Firms Heterogeneity and the Return to Cognitive Skills

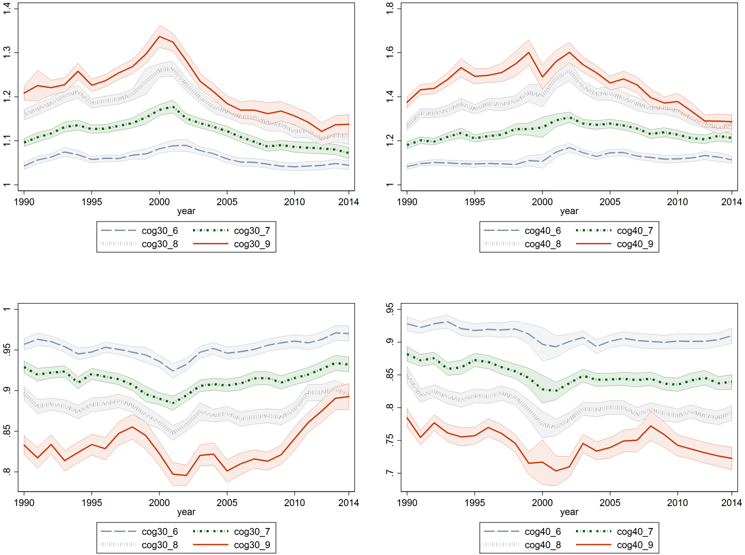
Michael Böhm (University of Bonn)

Khalil Esmkhani (UBC, Sauder Business School)

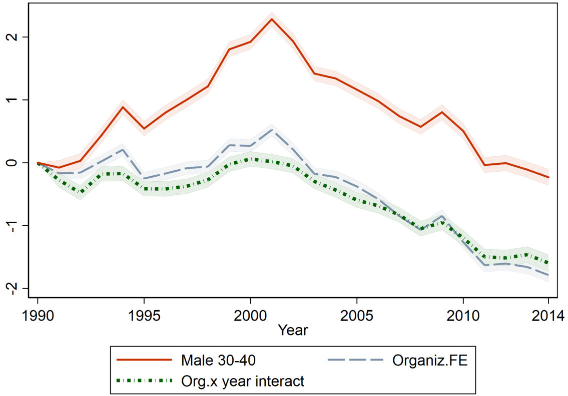
Giovanni Gallipoli (UBC, Vancouver School of Economics)

June 10, 2018

Earnings by Cognitive Skill: Non-Parametric Estimates



The Return to Cognitive Skills: Parametric Estimates



We estimate different versions of the general wage equation:

$\ln w_{ijt} = \alpha_t + \theta_{jt} + \lambda_t \ln(ST_i) + controls_t + \varepsilon_{ijt}$.

ST_i are cognitive skills, $\ln w_{ijt}$ log earnings, t is sample year. Sample: 30 and 40 year old males.

Motivation and Objectives

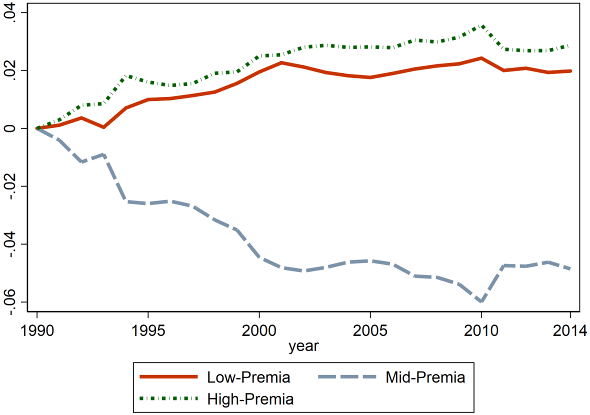
- Mounting evidence of run-up and subsequent slow down in return to cognitive skills over past decades.
    - Lively debate, motivated by surging economic inequality.
- Recurring question: what drives the evolution of skill premia and inequality?
- Large employment and wage shifts suggest structural change in labor market.
- Establishing what drives these shifts is hard because
    a. outcomes determined in equilibrium.
    b. many confounding factors and measurement issues.

We focus on two aspects of the general question above:

- Does firms heterogeneity matter for observed skill premia?
- If so, how does it interact with workers heterogeneity?

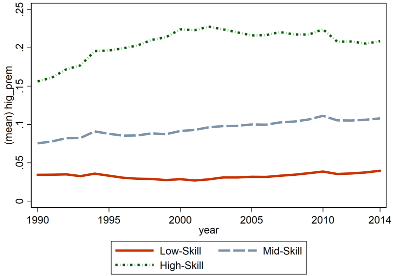
⇒ **Objective:** Move beyond fixed effects. Estimate skill sorting in a model with two-sided heterogeneity.

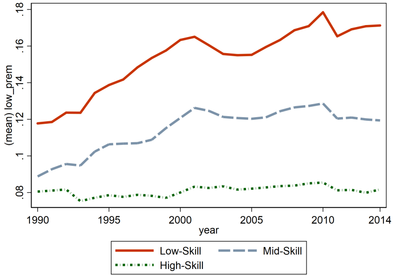
Employment Polarization: Parametric Estimates



Changes in employment shares in high-wage (one $sd(\theta)$ above mean in respective year), low-wage $sd(\theta)$ below), and middle-wage (all other) firms. ▸ Data Details

Workers Sorting: Cognitive Skills





The plots show the unconditional probability that high-skill (stanine cognitive 7–9), middle-skill (4–6), and low-skill (1–3) workers are employed in, respectively, high wage premia and low wage premia firms. ▸ Data Details

Key Facts: Summary

a. Economy wide return to skills had a boom-bust episode over the 1990-2014 interval.

  - Between-firm heterogeneity explain most of the run-up and a much smaller part (roughly 1/3) of the reversal.
  - Net of composition effects due to between-firm heterogeneity, the return to cognitive skills has been declining since 1990.
  - Decline in average within-firm return to cog. skills accelerated after 2000 and accounts for most of reversal.

b. Significant firm-specific wage premia persist after controlling for worker skills.

  - Skill-intensive firms exhibit higher wage premia on average.
  - Firms that employ many high-skill workers increased their hiring during the 1990s, compared to less skill-intensive firms.

A Model of Skill Demand by Heterogeneous Firms

We develop a model with two-sided heterogeneity (firms and workers) and imperfect competition in both input and output markets.

a. **Firms produce using the labor of workers with different skills.**

b. **Workers are ex-ante heterogeneous in:**
   1. productive skills;
   2. non-pecuniary preferences for different firms.

c. **Firms are heterogeneous in:**
   1. Production function: firm-specific returns to workers' skill inputs.
   2. Monopoly power (output market): Dixit-Stiglitz demand shares.
   3. Monopsony power (input market): non-pecuniary returns.

▸ Model Details: the Problem of the Firm

Implications of the Model

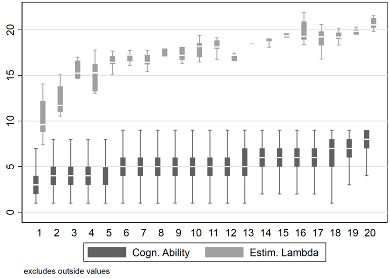
Some implications of the model:

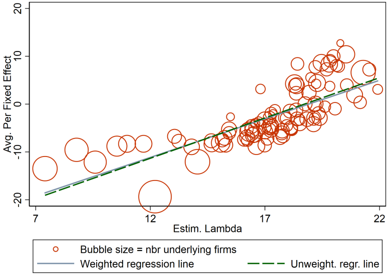
- Within-firm wage skill premia depend on relative skill differences between workers.
- Ceteris paribus, more productive firms pay higher wages.
- Given direct proxies of employees' cognitive skills, one can identify:
  1. firm-specific skill gradients;
  2. firm fixed effects (unconditional pay differentials).
- Workers with better skills sort proportionally more into firms with higher returns to skills. Hence, the average skill of workers in a firm is an increasing function of that firm return to skills.
- Model-based wage equation posits heterogeneous skill premia across firms:

$$\ln(w_{i,j,t}) = \alpha_s + \theta_{j,t} + \mu_i + \lambda_g \ln(s_i) + X_{it} + \varepsilon_{i,j,t}. \quad (1)$$

- Wage specification accommodates both firm and worker unobserved heterogeneity, like AKM model.

Firm-Specific Returns to Cog. Skills: Estimates





Left plot shows boxplots of distribution of workers' cognitive ability for different firm bin vingtiles (i.e., the variation we use for identification) and estimated skill return λ_gs in each vingtile. Right plot are estimated λ_gs against average person fixed effect in each of 100 firm bins.

Composition Effects and Average Return to Skills

	(1)	(2)	(3)	(4)	(5)	(6)
	Low Premium		Mid Premium		High Premium	
	S=0	S=1	S=0	S=1	S=0	S=1
			BASELINE LEVELS			
Initial period (1990-94)	**0.090**	**0.062**	**0.802**	**0.767**	**0.099**	**0.163**
			CHANGES FROM BASELINE LEVELS			
1995-99	0.084	0.049	0.786	0.685	0.125	0.262
2000-04	0.081	0.044	0.792	0.654	0.122	0.299
2005-09	0.081	0.043	0.792	0.664	0.122	0.290
2010-14	0.085	0.045	0.789	0.693	0.114	0.256

Table: Sorting into low, middle, and high (time-varying) premium firms. This table reports the baseline and subsequent changes in the share of workers, by skill type, work in high-premium (one $sd(\theta_{jt})$ above mean in respective year), low-premium (one $sd(\theta_{jt})$ below), and middle-premium (all other) firms where $sd(\theta_{jt})$ is estimated for each year separately.

Composition Effects and Average Return to Skills

	Gross Return	Skill Returns	Firm Sorting	Implied Drift
	(1)	(2)	(3)	(4)
		BASELINE LEVELS		
Initial period (1990-94)	**21.0**	**18.4**	**-0.7**	–
		CHANGES RELATIVE TO BASELINE		
difference in 1995-99	1.5	0.1	5.2	-3.8
difference in 2000-04	3.9	0.3	9.3	-5.7
difference in 2005-09	0.3	0.3	9.4	-9.5
difference in 2010-14	-2.9	0.4	11.2	-14.6

Table: Decomposition of Skill Returns, by Five-Year Periods. Returns to high-skill dummy in log points, i.e., column (1) is the result from regression of 100*Ln(Wage). *Skill Returns* is the contribution of $\lambda_g \ln(s_i)$ (i.e., average λ_g that workers face). *Firm Sorting* the contribution of sorting into high-premium firms $\frac{cov_t(\theta_{jt}, \ln(s_i))}{Var(\ln(s_i))}$. The last column reports the drift in underlying λ_t, that is, the difference from the sum of the other components.

▸ Log Wages: Variance Decomposition

Summary and Conclusions

- We estimate that:
  - the return to cognitive skills varies considerably across firms.
  - more able workers sort into firms with higher returns.
  - firm-specific wage premia exist after controlling for worker skills.
- We use our estimates to show that:
  - Shifts in composition of aggregate demand, favoring firms with high returns to cognitive skills, account for run-up in cognitive skills premia in 1990s.
  - Reverse pattern takes hold in 2000s, inducing economy-wide decline in return to skills.
  - Firm-level growth spurts shape evolution of economy-wide skill premium.
- Average skill premium, net of between-firm composition effects:
  - Average return to skills, net of employment composition effects, declined continuously between 1990 and 2014, and at increasingly fast pace.
  - In the first part of the sample period, this decline has been offset and masked by employment shifts towards high return firms

Decomposing the Variance of Log Wages.

	Log(W)	Pred.Log(W)	Skill Return	Firm Premia	Sorting	Xs
	(1)	(2)	(3)	(4)	(5)	(6)
			BASELINE LEVELS			
1990-94	**100**	**70.7**	**2.8**	**25.0**	**-0.2**	**3.4**
			CHANGES FROM BASELINE (1990-94)			
Δ 1995-99	13.01	5.67	0.00	0.38	1.71	2.34
Δ 2000-04	21.02	9.38	0.04	-0.83	3.09	0.88
Δ 2005-09	16.39	7.17	-0.04	-1.96	3.09	-1.33
Δ 2010-14	8.76	3.79	-0.04	0.00	3.59	-2.04

Table: Decomposition of the variance of log wages. Baseline levels (in 1990-1994) and changes relative to baseline in different five-year periods (Δ). Magnitudes expressed as share of Log(W) in 1990-94. *Ln(Wage)* is $\ln(w_{i,j,t})$, *Pred.Ln(Wage)* the model predicted wage $\ln(\hat{w}_{i,j,t})$, *Skill Return* the wage variation do to skill returns $\lambda_g \ln(s_i)$, *Firm Premia* the firm wage premium $\theta_{j,t}$, and *Sorting* the wage variation due to interaction between skills and firm wage premium $2 \times Cov(\lambda_g \ln(s_i), \theta_{j,t})$. Circa 1.9 million observations per five-year period.

▸ Return

Data Summary

Administrative records: all individuals age 16+, domicile in Sweden, 1990-2014:

1. Employment: firm & plant id, industry; occupation in 1990 and from 2001.
2. Demographics: age, education, gender.
3. Earnings: individual's total annual labor income, including end-of-year bonuses.
4. Cognitive ability: military enlistment tests at age 18–19. Available for most males, birth years 1951–1985.

Sample selection:

- Annual labor income above (*Prisbasbelopp*) that qualifies to earnings related part of public pension system (36,400 SEK in 1998).
- 25–54 year old prime age dependently employed in the private and non-primary sector.
- Ca 59 million individual-year observations.

Return to Intro Graphs

Labor Market and Income Measures

Dataset drawn from administrative records:

- All individuals age 16+ domiciled in Sweden, 1990–2014.
- Employment (firm&plant id, industry; occupation 1990 and after 2001, impute 1991–2000), demogr (age, educ, etc).
- Earnings are individual's total annual labor income including end-of-year bonus payments.
- Alternatives: total taxable income (including capital gains), labor income from largest source (in case multiple employers), hours from Swedish Labor Force Survey.

Check hourly wages from wage structure survey of employers (all public employees and sample of private sector employment). In paper, use evidence from Edin et al 2017.

▸ Return to Intro Graphs

Sample: Descriptives.

Both Sexes, Age 25–54	obs (tsd)	mean	sd	p5	p25	p50	p75	p95
Age	58,900	39.6	8.5	26	32	40	47	53
Female	58,900	0.47	0.50	0	0	0	1	1
Earnings (Tsd 2014SEK)	58,900	287.9	209.1	80.9	188.3	261.1	342.2	571.4
Cognitive Ability	21,500	5.2	1.9	2	4	5	7	8
Non-cog Ability	20,800	5.1	1.7	2	4	5	6	8
High-School GPA	31,500	50.6	28.6	5.7	26.2	50.7	75.3	95.1
Predict. Cogn. Ability	31,500	50.7	28.7	5.7	26.1	50.8	75.6	95.1
Years of School	58,600	12.1	2.7	9.0	10.5	12.0	13.5	16.0
Post-Second. Degree	58,600	0.37	0.48	0.00	0.00	0.00	1.00	1.00
University Degree	58,600	0.22	0.41	0.00	0.00	0.00	0.00	1.00
PhD Degree	58,600	0.01	0.10	0.00	0.00	0.00	0.00	0.00

Cognitive and non-cognitive ability

- From military enlistment at age 18–19. Available for most of male birth years 1951–1985.
- Cognitive good measure of general intelligence (more fluid than crystallized AFQT). Focus in analysis!
- Non-cognitive score based on 25-minute semi-structured interview by psychologist.
  - elicit, among others, willingness to assume responsibility, independence, outgoing character, persistence, emotional stability, and power of initiative
  - assesses ability to function in very demanding environment (combat) but also in a group (military squad)
- Stanine scale (integers 1–9, mean/median 5, stdev 2)

Skill measures

High-school grades

- Percentile in distribution within graduation year. For both genders from birth years 1955.
  - combination of cognitive achievement and personality traits such as conscientiousness
- Additionally compute **predicted cognitive ability** percentile
  - regress males' cognitives on grades x track x age at graduation for each graduation year. R-square $\approx$ .35, use predicted value.
- Mean/median 50, stdev 28.5.

Finally: detailed **level and subject of education**. Coding system changed in 2000, have to adjust for. Comparisons also difficult because of deteriorating selection.

Production: Firms Heterogeneity.

- A firm j produces according to a linear production function using different skill labor input

$$y_j = T_j \left[s_L^{\lambda_j} q_{js_L} + s_H^{\lambda_j} q_{js_H} \right].$$

- λ_j is firm-specific skill-bias. Assume two skill types for simplicity $h_H > h_L$. T_j subsumes firm productivity.
- Firm output is used as an intermediate input in the production of a final good:

$$Y = \left[\sum_{j=1}^{J} \phi_j y_j^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}.$$

- Parameter ϕ_j captures demand for firm j's output: relative 'value' of firm's intermediate output for aggregate production.

✓ Denoting the market price for firm j's output as p_j, the demand function for that firm's output is

$$p_j = \frac{\partial Y}{\partial y_j} = \phi_j Y^{\frac{1}{\sigma}} y_j^{\frac{-1}{\sigma}}.$$

Return to Main Slides

Workers: skills and non-pecuniary preferences

- For worker i with skill $s \in \{s_H, s_L\}$ we write the indirect utility of working at firm j as
$$u_{ijs} = \tilde{\beta}_s \ln(w_{js}) + \ln(\tilde{a}_{js}) + \tau_s \varepsilon_{ijs}.$$

- $\tilde{a}_{js}$ is a firm-specific amenity common to all workers in group s and ε_{ijs} captures idiosyncratic preferences for working at firm j with scaling parameter τ_s.

- **Firms cannot observe** ε_{ijh}**.** Firms post skill-specific wage w_{jh}.

– **Assumption.** Distribution of ε_{ijh} is Type 1 Extreme Value (nice interpretation: highest draw among all possible non-pecuniary payoffs).

⇒ Logit conditional probability of sorting with firm j. If number of firms (J) large enough (50k/yr in data), exponential probabilities.

□ **Result.** Upward-sloping firm-specific labor supply (linear in logs):

$$\ln(q_{js}) = \ln(N_s \xi_s) + \beta_s \ln(w_{js}) + \ln(a_{js}),$$

The Problem of the Firm and the Demand for Skills

- The firm's optimal wage choices solve the cost-minimization problem:

$$\min_{w_{js_H}, w_{js_L}} \quad q_{js_L} w_{js_L} + q_{js_H} w_{js_H} \tag{2}$$

$$s.t. \quad y_j = T_j \left[s_L^{\lambda_j} q_{js_L} + s_H^{\lambda_j} q_{js_H} \right]$$

$$\ln(q_{js}) = \ln(N_s \xi_s) + \beta \ln(w_{js}) + \ln(a_{js})$$

$\Rightarrow$ The first order condition for $w_{j,s}$ is

$$q_{j,s} + w_{j,s} \frac{\partial q_{j,s}}{\partial w_{j,s}} = \chi_j \frac{\partial y_j}{\partial q_{j,s}} \frac{\partial q_{j,s}}{\partial w_{j,s}}$$

- **Firm's choice.** χ_j is the shadow value of each marginal unit produced by firm j. The firm sets wages that equalize the marginal cost to its marginal value. χ_j corresponds, at the optimal choice, to the marginal revenue, which in turn depends on the relative demand for firm j's output.

- A firm's optimal behavior implies:

$$w_{js} = \underbrace{\frac{\beta}{1+\beta}}_{\text{Monops.Markdown}} \times \underbrace{\frac{\sigma-1}{\sigma} \phi_j T_j \left(\frac{Y}{y_j} \right)^{\frac{1}{\sigma}}}_{\text{Marg.Revenue}} \times \underbrace{s^{\lambda_j}}_{\text{Skill Productivity}}$$

Some Implications of the Model

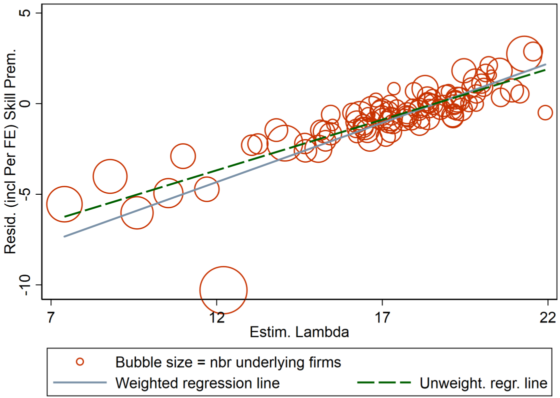
- **Proposition 2: Within-Firm Wage Premia.** Ceteris paribus, firms with higher λ_j exhibit higher skill premia. That is, $\ln(w_{js_H}) - \ln(w_{js_L}) = \lambda_j \left[\ln(s_H) - \ln(s_L)\right]$.
  - Corollary (Identification of Firm Skill-Bias): If direct measures of skills are available one can identify firm-specific skill-bias λ_j (up to a normalization) from the cross-section of wages, by projecting log wage differences on the difference of log skill measures.

- **Proposition 3: Between-Firm Wage Premia.** There exist between-firm differences in average wage premia. For given worker skill s and firm skill-bias λ_j), firms with higher output demand ϕ_j (and/or higher productivity T_j) pay higher wages.
  - Corollary (Identification of Firm Wage Premia): Firm-specific wage premia θ_j can be identified (up to a normalization) from variation in log wages across firms, conditional on employees' skill measures.

- **Proposition 4: Firm Size.** The output of a firm y_j and its total employment increase—ceteris paribus (i.e. for the same λ_j and a_{js})— in the firm's aggregate demand share ϕ_j (and/or productivity T_j).

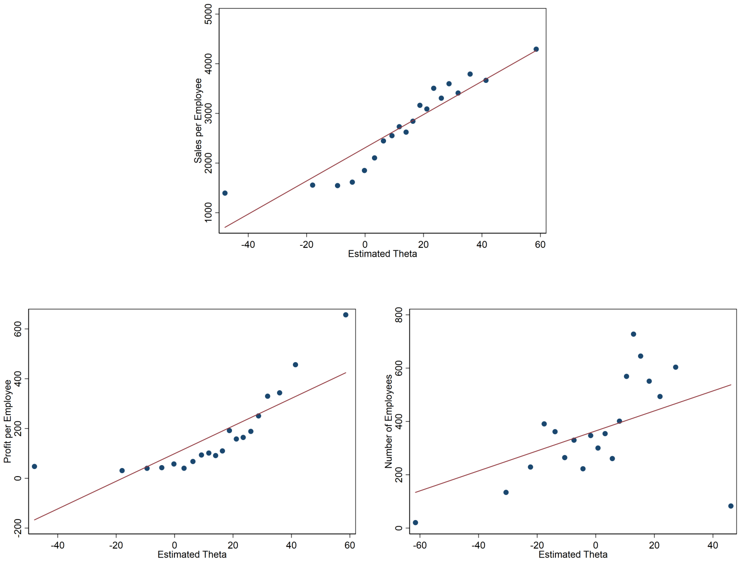
- **Proposition 5: Within-Firm Skill Composition.** Workers with higher skills sort proportionally more into firms with higher λ. Hence, the average skill of workers within a firm increases with that firm's λ.

Skill Premia across Firms: Parametric Estimates



Plot of the (residualized) skill premia for firms with different λ. Including person fixed effect.

Sales, Productivity and Headcounts across Firms



Bin scatter plots of the estimated θ_{jt} against different firm characteristics.
Sales/employee (Panel (a)), profits/employee (b), and number of employees (c)