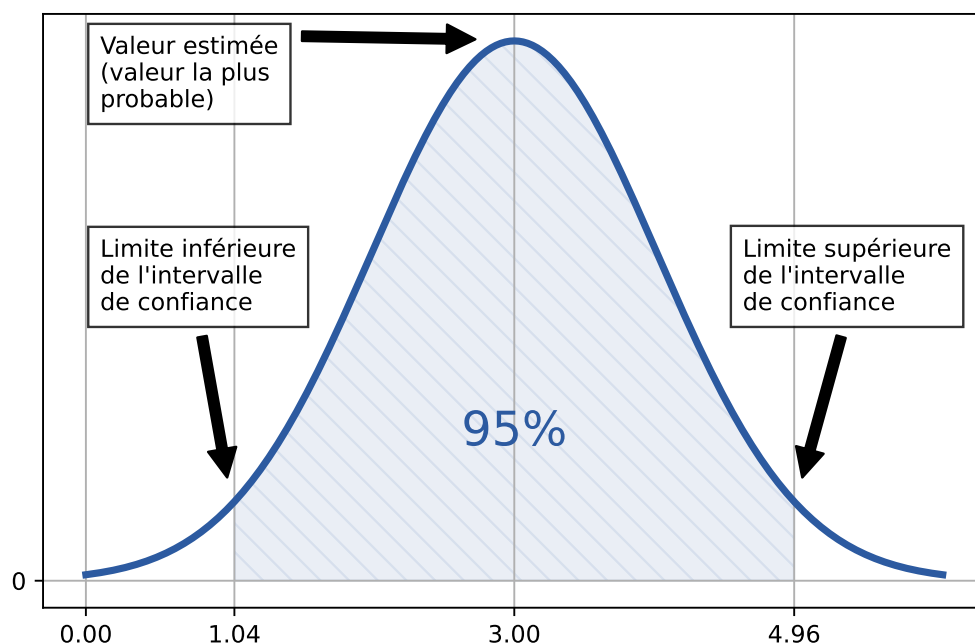
B.1 Test de significativité

La valeur estimée d'un coefficient par un modèle statistique (ou économétrique) fait toujours référence à la valeur la plus probable et est accompagnée d'une erreur standard qui mesure la précision de l'estimation. La combinaison de ces deux valeurs permet de construire un test statistique pour déterminer si l'effet estimé est différent d'une valeur prédéterminée. Normalement, les tests sont effectués par rapport à la valeur de zéro pour voir si le coefficient est différent de zéro, c'est-à-dire si l'effet existe réellement.

Indépendamment de la construction détaillée des différents tests, l'intervalle de confiance (IC) qui en résulte est toujours interprété de la même manière. Cet intervalle donne une gamme de valeurs possibles de l'estimateur pour un niveau de confiance choisi. Un niveau de confiance de 95 signifie qu'il est acceptable de se tromper dans 5 % des cas en qualifiant un effet de significatif alors qu'il ne l'est pas en réalité.

La figure 15 montre un exemple d'un coefficient estimé d'une valeur de 3,0 avec une erreur standard de 1,0.

FIGURE 15 – Illustration de l'intervalle de confiance



La courbe bleue fait référence à la densité¹⁸ qui est la plus élevée pour la valeur de 3,0. Il s'agit donc de la valeur la plus probable du coefficient réel. Des deux côtés, la probabilité diminue mais reste positive. La surface bleue indique 95 % de la surface sous la courbe, ce qui détermine l'intervalle de confiance (1.04 à 4.96).

18. Par souci de simplicité, cette valeur peut être interprétée comme étant approximativement la probabilité que la valeur réelle se trouve à un certain point.

Avec une confiance de 95 %, la valeur réelle du coefficient se situe dans cet intervalle entre 1,04 et 4,96. La valeur de zéro ne faisant pas partie de cet intervalle, on considère que le coefficient estimé est significativement différent de zéro.

La largeur de l'intervalle de confiance dépend, entre autres, de la taille de l'échantillon : plus il y a d'observations, plus la précision de l'estimation augmente et plus l'intervalle de confiance devient fin.

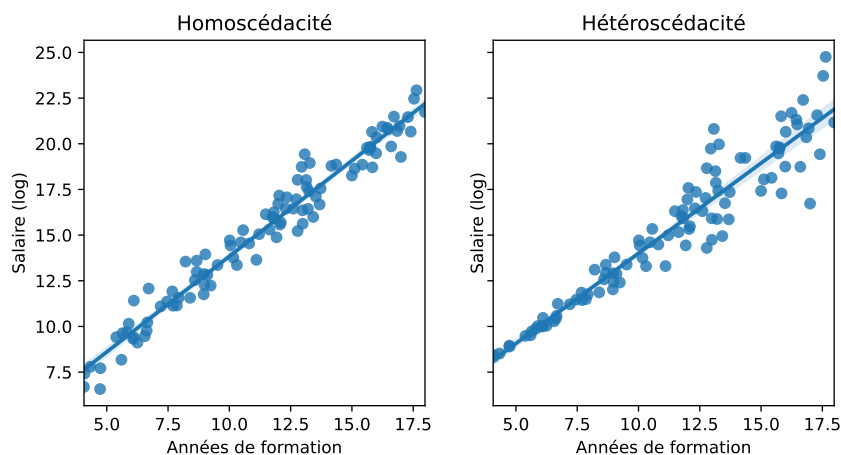
Points clés 8 (Test de significativité).

- La valeur estimée d'un coefficient fait référence à la valeur la plus probable, mais la valeur réelle peut se situer dans un intervalle autour de cette valeur.
- Les tests statistiques permettent de définir un niveau de confiance, qui détermine le risque de se tromper en identifiant comme significatif ce qui ne l'est pas.
- Même un effet relativement grand peut être non-significatif si la précision de l'estimation n'est pas suffisante.

B.2 Hétéroscédasticité

Pour que l'estimation des erreurs standard d'un modèle linéaire estimé par moindres carrés ordinaires soit valable, il faut que la distribution des résidus soit constante tout au long de la courbe de régression. La figure 16 illustre à gauche une situation où cette condition d'homoscédasticité est satisfaite. Les points (observations) se situent en dessus et en dessous de la ligne de régression et ceci de façon identique tout au long de la courbe. En revanche, le graphique à droite illustre

FIGURE 16 – Illustration de l'hétéroscédasticité



une situation d'hétérogénéité, car la variation des points autour de la courbe augmente avec la variable explicative (Années de formation).

Si la condition d'homoscédasticité n'est pas satisfaite, l'estimateur reste non biaisé, mais l'erreur standard estimée n'est pas valable (Cameron et Trivedi, 2005; Greene, 2005; Wooldridge, 2006). Si l'erreur standard n'est pas valable, l'intervalle de confiance qui en résulte peut être trop

large ou trop petit, ce qui peut changer la conclusion du test. Un simple test statistique qui est disponible dans tous les logiciels statistiques permet d'identifier une situation d'hétéroscédasticité et d'utiliser une version robuste des erreurs standard (White, 1980). Cette correction se fait très facilement et ne nécessite pas d'information ou de variables supplémentaires. Lorsqu'on travaille avec des données individuelles (personnes, entités employeuses, etc) une situation d'hétérogénéité est plutôt la norme que l'exception.

Points clés 9 (Heteroscédasticité).

- Lorsqu'on travaille avec des données individuelles (comme c'est le cas de Logib), il est fort probable que la condition d'homogénéité ne soit pas satisfaite
- Si la condition n'est pas satisfaite, l'inférence statistique avec les erreurs standard normales n'est pas valable.
- Il existe une version robuste des erreurs standard qui peut être calculée facilement et permet de faire l'inférence statistique correctement.

B.3 Spécification d'un modèle économétrique

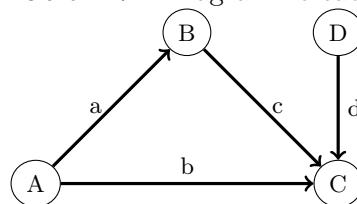
En économétrie, la spécification d'un modèle est aussi importante que le choix de l'estimateur. Dans cette annexe, l'outil visuel du diagramme causal est d'abord présenté, puis utilisé pour illustrer la discussion sur le type de variables à inclure ou non dans un modèle.

B.3.1 Diagramme causal et types d'effets

Les diagrammes causaux sont très utiles pour analyser la relation entre les facteurs indépendamment des variables disponibles ou du modèle à estimer. Ils sont utilisés pour identifier les variables à inclure dans un modèle lorsqu'un effet spécifique doit être estimé. Ils permettent également de comprendre comment l'inclusion ou l'exclusion d'une variable peut modifier l'interprétation de certains coefficients du modèle.

La figure 17 montre un diagramme causal présentant un exemple simple avec 4 variables génériques, qui permet d'illustrer les différents effets qu'un modèle peut estimer¹⁹.

FIGURE 17 – Diagramme causal



Les flèches font référence aux effets causaux, c'est-à-dire qu'il existe une relation cause-effet entre les deux variables. Dans notre exemple, la variable *A* a un effet causal sur les variables *B* et *C*,

19. Dans cette annexe, la version la plus simple de l'outil est présentée, car elle est suffisante pour la discussion sur le Logib. Une introduction plus détaillée peut être trouvée dans Morgan et Winship (2007)

mais pas sur D . La variable C est influencée par toutes les autres variables. Le tableau 8 montre les différents effets qu'on peut estimer, ainsi que le modèle qui permettrait de les estimer.

TABLEAU 8 – POSSIBLES EFFETS À ESTIMER AVEC UN MODÈLE ÉCONOMÉTRIQUE

Nom de l'effet	Taille de l'effet	Estimation de cet effet
Effet total de A sur C	$a + b \times c$	Coefficient β_1 du modèle* $C = \alpha + \beta_1 A + \varepsilon$
Effet direct de A sur C	a	Coefficient β_1 du modèle* $C = \alpha + \beta_1 A + \beta_2 B + \varepsilon$
Effet indirect de A via B sur C	$b \times c$	Impossible d'estimer avec un seul modèle. S'il n'y a pas d'autres effets, la différence des deux modèles ci-dessus peut être calculée. Sinon, un modèle économétrique plus avancé est nécessaire.

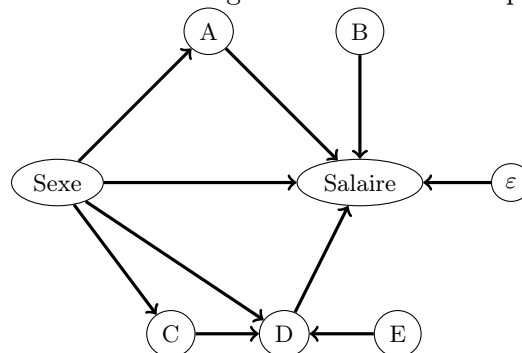
* = Inclure la variable D est optionnel, mais ne change rien à l'estimation de l'effet d'intérêt.

Le même coefficient est utilisé pour estimer à la fois l'effet total et l'effet direct, même si ces deux effets sont conceptuellement différents l'un de l'autre. L'inclusion d'une troisième variable (ici B) est responsable des différentes manières d'interpréter un coefficient. Cet exemple simple montre que l'inclusion d'un plus grand nombre de variables dans un modèle n'améliore pas seulement le pouvoir explicatif du modèle, mais peut également avoir des effets (involontaires) sur la façon dont le coefficient d'intérêt doit être interprété. Il est donc primordial de définir d'abord conceptuellement quel effet doit être estimé et ensuite de sélectionner les variables permettant d'estimer cet effet.

Pour Logib, la question qui se pose lors de l'inclusion d'une variable supplémentaire doit être de savoir comment cette inclusion pourrait affecter l'interprétation du coefficient d'intérêt (β_{Sexe}). Si la variable à inclure est (1) influencée par la variable sexe et (2) a un effet sur la variable dépendante (salaire), l'inclusion affectera le coefficient d'intérêt et probablement aussi la façon dont il sera interprété. Si l'une des conditions ci-dessus n'est pas réalisée, alors l'inclusion de la variable est moins problématique et il faut y prêter moins d'attention.

La figure 18 illustre cette problématique à l'aide d'un diagramme causal avec les variables *Sexe*, *salaire* ainsi que quelques variables génériques.

FIGURE 18 – Diagramme causal théorique



L'effet d'intérêt pour Logib est l'effet qui va de *Sexe* à *Salaire*. De ce fait, toutes les variables qui n'ont pas de lien avec *Sexe* ne sont en principe pas très importantes et leur inclusion ne devrait pas changer la valeur estimée ni l'interprétation du coefficient associé à la variable *Sexe*. Dans cet exemple, c'est le cas des variables B et E.

Depuis la variable *Sexe* il y a plusieurs chemins vers *Salaire* : d'une part il y a l'effet direct et d'autre part il y a des effets indirects via les variables A, C et D.

Pour choisir correctement la spécification d'un modèle micro-économétrique, il est important de bien comprendre les interactions possibles entre variables et surtout de bien définir l'effet qu'on souhaite estimer. Au sens de la LEg, l'interdiction de discrimination ne s'applique pas uniquement à la rémunération, mais également à la promotion et l'attribution de tâches entre autres. Il s'agit donc d'identifier aussi bien les effets directs (Strub (2004) utilise le terme *discrimination salariale directe*) comme les effets indirects via d'autres facteurs sous contrôle de l'entité employeuse.

Le modèle le plus simple possible serait

$$\ln(salaire_i) = \beta_0 + \beta_{Sexe} Sexe_i + \varepsilon_i \quad (3)$$

Dans ce modèle le coefficient β_{Sexe} représenterait l'effet total de la variable *Sexe* sur *salaire*, c'est-à-dire la somme de l'effet direct ($Sexe \rightarrow Salaire$) et les effets indirects ($Sexe \rightarrow A \rightarrow Salaires$ et $Sexe \rightarrow C \rightarrow D \rightarrow Salaires$). Ce modèle est le plus approprié pour calculer l'écart salarial total entre les femmes et les hommes.

L'équation suivante montre un modèle plus complet qui inclut également la variable B :

$$\ln(salaire_i) = \beta_0 + \beta_1 femme_i + \beta_2 B + \varepsilon_i \quad (4)$$

Ce deuxième modèle aura un pouvoir explicatif plus important, puisque la variable B aide à expliquer les variations du salaire. Par contre, au niveau de l'interprétation du coefficient d'intérêt (β_1) cette inclusion ne change absolument rien et même la valeur du coefficient n'est pas influencée par cette inclusion. Cet exemple fait référence directe à l'argument que le lien entre le pouvoir explicatif d'un modèle et le coefficient d'intérêt n'est de loin pas direct. Ceci ne veut pas dire que l'inclusion d'une variable supplémentaire ne change jamais la valeur ou l'interprétation du coefficient d'intérêt. Le modèle suivant inclut également la variable A :

$$\ln(salaire_i) = \beta_0 + \beta_1 femme_i + \beta_2 A + \beta_3 B + \varepsilon_i \quad (5)$$

L'inclusion de la variable A change l'interprétation du coefficient β_1 qui ne représenterait plus que la somme de l'effet direct ($Femme \rightarrow Salaire$) et l'effet indirect via C et D ($Femme \rightarrow C \rightarrow D \rightarrow Salaires$). Autrement dit, l'effet indirect via A a été enlevé de β_1 ou bien le chemin via A a été bloqué.

Est-ce que c'est bien d'exclure l'effet indirect via A ?

La question qui se pose maintenant est de savoir si l'exclusion d'un effet indirect est souhaitable ou non. La réponse à cette question réside dans la définition de l'effet à estimer. Dans le cadre des contrôles, il s'agit d'identifier les discriminations " sur le marché " pour reprendre la terminologie de Strub (2004) et non les discriminations " devant le marché ". En d'autres termes, il s'agit d'identifier uniquement les discriminations dont les entités employeuses sont responsables.

Si la variable fait référence à une variable qui ne relève pas de la responsabilité de l'entité employeuse, il convient de supprimer la partie de l'écart salarial qui lui est due. Un exemple typique d'une telle variable est la formation. Les entités employeuses ne peuvent et ne doivent pas être tenues pour responsables si les femmes ont moins d'années de formation que les hommes ²⁰.

Si, en revanche, la variable est sous la responsabilité des entités employeuses, son inclusion pourrait être problématique, voire contraire à l'objectif du modèle.

Pour illustrer cette possibilité, considérons une entité employeuse hypothétique accordant une gratification à tous les hommes mais pas aux femmes. L'inclusion de la gratification en tant que variable dans le modèle augmenterait sans aucun doute le pouvoir explicatif du modèle, mais elle serait toujours contraire à l'objectif du modèle. L'effet lié au sexe estimé ne capterait que la discrimination sans la gratification.

Bien sûr, cet exemple est extrême et dans la réalité, la frontière entre les variables à inclure et celles à ne pas inclure est beaucoup moins claire.

En règle générale, il est conseillé d'inclure les variables qui font référence à des différences entre les femmes et les hommes pour lesquelles les entités employeuses ne devraient pas être tenues responsables. En revanche, les variables qui peuvent inclure une discrimination de la part de l'entité employeuse sont plus délicates et il est conseillé de ne pas les inclure.

Il existe encore des variables qui ne peuvent être classées dans aucune des deux catégories. Par exemple, [Paschoud \(2015\)](#) mentionne qu'il peut y avoir des différences entre les femmes et les hommes en ce qui concerne les objectifs de carrière qui peuvent influencer les promotions. Si c'était le seul facteur influençant les promotions, il serait alors pertinent d'inclure la position hiérarchique dans le modèle, car les entités employeuses ne peuvent être tenues responsables d'éventuelles différences de comportement ou d'objectifs. En revanche, s'il existe en même temps une discrimination à l'égard des femmes dans les promotions, la position hiérarchique aurait une partie sous la responsabilité de l'entité employeuse et une partie hors de sa responsabilité. Dans ces cas, il est particulièrement difficile de savoir s'il est pertinent ou non d'inclure la variable, et il s'agit finalement d'un choix politique. Il est simplement important de garder à l'esprit que l'inclusion d'une variable peut changer la signification de l'effet estimé.

20. Cela ne signifie pas qu'il n'y a pas de discrimination au niveau de la société, mais seulement que cette discrimination n'est pas imputable aux entités employeuses ou au marché du travail



Bureau de l'égalité
entre les femmes
et les hommes - BEFH