Do Workers Sort to Firms or to Occupations?
ONLINE APPENDIX

Luke Heath Milsom*   Shihang Hou*

October 2, 2025

Contents

---

*Heath Milsom: Institute for Fiscal Studies and KU Leuven, luke.heathmilsom@kuleuven.be. Hou: Institute for Employment Research (IAB), shihang.hou@iab.de.

1 Summary Statistics for Other Periods

1.1 France

Table 1: Summary statistics, 2010-14

	Full data	Firms connected set	Firms LOO set	Jobs connected set	Jobs LOO set
2010-14					
N obs	46,103,347	39,467,309	37,286,700	33,154,058	28,775,013
N workers	12,664,742	9,400,612	8,890,944	7,934,805	7,034,652
N firms	1,146,870	580,256	329,961	502,751	222,969
N jobs	4,989,214	3,561,234	2,921,078	2,209,007	821,744
Mean log annual wage	10.34	10.37	10.38	10.36	10.38
Var log annual wage	0.24	0.23	0.23	0.23	0.23
Mean log hourly wage	2.89	2.91	2.92	2.91	2.92
Var log hourly wage	0.19	0.19	0.19	0.19	0.18
Var residualised log hourly wage	0.23	0.22	0.22	0.21	0.21
N moves	8,503,206	2,954,239	2,694,455	6,003,822	4,562,592
N firm moves	3,131,395	2,954,239	2,694,455	2,760,518	2,116,072
N occ moves	7,199,647	2,660,426	2,395,688	4,855,056	3,569,830
N firm + occ moves	1,827,836	1,735,055	1,550,929	1,611,752	1,123,310

Notes: This table shows the summary statistics for data from 2010 to 2014, cleaned in the same way as our main sample, which we use as a robustness check. The underlying data is from yearly BTS data files, French administrative matched employee-employer data. Individuals are mapped over time using the procedure and kindly provided programs in Babet et al. [2022]. Sample construction and restrictions are discussed in the main text.
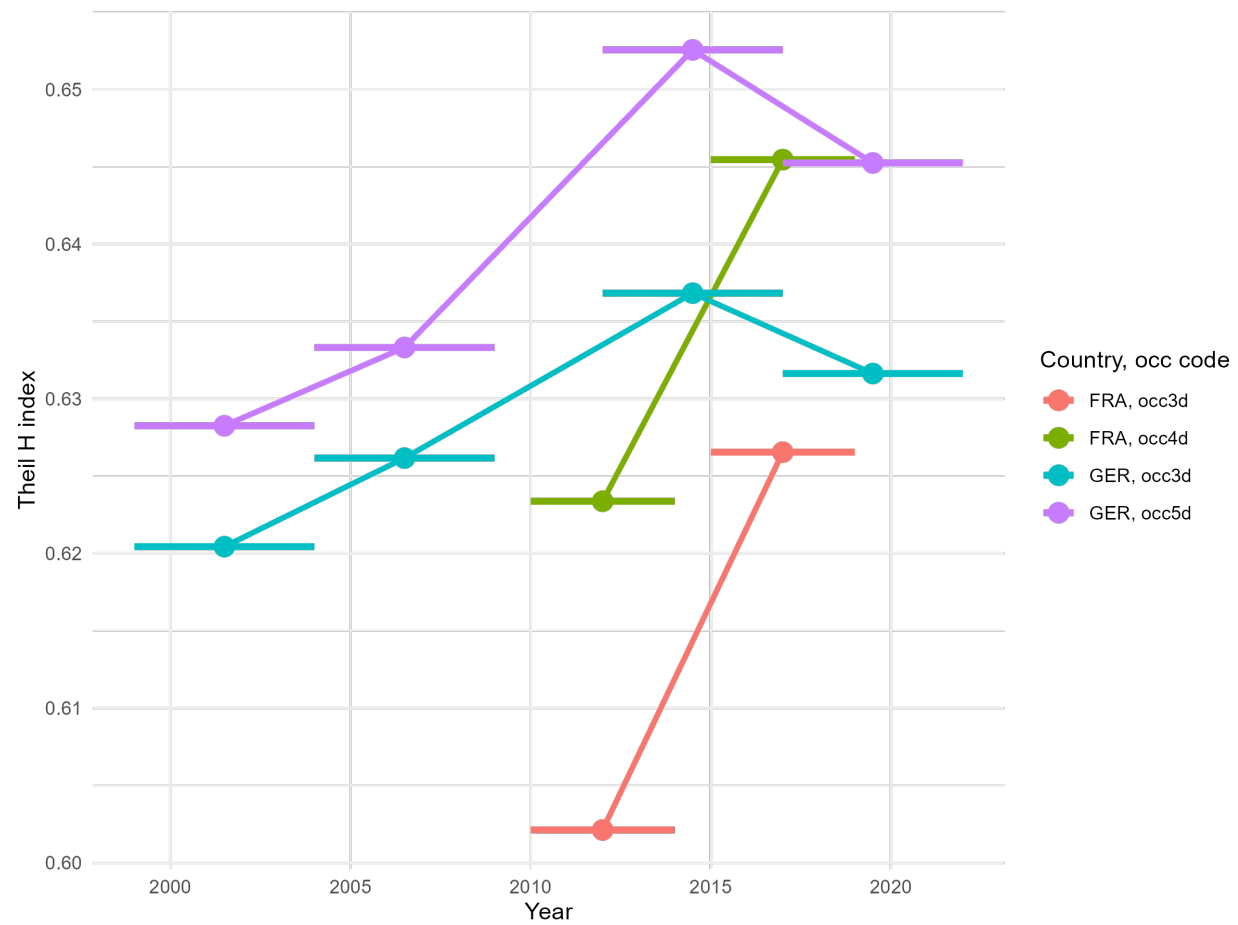
1.2 Germany

Table 2: Descriptive statistics for entire West German sample, in four periods from 1999-2022

	1999-2004	2004-09	2012-17	2017-22
Total observations (m)	60.34	57.78	62.47	54.56
Total workers (m)	13.62	12.81	14.03	14.22
Total firms (m)	1.56	1.49	1.40	1.35
Total jobs (m)	4.03	3.78	4.79	4.91
Average daily wages	4.77	4.74	4.75	4.78
SD daily wages	0.49	0.53	0.53	0.52
SD resid. daily wages	0.48	0.51	0.51	0.50
N moves	6.51	5.34	6.16	5.79
N firm moves	5.74	4.76	5.37	4.88
N occ moves	3.20	2.58	3.85	3.77
N firm + occ moves	2.43	2.00	3.06	2.86
Largest connected set				
Total observations (m)	48.69	44.81	47.24	40.16
Total workers (m)	9.89	8.98	9.56	9.44
Total firms (m)	0.99	0.87	0.85	0.78
Total jobs (m)	2.24	1.91	2.41	2.30
Average daily wages	4.80	4.78	4.77	4.80
SD daily wages	0.47	0.51	0.52	0.51
SD resid. daily wages	0.46	0.49	0.50	0.49
N moves	5.65	4.62	5.37	5.00
N firm moves	5.08	4.18	4.71	4.27
N occ moves	2.79	2.26	3.42	3.29
N firm + occ moves	2.23	1.82	2.76	2.56
Largest leave-out observation connected set				
Total observations (m)	42.22	38.06	39.32	32.96
Total workers (m)	8.73	7.76	8.17	7.95
Total firms (m)	0.54	0.45	0.44	0.39
Total jobs (m)	1.05	0.86	1.00	0.92
Average daily wages	4.81	4.81	4.79	4.82
SD daily wages	0.46	0.50	0.52	0.51
SD resid. daily wages	0.45	0.49	0.50	0.49
N moves	4.43	3.53	3.94	3.58
N firm moves	3.98	3.18	3.40	3.01
N occ moves	2.02	1.58	2.37	2.21
N firm + occ moves	1.57	1.22	1.82	1.64

Notes: This table presents summary statistics for all four periods for three samples - the entire population, the largest connected set of the main worker-job specification, and the largest leave-out observation set of the main worker-job specification. We present the total number of observations, the total number of unique workers, the total number of unique firms and the total number of unique jobs in millions, the total number of moves, as well as the average and standard deviation of the daily (imputed) wage as well as the standard deviation of the daily wage residualised on year fixed effects and a cubic age profile.

2 Theil index

Figure 1: Segregation of occupations across firms in studied periods

Notes: This figure presents the Theil H-Index in Germany and France at both the 3 digit level and at the most detailed level (5 digit for Germany, 4 digit for France).

3 Variance decomposition table

Table 3: Variance decomposition: France

Component		Firm model			Job model		
		Variance	Proportion of total variance	Proportion of component variance	Variance	Proportion of total variance	Proportion of component variance
Worker	Total	0.118	0.574	1	0.098	0.493	1
Job/ Firm	Total	0.013	0.064	1	0.021	0.106	1
	Within Occupation				0.011	0.06	0.52
	Between Occupation				0.010	0.05	0.48
Sorting	Total	0.019	0.091	1	0.042	0.210	1
	Within Occupation				0.008	0.04	0.19
	Between Occupation				0.034	0.17	0.81
Error & controls	Total	0.056	0.271	1	0.038	0.192	1

Notes: This table decomposes log wage variance in the French sample using the worker-job model described in the main text. It also shows the results from applying the law of total variance and law of total covariance to further decompose job-level components into occupation and firm effects.
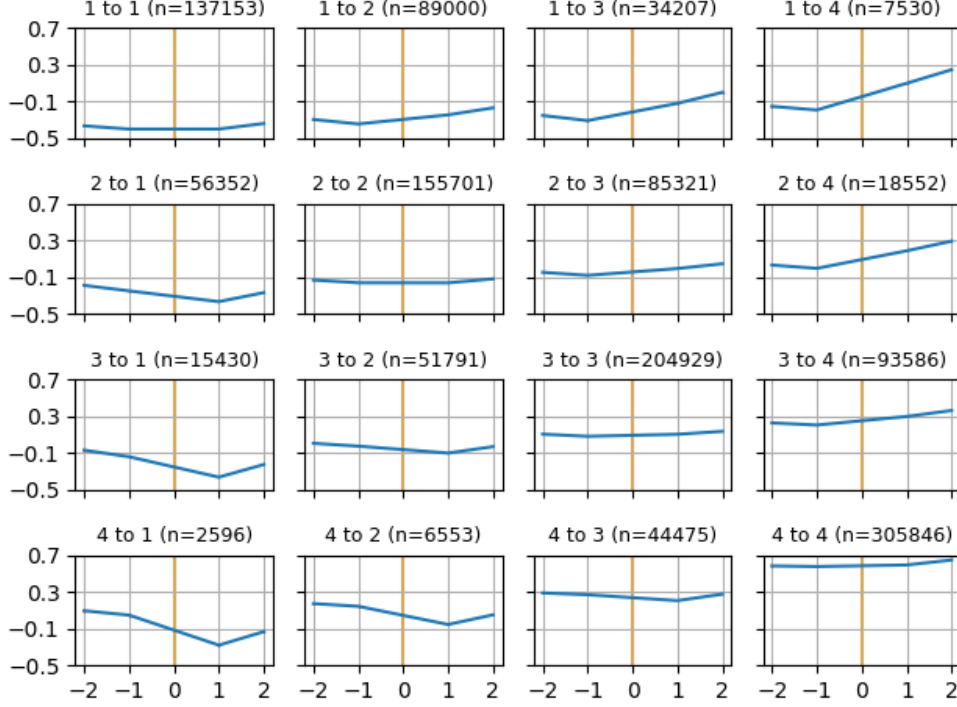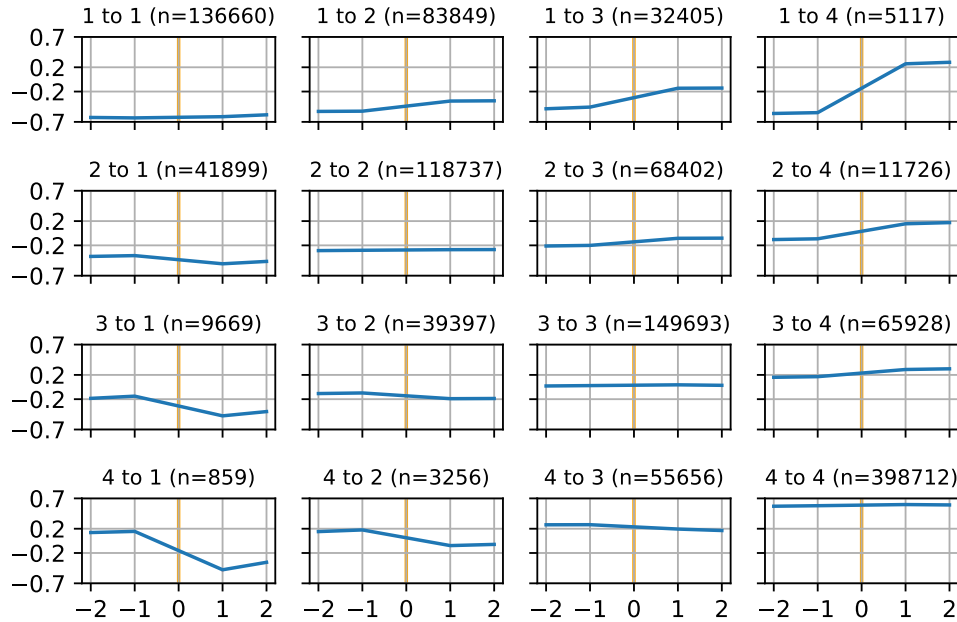
Table 4: Variance decomposition: Germany

Component		Firm model			Job model		
		Variance	Proportion of total variance	Proportion of component variance	Variance	Proportion of total variance	Proportion of component variance
Worker	Total	0.130	0.525	1	0.106	0.416	1
Job/ Firm	Total	0.027	0.107	1	0.047	0.184	1
	Within Occupation				0.030	0.12	0.64
	Between Occupation				0.017	0.07	0.36
Sorting	Total	0.035	0.143	1	0.054	0.213	1
	Within Occupation				0.002	0.05	0.05
	Between Occupation				0.051	0.20	0.95
Error & controls	Total	0.056	0.226	1	0.048	0.187	1

Notes: This table decomposes log wage variance in the German sample using the worker-job model described in the main text. It also shows the results from applying the law of total variance and law of total covariance to further decompose job-level components into occupation and firm effects.

4 Event studies

Figure 2: Event study around job moves, clustering by leave-out job mean wage
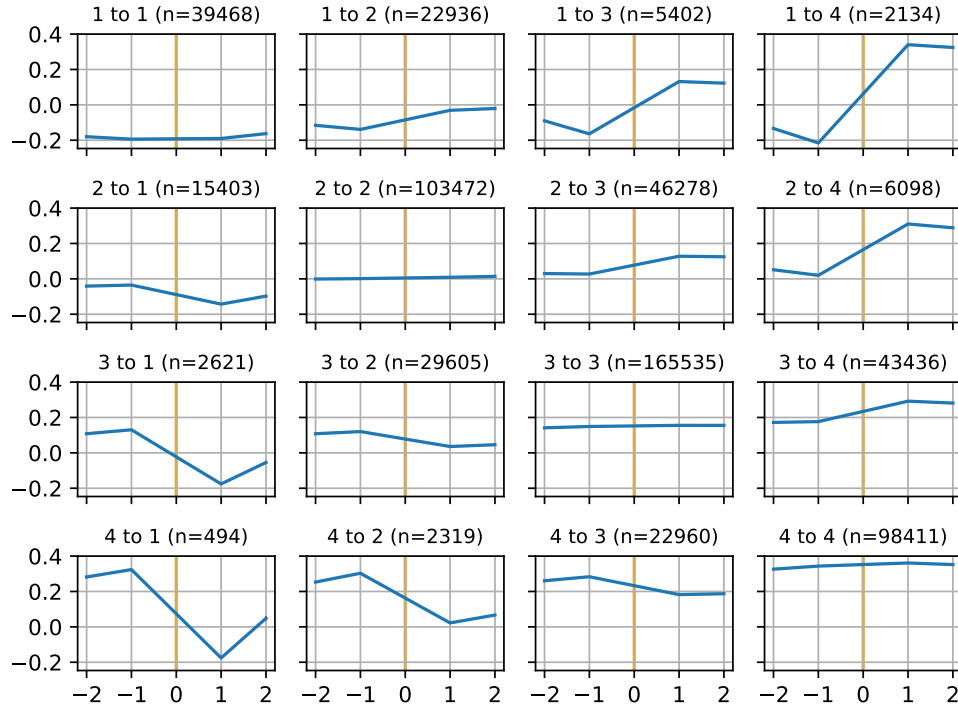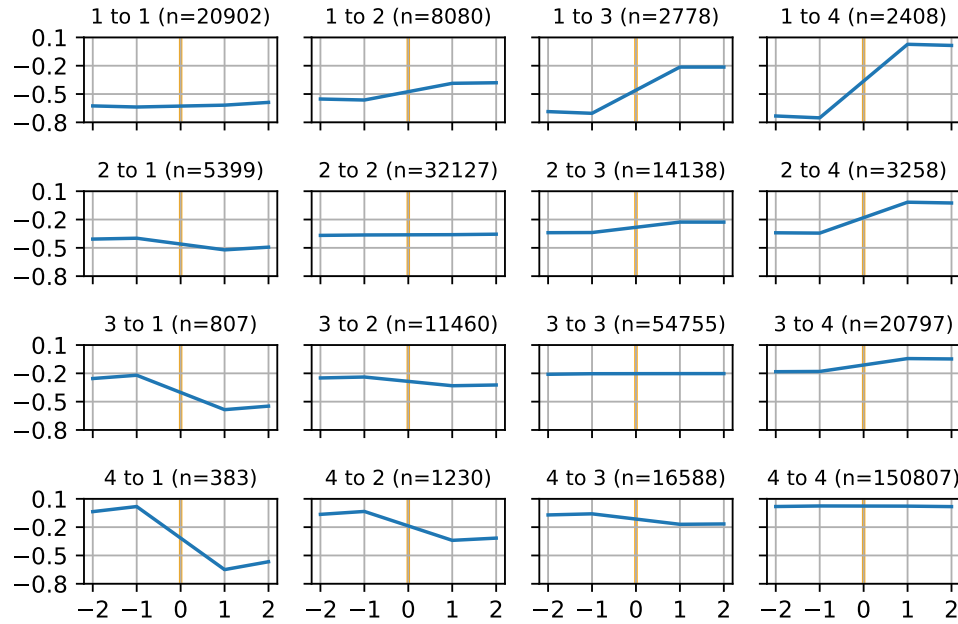
(a) France

(b) Germany

Notes: These figures show the impact on average wages around the event of job movement. Each cell shows the average wage change associated with a movement event from one quartile to another quartile of the average job wage distribution. Following Card et al. [2013], we cluster jobs into quartiles by computing the mean leave-out job wage. Only those who remain in their old job for two years before and their new jobs for two years after the move event are included. The number of switchers in each cell is given in the cell title. Panel (a) shows the results for France, and panel (b) shows the results for Germany.

Figure 3: Event study around job moves, including only moves across occupations within firms, clustering by estimated job fixed effect
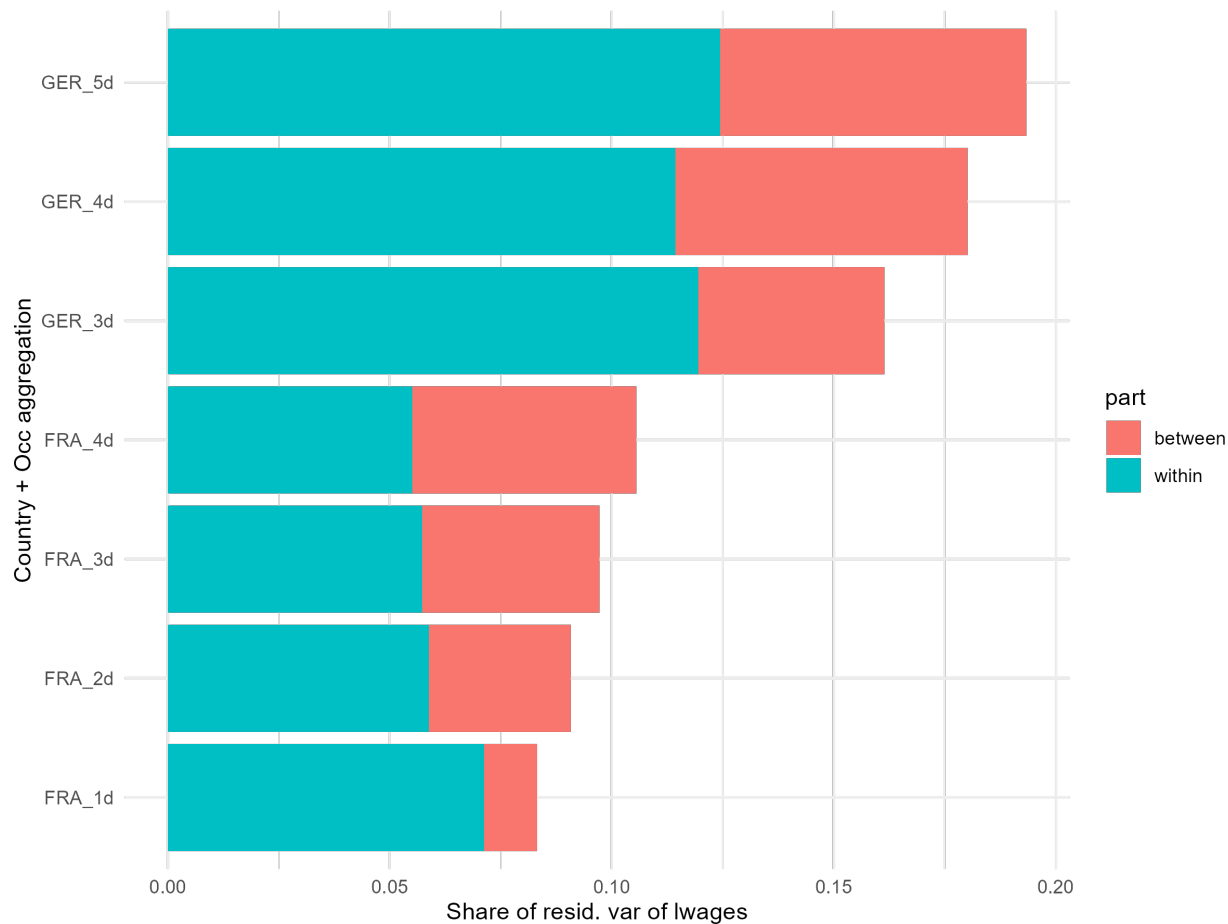
(a) France

(b) Germany

Notes: These figures show the impact on average wages around the event of job movement across occupations within firms. Each cell shows the average wage change associated with a movement event from one quartile to another quartile of the average job fixed effect distribution. Following Card et al. [2013], we cluster jobs into quartiles by computing the mean leave-out job fixed effect within the job excluding. Only those who remain in their old job for two years before and their new jobs for two years after the move event are included. The number of switchers in each cell is given in the cell title. Panel (a) shows the results for France, and panel (b) shows the results for Germany.
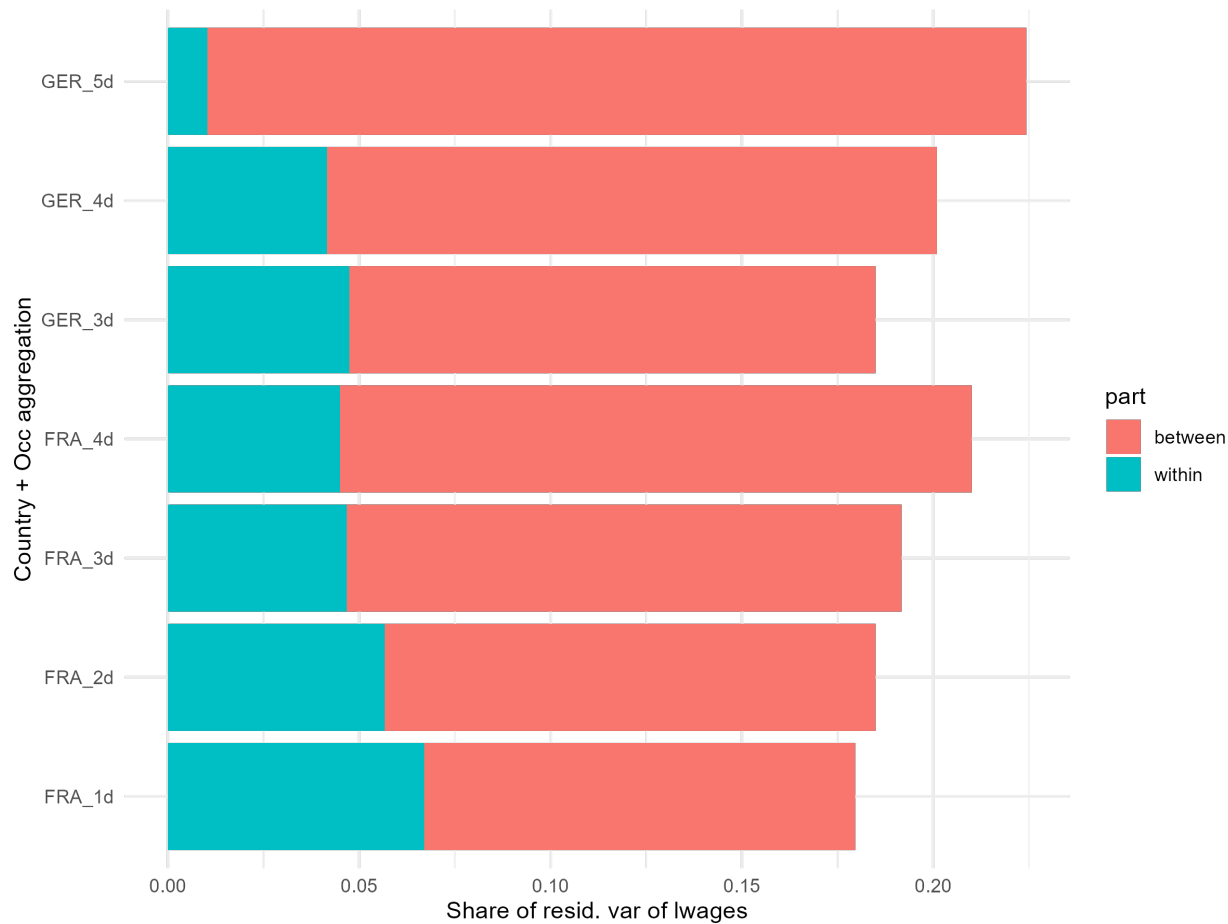
7

5 Granularity of occupation coding

Figure 4: Decomposing job variance into within and between occupation components by different levels of occupation granularity

Notes: This figure shows the decomposition of job variance into between and within occupation components for various definitions of occupation at different granularities. The dimension of granularity is indicated on the left-hand side of each bar.
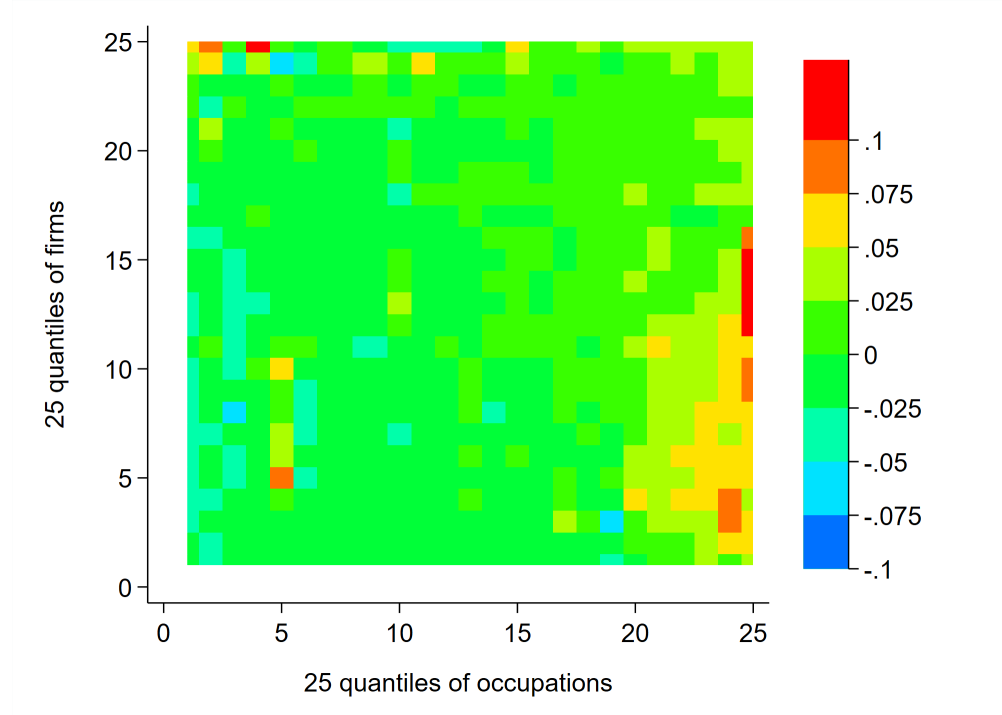
Figure 5: Decomposing worker-job covariance into within and between occupation components by different levels of occupation granularity

6 Linearity of worker-occupation effects

Figure 6: Contribution of worker-job match effects by firm and occupation mean pay quantiles

Notes: The figure shows a heatmap of the difference between the estimated job fixed effect in our main specification, and the sum of a firm fixed effect and an occupation fixed effect in an auxiliary specification, which we interpret as match effects between occupation and firms. Green squares imply that if the job fixed effect were replaced by the additive firm and occ fixed effects, the resulting predicted log wage would be within 0.05 log points of the true predicted log wage with job FE. Red implies that using the additive specification would underpredict the true wage by over 0.1 log points and blue implies that overprediction by over 0.075 log points. The x-axis indexes 25 quantiles of occupations, sorted by the mean wage within the occupation, and the y-axis indexes 25 quantiles of firms, sorted by the mean wage within the firm.

References

Damien Babet, Olivier Godechot, and Marco G Palladino. In the land of akm: Explaining the dynamics of wage inequality in france. 2022.

David Card, Jörg Heining, and Patrick Kline. Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, 128(3):967–1015, 2013.