

Dynamiques individuelles de rémunération: des écarts genrés ?

Pierre Pora Lionel Wilner

Insee-CREST

Journées de Méthodologie Statistique – Juin 2018

Introduction

- Une convergence considérable entre hommes et femmes en matière d'éducation et de participation au marché du travail au cours des dernières décennies (Goldin 2014)...
- Mais des écarts de rémunération qui reste importants et diminuent (très) lentement : en 2014 (1995) l'écart moyen (GPG) est de 24% (27%) en revenu salarial et 17% (18%) en salaire horaire
- Deux explications proposées pour cette persistance :
 - Les naissances accentuent les contraintes qui pèsent sur les femmes et déplacent leurs préférences vis-à-vis de la famille et du travail par rapport à celles des hommes, ce qui explique une part importante et croissante du GPG (Kleven et al. 2018)
 - La distribution des rémunération a des queues épaisses, de sorte que la sous-représentation des femmes dans un petit sous-ensemble de positions très bien rémunérées peut expliquer une part importante du GPG (Fortin et al. 2017)

Notre contribution

- Ces deux canaux sont-ils liés ?
- La sous-représentation des femmes au sein des positions les mieux rémunérées est-elle cohérente avec ce que l'on imagine être un *plafond de verre* ?
- Pour répondre à ces questions :
 - Des données longitudinales issues de sources administratives
 - Une approche descriptive et non-paramétrique initialement développée pour l'étude des dynamiques individuelles de revenu salarial

Notre contribution

- Il y a bien des écarts massifs entre hommes et femmes parmi les individus les mieux rémunérés, mais...
- En l'absence de naissance, parmi les individus les mieux rémunérés les femmes ont des progressions (en salaire horaire) aussi rapides que celles des hommes
- L'effet (défavorable) des enfants sur les carrières des femmes n'est pas plus élevé pour les femmes les mieux rémunérées que pour les autres
- Les différences entre hommes et femmes au sommet de la distribution de salaire apparaissent dès l'entrée sur le marché du travail
- Peut-être du fait du coût anticipé des interruptions de carrière

Revue de littérature

- Des écarts hétérogènes le long de la distribution des rémunérations :
 - Approche par les régressions quantiles : Albrecht et al. 2003 ; Gardeazabal et Ugidos 2005 ; Arulampalam et al. 2007 ; Rica et al. 2008 ; Etienne et Narcy 2010 ; Christofides et al. 2013 . En général le GPG est plus large au sommet : plafond de verre (statique)
 - Les positions les mieux rémunérées peuvent expliquer une part importante du GPG agrégé : Fortin et al. 2017
 - Au sommet de la distribution, les femmes ont des progressions moins rapides / des pertes plus sévères que les hommes : plafond de verre (dynamique) (Weinberger 2011) et plancher de papier (Guvenen et al. 2014)
- Des divergences en termes de rémunération et d'offre de travail liées aux naissances :
 - Plusieurs approches destinées à prendre en compte l'endogénéité des décisions de fertilité : Rosenzweig et Wolpin 1980 ; Korenman et Neumark 1992 ; Angrist et Evans 1998 ; Simonsen et Skipper 2006 ; Wilner 2016
 - Mais ce n'est peut-être pas si important que ça : Kleven et al. 2018
 - Ces divergences peuvent expliquer une grande part du GPG agrégé : Bertrand et al. 2010 ; Wilner 2016 ; Adda et al. 2017 ; Juhn et McCue 2017 ; Kleven et al. 2018

Les données : le panel DADS-EDP

- Des données longitudinales qui couvrent tous les salariés du secteur privé (hors agriculture)
- En deux parties :
 - Le panel DADS : c'est la version longitudinale des DADS. Renseigne sur le revenu salarial, le volume de travail, la catégorie socioprofessionnelle et le secteur d'activité
 - Le fichier biographique issu de l'EDP : données longitudinales issues des bulletins d'état-civil et des recensements (RP et EAR). Naissances, mariages, diplôme.
- Taux de sondage : 1/12 des salariés du secteur privé de 2002 à 2015
- Censure à 1/8 du Smic annuel
- Echantillon de travail : individus "relativement stables" (salariés en $t - 1$ et au moins deux années entre $t - 5$ et $t - 2$) : 6 millions d'observations au niveau individu-années qui correspondent à plus de 900 000 individus.

Méthode : normalisation

- Notre approche est adaptée de celle de Guvenen et al. 2016 ; Guvenen et al. 2017
- On commence par régresser le (log) salaire horaire sur des indicatrices d'âge, période et cohorte :

$$\ln(\tilde{w}_{it}) = \ln(\tilde{w}_0) + \sum_c \alpha_c \mathbb{1}_{\text{cohort}_i=c} + \sum_a \beta_a \mathbb{1}_{\text{age}_{it}=a} + \sum_j \gamma_j \mathbb{1}_{t=j} + \epsilon_{it}$$

- L'identification requiert une normalisation supplémentaire (Deaton et Paxson 1994 ; Deaton 1997)
- On fait de même pour le volume de travail
- Cela nous permet de considérer un salaire horaire, un volume de travail et un revenu salarial normalisés : $w_{it} = \tilde{w}_{it} / \exp(\beta_a^w)$ $l_{it} = \tilde{l}_{it} / \exp(\beta_a^l)$
 $y_{it} = \tilde{y}_{it} / \exp(\beta_a^w + \beta_a^l)$

Méthode : imputation du rang

- On introduit une mesure du salaire horaire récent :

$$W_{it} = \frac{\sum_{\tau=t-5}^{t-1} \tilde{w}_{it}}{\sum_{\tau=t-5}^{t-1} \exp(\hat{\beta}_a) \mathbf{1}_{age_{i,\tau}=a}}$$

- Pour chaque année croisée avec chaque cohorte : on ordonne les salariés selon W_{it} pour créer 21 groupes P0-P5, ..., P90-P95, P95-P99, P99-P100 (les très hauts salaires)
- On peut ainsi comparer des hommes et des femmes qui ont un passé récent sur le marché du travail relativement similaire

Méthode : comparaison entre hommes et femmes

- On compare le revenu salarial des hommes et des femmes au sein de chacune de ces cases en $t + k$ ($k = 0 \dots 10$)
- On inclut les individus en-dessous de 1/8 Smic annuel, ou absent des données, comme ayant un revenu salarial nul
- Une décomposition comptable du revenu salarial :

$$\begin{aligned} \ln(\mathbb{E}[y_{i,t+k}]) &= \underbrace{\ln(\mathbb{P}(d_{i,t+k} = 1))}_{\text{Participation}} + \underbrace{\ln(\mathbb{E}[l_{i,t+k} | d_{i,t+k} = 1])}_{\text{Volume de travail}} \\ &\quad + \underbrace{\ln\left(\frac{\mathbb{E}[w_{i,t+k} l_{i,t+k} | d_{i,t+k} = 1]}{\mathbb{E}[l_{i,t+k} | d_{i,t+k} = 1]}\right)}_{\text{Salaire horaire}} \end{aligned}$$

Résultats : part des femmes le long de la distribution de salaire récent

Au fil d'une carrière, les femmes sont de moins en moins (de plus en plus) nombreuses au sommet (en bas) de la distribution de salaire récent

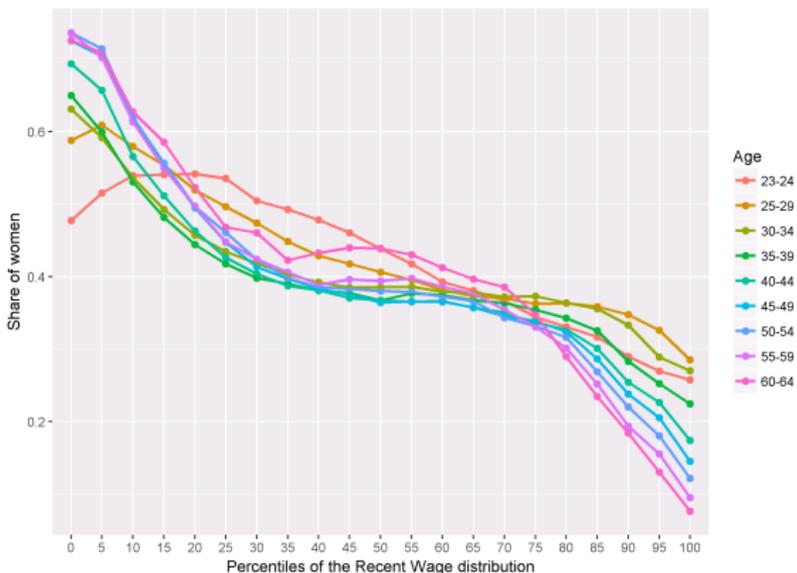


Figure 1 – Share of women along the recent wages distribution

Résultats : décomposition de l'écart de revenu salarial le long de la distribution

- Le GPG en revenu salarial est une fonction décroissance du salaire récent
- L'écart en salaire horaire a une forme en U qui s'accroît au cours du temps
- L'écart en salaire horaire est très élevé parmi les très hauts salaires (~ 20 log-points)

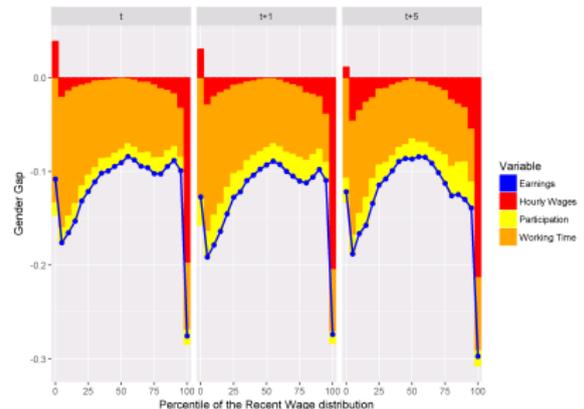


Figure 2 – Décomposition du GPG le long de la distribution de salaire récent

Résultats : progression de salaire et interruptions de carrière

- On compare en $t + k$ ($k > 0$) hommes et femmes selon qu'ils ont ou non participé au marché du travail en t
- On pousse plus loin la décomposition de l'écart en salaire horaire pour les participants : salaire horaire en t vs. progression de t à $t + k$
- Parmi les participants : de petites différences en salaire horaire sauf pour les très hauts salaires
- Parmi les très haut salaires les différences en $t \gg \gg$ les différences de progression
- Les sorties de l'emploi salarié sont plus persistantes pour les femmes en bas de la distribution que pour les hommes
- Très grand écart entre hommes et femmes pour les très hauts salaires ayant connu une interruption de carrière

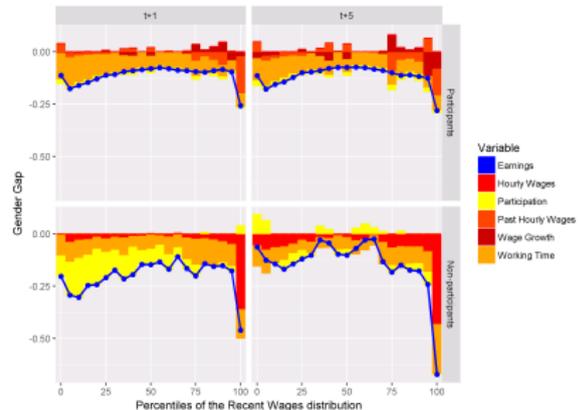


Figure 3 – Décomposition du GPG le long de la distribution de salaire récent : par participation

Résultats : les écarts autour d'une naissance

- On compare les salariés qui ont eu un enfant en t avec ceux qui n'en ont pas eu entre t and $t + k$ ($k = 0 \dots 10$)
- De très petits écarts pour les hommes sauf aux extrémités de la distribution
- De très grands écarts pour les femmes...
- Qui décroissent avec le salaire récent...
- Et sont en premier lieu tirés par la participation et le volume de travail
- L'écart est très petit parmi les femmes les mieux rémunérées

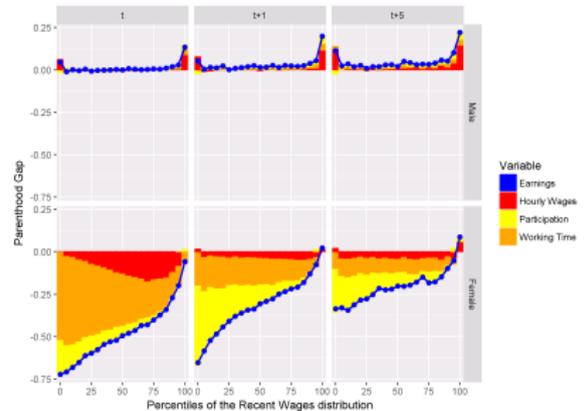


Figure 4 – Ecart autour d'une naissance le long de la distribution de salaire récent : par sexe

Résultats : les écarts autour d'une naissance

- Une autre visualisation du même résultat
- Parmi les très hauts salaire, le GPG est légèrement plus élevé parmi les parents récents que parmi les autres individus...
- Mais la différence semble trop petite pour expliquer l'ampleur de l'écart agrégé

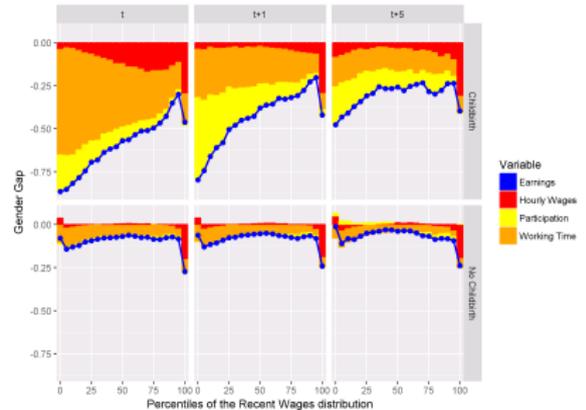


Figure 5 – Décomposition du GPG le long de la distribution de salaire récent : par naissance

Utiliser la distribution des progressions de salaire

- On regarde les individus qui ont participé en t et $t + k$ ($k > 0$)
- On définit la croissance (normalisée) de salaire horaire
$$\delta^k w_{it} = \ln(w_{i,t+k}) - \ln(w_{it})$$
- Dans chaque cellule on estime :

$$\begin{aligned} \delta^k w_{it} &= \zeta \text{ femme}_i \\ &+ \eta (1 - \text{femme}_i) \text{ naissance}_{i,t+1} \\ &+ \kappa \text{ femme}_i \text{ naissance}_{i,t+1} \\ &+ \lambda X_{i,t} + \nu_{it} \end{aligned}$$

- Aussi avec des régressions quantiles

Résultats : croissance à 1 an

- Les femmes qui n'ont pas donné naissance en $t + 1$ n'ont pas une progression plus lente que celle des hommes
- Y compris pour les mieux rémunérées : pas de plafond de verre (dynamique)
- Les hommes sont peu affecté par les naissances
- Pour les femmes, les pénalités liées aux naissances sont hétérogènes le long de la distribution de salaire
- Mais elles ne sont pas plus élevées parmi les très hauts salaires que pour les autres
- Résultats robustes au passage aux quantiles

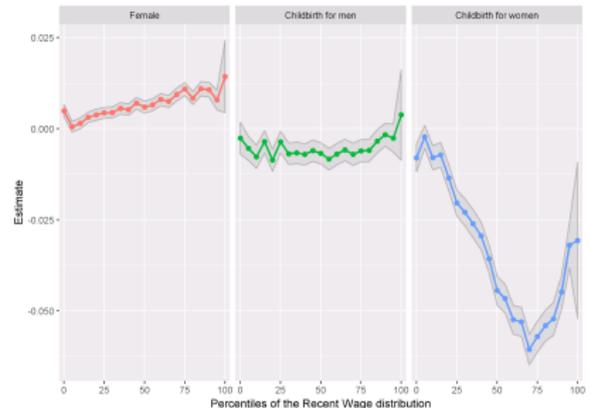


Figure 6 – Spécification OLS

Résultats : croissance à 1 an

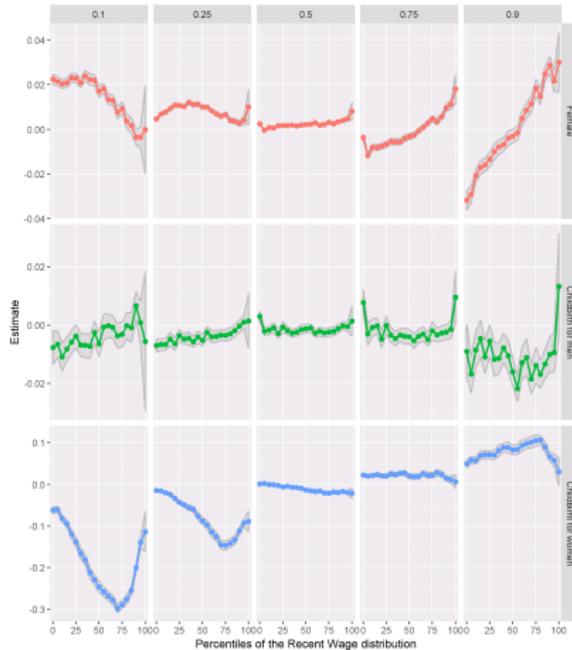


Figure 7 – Spécification QR

Résultats : croissance à 5 ans

- Une dimension selon laquelle les très hauts salaires diffèrent du reste de la distribution : le coût des interruptions de carrière
- Ces interruptions semblent plus coûteuses pour les très hauts salaires
- Régression quantiles : beaucoup d'hétérogénéité dans l'effet, vraisemblablement liée au motif de l'interruption
- ⚠ Ce n'est pas le vrai coût des interruptions de carrière

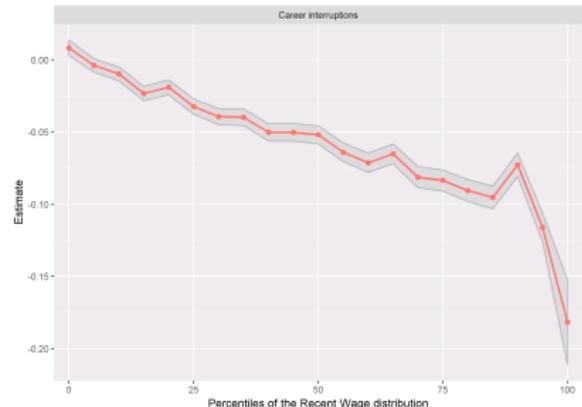


Figure 8 – Spécification OLS : interruptions de carrière

Résultats : croissance à 5 ans

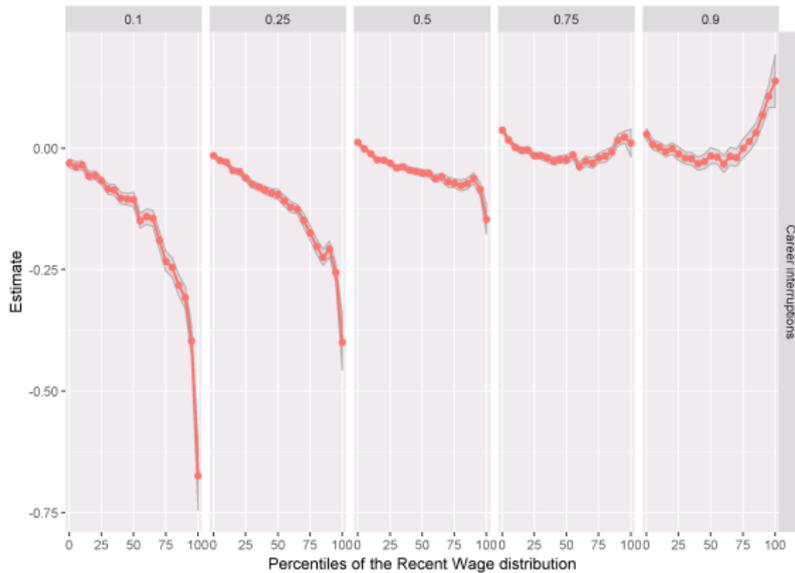


Figure 9 – Spécification QR : interruptions de carrière

Conclusion

- Les très hauts salaires sont différents : les écarts entre hommes et femmes en niveau y sont beaucoup plus prononcés
- Ce fait stylisé n'est vraisemblablement pas le résultat de progressions plus lentes des femmes, ou de pénalités liées au maternité plus prononcées qu'ailleurs
- Il émerge à (avant ?) l'entrée sur le marché du travail
- Peut-être du fait du coût anticipé des interruptions de carrière
- Cela peut expliquer les résultats décevants de certaines politiques publiques destinées à augmenter la représentation des femmes au sommet de la distribution de salaire

Merci pour votre attention