Workers, Workplaces, Sorting, and Wage Dispersion in Mexico

Jorge Pérez Pérez[1] José Nuño-Ledesma[2]

[1]Banco de México
[2]University of Guelph

LACEA LAMES 2023 - November 17th, 2023

Disclaimers

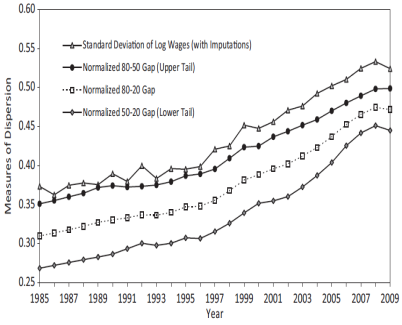
- The views and conclusions presented in this document are the exclusive responsibility of the authors and do not necessarily reflect those of Banco de Mexico.
- We acknowledge financial support from Colombia Cientifica – Alianza EFI 60185 contract FP44842- 220-2018, funded by The World Bank through the Scientific Ecosystems, managed by the Colombian Ministry of Science, Technology, and Innovation (MINCIENCIAS).
- The data were accessed through the Econlab at Banco de Mexico. The EconLab collected and processed the data as part of its effort to promote evidence-based research and foster ties between Banco de Mexico's research staff and the academic community. Inquiries regarding the terms under which the data can be accessed should be directed to econlab@banxico.org.mx

Summary

- **What do we do?**
  - Decompose wage variance in Mexico into components attributable to worker factors, establishment factors, and worker-workplace sorting.
  - Describe how these contributions vary over time and across regions.
- **How do we do it?**
  - Estimate models for log-wages with additive worker and workplace fixed effects.
  - Decompose the total log-wage variance using the model estimates.
- **What do we find?**
  - Increasing role of sorting and workplace effects in determining wage dispersion.
  - Workplace factors role: lowest in North (most-prosperous), and largest in South (least-developed).

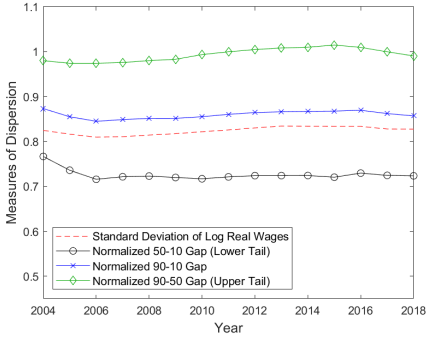
Motivation: Trends in Wage Dispersion. Germany vs Mexico

(a) **Germany**



Source: Card et al. (2013).

(b) **Mexico**

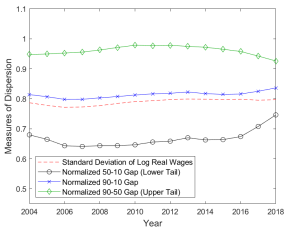


Source: Author's calculations with IMSS data.

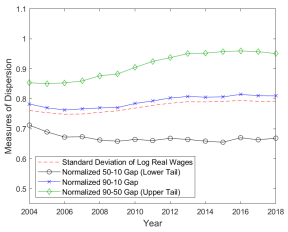
Motivation: Mexico and its Regions

Motivation: Regional Wage Dispersion

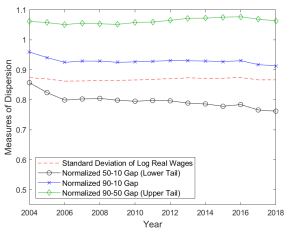
(a) **North**



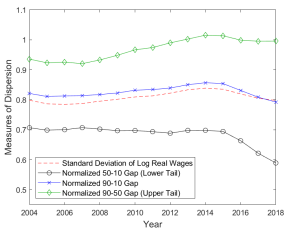
(b) **Center-North**



(c) **Center**



(d) **South**



Motivation: Workplace Effects and Development

- Firm effects explain 28% of total wage variance in Senegal and South Africa (Diallo et al. 2022; Bassier 2022).
- Bassier (2022): these estimates are significantly higher compared to estimates from developed European and American countries.
- Messina et al. (2019) document lower wage inequality across workplace as inequality falls in Latin American countries.
- **Does this pattern emerge within countries?**

Data

- Social security records from *Instituto Mexicano del Seguro Social* (IMSS).
- 83% of private-sector formal workers are affiliated with IMSS (as of 2022).
  - IMSS does not collect information from workers employed by the government or working in the informal economy.
- We use records registered between November 2004 and April 2018.
- Number of observations within the range of 12.8 million (November 2004) and 20.1 million (December 2018).

Data

- Key variables:
  - **Worker ID**: Social security number.
  - **Workplace ID**: *Registro patronal*.
  - **Wage**: Daily taxable income.
  - **Other**: Year of birth, gender.
- Data are bottom-coded (≈$5 USD per day) and top-coded (≈ $102 USD per day.)
- Only 1.9% of the observations are top coded.
- Data do not include part-time working status or education variables.
- One employee can be registered as working for more than one employer.
- An employer-employee pair can appear more than once in a month with different income.
- We pair only one job per worker: whichever reports the highest income.

Descriptive Statistics: Workers, Prime-Age men (25-54 y.o.), National Level

		Real wage		
	(1)	(2)	(3)	(4)
	Observations	Mean	Std. dev	Percent censored
2005	73,855,547	394.575	406.256	2.336
2009	80,069,659	394.594	403.065	2.359
2014	98,566,773	391.698	407.856	2.300
2018	113,516,335	397.765	410.689	2.626

Observations correspond to the sum of all the monthly observations in a year. Real wages using prices of July 2018. Percent censored is the percentage of observations with wages exactly equal to the upper wage limit.

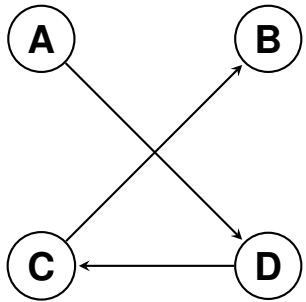
AKM model (Abowd, Kramarz, and Margolis 1999)

$$\ln W_{it} = \alpha_i + \psi_{\mathbf{J}(i,t)} + X'_{it}\beta + r_{it} \tag{1}$$

Where

- W_{it} is the real wage of worker i at period t.
- α_i are worker effects. Factors that are rewarded equally across employers giving rise to a worker-specific wage component.
- $\psi_{\mathbf{J}(i,t)}$ are establishment effects. Proportional wage premium (or discount) that is paid by firm J to all employees.
- x'_{it} is a vector of observable worker characteristics. We include age, age squared, age cube, and a time trend
- r_{it} is the error term.

Connected Set



Source: Fenizia (2019)

- The firm and worker effects in (1) are separately identified within a "connected set" of firms linked by worker mobility.
- We restrict analysis to the largest connected set in four time intervals: 2004-2008, 2009-2013, and 2014-2018.
- The ratio of observations in largest connected set to all observations ranges between 94.9% and 97.3% (and between 97.5% ad 98.6% individuals).

Connected Set

	All sample				Individuals in largest connected set			
			Log wage				Log wage	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Interval	All obs.	Individuals	Mean	Std. dev.	All obs.	Individuals	Mean	Std. dev.
Nov 2004-2008	158,543,931	5,721,179	5.525	0.772	150,458,370	5,576,345	5.556	0.772
Ratio: largest connected/all					94.91	97.51	100.61	100.01
2009-2013	226,528,652	7,072,043	5.487	0.791	216,360,702	6,920,461	5.515	0.792
Ratio: largest connected/all					95.51	97.91	100.51	100.11
2014-2018	297,395,413	9,069,558	5.488	0.793	288,394,833	8,941,908	5.507	0.793
Ratio: largest connected/all					97.01	98.61	100.31	100.01
Change from first to last interval			-0.0371	0.0211			-0.0491	0.0211

AKM Model Summary

	Interval1 2004-2008	Interval2 2009-2013	Interval3 2014-2018
Worker and workplace parameters			
Number of worker effects	5,576,345	6,920,461	8,941,908
Number of workplace effects	523,701	554,593	695,749
Summary of parameter estimates			
St. dev. of worker effects	0.504	0.486	0.472
St. dev. of workplace effects	0.444	0.479	0.487
Correlation worker/workplace effects	0.212	0.231	0.259
Correlation worker effects/Xb	-0.123	-0.074	-0.104
Correlation workplace effects/Xb	-0.051	-0.045	-0.048
Goodness of fit			
St. dev. of log wages	0.772	0.792	0.793
RMSE	0.238	0.237	0.233
R Squared	0.909	0.913	0.916
Adj. R Squared	0.905	0.910	0.913

Variance and Assortative Matching

The variance of wages can be decomposed as follows:

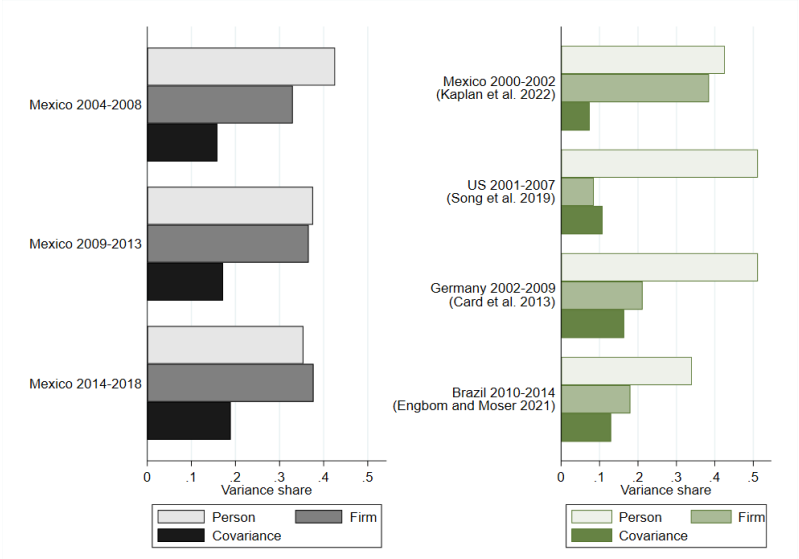
$$\begin{aligned}\text{Var}\left(\ln W_{it}\right) = &\underbrace{\text{Var}(\alpha_i)}_{\text{workers}} + \underbrace{\text{Var}(\psi_{\mathbf{J}(i,t)})}_{\text{firms}} + \text{Var}(x_{it}'\beta) + \text{Var}(r_{it}) \\ &+ 2\underbrace{\text{Cov}(\alpha_i, \psi_{\mathbf{J}(i,t)})}_{\text{sorting}} + 2\,\text{Cov}(\psi_{\mathbf{J}(i,t)}, x_{it}'\beta) + 2\,\text{Cov}(\alpha_i, x_{it}'\beta).\end{aligned} \tag{2}$$

Positive covariance of α_i and $\psi_{\mathbf{J}} \rightarrow$ greater wage inequality, i.e. high quality workers tend to be matched with high quality firms

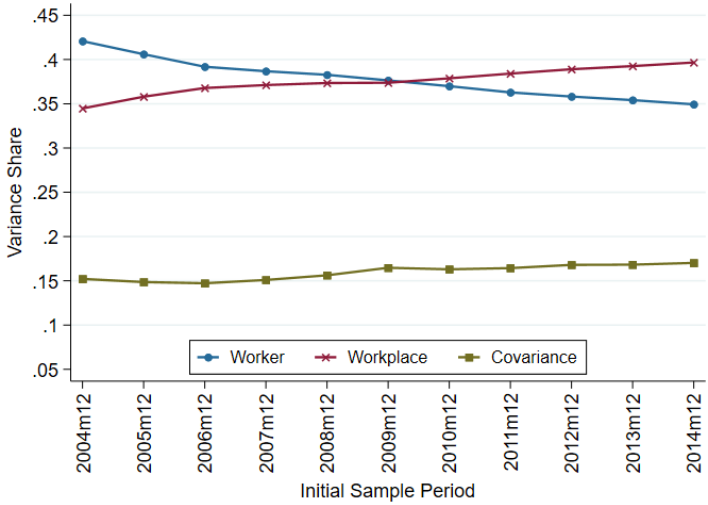
Wage Variance Decomposition, National Level

	Interval 1 2004-2008	Interval 2 2009-2013	Interval 3 2014-2018	Change from int. 1 to 3
Variance and covariance				
Total variance of log wages	0.596	0.627	0.628	0.032
Variance of worker effects	0.254	0.236	0.222	-0.032
Variance of workplace effects	0.197	0.230	0.237	0.040
Variance of covariates (Xb)	0.019	0.013	0.016	-0.004
Variance of residual	0.055	0.054	0.053	-0.002
2 Cov(worker effects, workplace effects)	0.095	0.108	0.119	0.024
2 Cov(worker effects, covariates)	-0.017	-0.008	-0.012	0.005
2 Cov(workplace effects, covariates)	-0.006	-0.005	-0.006	0.000
Variance shares				
Variance of worker effects	0.426	0.376	0.354	-0.073
Variance of workplace effects	0.330	0.366	0.377	0.047
Variance of covariates (Xb)	0.032	0.020	0.025	-0.006
Variance of residual	0.091	0.087	0.084	-0.008
2 Cov(worker effects, workplace effects)	0.159	0.172	0.189	0.030
2 Cov(worker effects, covariates)	-0.029	-0.013	-0.019	0.009
2 Cov(workplace effects, covariates)	-0.011	-0.008	-0.009	0.001
Counterfactuals for variance of log wages				
1. No rise in correl. of worker/firm effects	0.596	0.618	0.608	
2. No rise in var. of workplace effects	0.596	0.587	0.578	
3. Both 1 and 2	0.596	0.585	0.568	

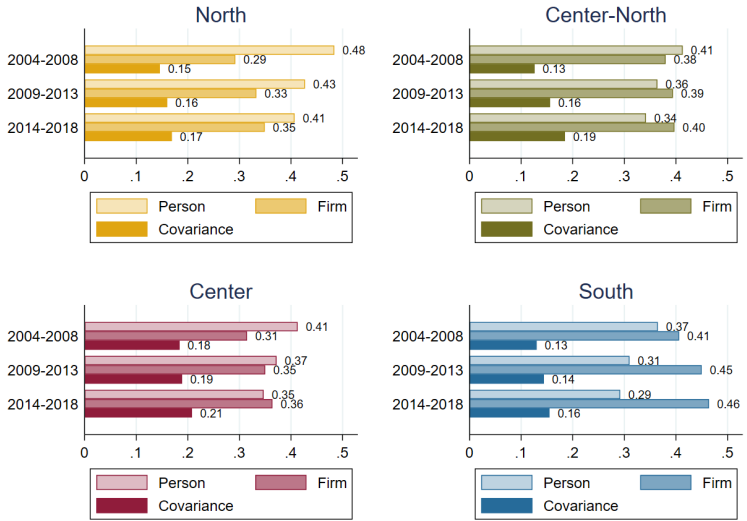
Comparing Estimated Worker and Workplace Contributions to Wage Dispersion



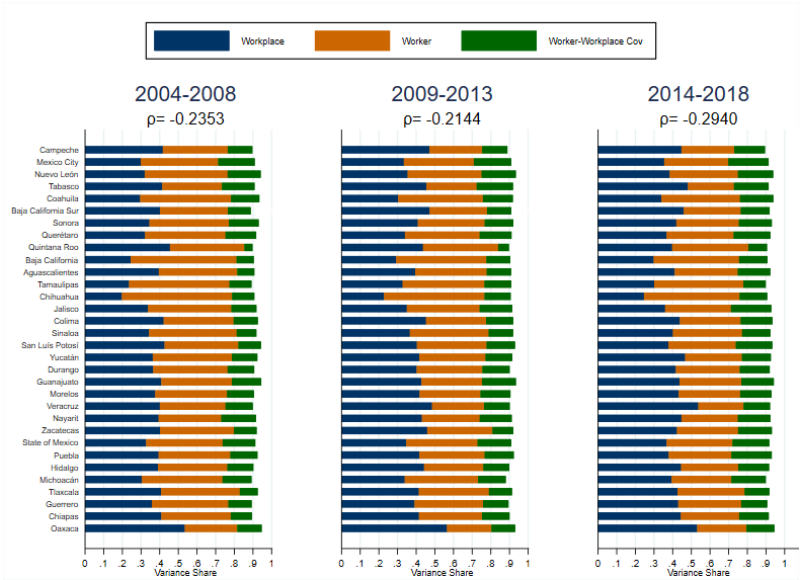
“Rolling” Variance Decomposition, National Level



Worker and Workplace Contributions to Wage Variability by Region



State Differences in Assortative Matching



Mechanisms behind regional differences

- **Educational attainment:** Lower in the south, can lead to lower variance of worker effects.
- **Labor market power:** Would expect more variance of firm premia if labor supply elasticities varies across firms. More labor market concentration in the south.
- **Industrial composition:** Better matching in the services sector located in the center.
- **Firm size:** Average firm size correlated with the share of variance due to assortative matching.
- **City size and Informality:** Weaker assortative matching in smaller cities with more informal labor. Pérez Pérez, Meléndez and Nuño-Ledesma (2023).

Concluding Remarks

- Workplace factors play a more important role in the South, less important role in the North.
- The importance of workplace factors decreases with the level of development.
- Regional heterogeneity in assortative matching.

jorgepp@banxico.org.mx
www.jorgeperezperez.com

Additional slides

Variance decomposition with the KSS estimator

	Interval 1 2004-2008	Interval 2 2009-2013	Interval 3 2014-2018
Connected set			
Total variance of log wages	0.596	0.627	0.628
Variance of worker effects	0.254	0.236	0.222
Variance of workplace effects	0.197	0.230	0.237
2 Cov(worker effects, workplace effects)	0.095	0.108	0.119
Leave-one-out connected set			
Total variance of log wages	0.596	0.627	0.628
Variance of worker effects	0.254	0.235	0.222
Variance of workplace effects	0.193	0.227	0.235
2 Cov(worker effects, workplace effects)	0.098	0.110	0.121
KSS corrected in leave-one-out connected set			
Total variance of log wages	0.596	0.627	0.628
Variance of worker effects	0.252	0.234	0.220
Variance of workplace effects	0.193	0.226	0.234
2 Cov(worker effects, workplace effects)	0.099	0.111	0.121

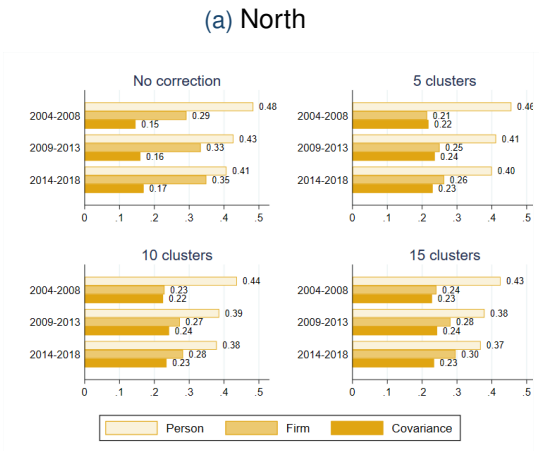
Source: Author's calculations. We use the "match" leave-one-out estimator, leaving out worker-workplace matches one at a time. To approximate the components, we use 50 iterations of the JILA algorithm.

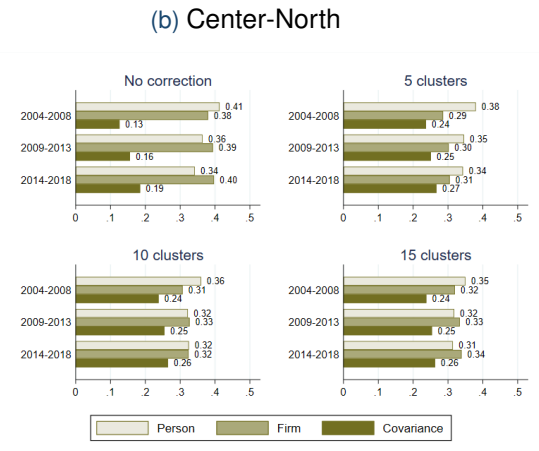
Variance Decomposition with Bonhomme et al's 2019 Correction for Limited Mobility Bias

	Interval 1 2004-2008	Interval 2 2009-2013	Interval 3 2014-2018
No clusters			
Variance and covariance			
Total variance of log wages	0.596	0.627	0.628
Variance of person effects	0.254	0.236	0.222
Variance of firm effects	0.197	0.230	0.237
2 Cov(person effects, firm effects)	0.095	0.108	0.119
Variance shares			
Variance of person effects	0.426	0.376	0.354
Variance of firm effects	0.330	0.366	0.377
2 Cov(person effects, firm effects)	0.159	0.172	0.189
5 clusters			
Variance and covariance			
Total variance of log wages	0.596	0.626	0.628
Variance of person effects	0.231	0.216	0.206
Variance of firm effects	0.159	0.188	0.197
2 Cov(person effects, firm effects)	0.144	0.154	0.161
Variance shares			
Variance of person effects	0.387	0.344	0.328
Variance of firm effects	0.266	0.301	0.313
2 Cov(person effects, firm effects)	0.241	0.247	0.256

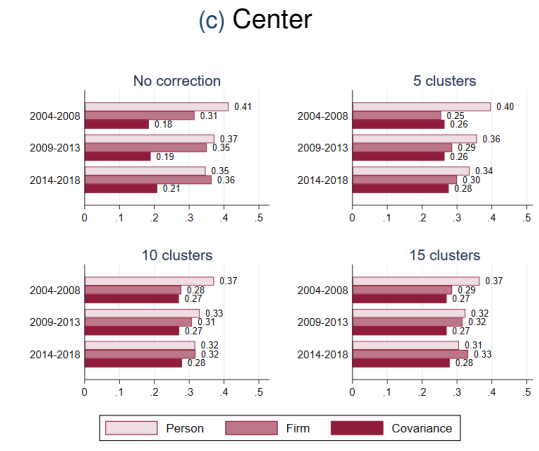
Source: Authors' calculations We use 20 percentiles of the within workplace log wage distribution to cluster workplaces in 5 groups to estimate the AKM model..

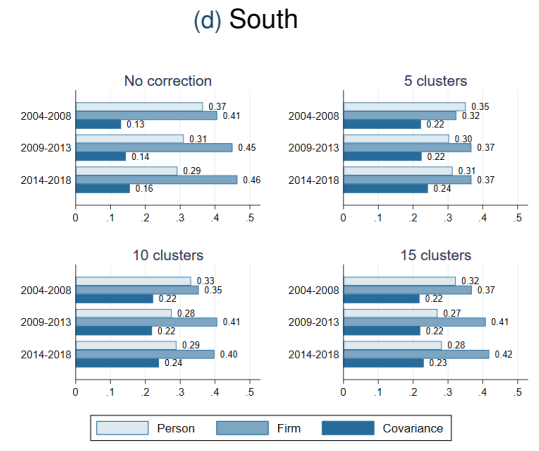
Estimated Worker and Workplace Contributions to Wage Variability by Region with Correction for Limited Mobility Bias





Estimated Worker and Workplace Contributions to Wage Variability by Region with Correction for Limited Mobility Bias





Average Worker and Workplace Fixed Effects by Region

	Average log wage	Average worker effect	Average workplace effect
National			
2004-2008	5.56	2.67	2.46
2009-2013	5.51	2.63	2.42
2014-2018	5.51	2.63	2.37
North			
2004-2008	5.56	2.65	2.50
2009-2013	5.51	2.63	2.45
2014-2018	5.54	2.62	2.39
Center-North			
2004-2008	5.49	2.64	2.48
2009-2013	5.44	2.62	2.43
2014-2018	5.43	2.60	2.36
Center			
2004-2008	5.64	2.72	2.45
2009-2013	5.59	2.69	2.43
2014-2018	5.57	2.66	2.39
South			
2004-2008	5.41	2.63	2.45
2009-2013	5.41	2.64	2.36
2014-2018	5.37	2.59	2.33

Source: Authors' calculations using IMSS data. Average log wages, worker fixed effects and workplace fixed effects for each region, using the estimates of the AKM model from equation (**??**).

Wage Variance Decomposition, National Level. Women Ages 25-54

	Interval 1 2004-2008	Interval 2 2009-2013	Interval 3 2014-2018	Change from int. 1 to 3
Variance and covariance				
Total variance of log wages	0.546	0.573	0.559	0.013
Variance of person effects	0.296	0.271	0.257	-0.039
Variance of firm effects	0.156	0.183	0.184	0.029
Variance of covariates (Xb)	0.021	0.014	0.022	0.001
Variance of residual	0.044	0.045	0.045	0.001
2 Cov(person effects, firm effects)	0.082	0.093	0.098	0.016
2 Cov(person effects, covariates)	-0.045	-0.027	-0.040	0.005
2 Cov(firm effects, covariates)	-0.008	-0.006	-0.007	0.001
Variance shares				
Variance of person effects	0.542	0.473	0.460	-0.082
Variance of firm effects	0.285	0.319	0.330	0.044
Variance of covariates (Xb)	0.039	0.024	0.039	0.001
Variance of residual	0.080	0.079	0.080	-0.000
2 Cov(person effects, firm effects)	0.150	0.163	0.176	0.025
2 Cov(person effects, covariates)	-0.082	-0.047	-0.072	0.011
2 Cov(firm effects, covariates)	-0.015	-0.011	-0.013	0.002
Counterfactuals for variance of log wages				
1. No rise in correl. of person/firm effects	0.546	0.563	0.545	
2. No rise in var. of firm effects	0.546	0.538	0.522	
3. Both 1 and 2	0.546	0.536	0.516	