

Discussion of "Firm Heterogeneity in Skill Returns" by Böhm, Esmkhani, and Gallipoli

Serdar Ozkan
University of Toronto

Vienna Macroworkshop
October 11, 2019

Chapter #1: Estimating Skill Bias Heterogeneity

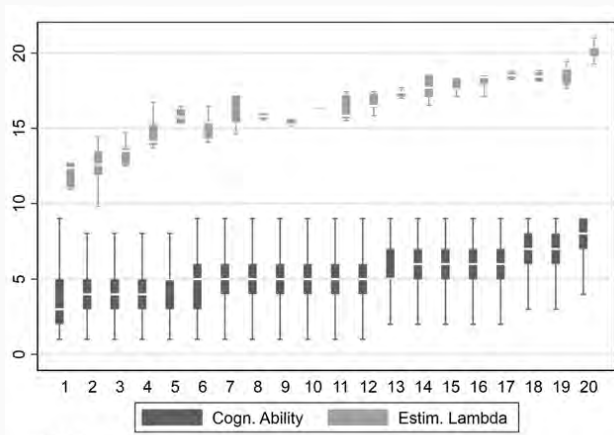
- **Research Question:** Do firms differ in the skill bias of their production technology and thereby in skill premium?
- **Methodology:** Estimate an Abowd *et al.* (1999) (AKM) type wage equation with skill return heterogeneity:

$$\ln(w_{i,j}) = \theta_j + \mu_i + \lambda_j \ln(s_i) + X_{i,t}b_t + e_{i,j,t}$$

- Motivated by a model of labor demand by heterogeneous firms.
- Use military enlistment cognitive tests scores for $s \in \{1, 2, \dots, 9\}$.
- **Identification:** λ_j is identified from within firm skill premium: $\frac{w_{js_H}}{w_{js_L}} = e^{\lambda_j} \left(\frac{s_H}{s_L} \right)$

Chapter #1: Estimating Skill Bias Heterogeneity

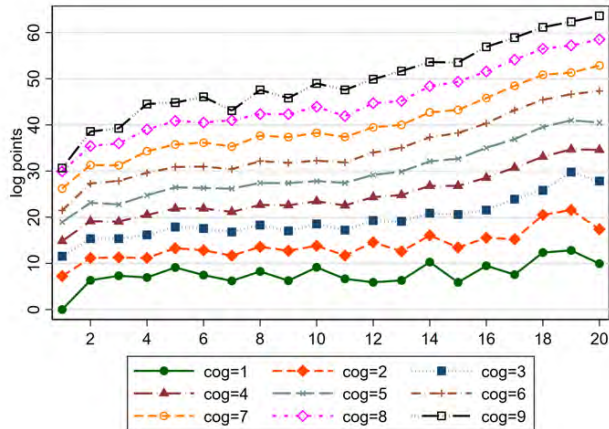
Benchmark findings. Estimated Skill Returns



- Larger skill returns in firms with more skilled labor: increase from below 12 log points to 20 log points.
- This result is not mechanical since λ identified from within firm skill premium.
- Robust to a battery of tests.

Chapter #1: Estimating Skill Bias Heterogeneity

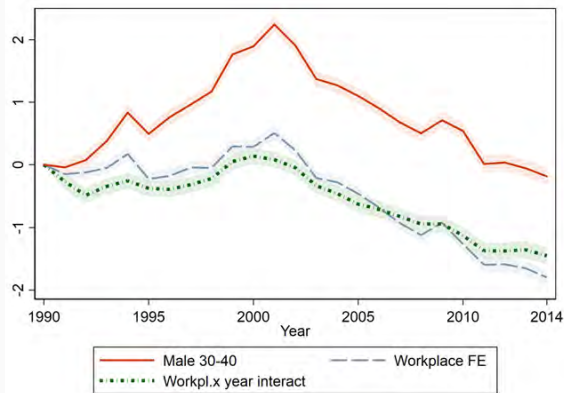
Benchmark findings. Skill Premia across Firms



- Shows large and systematic heterogeneity in skill returns.
- Average wages increase by average skill of firms.
- Within firm skill premium increase by average skill of firms.

Chapter #2: Role of Firm Heterogeneity Wage Inequality

Research Question: How much firm heterogeneity (in θ and λ) matter for inequality and its evolution over time?



- Shows the skill premium in Sweden: Increase until early 2000s and then taper off.
- When control for firm fixed effect (with or without year interacts) the increase goes away.

Chapter #2: Role of Firm Heterogeneity Wage Inequality

Methodology: Develop a quantitative model of labor demand.

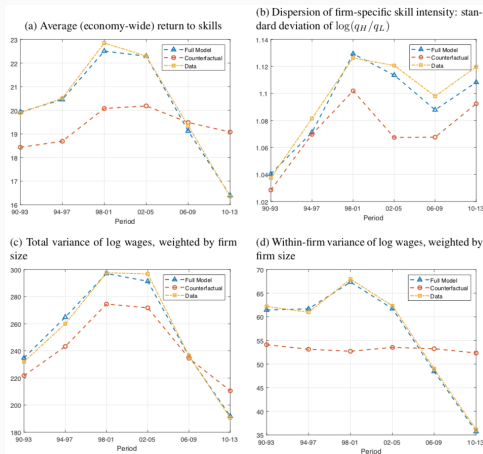
- Firms differ in:

- skill-bias** in production, λ_j : $y_j(q_{jSL}, q_{jSH}) = T_j [q_{jSL} + s_H^{\lambda_j} q_{jSH}]$, $s_H > 1$.
- monopoly** power in output market: $Y = \left[\sum_1^J \phi_j y_j^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$
- monopsony** power in labor market $\Rightarrow \ln(q_{js}) = \ln(N_s \xi_s) + \beta \ln(w_{js}) + \ln(a_{js})$.

- Equilibrium wages:** $\ln(w_{js}) = \underbrace{\ln\left(\frac{\beta}{1+\beta}\right)}_{\text{Mon. Markdown}} + \underbrace{\ln\left(\frac{\sigma-1}{\sigma} \phi_j T_j \left(\frac{Y}{y_j}\right)^{1/\sigma}\right)}_{\text{Marg. revenue}} + \underbrace{\lambda_j \ln(s)}_{\text{skill return}}$
- Model consistent with several features of the data, e.g., firm size, sorting.

Chapter #2: Role of Firm Heterogeneity Wage Inequality

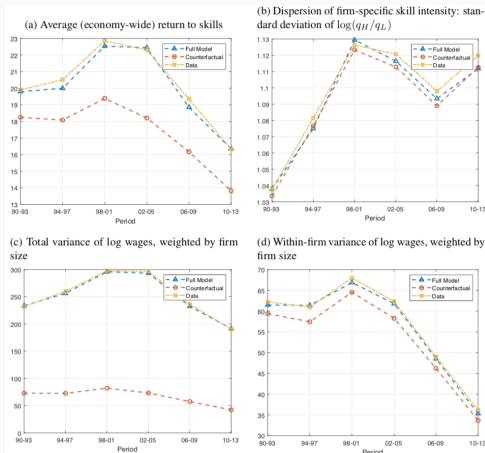
Counterfactual #1. Shutdown heterogeneity in λ



- λ heterogeneity explains ~10% of skill premium.
- Interestingly explains very little of variation in skill intensity across firms.
- Also very little role in overall wage dispersion (though I don't know what the y-axis scale means!).

Chapter #2: Role of Firm Heterogeneity Wage Inequality

Counterfactual #2. Shutdown heterogeneity in ϕ

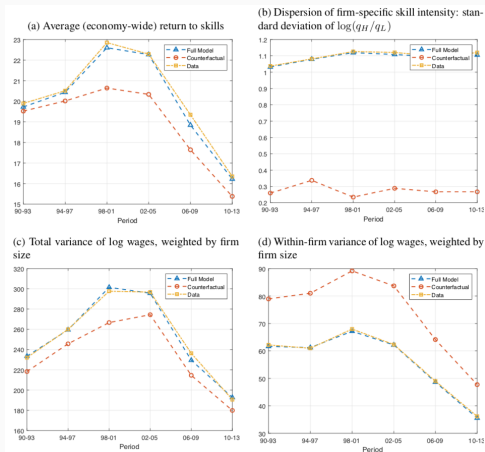


- ϕ heterogeneity explains another ~10% of skill premium because of positive correlation between λ .
- No effect on variance of $\log(q_H/q_L) \Rightarrow$ no sorting due to ϕ .
- Large effect on overall wage dispersion.

$$\ln(q_{js}) = \ln(N_s \xi_s) + \beta[\tilde{\beta} + \tilde{\phi}_j + \lambda_j \ln(s)] + \ln(a_{js}).$$

Chapter #2: Role of Firm Heterogeneity Wage Inequality

Counterfactual #3. Shutdown heterogeneity in firm amenities a



- a heterogeneity explains most of the variation in skill intensity across firms. Sorting due to a ?
- Significant effect on skill premium and within firm dispersion.

- Fantastic paper! Enjoyed reading it a lot.
- By investigating a new aspect of firm heterogeneity, this paper contributes to the literatures on
 - firm vs worker heterogeneity in wage dispersion (going back to AKM),
 - the role of firm skill distribution in the production function: “Superstar” effect.
- Very careful work, a battery of robustness checks.
- Long paper, lot's of things going on!

Remark #1: Sorting in labor markets (or lack of)

- There is sorting everywhere in the economy: Marriage, neighborhoods, schools, etc.

Remark #1: Sorting in labor markets (or lack of)

- There is sorting everywhere in the economy: Marriage, neighborhoods, schools, etc.
- Except in labor markets (at least in the AKM world).

$$\ln(w_{i,j}) = \theta_j + \mu_i + X_{i,t}b_t + e_{i,j,t}$$

- Estimates of $\text{corr}(\theta_j, \mu_i)$ are close to zero and sometimes negative.

Remark #1: Sorting in labor markets (or lack of)

- There is sorting everywhere in the economy: Marriage, neighborhoods, schools, etc.
- Except in labor markets (at least in the AKM world).

$$\ln(w_{i,j}) = \theta_j + \mu_i + X_{i,t}b_t + e_{i,j,t}$$

- Estimates of $\text{corr}(\theta_j, \mu_i)$ are close to zero and sometimes negative.
- People have blamed incidental parameter bias i.e., “limited mobility problem” (e.g., Andrews *et al.* (2008); Bonhomme *et al.* (2019); Borovičková and Shimer (2017)).

Remark #1: Sorting in labor markets (or lack of)

- There is sorting everywhere in the economy: Marriage, neighborhoods, schools, etc.
- Except in labor markets (at least in the AKM world).

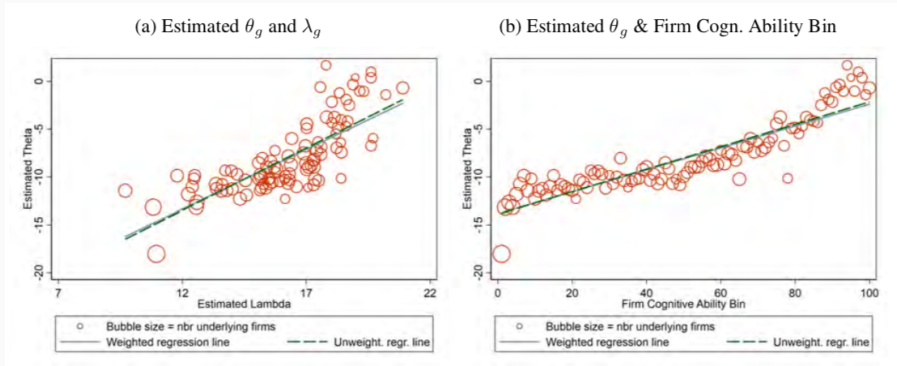
$$\ln(w_{i,j}) = \theta_j + \mu_i + X_{i,t}b_t + e_{i,j,t}$$

- Estimates of $\text{corr}(\theta_j, \mu_i)$ are close to zero and sometimes negative.
- People have blamed incidental parameter bias i.e., “limited mobility problem” (e.g., Andrews *et al.* (2008); Bonhomme *et al.* (2019); Borovičková and Shimer (2017)).
- Or have come up with theoretical models which does not generate log linear wage functions (e.g., Eeckhout and Kircher (2011); Shimer (2005)).

Remark #1: Can this paper help finding a stronger sorting?

- The identification of θ_j in AKM comes from workers' mobility between firms.
- **Simple Example:** Suppose two workers with $s_L = 0$, $s_H = 1$, and $\mu_H > \mu_L$ move from firm $j = 1$ (with normalization $\theta_1 = 0$ and $\lambda_1 = 0$) to $j = 2$ (with θ_2 and λ_2), then:
 - $\Delta w_{S_L} = (\theta_2 - \theta_1) = \theta_2$ and $\Delta w_{S_H} = (\theta_2 - \theta_1) + (\lambda_2 - \lambda_1) = \theta_2 + \lambda_2$
 - If we ignore λ s as in the original AKM then the biased firm fixed effect is $\tilde{\theta}_2 = \theta_2 + \lambda_2/2$ for both workers.
 - If not misspecified, then the true firm fixed effect for s_L is θ_2 and for s_H is $\theta_2 + \lambda_2$.
- Potential for improving the measurement of sorting:
 - Correlation of (μ_L, θ_2) and $(\mu_H, \theta_2 + \lambda_2)$ vs $(\mu_L, \theta_2 + \lambda_2/2)$ and $(\mu_H, \theta_2 + \lambda_2/2)$.

Remark #1: Can this paper help finding a stronger sorting?



- They say firm-specific skill biases play a key role for worker sorting but they don't actually quantify it.
- What's the $\text{corr}(\theta_j + \lambda_j s_i, \mu_i)$? How is this different than if we assume $\lambda = 0$? Without grouping firms (with potentially "incidental parameter bias")?

Remark #2: Could $\sigma_{\lambda}^2 > 0$ be due to within skill heterogeneity?

- The skill measure is a binary variable: $s = 1$ if Stanine scores 7–9; $s = 0$ o/w.
 - They also do it with $s \in \{1, 2, \dots, 9\}$ but the same idea applies.
- So lot's of within skill group heterogeneity, which may lead to heterogeneity in λs .
- **Example**
 - Suppose computer scientists (CS) and art historians (AH) are $s = 1$ and janitors (J) are $s = 0$.
 - Suppose firm G (software company) only employs CS and J and firm M (museum) only employs AH and J .
 - Then $\lambda_G > \lambda_M$ would simply capture the wage differential between CS and AH .

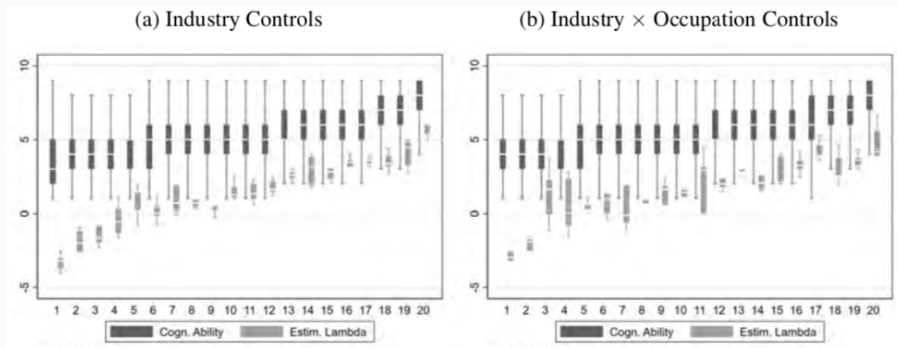
Remark #2: Could $\sigma_\lambda^2 > 0$ be due to within skill heterogeneity?

How do they address this issue? They control for individual fixed effects μ_i to capture unobserved differences b/w workers. But this would not work..

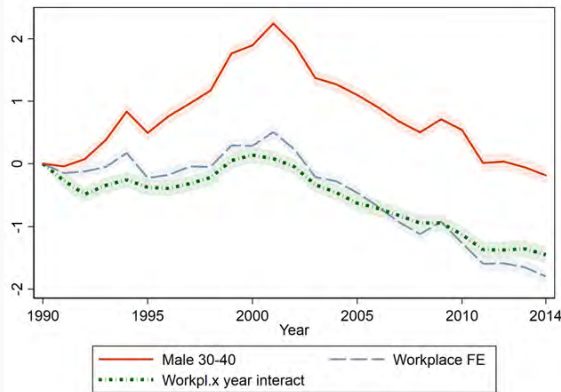
- If *CS* work for *only* software companies and *AH* work for *only* museums, because:
 - μ s are net of average λ of software companies for *CS* and that of museums for *AH*.
- They use data on 30-40 year olds who don't move around much.
- How about using a function of μ_i as the measure of skill and interact with λ ?
$$\ln(w_{i,j}) = \theta_j + \mu_i^{k+1} + f(\mu_i^k, \lambda_j) + X_{i,t}b_t + e_{i,j,t}$$
 - Estimation complicated but may be feasible by employing an iterative algorithm.
 - For given μ_i s estimate the wage equation and then update μ s iteratively until converges, i.e., $\mu^{k+1} \approx \mu^k$.

Remark #2: Could $\sigma_{\lambda}^2 > 0$ be due to within skill heterogeneity?

How do they address this issue? In a richer specification (which I like better) they control for **occupation and industry effects** and **they find smaller λ heterogeneity**.

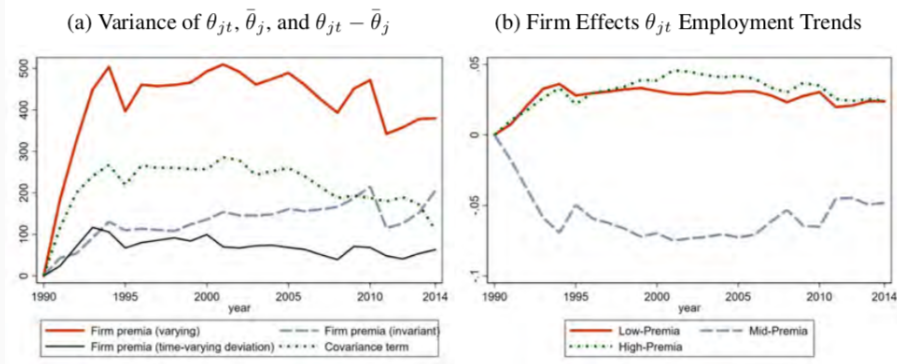


Remark #3: What drives the changes in skill premium?



- When control for **time-invariant** firm fixed effect, the increase in skill premium b/w 1990-2000 goes away.
- So changes in firm effects do not matter much?
- What's happening?

Remark #3: What drives the changes in skill premium?



- A classic example of job polarization: Workers in the middle paying firms move to low and high paying firms.
- If you're working in a high θ firm in the 90s, the SBTC doesn't affect your firm premium.

Remark #3: What drives the changes in skill premium?

- But this seems to be a different story than varying firm characteristics in the quantitative exercise.
- Do the high and low paying firms absorb employees of middle paying firms in the model (without changing AKM firm fixed effects)?
- Cheap shot (sorry)! They compare 6 steady states, each of which is calibrated to match the averages over a 4-year period.
 - The economy probably doesn't transition into a new steady state in 4 years.

- Super interesting, thought provoking paper, important contribution to the literature.
 - Learned a lot!
 - Strongly recommend reading it.
- More emphasis on their sorting result which may be possibly very interesting.
- May be, it is a bit too long; there can be two separate papers.

References i

- ABOWD, J. M., KRAMARZ, F. and MARGOLIS, D. N. (1999). High wage workers and high wage firms. *Econometrica*, **67** (2), 251–333.
- ANDREWS, M. J., GILL, L., SCHANK, T. and UPWARD, R. (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **171** (3), 673–697.
- BONHOMME, S., LAMADON, T. and MANRESA, E. (2019). A distributional framework for matched employer employee data. *Econometrica*, **87** (3), 699–739.
- BOROVÍČKOVÁ, K. and SHIMER, R. (2017). *High wage workers work for high wage firms*. Tech. rep., National Bureau of Economic Research.
- ECKHOUT, J. and KIRCHER, P. (2011). Identifying sorting in theory. *The Review of Economic Studies*, **78** (3), 872–906.

SHIMER, R. (2005). The assignment of workers to jobs in an economy with coordination frictions. *Journal of Political Economy*, **113** (5), 996–1025.