Salaires au long du cycle de vie et Capital Humain: Sélection Données Manquantes

Laurent Gobillon *(Paris School of Economics, CNRS)*

Thierry Magnac *(Toulouse School of Economics)*

Sébastien Roux *(Insee, Ined and Crest)*

JMS, Juin 2018

Travail préliminitaire, Commentaires bienvenus

Motivation

- Les dynamiques de revenus au long du cycle de vie affectent les inégalités de long terme
- Développement récents et importants sur les dynamiques de revenus à la Mincer (1974) incluant beaucoup d'hétérogénéité (Browning et al., 2012, Polachek et al., 2015, Magnac et al., 2018)

- Une question négligée : Données manquantes, les travailleurs peuvent sortir temporairement ou définitivement de l'emploi

  - Problème de spécification :
    - une accumulation différente du capital humain en non-emploi
  - Problème de sélection :
    - panel équilibré : Résultats spécifiques à une sous-population
    - panel non équilibré: Hypothèse "manquant au hasard"

Contributions

- Contribution théorique: Extension du modèle d'investissement en capital humain post-éducation de Magnac, Pistolesi and Roux (2018, JPE), fondé sur Ben Porath (1967)
  - Prise en compte dans les équations de salaires des investissements en capital humain effectués en dehors de l'emploi dans le secteur privé

- Contribution empirique:
  - Utilisation d'un modèle à facteurs linéaires individuels pour contrôler de la sélection, Hypothèse "Manquant au hasard conditionnellement aux facteurs"(MCFAR)
  - Hétérogénéité individuelle multi-dimensionnelle

  (à la fois avantages et limites)

Résultats

- Application:
  - Panel sur longue période, observation des individus sur leurs 20 premières années dans le secteur privé
  - Données manquantes lorsques les individus n'y sont pas observés

- Résultats:
  - Biais substantiel des rendements de l'expérience due à l'erreur de spécification et aux effets de sélection
  - Biais principalement lié à la non prise en compte de l'évolution en dehors de l'emploi (erreur de spécification)
  - Indicateurs de dispersion plus élevés lorsqu'on corrige l'erreur de spécification et les effets de sélection
  - Estimations des profils de salaires et d'indicateurs d'inégalité pour différents groupes de travailleurs (stratifiés par éducation ou nombre d'interruptions)

Littérature

- Investissement en capital humain (Ben Porath, 1967): intérprétation structurelle aux équations de salaire
  Non linéaire: Polachek et al (2015); linéaire: Magnac et al (2018)

  Identification de paramètres reflétant les capacités de gain (to earn) et d'apprentissage (to learn) (Browning et al., 1999, and Rubinstein and Weiss, 2006). Ici, spécifiques à chaque individu.

- Les processus d'accumulation du capital humain diffèrent selon le secteur
  (see Blundell et al., 2016, for part-time/full-time evidence or Mincer and Polachek, 1974 for labor supply of women)

- Modèle à facteurs et sélection:
  Aakvik et al. (2005), Gobillon and Magnac (2016) pour une généralisation des méthodes de différence en différence.

Plan

1. Données et Faits stylisés
2. Modèle d'investissement en capital humain post-éducation avec deux secteurs
3. Modèle empirique et Stratégie d'estimation
4. Résultats
5. Conclusion et perspectives

Données

- *Panel DADS Grand Format - EDP*: Individus nés les 4 premiers jours d'Octobre d'une année paire.
  Mélange de données administrative et du recensement.
- Emplois dans le secteur privé de 1985 à 2011. Observations manquantes en 1990.
- Information sur les caractéristiques de l'emploi (Rémunération, jours rémunérés, Temps Complet), **Salaire journalier**.
- Éducation mesurée à partir de l'EDP, regroupée en 4 catégories:
  Inférieur au bac, niveau bac, niveau licence, and master et au-delà.

Sélection des observations

- Chaque année, les individus peuvent être employés dans le secteur e ou être dans le secteur n, qui agrège le non-emploi, le travail indépendant, l'emploi à temps partiel ou dans le secteur public.
- Un travailleur est dans le secteur n (i-e pas dans e) l'année t si son salaire journalier est inférieur à 80% du Smic, s'il occupait un emploi à temps partiel ou s'il a été rémunéré moins de 6 mois

Sélection des individus

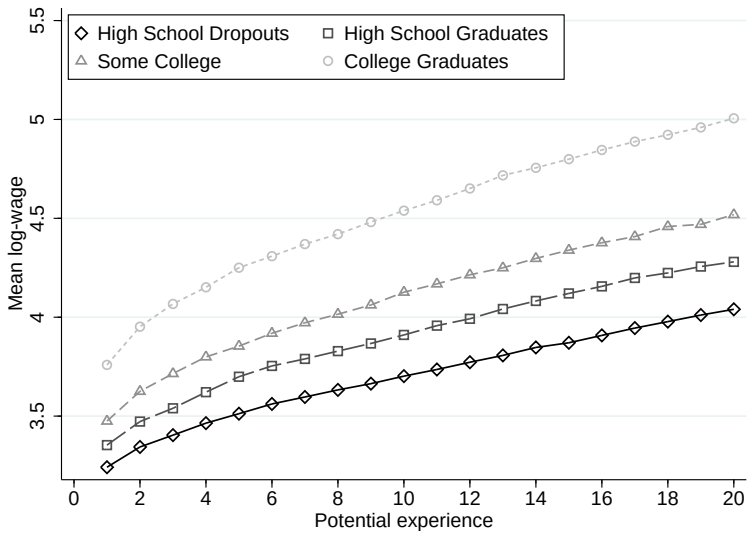
- Hommes âgés de 16 à 30 à leur entrée sur le marché du travail entre 1985 et 1992 (hors 1990):
  178,098 observations pour 12,212 hommes
- Individus dont les salaires dans le secteur e sont observés au moins 15 années: 137,315 observations pour 7,004 hommes.
- A noter que seulement 1,219 hommes sont continûment présents sur la période totale (i-e de leur entrée à 2011).

Statistiques descriptives sur les interruptions

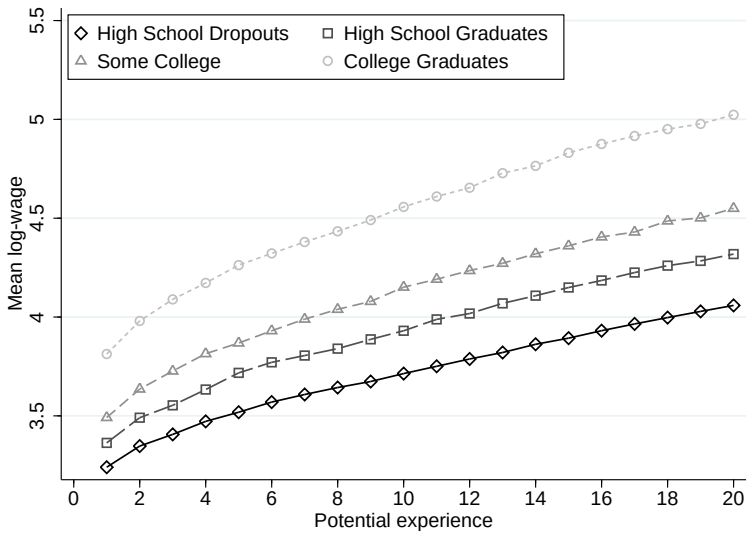
Nombre d' interruptions	Nombre d' individus	Proportion en interruption	Durée cumulée en interruption	Nombre moyen d' interruptions
Tous	7004	0.154	3.7	1.44
0	1219	0.000	0.0	0.00
1	2279	0.110	2.6	1.00
2	1933	0.196	4.7	2.00
3	1050	0.261	6.3	3.00
4	383	0.321	7.7	4.00
5	118	0.355	8.5	5.00
6	22	0.378	9.3	6.00

Log-Salaire fonction de l'expérience potentielle

All

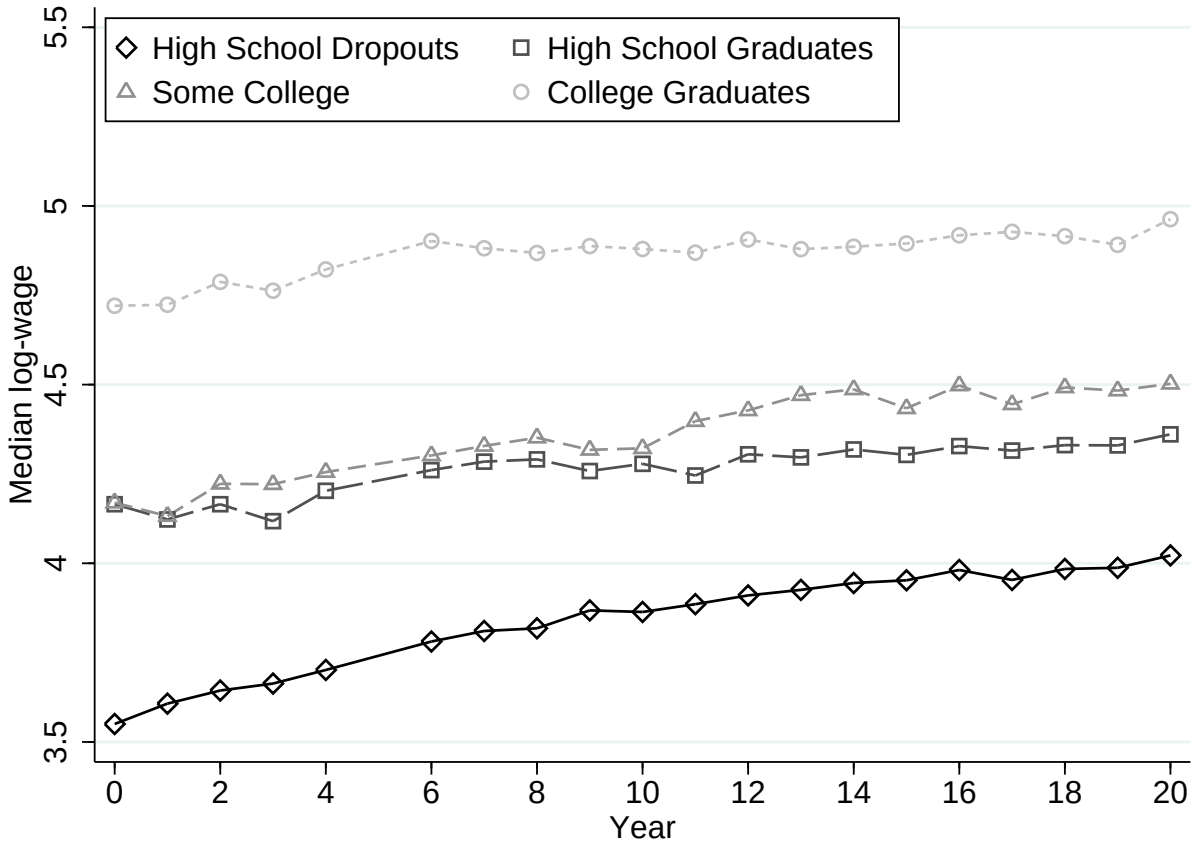


Selected



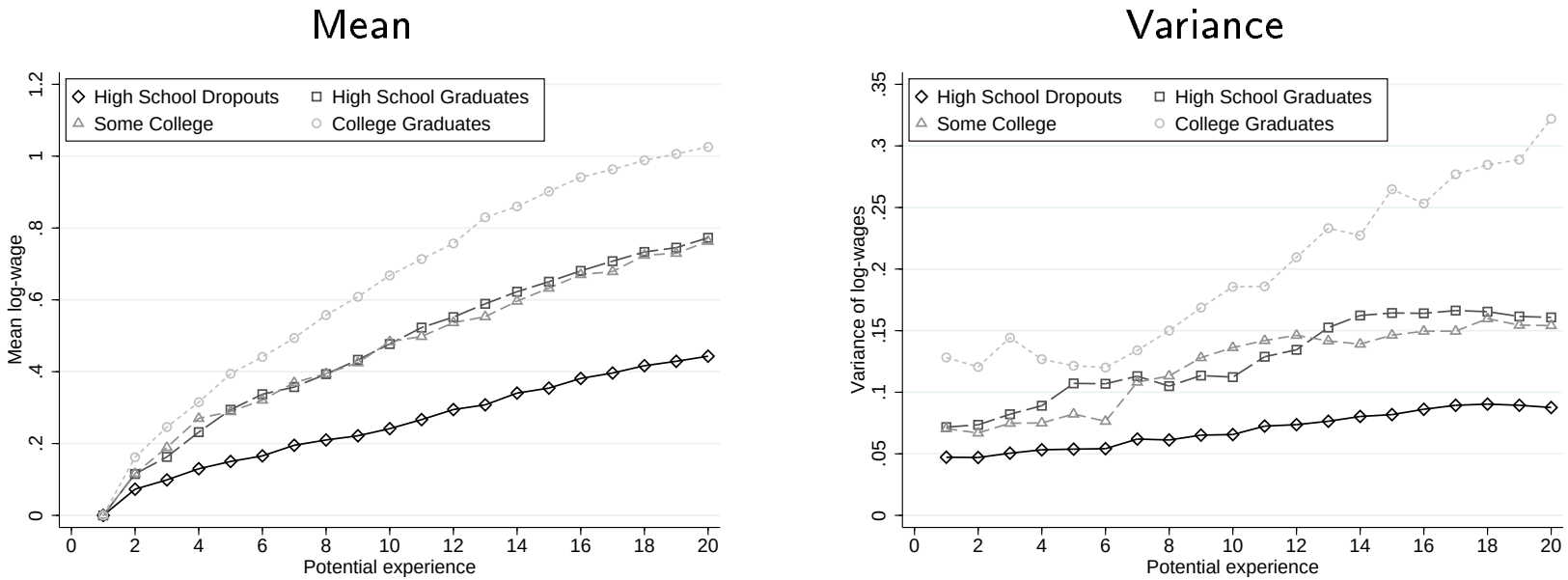
Note: "All": all individuals of our seven cohorts; "Selected": individuals of our seven cohorts who are employed at least 15 years during their timespan.
Wages are here deflated by the World Bank consumer price index.

Prix du capital humain par diplôme



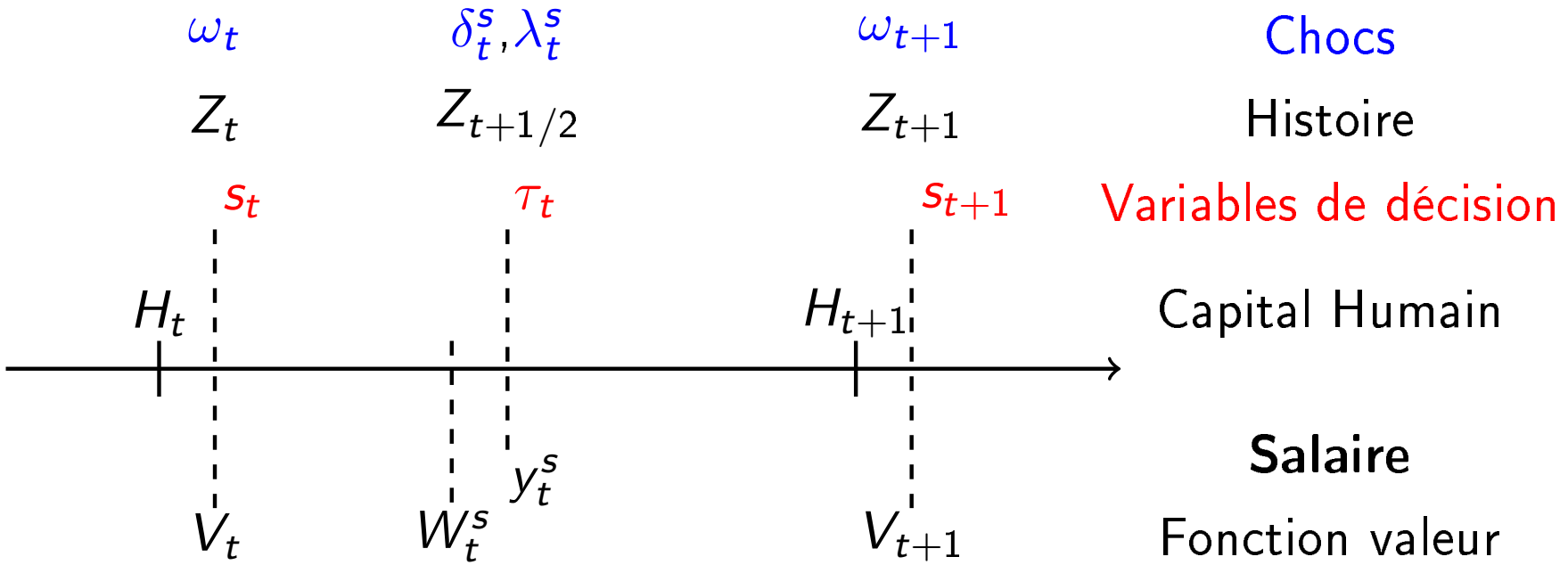
Note: Median log-wage computed on the subsample of observations such that individuals are 50-55 years old, cf. Bowlus and Robinson (2015).

Moyenne et variance des log-salaires déflatés par les prix du capital humain, par diplôme



Note: individuals entering the labour market between 1985 and 1992 who are employed at least 15 years in our panel data. Means are normalized at zero at entry.

Modèle d'investissement en capital humain



$$\begin{aligned} \ln y_t^s &= \ln H_t + \delta_t^s - \tau_t \\ \ln H_{t+1} &= \ln H_t + \rho^s \tau_t - \lambda_t^s \\ u_t &= \ln y_t^s - c\frac{(\tau_t)^2}{2} + \omega_t \mathbf{1}\{s = e\} \end{aligned}$$

β taux d'escompte

Résolution par induction inverse, à partir de la valorisation κ du capital humain à une date arbitraire dans le futur ▸ Modèle Détaillé

La forme réduite

Sous l'hypothèse que $\tau_t^{s_t} > 0$ sur la période considérée, le log-salaire s'écrit :

$$\ln y_t = \eta_0 + \eta_1 t + \eta_2 \beta^{-t} + \eta_3 x_t^{(3)} + \eta_4 x_t^{(4)} + v_t$$

où $(\eta_0, \eta_1, \eta_2, \eta_3, \eta_4)$ sont fonctions des paramètres individuels $(\ln H_{t_0}, \rho^e, \rho^s, \kappa, c)$
et $x_t^{(3)}$, $x_t^{(4)}$ et v_t sont définis par :

$$\begin{aligned} x_t^{(3)} &= \sum_{k=0}^{K_t-1} \left(t_{2k+2} - t_{2k+1}\right) \\ x_t^{(4)} &= \sum_{k=0}^{K_t-1} \left(\beta^{-t_{2k+2}+1} - \beta^{-t_{2k+1}+1}\right) \\ v_t &= \delta^{s(t)}(t) - \sum_{l=t_0}^{t-1} \lambda^{s(l)}(l) \end{aligned}$$

▸ Formalisation Trajectoires

Interprétations

- $\eta_0 + \eta_1 t + \eta_2 \beta^{-t}$ est l'évolution des salaires potentiels, si le travailleur était resté dans le secteur e
- $\eta_3 x_t^{(3)} + \eta_4 x_t^{(4)}$ reflète la différence d'accumulation du capital humain résultant du temps passé hors du secteur e
  - $x_t^{(3)}$ est le temps passé dans le secteur n avant t
  - $x_t^{(4)}$ est le temps passé en n avant t, chaque date passée l étant pondérée par $1/\beta^l$
- Les chocs v_t reflètent l'évolution des prix du capital humain et les chocs de dépréciation

Spécification

- Dans les données, les salaires ne sont observée que dans le secteur e aux dates $1, ..., T$.
- Les paramètres structurels sont spécifiques à chaque individu, sauf β, homogène.
- Equation de salaire:

$$\begin{aligned} \ln y_{it} &= \eta_{i0} + \eta_{i1} t + \eta_{i2} \beta^{-t} + \eta_{i3} x_{it}^{(3)} + \eta_{i4} x_{it}^{(4)} + v_{it}, \\ &= x_{it} \eta_i + v_{it}, \end{aligned}$$

où $v_{it} = \delta_{it} - \sum_{l=t_0}^{t-1} \lambda_l^{s_{il}}$, s_{il} étant le secteur choisi par i à la date l, $x_{it} = (1, t, \beta^{-t}, x_{it}^{(3)}, x_{it}^{(4)})$ et $\eta_i = (\eta_{ij})_{j=0,.,4}$.

Identification sous sélection exogène

- Identification de η_{i3} et η_{i4}: mobilité entre secteurs requise

  - Sans ou seulement une transition $e \rightarrow n$, $x_{it}^{(3)} = x_{it}^{(4)} = 0$:
  η_{i3} et η_{i4} **ne sont pas identifiés**, fixés à 0.

  - Avec un seul aller-retour $e \leftrightarrow n$ (i-e 2 ou 3 transitions), $x_{it}^{(3)}$ et $x_{it}^{(4)}$ proportionnels:
  η_{i3} et η_{i4} **ne sont pas séparément identifiés**. η_{i4} fixé à 0.

  - Besoin de 2 aller-retours au moins $e \leftrightarrow n$ (i-e 4 transitions) pour identifier séparément η_{i3} et η_{i4} .

- Identification de η_{i0}, η_{i1} et η_{i2} requiert suffisamment d'observations en emploi (seulement), non affecté par la sous-estimation de η_{i3} et η_{i4}.

Manquant au hasard conditionnellement aux facteurs

- Spécification des chocs sous forme de facteurs linéaires

$$\begin{aligned}\omega_{it} &= \varphi_t^{(\omega)}\theta_i^{(\omega)} + \widetilde{\omega}_{it},\\ \delta_{it}^s &= \varphi_t^{(\delta),s}\theta_i^{(\delta),s} + \widetilde{\delta}_{it}^s,\\ \lambda_{it}^s &= \varphi_t^{(\lambda),s}\theta_i^{(\lambda),s} + \widetilde{\lambda}_{it}^s.\end{aligned}$$

- Supposant $\theta_i^{(\lambda),e} = \theta_i^{(\lambda),n} = \theta_i^{(\lambda)}$:

$$\ln y_{it} = x_{it}\eta_i + \varphi_t\theta_i + \tilde{v}_{it}$$

avec $\theta_i = (\theta_i^{(\delta)}, \theta_i^{(\lambda)})$

- Manquant au hasard conditionnellement aux facteurs (MCFAR):

- Sélection exogène conditionnellement aux facteurs individuels
- Variables explicatives exogènes conditionnellement aux facteurs individuels

▸ MCFAR details

Stratégie d'estimation

- Équation en forme réduite:

$$\ln y_{it} = x_{it}\eta_i + \varphi_t\theta_i + \tilde{v}_{it}$$

où:

$$\begin{aligned} x_{it} &= \left(1, t, \beta^{-t}, x_{it}^{(3)}, x_{it}^{(4)}\right) \\ \eta_i &= \left\{\eta_{i0}, \eta_{i1}, \eta_{i2}, \eta_{i3}, \eta_{i4}\right\}' \end{aligned}$$

- Maximisation de la pseudo-vraisemblance, équivalente à minimiser

$$C(\theta, \varphi, \eta) = \sum_{i,t|s_{it}=e} \left(y_{it} - x_{it}\eta_i - \varphi_t\theta_i\right)^2$$

- Restrictions de normalisation:

$$\varphi \perp \left(e_T, x^{(1)}, x^{(2)}\right), \frac{\varphi'\varphi}{T} = I$$

où $x^{(1)} = (1, ..., T)'$ and $x^{(2)} = (1, ..., \beta^{-T})'$

Algorithme EM par Itération

Étape k:

1. Regresse y_{it} sur $\left(x_{it}, \varphi_t^{(k-1)}\right)$ pour chaque i où $s_{it} = e$
   $\Longrightarrow \eta_i^{(k)}, \theta_i^{(k)}$.
2. $\widehat{y}_{it} = x_{it}\eta_i^{(k)} + \varphi_t^{(k-1)}\theta_i^{(k)}$ où $s_{it} = n$, $\widehat{y}_{it} = y_{it}$ où $s_{it} = e$
   Étape Espérance
3. Modèle à facteur: $\widehat{y}_{it} - x_{it}\eta_i^{(k)} = \varphi_t\theta_i + \widetilde{\widetilde{v}}_{it}$
   $\Longrightarrow \varphi_t^{(k)}$ à partir de Bai (2009)

À chaque étape, on a

$$C\left(\theta^{(k)}, \varphi^{(k)}, \eta^{(k)}\right) < C\left(\theta^{(k-1)}, \varphi^{(k-1)}, \eta^{(k-1)}\right)$$

ce qui assure sa convergence

▸ Stop Criterium

Décomposition de la variance

Spécification: $\left(1, t, \beta^{-t}, x_{it}^{(3)}, x_{it}^{(4)}\right)$, 2 factors, $\beta = 0.95$.

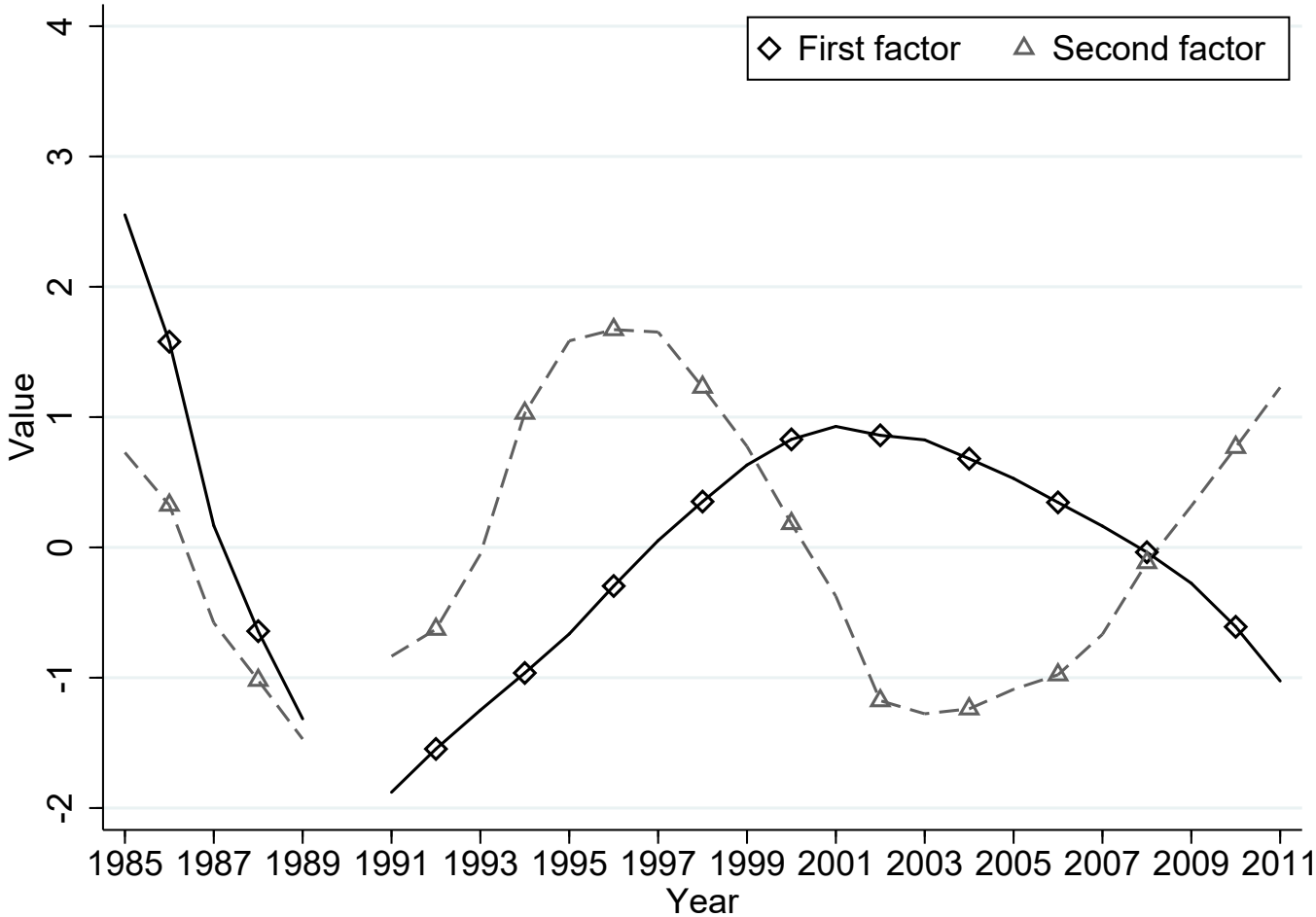
4 composantes:

- Expérience potentielle: $\eta_{i0} + \eta_{i1} t + \eta_{i2} \beta^{-t}$
- Effets des interruptions: $\eta_{i3} x_{it}^{(3)} + \eta_{i4} x_{it}^{(4)}$
- Facteurs inobservés: $\varphi_t \theta_i$
- Résidu $\tilde{v}_{it}$

	Variance	Correlation				
		Log-wage	Potential experience effect	Interruption effect	Effect of factors	Residual
Log-wage	0.147	1.000				
Pot. exp.	0.671	0.425	1.000			
Interruptions	0.501	0.004	-0.848	1.000		
Factors	0.059	0.051	-0.233	-0.046	1.000	
Residual	0.008	0.229	0.000	-0.000	0.000	1.000

Factors

Males: Value of factors as a function of time, main specification



Note: Factors are function of calendar time, not potential experience.

Profil en cycle de vie des salaires potentiels

- Focus sur les effets de l'expérience potentielle sur les salaires:

$$y_{it}^{c} = \widehat{\eta}_{i0}^{c} + \widehat{\eta}_{i1}^{c}\, t + \widehat{\eta}_{i2}^{c} \beta^{-t}$$

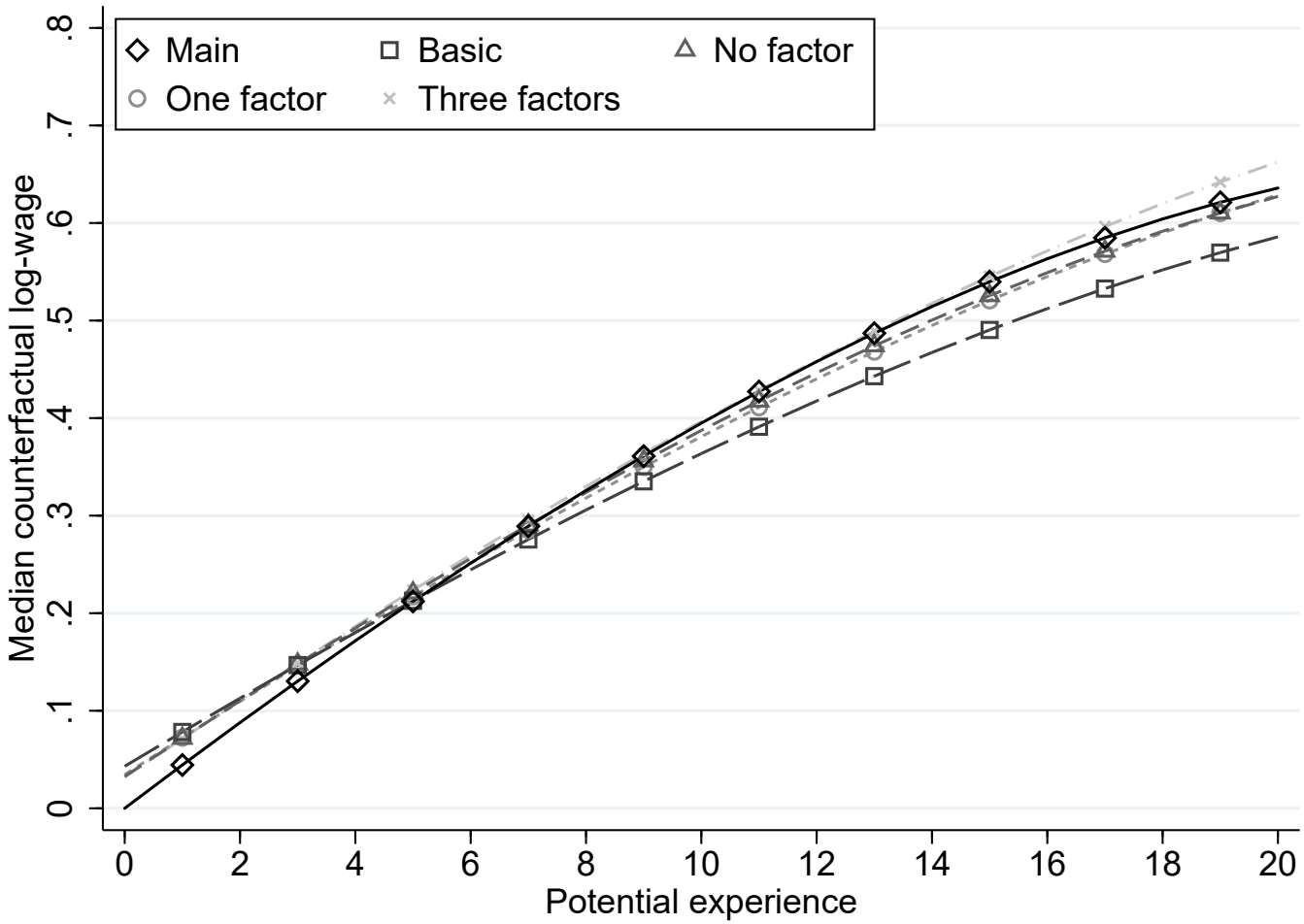
- Pour rendre comparables les différentes cohortes
$(\eta_{i0}^{c}, \eta_{i1}^{c}, \eta_{i2}^{c}) = (\eta_{i0} + \eta_{i1}\left(T - t_{0i}\right), \eta_{i1}, \eta_{i2} \exp\left[\beta^{-(T - t_{0i})}\right])$

- Différente spécifications sont explorées:
  1. basic: pas d'effets des interruptions, pas de facteur. Descriptive.
  2. no factor: effets des interruptions, pas de facteurs.
  3. one factor: effets des interruptions, 1 facteur.
  4. main: effets des interruptions, 2 factors.

- La moyenne de $\hat{\eta}_{i0}^{c}$ est normalisée pour que $\bar{y}_{i0}^{c}$ soit 0 pour la spécification principale (main).

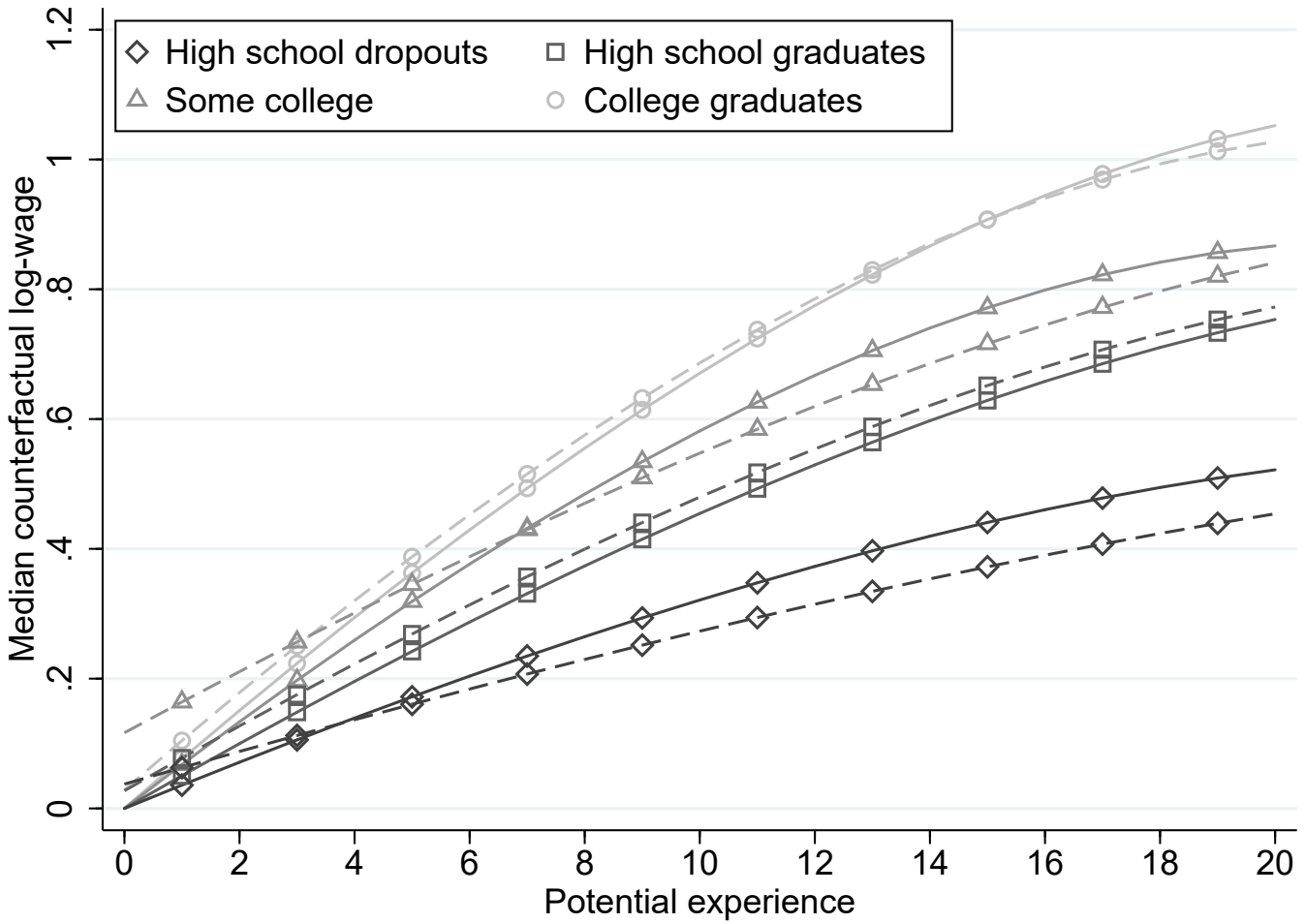
Profil des salaires potentiels - par spécification

Median counterfactual log-wage as a function of potential experience



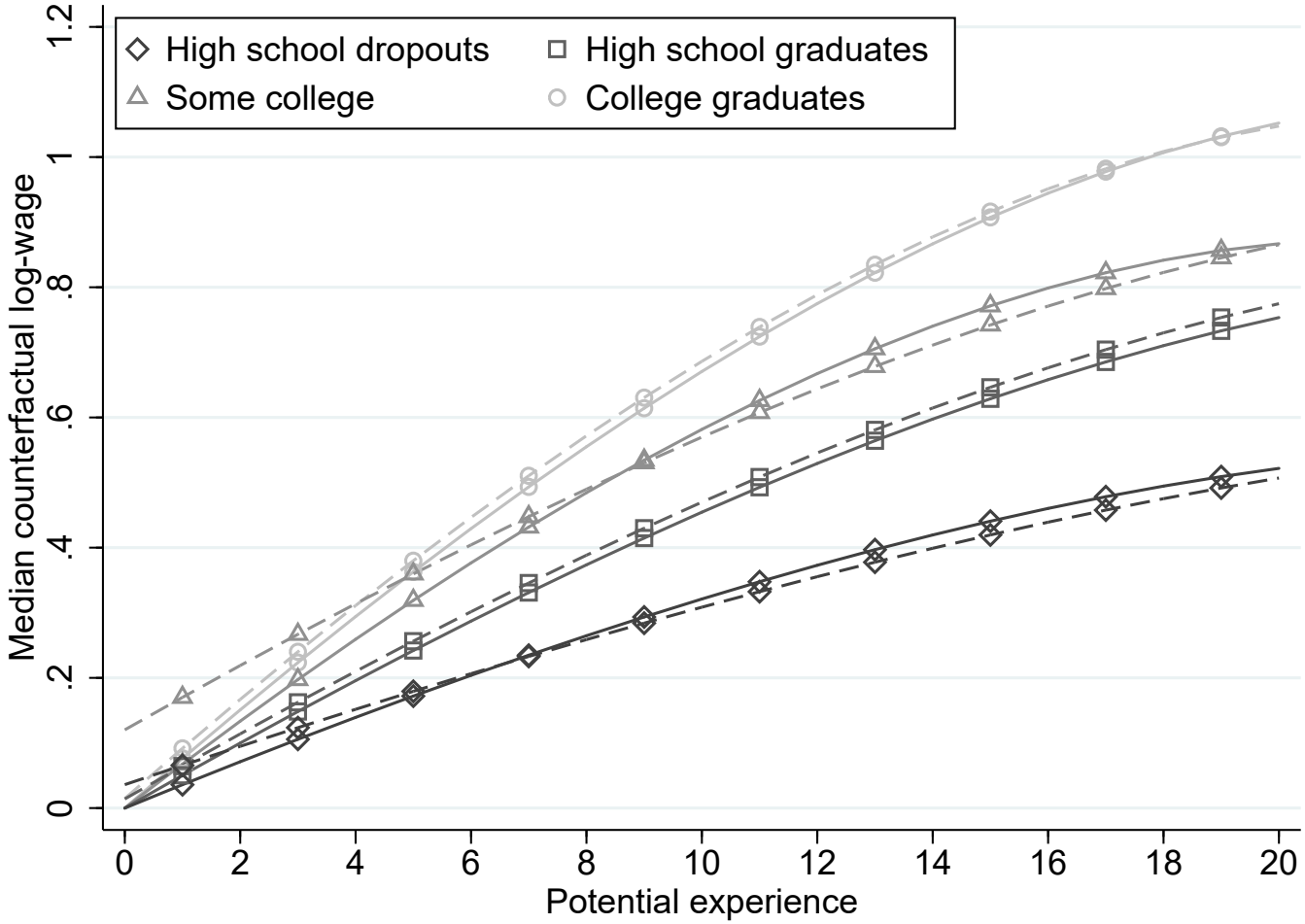
Profil des salaires potentiels - par éducation

Median counterfactual log-wage as a function of potential experience by education level, main and basic



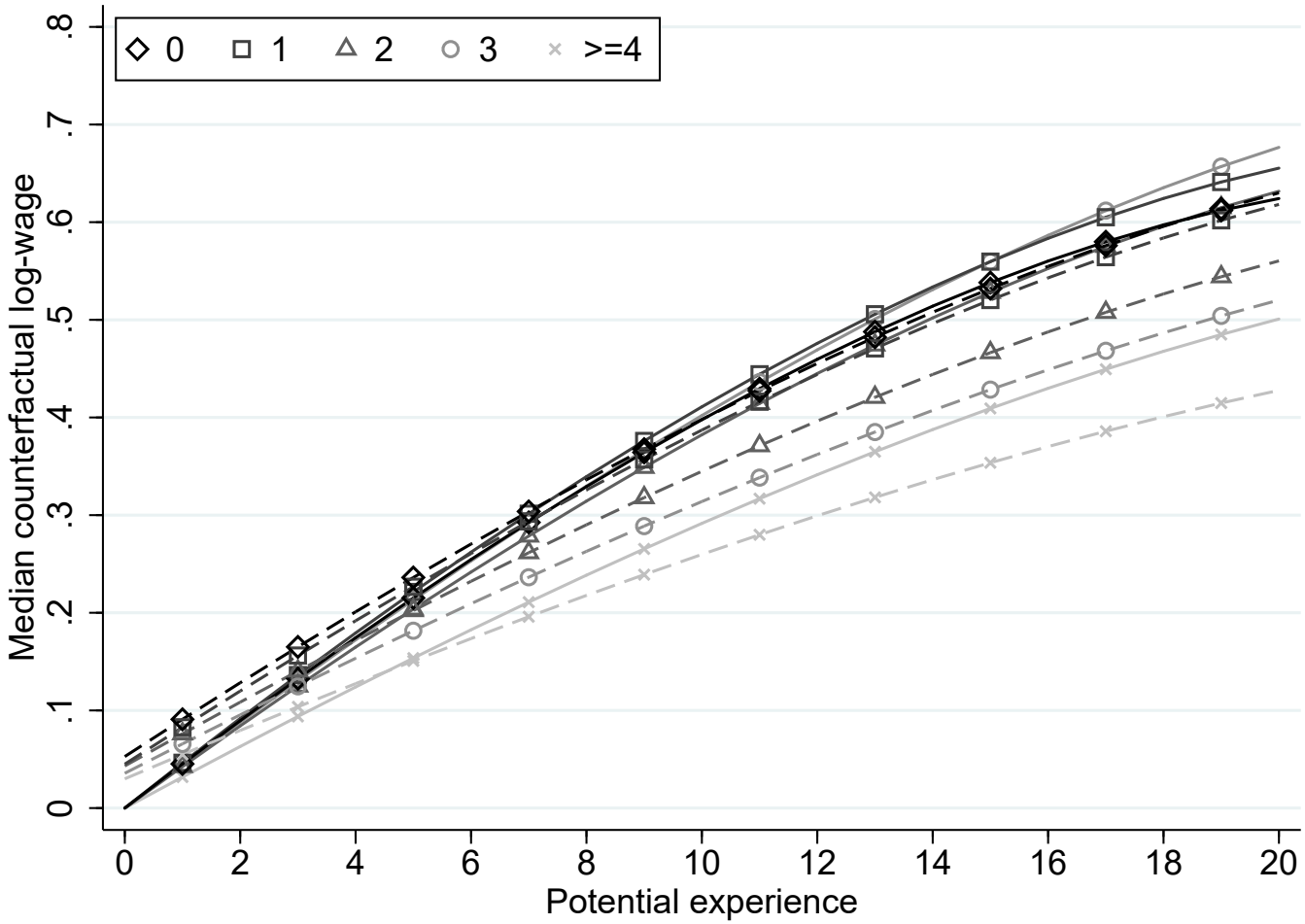
Profil des salaires potentiels - par éducation

Median counterfactual log-wage as a function of potential experience by education level, main and with no factor



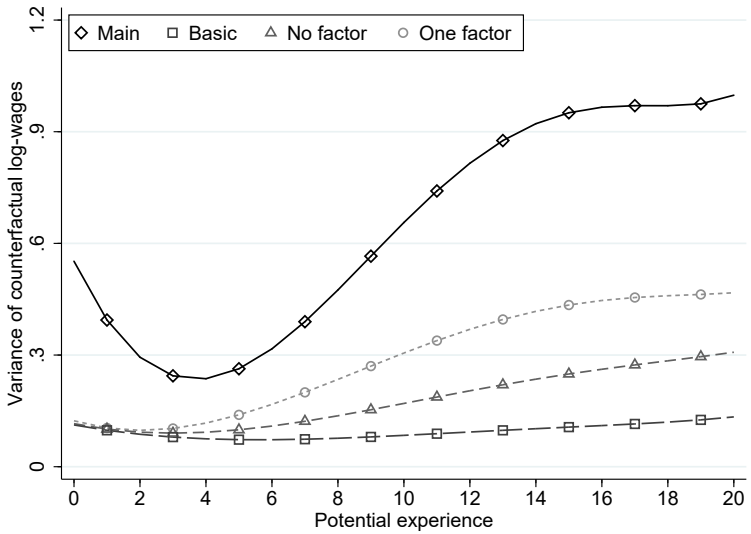
Profil des salaires potentiels - par nombre d'interruptions

Median counterfactual log-wage as a function of potential experience by education level, main and basic

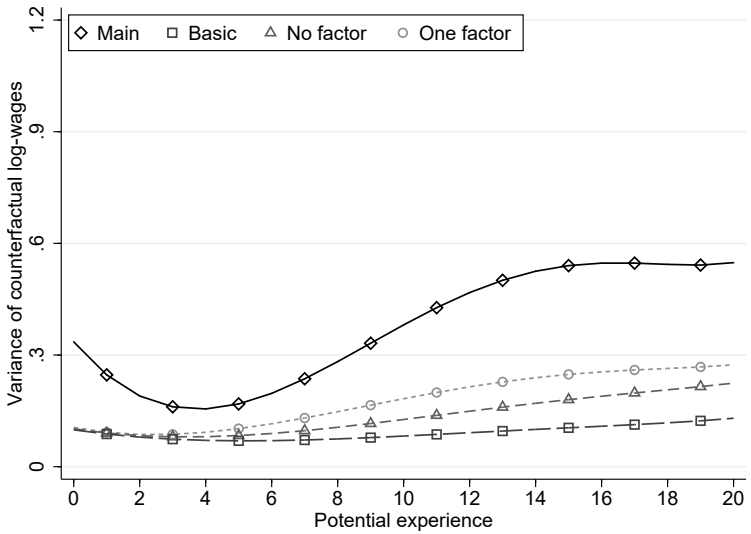


Variance des salaires potentiels - par spécification

Raw



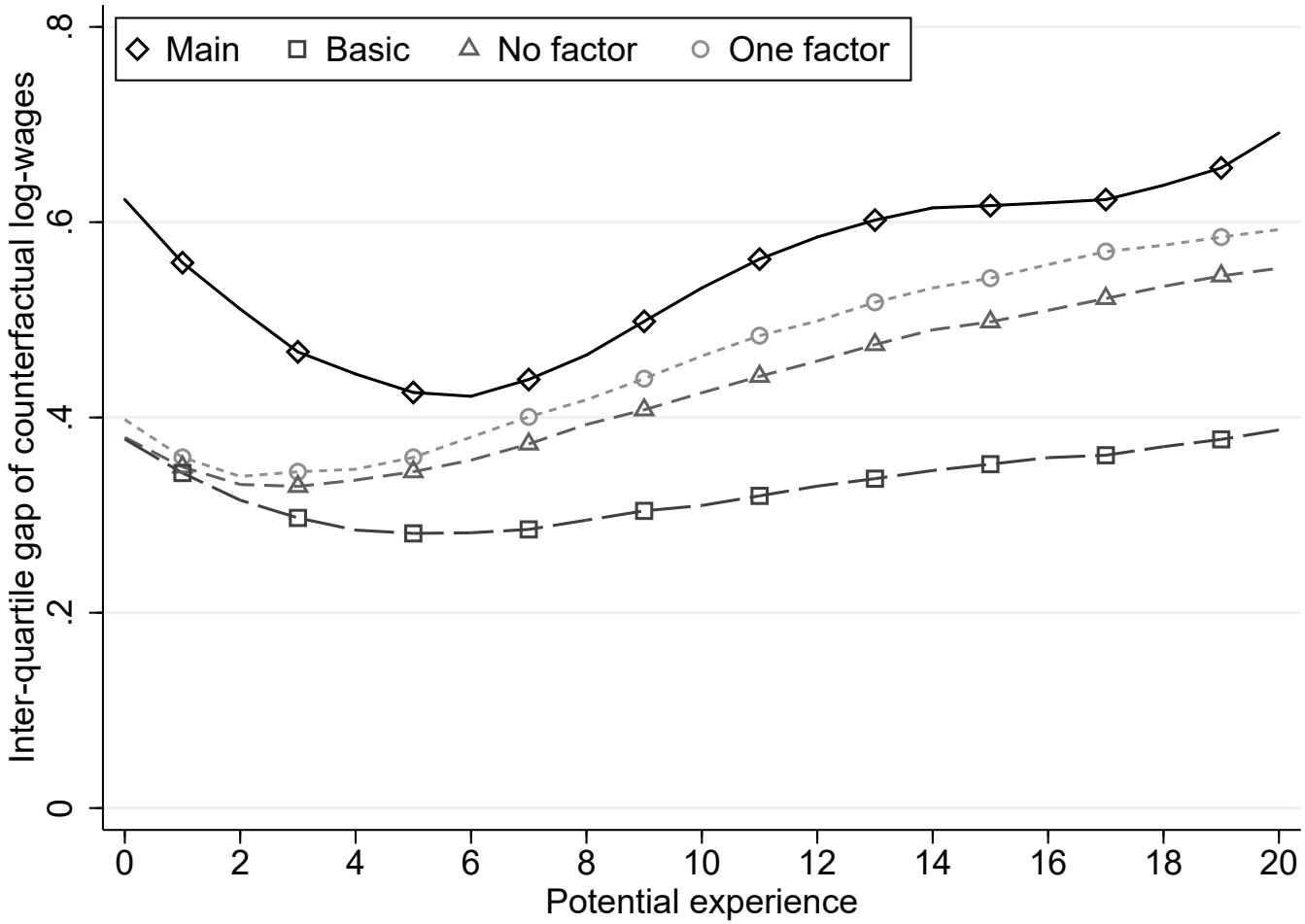
Corrected



Note: Correction for the sampling bias, see ▸ Bias-Corrected Variances

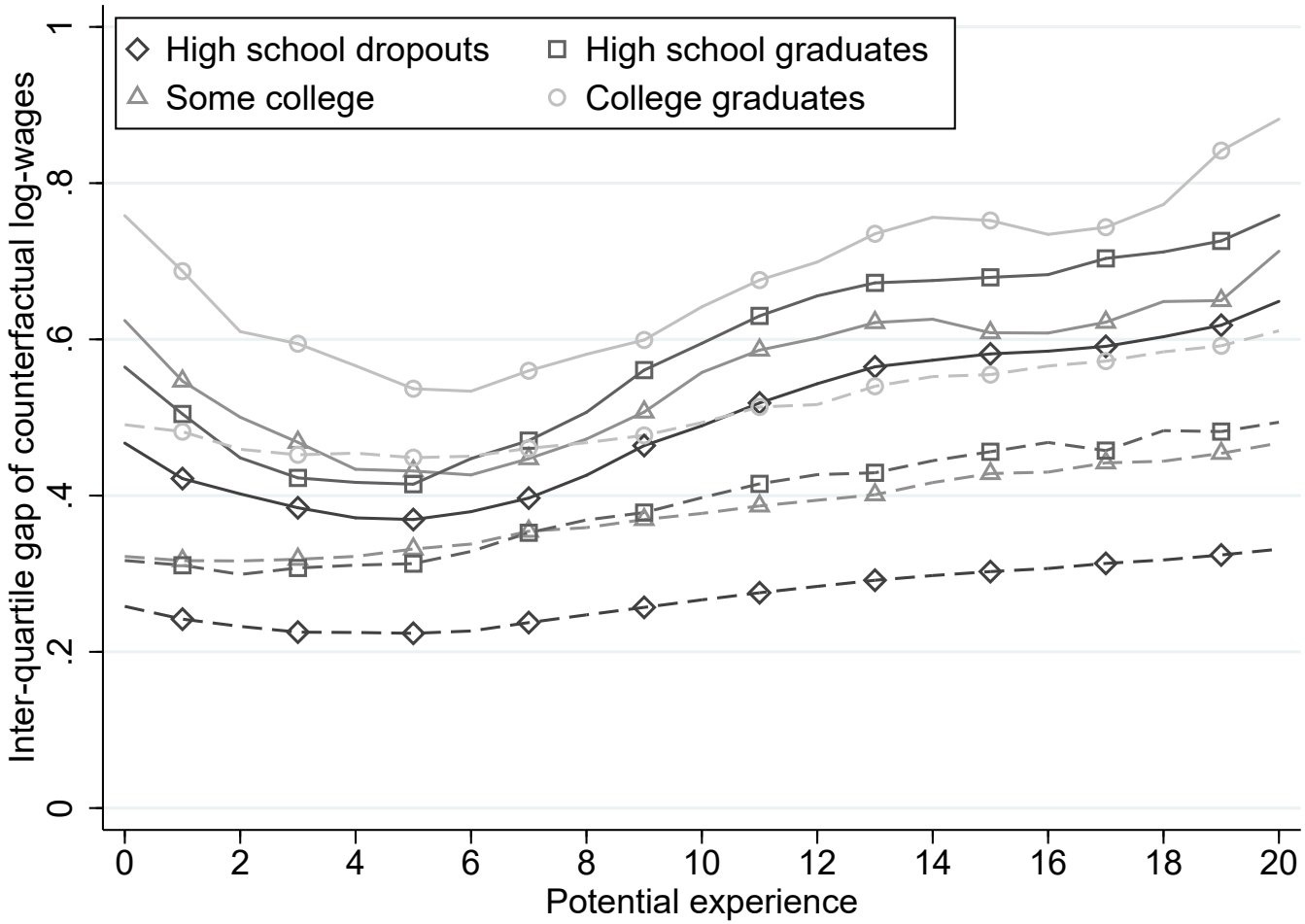
Interquartile range - by specification

Interquartile gap of counterfactual log-wages as a function of potential experience



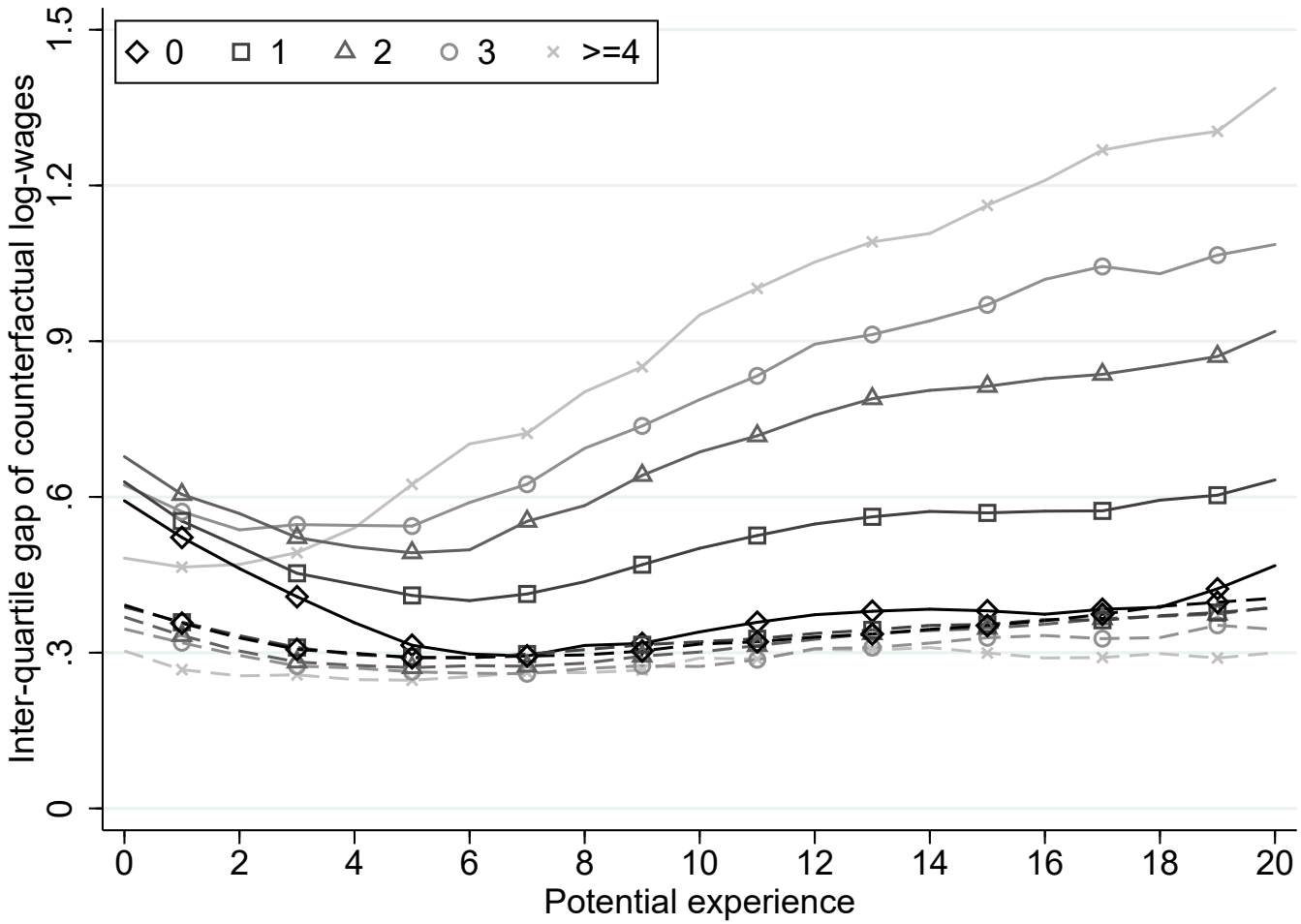
Interquartile range - by education

Interquartile gap of counterfactual log-wages as a function of potential experience, by education, main and basic



Interquartile range - by number of interruptions

Interquartile gap of counterfactual log-wages as a function of potential experience, by number of interruptions, main and basic



Conclusion

- Prendre en compte les période hors de l'emploi et les facteurs amène:
  - Des rendements plus élevés de l'expérience potentielle
  - Une plus grande dispersion des profils de carrière salariale

  Les problèmes de spécification sont plus importants que ceux de sélection
- Application aux femmes: pertinent de par leur plus faible participation
  Difficulté: des effets de sélection plus forts du fait de la restriction aux individus en emploi pendant au moins 15 ans.
- Des difficultés empiriques d'identification du fait de l'hétérogénéité individuelle massive
  $\Rightarrow$ Utilisation de procédures d'aggrégation (e.g. Bonhomme, Lamadon and Manresa, 2017) ?

Proportion of individuals in a given cohort occupying a job at potential experience t

Cohort	N	Potential experience										
		1	2	3	4	5	10	15	20	21	24	27
1985	897	1.00	0.73	0.70	0.76	0.83	0.69	0.91	0.84	0.86	0.79	0.63
1986	818	1.00	0.73	0.75	0.80	0.80	0.87	0.91	0.88	0.87	0.80	0.00
1987	896	1.00	0.70	0.77	0.77	0.82	0.89	0.91	0.89	0.88	0.77	0.00
1988	905	1.00	0.79	0.79	0.81	0.86	0.90	0.92	0.90	0.88	0.77	0.00
1989	1123	1.00	1.00	0.80	0.85	0.82	0.92	0.87	0.93	0.90	0.00	0.00
1991	1469	1.00	0.86	0.82	0.68	0.87	0.93	0.91	0.87	0.86	0.00	0.00
1992	896	1.00	0.79	0.64	0.87	0.86	0.94	0.95	0.88	0.00	0.00	0.00

Correlation of log-wages deflated with prices of human capital at two values of potential experience

	1	2	3	4	5	9	10	14	15	19	20
1	1										
2	.82	1									
3	.75	.84	1								
4	.70	.78	.86	1							
5	.64	.73	.79	.85	1						
9	.49	.58	.62	.67	.73	1					
10	.45	.53	.57	.62	.68	.87	1				
14	.37	.44	.50	.55	.59	.76	.79	1			
15	.33	.41	.45	.51	.55	.73	.75	.90	1		
19	.28	.36	.42	.48	.51	.67	.69	.81	.85	1	
20	.28	.34	.41	.48	.50	.65	.68	.80	.83	.93	1

$cov\left(w_t, w_{t-1}\right)$

Wage setting and human capital accumulation

- All parameters are individual-specific
- Individual earnings in sector s at time t:

$$y_t^s = \exp(\delta_t^s) H_t \exp(-\tau_t)$$

where:

  - δ_t^s: "price" of human capital in sector s at period t, shock
  - H_t: level of human capital at period t
  - Individuals decide τ_t at period t: investment in human capital

- Technology of production of Human capital:

$$H_{t+1} = H_t \exp\left[\rho^s \tau_t - \lambda_t^s\right]$$

where:

  - ρ^s: rate of return of human capital investments in sector s
  - λ_t^s: depreciation of human capital in sector s at period t

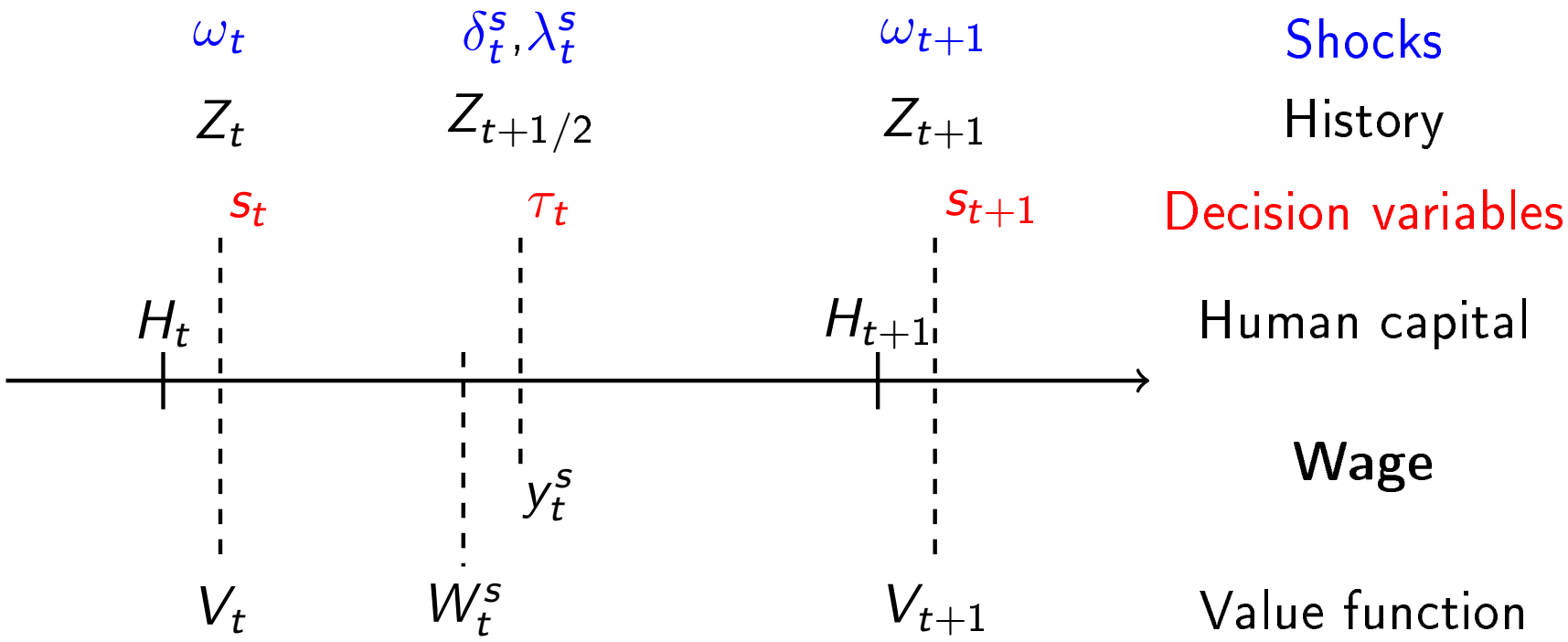
Decisions

- Period-t utility

$$\ln y_t^s - c\frac{(\tau_t)^2}{2} + \omega_t \mathbf{1}\{s = e\}$$

  - c: cost of human capital investment (on top of wage loss)
  - ω_t: relative preference for working in sector e
- Value function: intertemporal sum of utilities with discount factor β
  - H_t results from the past shocks and past decisions
  - Choice of sector s_t is conditional on $Z_t = \{\omega_t\} \cup Z_{t-1/2}$
  - Choice of HC investment τ_t is conditional on $Z_{t+1/2} = \{\delta_t^s, \lambda_t^s\} \cup Z_t$
- Backward solution: the terminal value of human capital stocks κ at an arbitrary date in the future is fixed.

Timing of decisions



Exogeneity assumption: the distribution of future shocks $(\omega_l, \delta_l^s, \lambda_l^s)_{l \geq t}$ conditionally on $Z_{t-1/2}$ does not depend on the state variable history $H_t, H_{t-1}, ., H_1$.

Theoretical Results

- The sequence of potential investments between $t = t_0$ and $t = t_0 + d$ in each sector s is:

$$\tau_t^s = \max\left\{0, \frac{1}{c}\left(\rho^s \beta \kappa_{t+1} - 1\right)\right\}$$

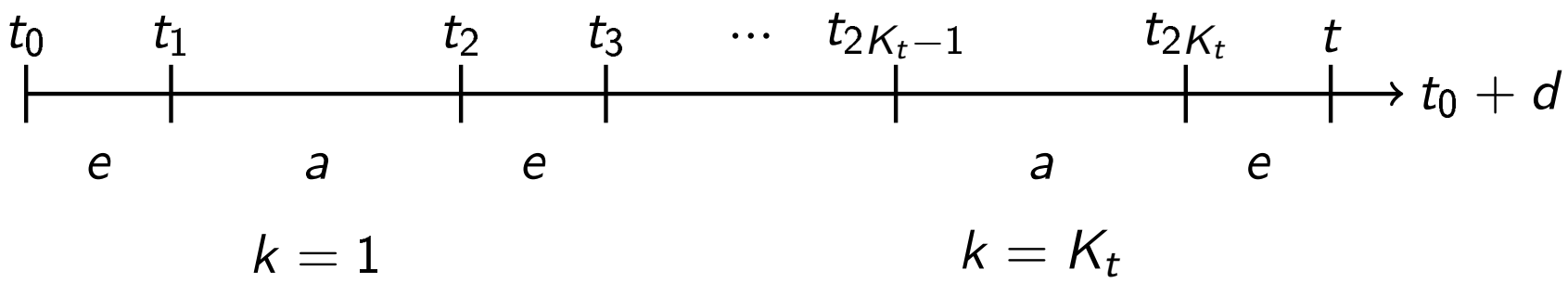
where $\kappa_t = \frac{1}{1-\beta} + \beta^{t_0+d+1-t}\left(\kappa - \frac{1}{1-\beta}\right)$.
Possibility of flat spots (see Magnac et al., 2018)

- The sector choice is determined by:

$$s_t = e \text{ iff}$$
$$\omega_t + \mathbb{E}_t\left(\delta_t^e - \beta\kappa_{t+1}\lambda_t^e + c\frac{(\tau_t^e)^2}{2}\right)$$
$$\geq \mathbb{E}_t\left(\delta_t^n - \beta\kappa_{t+1}\lambda_t^n + c\frac{(\tau_t^n)^2}{2}\right)$$

Return

Trajectories



- A worker enters the sector e at period t_0 and is observed up to $t_0 + d$, $s^{t_0} = e$.
- t_1: first entry date into into n, t_2 reentry into sector e,...
- t_{2k-1}: date of entry into spell k in sector n out of e
- t_{2k}: date of exit out of the spell k in sector n into sector e
- K_t: number of spells in sector n before period t

Return

Factor representation of residuals

- Specification of the shocks as linear factors

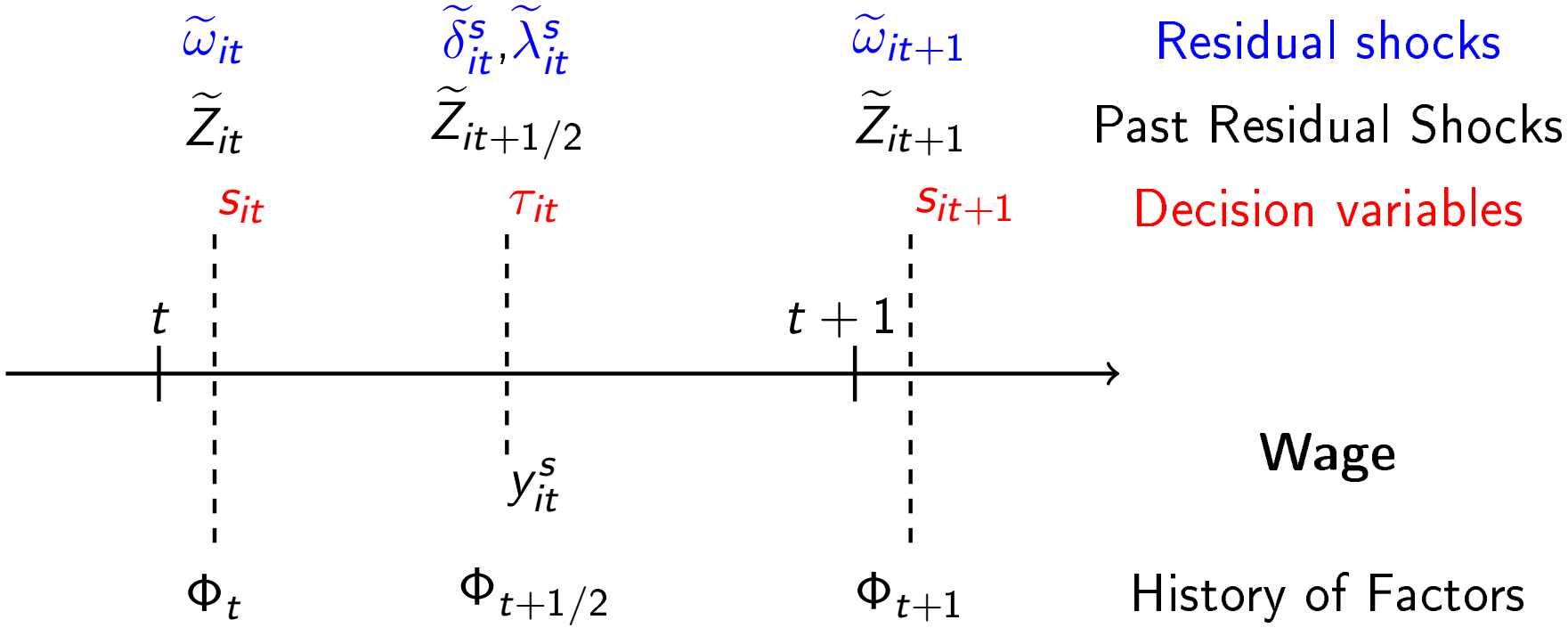
$$\begin{aligned}
\omega_{it} &= \varphi_t^{(\omega)}\theta_i^{(\omega)} + \widetilde{\omega}_{it}, \\
\delta_{it}^s &= \varphi_t^{(\delta),s}\theta_i^{(\delta),s} + \widetilde{\delta}_{it}^s, \\
\lambda_{it}^s &= \varphi_t^{(\lambda),s}\theta_i^{(\lambda),s} + \widetilde{\lambda}_{it}^s.
\end{aligned}$$

Standard orthogonality restrictions on the residual random shocks.

- Define:
  - $\tilde{\theta}_i = \left\{\theta_i^{(\omega)}, \theta_i^{(\delta),e}, \theta_i^{(\delta),n}, \theta_i^{(\lambda),e}, \theta_i^{(\lambda),n}\right\}$
  - $\Phi_t = \left\{\varphi_t^{(\omega)}, \Phi_{t-1/2}\right\}$
  - $\Phi_{t+1/2} = \left\{\varphi_t^{(\delta),e}, \varphi_t^{(\delta),s}, \varphi_t^{(\lambda),e}, \varphi_t^{(\lambda),n}, \Phi_t\right\}$
  - $\tilde{Z}_{it}$ and $\tilde{Z}_{it+1/2}$: histories of residual random shocks $\widetilde{\omega}_{it}$, $\widetilde{\delta}_{it}^s$ and $\widetilde{\lambda}_{it}^s$, which mimic Z_{it} and $Z_{it+1/2}$

Timing of shocks

$\tilde{\theta}_i$ fixed



Assumption Missing Conditionally on Factors at Random (AMCFAR)

- Selection:

$$\widetilde{\omega}_{it} \perp \tilde{Z}_{it-1/2} \Big| \Phi_t, \tilde{\theta}_i$$

- Human Capital accumulation:

$$(\widetilde{\delta}^s_{it}, \widetilde{\lambda}^s_{it}) \perp \tilde{Z}_t \Big| \Phi_{t+1/2}, \tilde{\theta}_i$$

- Implications:
  - Log-earnings equation: factor specification

$$\ln y_{it} = x_{it}\eta_i + \varphi_t \theta_i + \tilde{v}_{it}$$

where $\tilde{v}_{it}$ function of $\left(\tilde{\delta}_{it}, \tilde{\lambda}_{il}, l < t\right)$, $\theta_i = \left\{\theta_i^{\delta}, \theta_i^{(\lambda)}\right\}$, assuming $\theta_i^{(\lambda),e} = \theta_i^{(\lambda),n} = \theta_i^{(\lambda)}$, φ_t accordingly.

  - Exogeneity of selection and of $\left(x_{it}^{(3)}, x_{it}^{(4)}\right)$

$$E\left(\tilde{v}_{it} \left| x_{it}^{(3)}, x_{it}^{(4)}, \Phi_{t+1/2}, \tilde{\theta}_i \right.\right) = 0.$$

Return

Stop criterium

- For the factors: $C_1 \equiv \left\| M_{\varphi^{(k-1)}\varphi^{(k)}} \right\| / RT$
- For the individual coefficients:

$$c_1 = N\left(\overline{\theta}^{(k)} - \overline{\theta}^{(k-1)}\right)' V\left(\overline{\theta}^{(k-1)}\right)^{-1}\left(\overline{\theta}^{(k)} - \overline{\theta}^{(k-1)}\right)$$

$$c_2 = \frac{tr\left[\left(V\left(\overline{\theta}^{(k)}\right) - V\left(\overline{\theta}^{(k-1)}\right)\right)\left(V\left(\overline{\theta}^{(k)}\right) - V\left(\overline{\theta}^{(k-1)}\right)\right)'\right]}{tr\left[V\left(\overline{\theta}^{(k-1)}\right)\right]}$$

$$C_2 \equiv \min\left(c_1, c_2\right)$$

▸ Return

Bias-Corrected Variances

- Denote $\phi_i = \left(\eta_i', \theta_i'\right)'$
- Unfeasible estimator ϕ_i^u:

$$\phi_i^u = \phi_i + \left(Z_i' Z_i\right)^{-1} Z_i' \widetilde{v}_i$$

where $Z_i = \left(1_{T_i}, X_i, \varphi_i\right)$ explanatory variables observed on the T_i periods the individual i is observed in the panel.

- Consider $V\left(\widetilde{v}_{it}\right) = \sigma_i^2 \omega_t^2$, $\sum_{t=1}^{T} \omega_t^2 = T$ for normalization, the conditional variance of the unfeasible solution writes:

$$V\left(\phi_i^u \,|Z_i\right) = V\left(\phi_i \,|Z_i\right) + E(\sigma_i^2 \mid Z_i) \left(Z_i' Z_i\right)^{-1} \left(Z_i \Sigma_i Z_i\right) \left(Z_i' Z_i\right)^{-1}$$

where $\Sigma_i = diag\left(\omega_{t_{i1}}^2, ..., \omega_{t_{iTi}}^2\right)$

- Estimates of σ_i and ω_t recovered by iteration.
- We get:

$$V\left(\phi_i\right) = V\left(\phi_i^u\right) - E\left[\sigma_i^2 \left(Z_i' Z_i\right)^{-1} \left(Z_i \Sigma_i Z_i\right) \left(Z_i' Z_i\right)^{-1}\right]$$

replacing all terms by their estimated values.

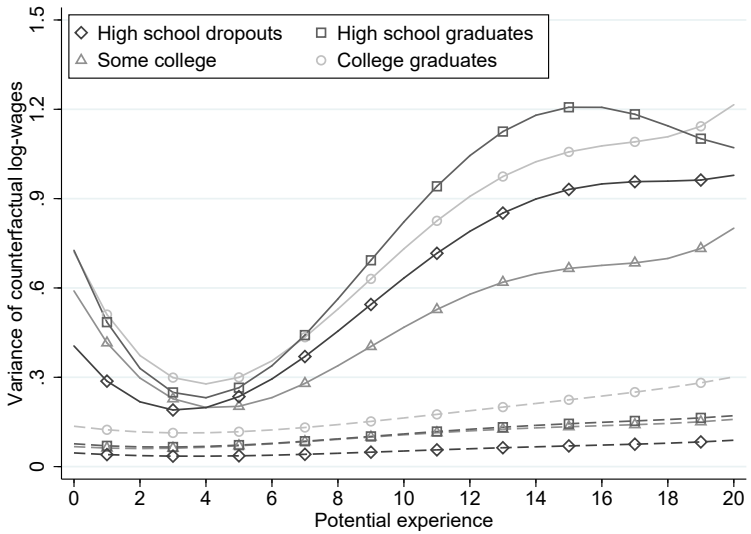
- $V\left(\eta_{i0} + \eta_{i1} t + \eta_{i2} \beta^{-t}\right)$ is deduced from $V\left(\phi_i\right)$

Return

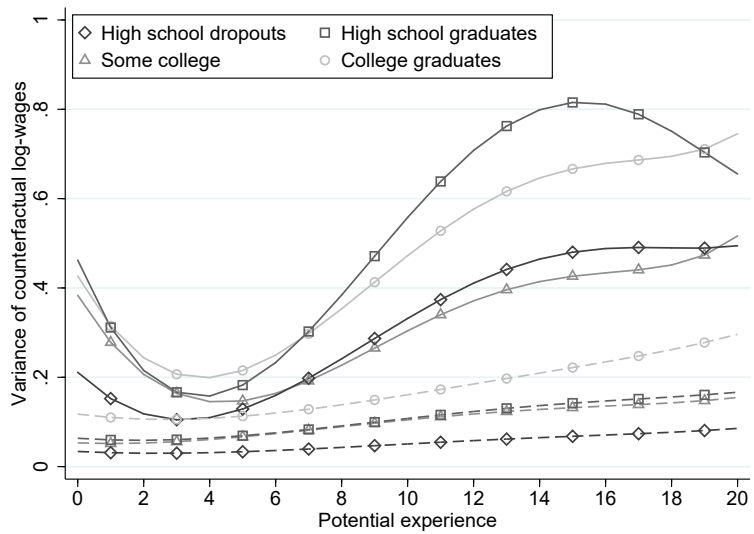
Variance of wage profiles - by education

Variance of counterfactual log-wages as a function of potential experience, by education, main and basic

Raw



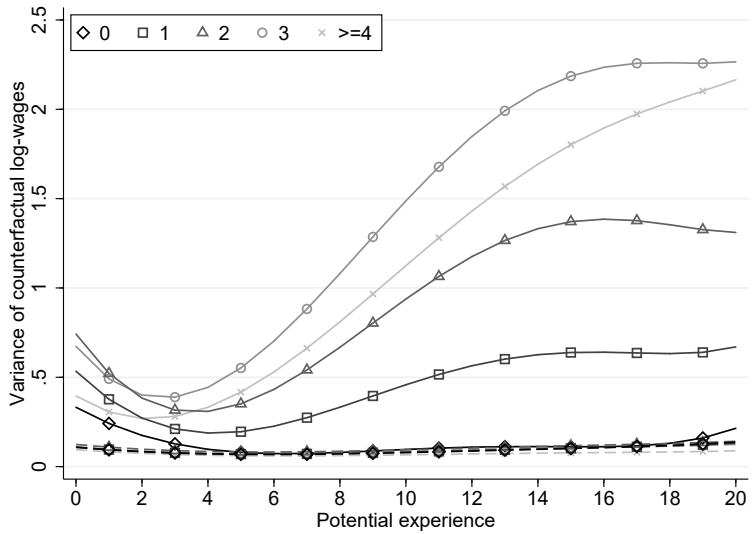
Corrected



Variance of wage profiles - by number of interruptions

Variance of counterfactual log-wages as a function of potential experience, by number of interruptions, main and basic

Raw



Corrected