The returns to education unveiled

Ana Rute Cardoso (IAE-CSIC and Barcelona GSE)
Pedro Portugal (Banco de Portugal and NOVA SBE)
Paulo Guimarães (Banco de Portugal)
Hugo Reis (Banco de Portugal)

Banco de Espanha, June 2016

Motivation 1: "The return of the firm to labor economics"

- Firm-side explanations for wage differentials: a profusion of theories
  - why firms find it profitable to pay non-competitive wages
  - they design incentive schemes to attract better workers, retain them, and enhance their productivity
  - efficiency wages; implicit contracts; rent-sharing; insider-outsider; labor market frictions & job search and matching; ...
- Empirically
  - early case studies in 1950s: firms' pay standard diverges within narrowly defined regions and industries (Lester, 1952) (Reynolds, 1951)
  - in 1990s, firm back on the spotlight: intra-industry wage differentials or gender pay gap (Groshen, 1991, 1991a)
  - the boom that followed Abowd, Kramarz and Margolis (1999) estimation of firm and worker fixed effects in wage regressions
- Interest in the role of the firm has been growing
  - increasingly richer linked employer-employee longitudinal datasets
  - topics such as: trends in wage dispersion; matching of workers and firms (Torres et al., 2012); gender pay gap (Cardoso et al., 2016; Card

Motivation 2: Measuring the wage returns to education

- Fits into a long tradition of explaining wage differentials by looking at worker-side attributes
  - mainly competitive background (Becker, 1962): wage differentials reflect differences in workers' productivity
  - which depends on workers' skills: education, training, seniority, experience, motivation, ability, ...
  - key role of education
- Empirically, major progress over past decades
- Line of literature fully detached from that on the firm shaping wages
  - possibly constrained, for a long time, by the availability of data from household surveys only
  - time to redress this neglect?

Motivation 3: The role of peer and spillover effects

- A new branch of literature based on peer and spillover effects
  - Theoretical models allowing for the role of learning networks
  - Having in mind assortative matching (Gary Becker, again)
  - Key role of education, again (Acemoglu and Angrist, 2000)
- Empirically, some progress recently advanced
- Attempts made to measure peer and spillover effects
  - Along the geographical dimension (Moretti, 2004)
  - Focused on the spillovers from education
  - Despite the perils of peer effect (Angrist, 2014) and the curse of reflection effects
  - An elegant way to overcome reflection effects (Arcidiacono et al., 2012)
  - An extension to the sources of wage variation (Cornelissen et al., 2013; Batisti, 2013)

Aims

- How much of the economy-wide return to education operates through the allocation of workers with different schooling levels to firms with different pay standards?
  - if firms are heterogeneous and workers are not randomly allocated,
  - part of the overall return to education could indeed be a return to working for a good firm
- Assess the returns to education taking into account:
  - *who* the worker is —observed and unobserved ability
  - *what* he does —detailed task (job)
  - *for whom* —the employer
  - Measure the role of co-worker education (in the same firm and job-title)
  - Measure peer effects
  - How the return on co-worker education relates with peer effects
- A decomposition exercise on the sources of the returns to education

Institutional wage setting

Wage bargaining system in Portugal prevailing over the sample period:

- Mandatory national minimum wage
- Collective bargaining takes place mainly at a sectoral level, and voluntary or mandatory extensions are commonplace
- Around 30,000 job-title wage floors are settled each year
- Despite very low unionization rates (around 10 percent)

Data

- Longitudinal linked employer-employee data: *Quadros de Pessoal*
- Years: 1986 to 2013 (1990 and 2001 not available)
- Variables:
  - worker's gender, birth date, schooling, occupation, date of hire, earnings (several components), hours of work (normal and overtime), collective bargaining agreement, worker category in the agreement
  - firm's industry, location, etc.
  - unique identifiers for workers, firms and job titles
- Final dataset (after constraints full-timer, base wage at least 80% min wage, non-missing relevant data, etc; connected set):
  - 28 million observations
- Hourly wage = (overall monthly earnings, incl, overtime)/(sum of normal and overtime hours)

A wage regression equation with worker and firm fixed effects

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \lambda_i + \theta_{f(i,t)} + \gamma_{j(i,t)} + \mu_t + \varepsilon_{it} \quad , \tag{1}$$

- y_{it} is the logarithm of the hourly wage for each worker i at year t;
- $\mathbf{x}_{it}$ is a vector of observed time-varying characteristics of workers and firms;
- λ_i is a time-invariant worker fixed effect;
- $\theta_{f(i,t)}$ is a firm specific time-invariant fixed effect;
- $\gamma_{j(i,t)}$ is a job-title specific time-invariant fixed effect;
- μ_t are time fixed effects;
- ϵ_{it} is assumed to follow the conventional assumptions

Estimation algorithm, Guimarães and Portugal (2010)

C. *Estimation Strategy*

Controlling simultaneously for worker, firm, and job title-specific effects requires the introduction of three high-dimensional fixed effects in the linear regression model. To illustrate our estimation strategy, consider the following linear regression model in matrix form:

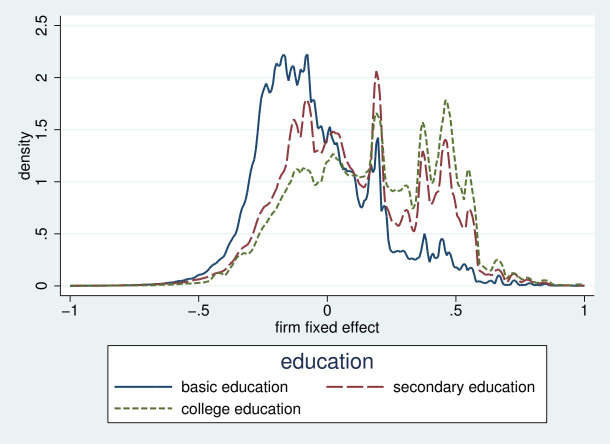
$$\mathbf{Y} = \mathbf{Z}\beta + \mathbf{D}_1\lambda + \mathbf{D}_2\theta + \mathbf{D}_3\gamma + u, \tag{2}$$

where $\mathbf{Z}$ is a matrix of time-varying explanatory variables and $\mathbf{D}_1$, $\mathbf{D}_2$, and $\mathbf{D}_3$ are high-dimensional matrices for the fixed effects. The normal equations may be rewritten as

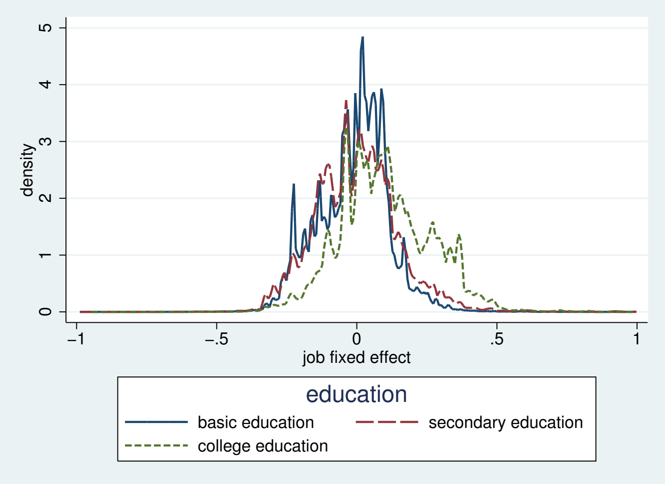
$$\begin{bmatrix} \beta = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'(\mathbf{Y} - \mathbf{D}_1\lambda - \mathbf{D}_2\gamma - \mathbf{D}_3\theta) \\ \lambda = (\mathbf{D}_1'\mathbf{D}_1)^{-1}\mathbf{D}_1'(\mathbf{Y} - \mathbf{Z}\beta - \mathbf{D}_2\gamma - \mathbf{D}_3\theta) \\ \theta = (\mathbf{D}_3'\mathbf{D}_3)^{-1}\mathbf{D}_3'(\mathbf{Y} - \mathbf{Z}\beta - \mathbf{D}_1\lambda - \mathbf{D}_2\gamma) \\ \gamma = (\mathbf{D}_2'\mathbf{D}_2)^{-1}\mathbf{D}_2'(\mathbf{Y} - \mathbf{Z}\beta - \mathbf{D}_1\lambda - \mathbf{D}_3\theta) \end{bmatrix},$$

suggesting an iterative solution that alternates between estimation of β, λ, θ, and γ.

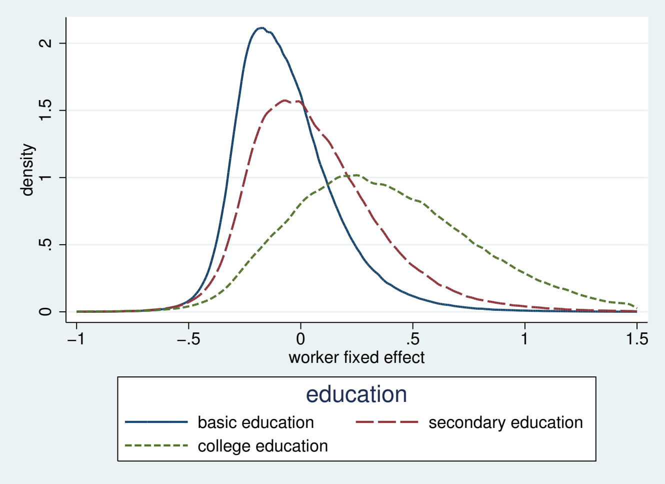
Graphical overview of the fixed effects



Firm fixed effects: wide diversity of pay standards; peaks for large firms; basic education more concentrated in worse-paying firms



Job fixed effects: narrower range and higher peak; concentration of job wage floors consistent with egalitarian aims of trade unions



Dispersion of worker abilities larger among university educated workers

What role do firms and jobs play shaping the returns to education? Gelbach (2016) decomposition

The idea is quite simple. It starts out with the well-known omitted variable bias formula:

- Assume there are two sets of right-hand side variables in the full model, X_1 and X_2
- X_1 contains the regressor of interest, worker's level of education (plus basic controls)
- X_2 contains covariates traditionally omitted in wage regressions, the firm-, the job- and the worker-fixed effects
- Consider the base regression of (log) wages on X_1 only:

$$\mathbf{Y} = \mathbf{X_1} b_1^{base} + \varepsilon \tag{2}$$

$$\begin{aligned} \widehat{b}_1^{base} &= \widehat{\beta}_1^{full} + (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2\widehat{\beta}_2 \qquad (3) \\ &= \widehat{\beta}_1^{full} + \boldsymbol{\delta} \qquad (4) \end{aligned}$$

Following Gelbach (2016) decomposition of the difference in the estimated coefficients of X_1 between the base and the full models, we know that

$$\widehat{b}^{base} - \widehat{\beta}^{full} = \mathbf{P_{X_1}} \mathbf{D}_f \widehat{\boldsymbol{\theta}} + \mathbf{P_{X_1}} \mathbf{D}_j \widehat{\gamma} + \mathbf{P_{X_1}} \mathbf{D}_i \widehat{\boldsymbol{\lambda}} \quad , \tag{5}$$

where $\mathbf{P_{X_1}} = (\mathbf{X_1}'\mathbf{X_1})^{-1}\mathbf{X_1}'$ and (taking the firm effects as illustration):

- $\mathbf{D}_f$ is a design matrix for the firm effects (n x F)
- $\widehat{\boldsymbol{\theta}}$ is a vector of the estimates of the firm fixed-effects in the full model (F x 1)
- therefore, $\mathbf{D}_f \widehat{\boldsymbol{\theta}}$ is a vector with the estimates of the firm fixed effects
- and $\mathbf{P_{X_1}} \mathbf{D}_f \widehat{\boldsymbol{\theta}}$ simply the coefficients of the regression of the firm fixed effects on the variables $\mathbf{X}_1$ in the base model

Thus, we can rewrite the above equation more succinctly as

$$\widehat{b}_{base} - \widehat{\beta}_{full} = \widehat{\boldsymbol{\delta}}_{\boldsymbol{\lambda}} + \widehat{\boldsymbol{\delta}}_{\boldsymbol{\theta}} + \widehat{\boldsymbol{\delta}}_{\gamma} \tag{6}$$

- The change in the coefficient of interest is partitioned into the role of the different additional covariates
- Conditional decomposition is invariant to the order of inclusion of additional covariates (because based on the full model)

Our application of Gelbach (2016) decomposition

Base regressor of interest in X_1:

- education

Additional covariates in X_2

- firm fixed effects
- job fixed effects

Purpose – identify the conditional contribution of the firm and the job to the returns to schooling

Decomposing the returns to education, a first pass

Table : Conditional Decomposition of the Return to Education

Benchmark regression	Worker FE	Firm FE	Job FE
0.0837			
(0.0008)			
	0.0474	0.0227	0.0136
	(0.0004)	(0.0007)	(0.0003)

Note: Decompositions based on Gelbach (2016).

Decomposing the returns to education: Accounting for Co-worker Education

Table : Conditional Decomposition of the Return to Education

	Benchmark regression	Worker FE	Firm FE	Full Regression
Own Education	0.0442			
	(0.0000)			
Co-worker education	0.0607			
	(0.0000)			
Own Education		0.0352	0.0090	0.0000
		(0.0000)	(0.0000)	(0.0000)
Co-worker education		0.0342	0.0168	0.0097
		(0.0000)	(0.0000)	(0.0000)

Note: Decompositions based on Gelbach (2016).

A wage regression equation with worker, firm, and peer effects

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \lambda_i + \delta\overline{\lambda}_{-it} + \theta_{f(i,t)} + \mu_t + \varepsilon_{it} \quad , \tag{7}$$

where $\overline{\lambda}_{-it}$ is the mean of the fixed effects of the peer group of worker i at time t and δ is the associated coefficient.

The role of Peer Effects

	(1)	(2)	(3)	(4)	(5)
Age	0.0375	0.0197	0.0176	0.0184	0.0149
	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Age squared	-0.0003	-0.0002	-0.0002	-0.0002	-0.0002
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Tenure	0.0185	.0095	0.0081	0.0069	0.0064
	(0.0002)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Tenure squared	-0.0003	0-.0002	-0.0002	-0.0002	-0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Firm size (in logs)	0.0479	.0289	0.0351	0.0333	0.0272
	(0.0006)	(0.0001)	(0.0001)	(0.0000)	(0.0001)
Schooling	0.0442				
	(0.0001)				
Gender (Female=1)	-0.2586				
	(0.0009)				
Peer schooling	0.0607	-0.0061	-0.0034	-0.0035	-0.0021
	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Co-worker quality		0.7187	0.5515	0.4450	0.3210
		(0.0007)	(0.0007)	(0.0012)	(0.0005)
Time effects	✓	✓	✓	✓	✓
Worker effects		✓	✓	✓	✓
Firm effects			✓	✓	
Job title effects				✓	
Firm/Job title effects					✓
N	27 961 816	27 961 816	27 961 816	27 961 816	27 961 816
R Squared	0.6021		0.9273	0.9329	0.9487

Fixed effects

Table : Basic statistics

	Model 4	Model 5
σ_{λ}	0.296	0.298
σ_{θ}	0.222	0.268
$\sigma_{delta\overline{\lambda}}$	0.259	0.288

Correlations

Table : Basic statistics

	Model 4	Model 5
$Corr(\lambda, \theta)$	0.175	0.075
$Corr(\lambda, \overline{\lambda})$	0.729	0.762
$Corr(\theta, \overline{\lambda})$	0.197	0.083

Wage Variance Decomposition

Table : Wage Variance Decomposition

	Model 4	Model 5
λ	0.4088	0.4157
θ	0.2202	0.2780
$\delta\overline{\lambda}$	0.1856	0.1522

New Decomposition Accounting for Peer Effects

	Worker Fixed Effect	Firm Fixed Effect	Peer effects
Specification (5)			
Schooling	0.0242	0.0154	0.0045
	(0.0001)	(0.0000)	(0.0000)
Peer Schooling	0.0230	0.0296	0.0102
	(0.0001)	(0.0001)	(0.0000)

Interpreting the decomposition

Breaking down the returns to education (in the preferred specification 4)

- return to own education
  - 0.0242 - "true" return to schooling" (including unobserved ability - upper bound)
  - 0.0154 - more educated workers sort themselves into firms with more generous wage policies
  - 0.0045 - more educated workers match with high skilled workers
- return to co-worker education
  - 0.0102 - "true" spillover effect
  - 0.0296 - more educated peers match with firms with more generous wage policies
  - 0.0280 - largely the refection effect (homophily)

Heterogenous Peer Effects: Average effects

Table : College versus Non-college

	Coefficient	Standard error
Non-College	0.2222	0.0007
College	0.1182	0.0006

Heterogenous Peer Effects: Interaction effects

Table : College versus Non-college

	Coefficient	Standard error
Non-College	0.4099	0.0008
College	0.3172	0.0008

Conclusion

- Quantified the sources of the returns to education: the firm and co-worker channels (and the job channels)
  - Mincerian-type regressions point to an overall 8.4 percent return per year of schooling in Portugal over 1986-2013
  - allocation of better educated workers into higher paying firms accounts for 26% of the overall conditional return on education
- Relevant role of peer effects driving wages
- Co-worker education directly influences wages
- Other peer effects drive wages through channels other than education
- Under random assignment of worker to firms the return to own education would have been 2.8 percent, and the benefit from one additional year of co-worker education 1.7 percent

The way ahead

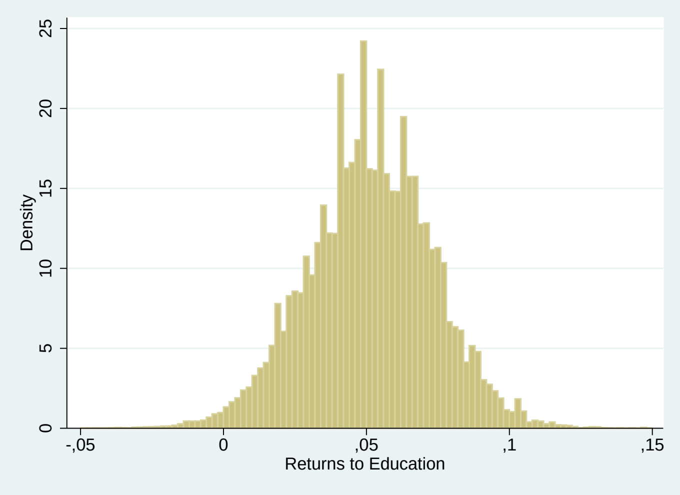
- Firms reward education at very different rates: Why?
- Potential mechanism we're exploring —quality of the peers
  - can peer quality be better captured by worker fixed effects (average for peers will capture not only their education, but a set of unobservables)? Discussion in the literature
- Back to the aims defined - to assess the returns to education taking into account: what the worker does (job); for whom (firm); *with whom* (peers in the firm)

Heterogenous Returns to Education

$$y_{ift} = \theta_f + \alpha_t + \beta X_{ift} + \phi_f Education_i + \epsilon_{ift} \quad (8)$$

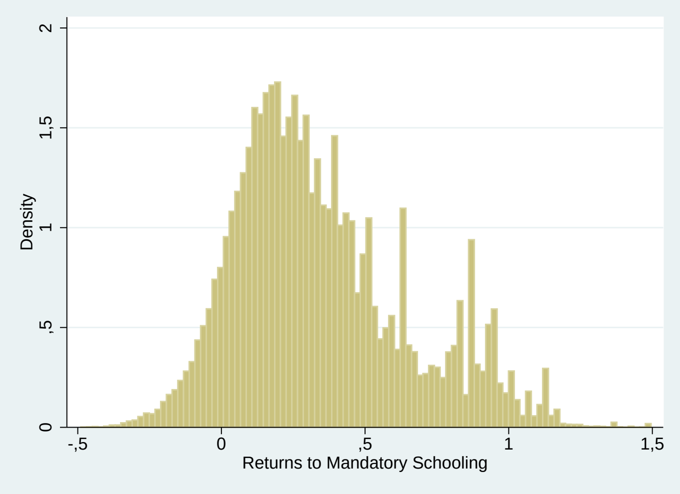
- y_{ift} represents the logarithm of the hourly wage for each individual
- ϕ_f is a firm specific coefficient of education
- X_{ift} represents other observed time-varying characteristics of individual i and firm j in year t
- θ_f is a firm fixed effect
- ϵ_{ift} is assumed to follow the conventional assumptions

Figure : **The distribution of the returns to education**



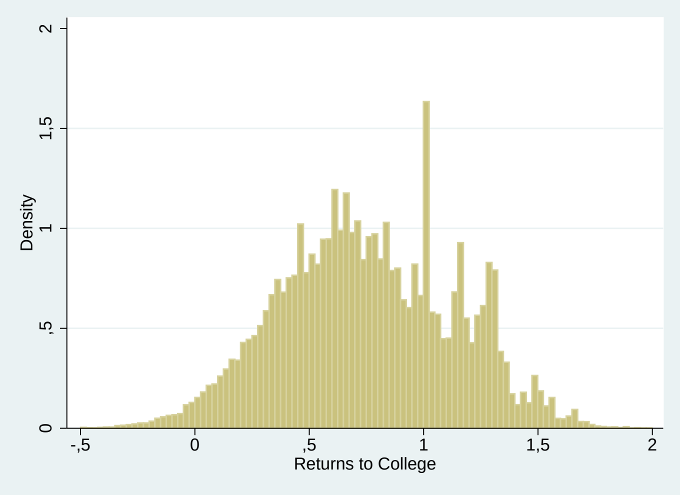
Heterogeneity on the returns to education

Figure : **The distribution of the returns to mandatory schooling**



Heterogeneity on the returns to education

Figure : **The distribution of the returns to college education**



The connection with labor productivity

Table : Determinants of (log of) labor productivity

	(1)	(2)	(3)
Returns to schooling	9.9453	7.0520	5.262
	(0.612)	(0.579)	(0.5662)
Average schooling		0.1669	0.1518
		(0.0072)	0.0138
Schooling dispersion			0.1703
			(0.0075)
Industry dummies	yes	yes	yes
Year Dummies	yes	yes	yes
R^2	0.4176	0.4791	0.4918

Note: Cluster standard errors in parenthesis.

BPlim

- The Bank of Portugal is launching a new research unit to promote the analysis of micro data.
  - The researchers of the unit are specialized in the production of micro econometric estimators and techniques.
  - The datasets used and produced at the Bank of Portugal are going to be available to the international academic community.
  - Through a network of powerful virtual computers (sand boxes) that can be accessed anywhere in the world.
  - Provided that a research project was submitted and approved by the head of the research unit.
  - Issues of statistical secret are dealt on a case by case basis.
  - Portuguese micro data is very, very rich.

Motivation: Worker heterogeneity

WORKER-SIDE explanations for the wage differentials:

- Perfect competition: wage differentials reflect differences in workers' productivity, which depends on:
  - Workers' skills (observed or not)
  - Motivation
  - Ability
  - Risk aversion
- Role of education and other human capital variables (seniority, experience, age, etc.): Becker (1962), Spence (1973),
- Over and under-education
- Mismatch
- Assessment by means of earnings Mincer-type regressions)

Motivation: Firm heterogeneity

FIRM-SIDE explanations for the wage differentials:

- Theories that explain why firms find it profitable to pay non-competitive wages
- Firms design incentive schemes to retain their workers, attract better workers, and enhance their productivity (compensation and retention policies)
- Examples: implicit contracts, principal-agent, efficiency-wages, rent-sharing, and insider-outsider considerations
- Labor market frictions explanations for the wage differentials: job search and matching literature
- Third wave: longitudinal linked employer-employee data, to account for both workers and firms characteristics (observed and unobserved)

Motivation: JOB TITLE heterogeneity

- Third important dimension of wage formation: JOB TITLE HETEROGENEITY
- There are compensating differentials for certain occupations involving:
  - Risks of accidents/injuries
  - Stressful working conditions
  - Complexity of tasks (requiring specific training or unusual skills)
  - Excess labour demand
  - Possibility to inflict losses on employers/society (unions; industrial action)
  - Entry barriers
- Detailed information required on the kind of jobs being undertaken by individuals $\rightarrow$ Even very detailed information on occupations not enough for this purpose (why?)
- Advantages of including job title fixed effects in the wage regression:
  - Evaluation of the contribution of job title heterogeneity, alongside with worker and firm heterogeneity, to the wage variation
  - Production of refined estimates of worker and firm fixed effects, as they are filtered from the job title fixed effects