The Micro and Macro Implications of Multidimensional Skill Uncertainty

Piero De Dominicis

Bocconi University

NBER Summer Institute 2026

The Micro and Macro Perspectives of the Aggregate Labor Market

20 July 2026

Motivation

- Workers possess a portfolio of **skills** whose payoffs depend on the occupation.
- Early in their career, workers may have imperfect information about their own skills.
- **Mismatch:** workers sort into occupations based on their perceived comparative advantage.

Motivation

- Workers possess a portfolio of **skills** whose payoffs depend on the occupation.
- Early in their career, workers may have imperfect information about their own skills.
- **Mismatch:** workers sort into occupations based on their perceived comparative advantage.
- Occupational choice is **path dependent**, where you work determines:
  - what you *learn* about your comparative advantage;
  - what type of human capital you *accumulate*.

Motivation

- Workers possess a portfolio of **skills** whose payoffs depend on the occupation.
- Early in their career, workers may have imperfect information about their own skills.
- **Mismatch:** workers sort into occupations based on their perceived comparative advantage.
- Occupational choice is **path dependent**, where you work determines:
  - what you *learn* about your comparative advantage;
  - what type of human capital you *accumulate*.

How does skill uncertainty shape occupational mobility?
How costly is it for output and wages?

This Paper

1. **Dynamic Roy model with multidimensional skill learning.**
   - **Occupations:** bundles of tasks (Gathmann & Schönberg, 2010)
     - each task can be performed using skills and experience, heterogeneous production technologies
   - **Workers:** supply skills and experience
     - *learn* unobserved permanent multidimensional skills
     - *accumulate* task-specific experience at different rates

This Paper

1. **Dynamic Roy model with multidimensional skill learning.**
   - **Occupations:** bundles of tasks (Gathmann & Schönberg, 2010)
     - each task can be performed using skills and experience, heterogeneous production technologies
   - **Workers:** supply skills and experience
     - *learn* unobserved permanent multidimensional skills
     - *accumulate* task-specific experience at different rates

2. **Estimate the model on Portuguese administrative data.**
   - Estimate parameters by two-step MLE, accounting for **dynamic selection on beliefs**.

This Paper

1. **Dynamic Roy model with multidimensional skill learning.**
   - **Occupations:** bundles of tasks (Gathmann & Schönberg, 2010)
     - each task can be performed using skills and experience, heterogeneous production technologies
   - **Workers:** supply skills and experience
     - *learn* unobserved permanent multidimensional skills
     - *accumulate* task-specific experience at different rates

2. **Estimate the model on Portuguese administrative data.**
   - Estimate parameters by two-step MLE, accounting for **dynamic selection on beliefs**.

3. **Simulate counterfactual scenarios.**
   - Which moves reflect learning about comparative advantage?
   - What is the cost of imperfect information about skills? How can we reduce it?
   - How does job transformation (AI) reshape the cost of skill uncertainty? **(not today)**

Preview of the Results

- **Worker-occupation match quality** predicts who switches and where they move:
  - **Learning about comparative advantage** drives mobility low-match-quality
- **Removing information frictions** about skills raises aggregate output by **5.2%**
- Wage gains from **removing frictions** largest at **low match quality** ($>$**20%** vs **3%**)
- A **calibrated information treatment** at entry raises output by **0.9%**
- Job transformation (AI) makes skill uncertainty **more costly** **(not today)**
  - AI increases output, but gain from removing frictions rises **5.2%** $\rightarrow$ **5.9%**

Contribution to the Literature

- **Sources of Wage Growth and Human Capital Misallocation:** Topel and Ward 1992, Neal 1999, Pavan 2011, Hshieh et al. 2019, Guvenen et al. 2020, Lise and Postel-Vinay 2020, Adda and Dustmann 2023, Kiss et al. 2025, Goller et al. 2026
  $\Rightarrow$ Quantify the micro and macro consequences of multidimensional skill uncertainty

- **Learning and Job Mobility:** Jovanovic 1979, Miller 1984, Gibbons et al. 2005, Antonovics and Golan 2012, James 2012, Papageorgiou 2014, Sanders 2014, Groes et al. 2015, Guvenen et al. 2020, Baley et al. 2022, Arcidiacono et al. 2024, Pastorino 2024
  $\Rightarrow$ Estimation of structural model of learning about comparative advantage in tasks

- **Task-Based Approach to Human Capital:** Gathmann and Schönberg 2010, Acemoglu and Autor 2011, Sanders et Taber 2012, Yamaguchi 2012, Autor and Handel 2013, Lindelaub 2017, Guvenen et al. 2020, Lise and Postel-Vinay 2020, Grigsby 2024, Freund and Mann 2025, Althoff and Reichart 2026
  $\Rightarrow$ Study how job transformation (AI) + skill uncertainty affect occupation reallocation

Outline for Today

1. **Structural Model**

2. **Data, Identification & Estimation**

3. **The Implications of Multidimensional Skill Uncertainty**

Skills, Experience and Tasks

- Discrete time, continuum of workers and firms, occupations $j \in O \equiv \{1, \dots, N_o\}$
- Worker i characterized by **multidimensional skills and experience**:
  1. $\boldsymbol{\theta}_i = (\theta_{i1}, \dots, \theta_{iK})'$ time-invariant task-specific skills
  2. $\boldsymbol{\tau}_{it} = (\tau_{i1t}, \dots, \tau_{iKt})'$ task-specific tenure (experience)
- Skills drawn at entry from a group-specific distribution: $\boldsymbol{\theta}_i(g) \sim N\big(\boldsymbol{\mu}(g), \boldsymbol{\Sigma}(g)\big)$

Skills, Experience and Tasks

- Discrete time, continuum of workers and firms, occupations $j \in O \equiv \{1, \ldots, N_o\}$
- Worker i characterized by **multidimensional skills and experience**:
  1. $\boldsymbol{\theta}_i = (\theta_{i1}, \ldots, \theta_{iK})'$ time-invariant task-specific skills
  2. $\boldsymbol{\tau}_{it} = (\tau_{i1t}, \ldots, \tau_{iKt})'$ task-specific tenure (experience)
- Skills drawn at entry from a group-specific distribution: $\boldsymbol{\theta}_i(g) \sim N(\boldsymbol{\mu}(g), \boldsymbol{\Sigma}(g))$

- Workers have to perform K tasks when working in an occupation
- Each task k in occupation j requires **time**, **skill**, and **experience**:

$$Y_{ijkt} = \underbrace{\ell_{ijkt}}_{\text{time}} \cdot \underbrace{\theta_{ik}^{\lambda_{jk}^0}}_{\text{skill}} \cdot \underbrace{H_{jk}(\tau_{ikt})}_{\text{experience}}$$

λ_{jk}^0: skill elasticity; $H_{jk}(\tau_{ikt})$ concave in τ_{ikt}

From Tasks to Output

Task importance ω_{jk} determines each task k's contribution in occupation j, with $\sum_k \omega_{jk} = 1$

Output aggregates via **Cobb-Douglas**:

$$Y_{ijt} = \prod_{k=1}^{K} \left(\frac{Y_{ijkt}}{\omega_{jk}} \right)^{\omega_{jk}} \exp(\varepsilon_{ijt})$$

where $\varepsilon_{ijt} \sim N(0, \varsigma_j^2)$

From Tasks to Output

Task importance ω_{jk} determines each task k's contribution in occupation j, with $\sum_k \omega_{jk} = 1$

Output aggregates via **Cobb-Douglas**:

$$Y_{ijt} = \prod_{k=1}^{K} \left(\frac{Y_{ijkt}}{\omega_{jk}} \right)^{\omega_{jk}} \exp(\varepsilon_{ijt})$$

where $\varepsilon_{ijt} \sim N(0, \varsigma_j^2)$

Optimal time allocation

$$\ell^*_{ijkt} = \omega_{jk}$$

independent of worker type

Time-allocation problem

Substituting $\ell^*_{jk} = \omega_{jk}$, output given by:

$$Y_{ijt} = \underbrace{\prod_{k \in \mathcal{K}} \theta_{ik}^{\lambda_{jk}}}_{\text{unobserved skills bundle}} \cdot \underbrace{H_j(\boldsymbol{\tau}_{it})}_{\text{human-capital bundle}} \cdot \underbrace{\exp(\varepsilon_{ijt})}_{\text{idiosyncratic shock}}$$

$\lambda_{jk} = \omega_{jk} \lambda^0_{jk}$; human-capital bundle $H_j(\boldsymbol{\tau}_{it}) = \prod_{k \in \mathcal{K}} H_{jk}(\tau_{ikt})^{\omega_{jk}}$, log-concave in tenure

Production and Wage Equation

- **Perfect competition + piece-rate contracts** $\Longrightarrow W_{ijt} = P_{jt} Y_{ijt}$

Log-wage equation, $w_{ijt} = \log W_{ijt}$:

$$w_{ijt} = \underbrace{p_{jt}}_{\text{occ. price}} + \underbrace{\lambda_j' \xi_i}_{\substack{\text{multidimensional} \\ \text{skill match}}} + \underbrace{h_j(\tau_{it})}_{\text{task experience}} + \underbrace{\varepsilon_{ijt}}_{\text{id. shock}}$$

1. $p_{jt} = \log P_{jt}$: occupation price, determined in general equilibrium
2. $\xi_i = \log \theta_i$: latent log-skills, λ_j: occ.-specific skill loadings
3. $h_j(\tau_{it}) = \log H_j(\tau_{it})$: concave returns to task tenure
4. $\varepsilon_{ijt} \sim N(0, \varsigma_j^2)$: occ.-specific idiosyncratic productivity shock

Details: Production Function Details: Labor Demand

Production and Wage Equation

- **Perfect competition + piece-rate contracts** $\implies W_{ijt} = P_{jt} Y_{ijt}$

Log-wage equation, $w_{ijt} = \log W_{ijt}$:

$$w_{ijt} = \underbrace{p_{jt}}_{\text{occ. price}} + \underbrace{\lambda_j' \xi_i}_{\text{multidimensional skill match}} + \underbrace{h_j(\tau_{it})}_{\text{task experience}} + \underbrace{\varepsilon_{ijt}}_{\text{id. shock}}$$

1. $p_{jt} = \log P_{jt}$: occupation price, determined in general equilibrium
2. $\xi_i = \log \theta_i$: latent log-skills, λ_j: occ.-specific skill loadings
3. $h_j(\tau_{it}) = \log H_j(\tau_{it})$: concave returns to task tenure
4. $\varepsilon_{ijt} \sim N(0, \varsigma_j^2)$: occ.-specific idiosyncratic productivity shock

Details: Production Function Details: Labor Demand

Production and Wage Equation

- **Perfect competition** + **piece-rate contracts** $\implies W_{ijt} = P_{jt} Y_{ijt}$

Log-wage equation, $w_{ijt} = \log W_{ijt}$:

$$w_{ijt} = \underbrace{p_{jt}}_{\text{occ. price}} + \underbrace{\lambda_j' \xi_i}_{\text{multidimensional skill match}} + \underbrace{h_j(\tau_{it})}_{\text{task experience}} + \underbrace{\varepsilon_{ijt}}_{\text{id. shock}}$$

1. $p_{jt} = \log P_{jt}$: occupation price, determined in general equilibrium
2. $\xi_i = \log \theta_i$: latent log-skills, λ_j: occ.-specific skill loadings
3. $h_j(\tau_{it}) = \log H_j(\tau_{it})$: concave returns to task tenure
4. $\varepsilon_{ijt} \sim N(0, \varsigma_j^2)$: occ.-specific idiosyncratic productivity shock

Details: Production Function Details: Labor Demand

Production and Wage Equation

- **Perfect competition + piece-rate contracts** $\implies W_{ijt} = P_{jt} Y_{ijt}$

Log-wage equation, $w_{ijt} = \log W_{ijt}$:

$$w_{ijt} = \underbrace{p_{jt}}_{\text{occ. price}} + \underbrace{\lambda_j' \xi_i}_{\substack{\text{multidimensional} \\ \text{skill match}}} + \underbrace{h_j(\tau_{it})}_{\text{task experience}} + \underbrace{\varepsilon_{ijt}}_{\text{id. shock}}$$

1. $p_{jt} = \log P_{jt}$: occupation price, determined in general equilibrium
2. $\xi_i = \log \theta_i$: latent log-skills, λ_j: occ.-specific skill loadings
3. $h_j(\tau_{it}) = \log H_j(\tau_{it})$: concave returns to task tenure
4. $\varepsilon_{ijt} \sim N(0, \varsigma_j^2)$: occ.-specific idiosyncratic productivity shock

Details: Production Function Details: Labor Demand

Production and Wage Equation

- **Perfect competition** + **piece-rate contracts** $\implies W_{ijt} = P_{jt} Y_{ijt}$

Log-wage equation, $w_{ijt} = \log W_{ijt}$:

$$w_{ijt} = \underbrace{p_{jt}}_{\text{occ. price}} + \underbrace{\lambda_j' \xi_i}_{\substack{\text{multidimensional} \\ \text{skill match}}} + \underbrace{h_j(\tau_{it})}_{\text{task experience}} + \underbrace{\varepsilon_{ijt}}_{\text{id. shock}}$$

1. $p_{jt} = \log P_{jt}$: occupation price, determined in general equilibrium
2. $\xi_i = \log \theta_i$: latent log-skills, λ_j: occ.-specific skill loadings
3. $h_j(\tau_{it}) = \log H_j(\tau_{it})$: concave returns to task tenure
4. $\varepsilon_{ijt} \sim N(0, \varsigma_j^2)$: occ.-specific idiosyncratic productivity shock

Details: Production Function Details: Labor Demand

Information Structure and Beliefs Updating

- Workers have **imperfect information** about ξ_i; beliefs at time t: $\tilde{\mathcal{F}}_{it}$
- **Assumption**: workers are *rational*: $\tilde{\mathcal{F}}_0(g) = N(\boldsymbol{\mu}(g), \boldsymbol{\Sigma}(g))$
- **Signal:** After output realization, workers observe:

$$\hat{s}_{jt} = w_{jt} - p_{jt} - h_j(\tau_t) = \boldsymbol{\lambda}_j'\boldsymbol{\xi} + \varepsilon_{jt}$$

Information Structure and Beliefs Updating

- Workers have **imperfect information** about ξ_i; beliefs at time t: $\tilde{\mathcal{F}}_{it}$
- **Assumption**: workers are *rational*: $\tilde{\mathcal{F}}_0(g) = N(\boldsymbol{\mu}(g), \Sigma(g))$
- **Signal:** After output realization, workers observe:

$$\hat{s}_{jt} = w_{jt} - p_{jt} - h_j(\boldsymbol{\tau}_t) = \boldsymbol{\lambda}_j'\boldsymbol{\xi} + \varepsilon_{jt}$$

- **Beliefs updating:** Workers are Bayesian updaters, $\tilde{\mathcal{F}}_{t+1} = N\left(\tilde{\xi}_{t+1}, \tilde{\Sigma}_{t+1}\right)$:

$$\begin{aligned}\tilde{\xi}_{t+1} &= \tilde{\xi}_t + \boldsymbol{K}_{jt} \cdot \left(\hat{s}_{jt} - \tilde{\xi}_t'\boldsymbol{\lambda}_j\right)\\ \tilde{\Sigma}_{t+1} &= \tilde{\Sigma}_t - \boldsymbol{K}_{jt}\boldsymbol{\lambda}_j'\tilde{\Sigma}_t\end{aligned}$$

where $\boldsymbol{K}_{jt} = \tilde{\Sigma}_t\boldsymbol{\lambda}_j\left(\boldsymbol{\lambda}_j'\tilde{\Sigma}_t\boldsymbol{\lambda}_j + \varsigma_j^2\right)^{-1}$ is the *Kalman gain*

Dynamic Occupational Choice – Bellman Equation

- Each period worker chooses $j \in O$ to maximize lifetime utility:

$$V_g(\tilde{\xi}_t, \tilde{\Sigma}_t, \boldsymbol{\tau}_t, d_{t-1}, \boldsymbol{p}_t) = \max_{j \in O} \Big\{ \underbrace{p_{jt} + \tilde{\xi}_t' \boldsymbol{\lambda}_j + h_j(\boldsymbol{\tau}_t)}_{\text{expected wage}} + \underbrace{\eta_{jg}}_{\text{amenity}} - \underbrace{\kappa_0 \, \mathbf{1}\{j \neq d_{t-1}\}}_{\text{switching cost}} + \underbrace{\varrho \, \epsilon_{jt}}_{\text{taste shock}}$$

$$+ \beta \, \underbrace{\mathbb{E}\big[V_g(\tilde{\xi}_{t+1}, \tilde{\Sigma}_{t+1}, \boldsymbol{\tau}_{t+1}, j, \boldsymbol{p}_{t+1}) \big]}_{\text{continuation value}} \Big\}$$

$$\begin{aligned} \text{s.t.} \quad & \tilde{\xi}_{t+1} = \tilde{\xi}_t + \boldsymbol{K}_{jt}(\hat{s}_{jt} - \tilde{\xi}_t' \boldsymbol{\lambda}_j) \\ & \tilde{\Sigma}_{t+1} = \tilde{\Sigma}_t - \boldsymbol{K}_{jt} \boldsymbol{\lambda}_j' \tilde{\Sigma}_t \\ & \boldsymbol{\tau}_{t+1} = \boldsymbol{\tau}_t + \boldsymbol{\omega}_j \\ \text{given} \quad & (\tilde{\xi}_0, \tilde{\Sigma}_0, \boldsymbol{\tau}_0, d_{-1}) = (\boldsymbol{\mu}_g, \Sigma_g, \mathbf{0}, \varnothing) \end{aligned}$$

- Workers take **occupation prices** $\{p_j\}$ as given (perfect foresight)
- Use Value Function Approximation with Smolyak Method to compute the fixed point of $\widehat{EV}$

Model Dynamics · Blackwell sufficiency conditions · Solution Method

Closing the Model: General Equilibrium

- **Final goods producer** aggregates occupation outputs via CES:

$$Y_t = \left(\sum_{j \in \mathcal{O}} \alpha_j Y_{jt}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \qquad \alpha_j > 0$$

- **Occupation prices** from inverse demand, with the aggregate CES price index as numeraire:

$$P_{jt} = \alpha_j \left(\frac{Y_t}{Y_{jt}} \right)^{\frac{1}{\sigma}}$$

- **Stationary Competitive Equilibrium:** constant prices $\{P_j^*\}$ and invariant distribution π^*:
  1. Workers solve the Bellman equation taking $\{P_j^*\}$ as given, and die and are reborn with probability h
  2. Firms pay piece-rate wages: $W_{ijt} = P_j^* \, Y_{ijt}$
  3. Prices clear goods markets: $P_j^* = \alpha_j (Y / Y_j)^{1/\sigma}$
  4. π^* is invariant under optimal choices and state transitions

Data

- Portuguese administrative data (**Quadros de Pessoal**), 1995–2019
  - **8 occupation clusters** ($J = 8$): minimize **within-cluster task dispersion**, preserve **mobility flows**
  - **2 skill dimensions** ($K = 2$): **Cognitive, Manual** (Mihaylov & Tijdens 2019)
- **Large Sample** ($\approx$ 29m obs): mobility patterns for 8 clusters and 3-digit occupations + ESCO
- **Estimation Sample** ($\approx$ 2.3m obs): unbalanced worker panel for structural estimation
- **Two measures of wages:**
  - **Real hourly wages**
  - **Worker $\times$ occupation wage fixed effect** φ_{ij} to isolate worker-occupation match quality:

$$\log w_{ijt} = \underbrace{\varphi_{ij}}_{\text{match quality}} + f_j(\text{occ. tenure}_{ijt}) + \gamma_t + \varepsilon_{ijt}$$

Sample Selection Transitions are Similar

Switches by Experience Switches by Ability

Identification — Step 1: Wage Parameters under Learning

Wage parameters Γ. Arcidiacono et al. (2025), Bunting et al. (2025)

- Conditional on history, there is selection on observables (**Key: piece-rate**):

$$D_{it} \perp\!\!\!\perp \xi_i \mid \mathcal{H}_{it} \qquad \mathcal{H}_{it} \equiv \left(g_i,\ \{d_{is}, w_{is}\}_{s<t},\ \boldsymbol{\tau}_{it}\right)$$

- *Control function* approach (project skills on the observed history):

$$w_{ijt} \;=\; p_j + \boldsymbol{\lambda}_j' \,\mathbb{E}[\xi_i \mid \mathcal{H}_{it}] + h_j(\boldsymbol{\tau}_{it}) + \varepsilon_{ijt}$$

- Need to fix anchors, enough movers across occupations and variation in task-tenure

$$\Rightarrow \Gamma = \Big\{ \underbrace{(p_j, \boldsymbol{\lambda}_j, h_j, \varsigma^2_{\varepsilon j})_j,}_{\text{wage equation parameters}} \quad \underbrace{(\boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g)_g}_{\text{multidimensional skill distribution}} \Big\} \text{ are identified}$$

▸▸ Identification Details

Identification — Step 2: Dynamic Choice Parameters

Dynamic choice parameters $\Delta = \{\eta, \kappa_0\}$: Hotz–Miller (1993) inversion

- $\widehat{\Gamma}$ + rationality $(\tilde{\xi}_{i0}, \tilde{\Sigma}_{i0}) = (\mu_{g_i}, \Sigma_{g_i}) \Rightarrow$ sequence of beliefs $\Rightarrow$ **CCPs** $P(j \mid d, \boldsymbol{x})$
- Identifying content: the **level and composition of mobility** across destinations
- $\varrho = 0.271$ calibrated to the **overall mobility level**

Identification Details Estimation Details

Identification — Step 2: Dynamic Choice Parameters

Dynamic choice parameters $\Delta = \{\eta, \kappa_0\}$: Hotz–Miller (1993) inversion

- $\widehat{\Gamma}$ + rationality $(\tilde{\xi}_{i0}, \tilde{\Sigma}_{i0}) = (\mu_{g_i}, \Sigma_{g_i}) \Rightarrow$ sequence of beliefs $\Rightarrow$ **CCPs** $P(j \mid d, \boldsymbol{x})$
- Identifying content: the **level and composition of mobility** across destinations
- $\varrho = 0.271$ calibrated to the **overall mobility level**

Amenities — two destinations $j, m \neq d$ from a common origin ($\eta_{0g} = 0$):

$$\eta_{jg} - \eta_{mg} = (1-\beta)\Big[\varrho \log \frac{P(j|d)}{P(m|d)} - (\tilde{w}_j - \tilde{w}_m) - \beta\big(\Psi_j(\tilde{w}, P) - \Psi_m(\tilde{w}, P)\big)\Big] \quad (j, m \neq d)$$

Identification Details Estimation Details

Identification — Step 2: Dynamic Choice Parameters

Dynamic choice parameters $\Delta = \{\eta, \kappa_0\}$: Hotz–Miller (1993) inversion

- $\widehat{\Gamma}$ + rationality $(\tilde{\xi}_{i0}, \tilde{\Sigma}_{i0}) = (\mu_{g_i}, \Sigma_{g_i}) \Rightarrow$ sequence of beliefs $\Rightarrow$ **CCPs** $P(j \mid d, \boldsymbol{x})$
- Identifying content: the **level and composition of mobility** across destinations
- $\varrho = 0.271$ calibrated to the **overall mobility level**

Amenities — two destinations $j, m \neq d$ from a common origin ($\eta_{0g} = 0$):

$$\eta_{jg} - \eta_{mg} \;=\; (1-\beta)\Big[\varrho \log \frac{P(j|d)}{P(m|d)} - (\tilde{w}_j - \tilde{w}_m) - \beta\big(\Psi_j(\tilde{w}, P) - \Psi_m(\tilde{w}, P)\big)\Big] \quad (j, m \neq d)$$

Switching cost — move j vs stay d from the same origin:

$$\kappa_0 \;=\; (\tilde{w}_j - \tilde{w}_d) + \beta\big(\Psi_j(\tilde{w}, P) - \Psi_d(\tilde{w}, P)\big) + \frac{\eta_{jg} - \eta_{dg}}{1-\beta} - \varrho \log \frac{P(j|d)}{P(d|d)} \quad (j \text{ vs. stay } d)$$

Identification Details Estimation Details

Estimation: Two Stages EM-NFXP

Stage 1 — Expectation-Maximization (Dempster et al. 1977) for wage parameters Γ

- **E-step:** given observed history, we compute $\tilde{\xi}_i^q$ posterior (control) for each worker;
- **M-step:** regress w_{ijt} using posterior from E-step $\tilde{\xi}_i^q$ as a control

Skill Distribution Tenure Returns Value of Learning

Estimation: Two Stages EM-NFXP

Stage 1 — Expectation-Maximization (Dempster et al. 1977) for wage parameters Γ

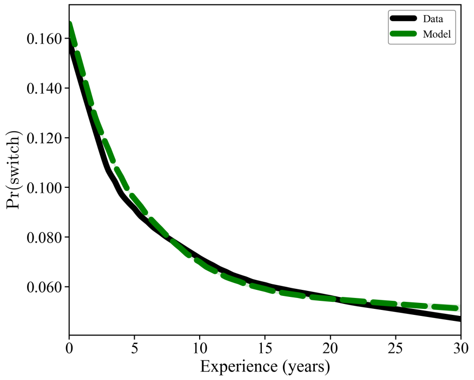
- **E-step:** given observed history, we compute $\tilde{\xi}_i^q$ posterior (control) for each worker;
- **M-step:** regress w_{ijt} using posterior from E-step $\tilde{\xi}_i^q$ as a control

Stage 2 — Nested Fixed Point (NFXP, Rust 1987) for dynamic choice parameters Δ given $\widehat{\Gamma}$

- **Inner loop:** at each Δ, solve the integrated Bellman fixed point for $\widehat{EV}(\boldsymbol{x})$;
- **Outer loop:** update Δ from the implied logit CCPs until convergence.

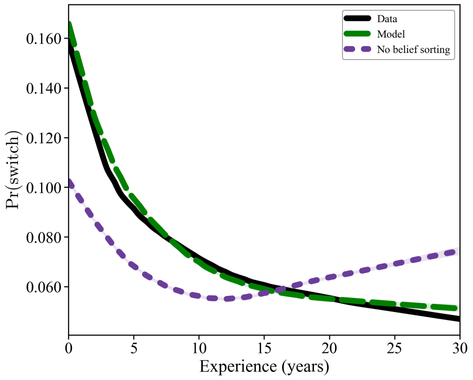
▶▶ Skill Distribution ▶▶ Tenure Returns ▶▶ Value of Learning

The Declining Hazard Rate of Occupational Mobility



- **Learning** and **task-specific experience** generate the declining probability of switching

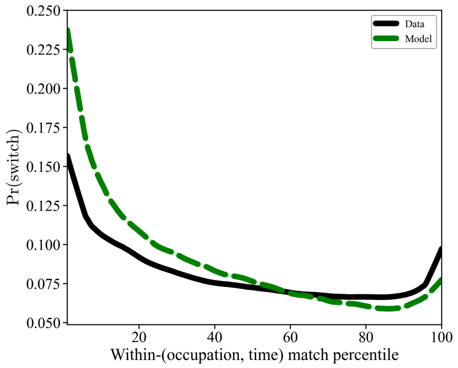
The Declining Hazard Rate of Occupational Mobility: the Role of Belief Sorting



- **Learning** and **task-specific experience** generate the declining probability of switching
- **Learning and beliefs sorting** generates mobility early in the career, and **dampens** it later

The U-Shape of Occupational Mobility

Occupational Mobility

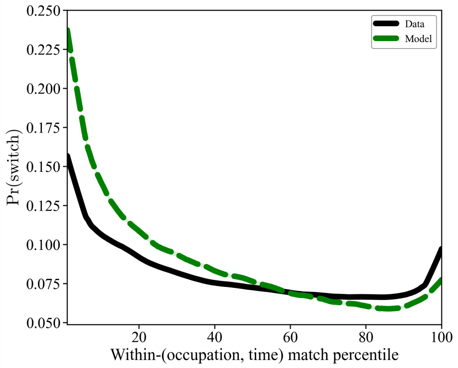


- Low- and high-match quality workers more likely to switch (Groes, Kircher and Manovskii, 2015)

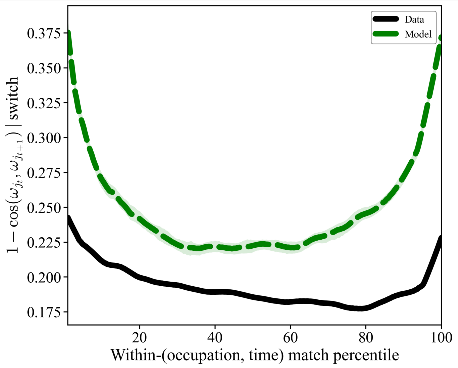
Directionality of Change | Task Distance Density | Directionality | By Experience

The U-Shape of Occupational Mobility and Task-Content Change

Occupational Mobility



Task-Content Change

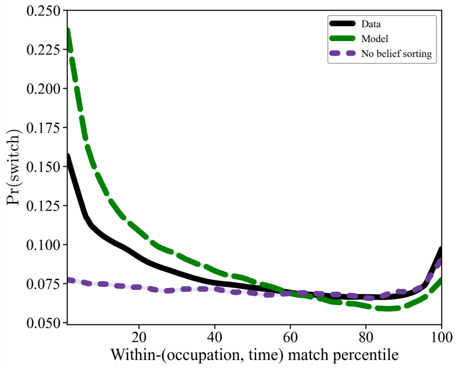


- Low- and high-match quality workers more likely to switch (Groes, Kircher and Manovskii, 2015)
- They also change **task content** more: task distance is **U-shaped** in match quality

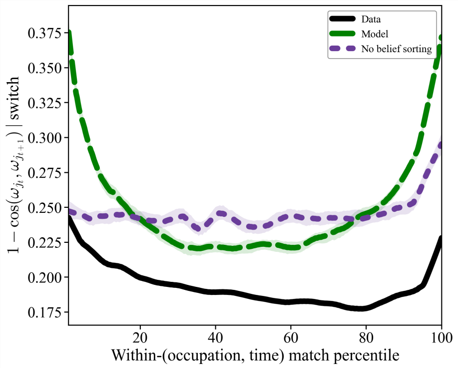
▸▸ Directionality of Change ▸▸ Task Distance Density ▸▸ Directionality ▸▸ By Experience

Occupational Mobility and Task-Content Change: the Role of Belief Sorting

Occupational Mobility



Task-Content Change



- Without **belief sorting** we miss the **left-skewness** of the U-shape
- Without **belief sorting** task-content change becomes **flat** for low-match quality workers

Directionality of Change Task Distance Density Directionality By Experience

The Aggregate Gains of Removing Information Frictions

Two career paths. Same worker i, same true skills $\boldsymbol{\xi}_i$, two information regimes:

$$d_{it}^B = \arg\max_{j \in \mathcal{O}} v_j(\widetilde{\boldsymbol{\xi}}_{it}, \widetilde{\boldsymbol{\Sigma}}_{it}, \boldsymbol{x}_{it}) + \varrho\epsilon_{jt} \qquad \text{(imperfect information)}$$

$$d_{it}^T = \arg\max_{j \in \mathcal{O}} v_j(\boldsymbol{\xi}_i, \mathbf{0}, \boldsymbol{x}_{it}) + \varrho\epsilon_{jt} \qquad \text{(perfect information)}$$

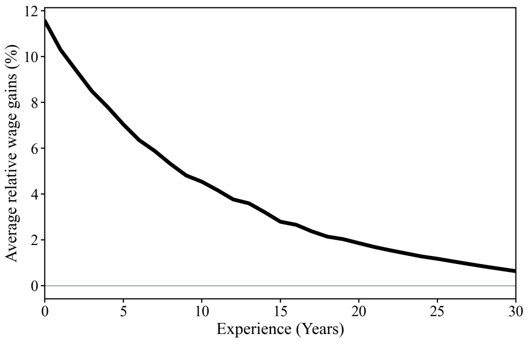
Aggregate output (CES final good, $\sigma = 1.57$):

$$Y = \left(\textstyle\sum_{j \in \mathcal{O}} \alpha_j Y_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \qquad Y_j = \textstyle\int y_j(\boldsymbol{\xi}, \boldsymbol{x})\, \mathbf{1}\{d = j\}\, \mathrm{d}\pi^*.$$

Output gain from removing frictions: $\mathcal{G}_Y = \dfrac{Y^T - Y^B}{Y^B} = \mathbf{5.2\%}$

The Wage Gains of Removing Information Frictions

Lifecycle Wage Gains

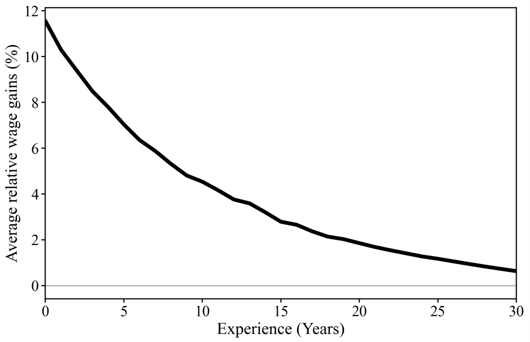


- **Large wage gains early** (≈ 11.6% at entry), **decaying** over the lifecycle
- Gains concentrated among the **worst-matched** workers (> 20% at the bottom vs 3% at the top)

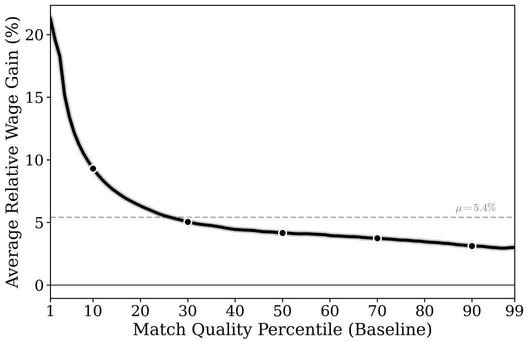
⏩ Decomposition ⏩ Wage Gains by Percentile ⏩ Specialists vs Generalists ⏩ Sorting on Comp. Advantage

The Wage Gains of Removing Information Frictions

Lifecycle Wage Gains



Wage Gains by Match Quality



- **Large wage gains early** ($\approx$ 11.6% at entry), **decaying** over the lifecycle
- Gains concentrated among the **worst-matched** workers ($>$ 20% at the bottom vs 3% at the top)

⏩ Decomposition ⏩ Wage Gains by Percentile ⏩ Specialists vs Generalists ⏩ Sorting on Comp. Advantage

Correcting Beliefs about Comparative Advantage

Counterfactual: Information provision about comparative advantage (Kiss, Garlick, Orkin, and Hensel 2025)

- **Treatment:** skills assessment + report on relative performance
- **In the model:** workers receive noisy signals about latent skills

Signals: K independent Gaussian signals at t:

$$s_{ikt} = \xi_{ik} + \zeta_{ikt}, \qquad \zeta_{ikt} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\zeta^2), \qquad k = C, M$$

Correcting Beliefs about Comparative Advantage

Counterfactual: Information provision about comparative advantage (Kiss, Garlick, Orkin, and Hensel 2025)

- **Treatment:** skills assessment + report on relative performance
- **In the model:** workers receive noisy signals about latent skills

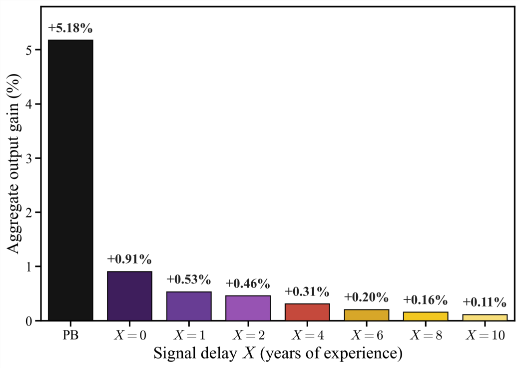
Signals: K independent Gaussian signals at t:

$$s_{ikt} = \xi_{ik} + \zeta_{ikt}, \qquad \zeta_{ikt} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\zeta^2), \qquad k = C, M$$

Calibrate signal variance σ_ζ^2 to match Kiss et al. 2025 Average Treatment Effect:

- Target ATE: $+13.5$ pp in the share with beliefs aligned with assessed comparative advantage

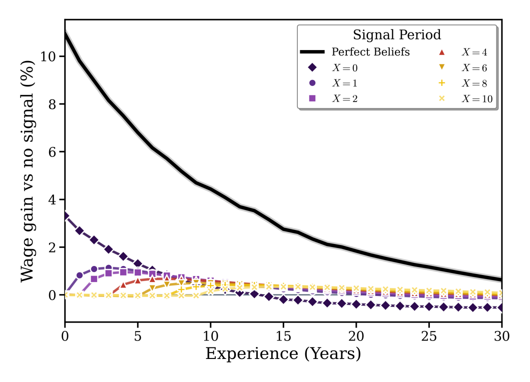
Output Gains from the Information Treatment at Different Career Stage



- **At entry**: output rises **0.9%** — ≈**18%** of the **5.2%** perfect-information gain
- **Decays sharply if delayed**: **half** after two years, **nearly gone** after a decade

Effect of the Information Provision at Different Career Stages

Lifecycle Wage Gains

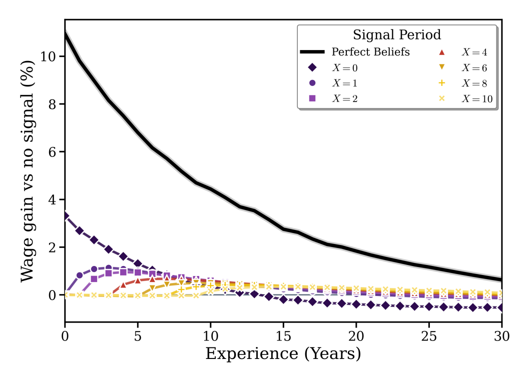


- Gains are **front-loaded** (3.3% at entry) and **decay sharply** if the signal is delayed
- Worst-matched gain ≈ 9%: **half** the perfect-information gain; above the median, **nothing**

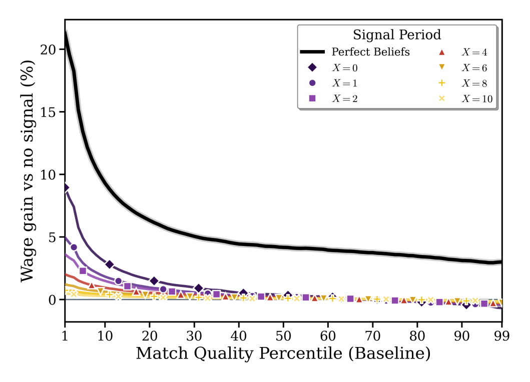
Reducing Switching Costs

Effect of the Information Provision at Different Career Stages

Lifecycle Wage Gains



Wage Gains by Match Quality



- Gains are **front-loaded** (3.3% at entry) and **decay sharply** if the signal is delayed
- Worst-matched gain ≈ 9%: **half** the perfect-information gain; above the median, **nothing**

▶▶ Reducing Switching Costs

Conclusion

Main Contribution: Structural Dynamic Roy Model with **multidimensional skills learning**.

Key Findings:

- **Learning** drives mobility of **low-match-quality** workers early in the career
- **Removing information frictions** about skills raises aggregate output by **5.2%**
- Wage gains concentrated among **low-match-quality workers**
- A **calibrated information treatment** at entry raises output by **0.9%**

In the paper:

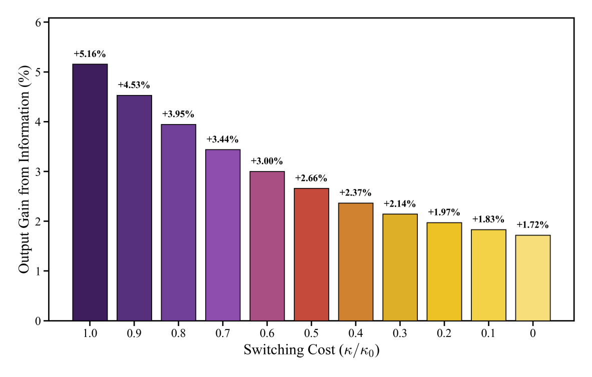
- I use the model to study the interaction of **skill uncertainty** with **job transformation**.
- While output increases, the **cost of skill uncertainty increases**.

Reducing Switching Costs AI Counterfactual AI & Wage Cost

Thank you!

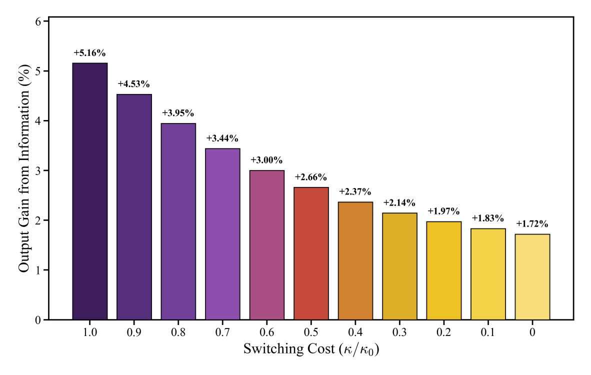
Appendix: Reducing Switching Costs Decreases Cost of Mismatch

Counterfactual (Appendix): Reduce switching costs (κ_0, κ_1) by 10%–90%



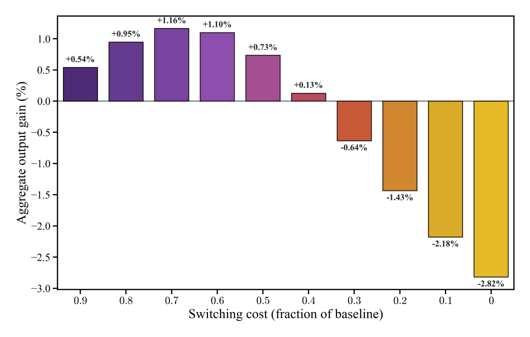
Appendix: Reducing Switching Costs Decreases Cost of Mismatch

Counterfactual (Appendix): Reduce switching costs (κ_0, κ_1) by 10%–90%

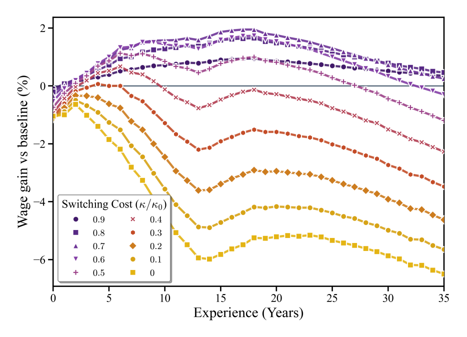


Appendix: Reducing Switching Costs Increases Output Non-Monotonically

Aggregate Output



Lifecycle Wage Gains



- Aggregate output peaks at -50% reduction ($+2.4\%$), then declines
- **Two opposing forces:** better matching vs. excessive churning that destroys human capital

Back

Appendix: Optimal Experimentation at the Beginning of the Career

- **Counterfactual:** assign entrants to occupation that maximizes **information gain**

Informative Occupations by Group

Back

Appendix: Optimal Experimentation at the Beginning of the Career

- **Counterfactual:** assign entrants to occupation that maximizes **information gain**

Information Gain:

$$IG_{jg}^{\text{entry}} \equiv \underbrace{I(\xi_i;\, s_{ij1} \mid g)}_{\text{Information Gain}} = \underbrace{h(s_{ij1} \mid g)}_{\text{Total Entropy}} - \underbrace{h(s_{ij1} \mid \xi_i, g)}_{\text{Noise Entropy}} = \frac{1}{2} \log\left(1 + \underbrace{\frac{\boldsymbol{\lambda}_j' \Sigma_g \boldsymbol{\lambda}_j}{\varsigma_{\varepsilon j}^2}}_{\text{Signal-to-Noise Ratio}}\right)$$

Intuition:

- **High Signal**: Occupation loads strongly ($\boldsymbol{\lambda}_j$) on dimensions with **high prior uncertainty** (Σ_g)
- **Low Noise**: Wages are a **precise signal** of skills (low residual variance $\varsigma_{\varepsilon j}^2$)

Informative Occupations by Group

Back

Appendix: Optimal Experimentation at the Beginning of the Career

- **Counterfactual:** assign entrants to occupation that maximizes **information gain**

Information Gain:

$$IG_{jg}^{\text{entry}} \equiv \underbrace{I(\xi_i; s_{ij1} \mid g)}_{\text{Information Gain}} = \underbrace{h(s_{ij1} \mid g)}_{\text{Total Entropy}} - \underbrace{h(s_{ij1} \mid \xi_i, g)}_{\text{Noise Entropy}} = \frac{1}{2} \log \left(1 + \underbrace{\frac{\boldsymbol{\lambda}_j' \boldsymbol{\Sigma}_g \boldsymbol{\lambda}_j}{\varsigma_{\varepsilon j}^2}}_{\text{Signal-to-Noise Ratio}} \right)$$

Intuition:

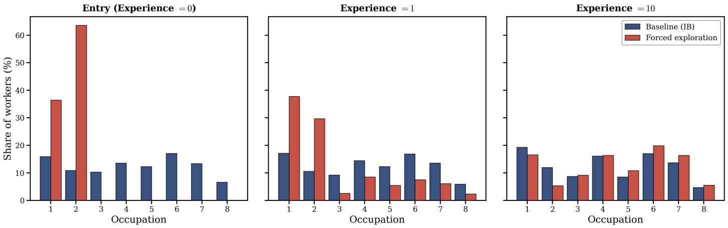
- **High Signal**: Occupation loads strongly ($\boldsymbol{\lambda}_j$) on dimensions with **high prior uncertainty** ($\boldsymbol{\Sigma}_g$)
- **Low Noise**: Wages are a **precise signal** of skills (low residual variance $\varsigma_{\varepsilon j}^2$)

Assignment rule: Entrants assigned to $\arg\max_j \; IG_{jg}^{\text{entry}}$ for 1 period, then free to choose

Informative Occupations by Group

Back

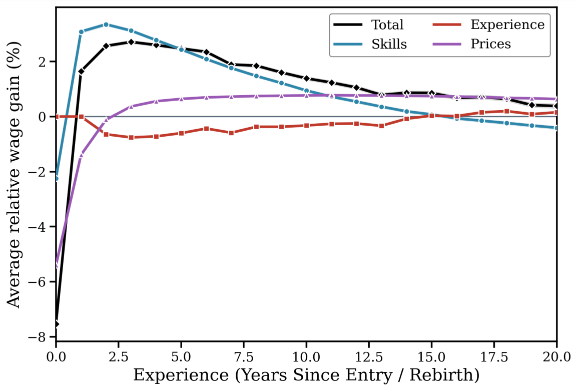
Appendix: Occupational Allocation under Forced Exploration



- Period 1: workers concentrated in the most **informationally rich** occupations
- By period 10, occupational distribution **converges back** to baseline — but workers are better matched!

▸ Back

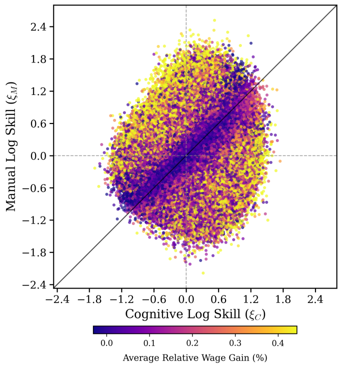
Appendix: Wage Gains from Forced Exploration



- **Large initial wage losses** followed by **persistent wage gains** from better sorting on skills
- Output only increases by 0.7%

▸ Back

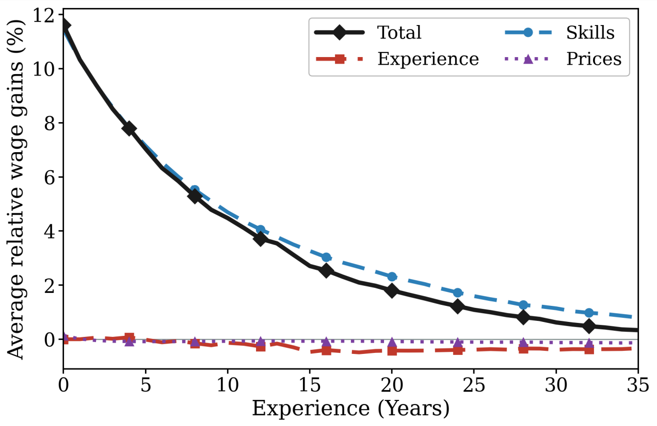
Appendix: Who Benefits Most? Specialists vs Generalists



- **30%** choose a different initial occupation
  50% switch at some point (**17.5%** gains)
- **Specialists** (away from 45° line)
  gain most from perfect information
- **Generalists** (near 45° line)
  lose little from uncertainty

▸ Back

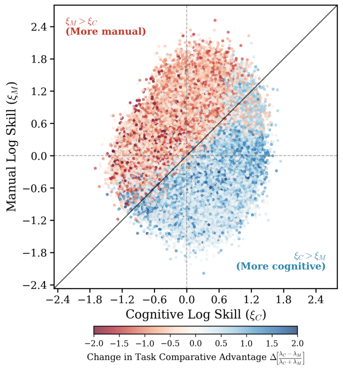
Appendix: Decomposition of the Wage Gains from Removing Frictions



- Gains come almost entirely from **better sorting on skills**
- **Experience** and **price** channels are negligible

Back

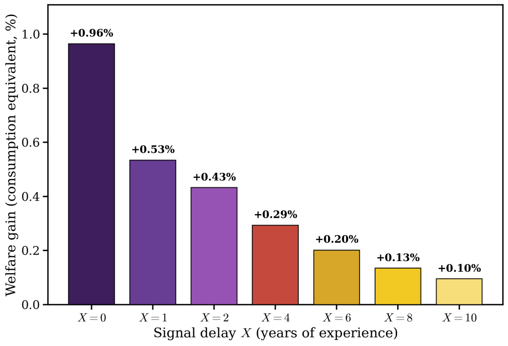
Appendix: Removing Information Frictions Improves Sorting

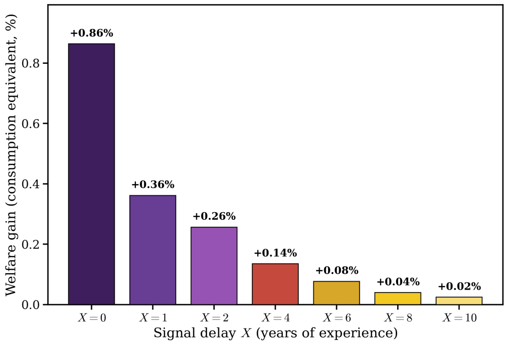


- Cognitive specialists ($\xi_C > \xi_M$) move to more **cognitive-intensive** occupations
- Manual specialists ($\xi_M > \xi_C$) move to more **manual-intensive** occupations
- For low-skilled workers, **tenure** matters more; for high-skilled, **skills** matter more

Tenure & Skill Returns

Appendix: Welfare Gains from Information Provision by Career Stage





- At entry, information provision raises aggregate welfare by $+0.022$ **utils**, decaying with delay
- Welfare gains are **front-loaded** early in the career and fade for late treatment

▸ Back

Appendix: AI Shock – LLM Exposure Framework (Eloundou et al. 2024)

How does AI change task importance within occupations?

Classify each ESCO skill by **LLM exposure** (Eloundou et al. 2024):

E0: no exposure **E1**: direct exposure **E2**: LLM+ exposure

Key insight: all Manual skills → E0 | Cognitive skills → E0, E1, or E2

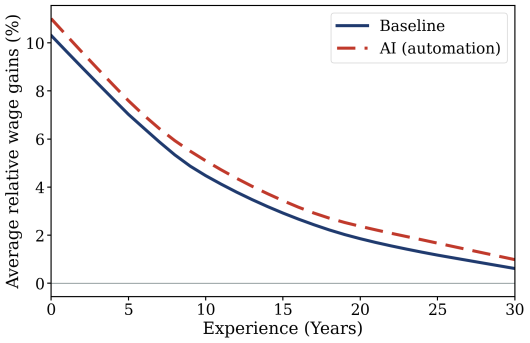
Exposure → task importance shock: α_{jC} = wtd. avg. exposure of C-skills in occ. j (E0→0, E1→1, E2→0.5)

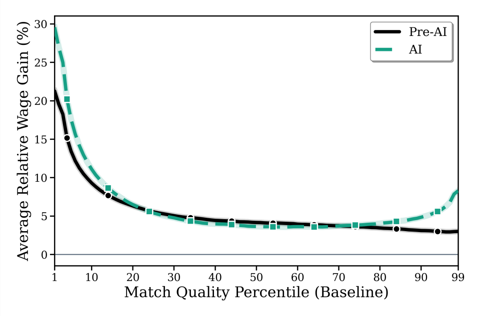
$$\omega_{jC}^{\text{new}} = \omega_{jC}\,(1-\alpha_{jC}), \qquad \omega_{jM}^{\text{new}} = \omega_{jM}, \qquad \text{renormalize to } \textstyle\sum_k \omega_{jk}^{\text{new}} = 1$$

Classification Examples

Manual Skills Revaluation

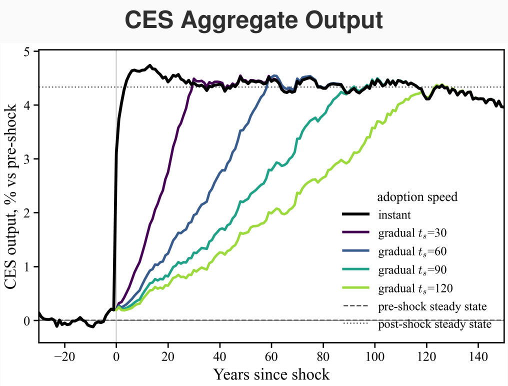
Appendix: AI Increases the Wage Cost of Imperfect Information

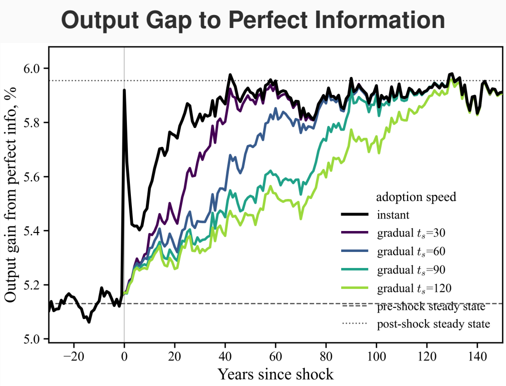




- AI raises aggregate output, but **amplifies the output cost of mismatch** from **5.2%** to **5.9%**
- Increase is largest **early in the career** and among the **worst-matched** workers

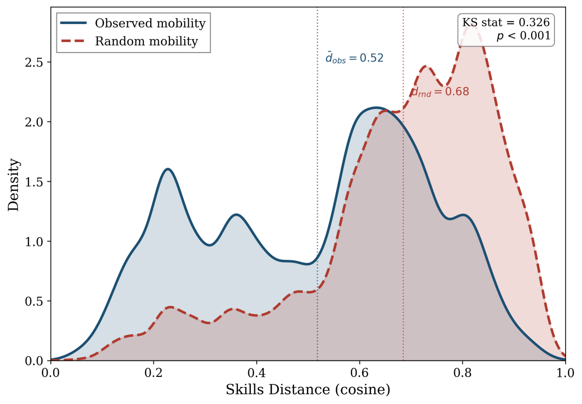
Appendix: AI Along the Transition Path





- AI raises long-run output by $+4.3\%$; faster adoption **front-loads** the gains
- AI **widens** the gap to perfect information ($5.2\% \to 5.9\%$): uncertainty is **more costly**

Appendix: Occupational Transitions are Similar



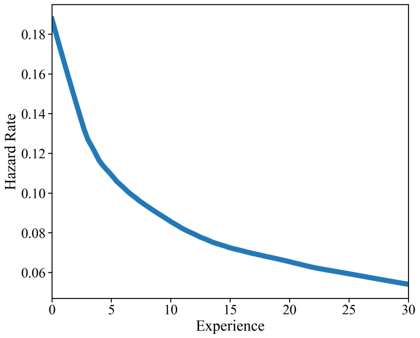
Skills distance (Gathmann & Schönberg 2010):

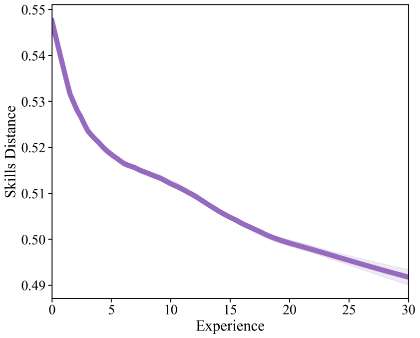
$$d(j, j') = 1 - \frac{\omega_j^\top \omega_{j'}}{\|\omega_j\| \, \|\omega_{j'}\|}$$

- Mean distance: **0.52** (s.d. 0.22). Random counterfactual: **0.68** (KS test $p < 0.001$)
- Closest common switches: cashiers $\leftrightarrow$ shop salespersons ($d = 0.24$)
- Most distant switch: mathematician $\rightarrow$ domestic helper ($d = 0.99$)

⏩ Back

Appendix: Switching Probability and Skills Distance by Experience

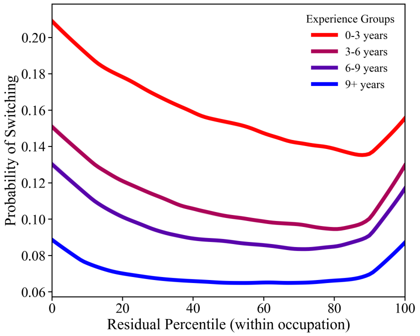


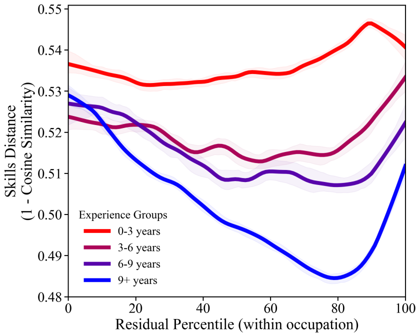


- Probability of switching occupations declines with experience in labor market
- Switchers move to occupations requiring more different skills early in their career
  - Task distance drops by $\approx$ **0.23 standard deviations** over the first 15 years

Back

Appendix: Switching by Ability and Experience





- **"U-shape"**: bad and good matches more likely to switch across experience levels
- Conditional on moving, **intensity of skills distance depends on match quality**

▸▸ Back

Appendix: From Tasks to Output

Task importance ω_{jk} determines each task k's contribution in occupation j, with $\sum_k \omega_{jk} = 1$

Output aggregates via **Cobb-Douglas**:

$$Y_{ijt} = \prod_{k=1}^{K} \left(\frac{Y_{ijkt}}{\omega_{jk}} \right)^{\omega_{jk}}$$

Optimal time allocation

$$\ell^*_{jk} = \omega_{jk}$$

independent of worker type

Time-allocation problem

Substituting $\ell^*_{jk} = \omega_{jk}$, output reduces to:

$$Y_{ijt} = \underbrace{\prod_{k=1}^{K} \theta_{ik}^{\lambda_{jk}}}_{\text{skill bundle}} \cdot \underbrace{\prod_{k=1}^{K} \tau_{ikt}^{\beta_{jk}}}_{\text{experience bundle}} = \boldsymbol{\theta}_i^{\boldsymbol{\lambda}_j} \boldsymbol{\tau}_{it}^{\boldsymbol{\beta}_j}$$

where $\lambda_{jk} = \omega_{jk}\lambda^0_{jk}$ and $\beta_{jk} = \omega_{jk}\beta^0_{jk}$ are *effective* occupation-task elasticities

▸ Back

Appendix: Labor Demand and Wage Equation

- Taking P_{jt} as given, firm maximizes profits:

$$\max_{\{n_{jt}\}} \Pi_{jt} = P_{jt} Y_{jt} - \int W_{jt}(\boldsymbol{\theta}, \boldsymbol{\tau}_t)\, n_{jt}(\boldsymbol{\theta}, \boldsymbol{\tau}_t)\, dG_t$$

- Output subject to idiosyncratic shock: $\exp(\varepsilon_{ijt}) \cdot Y_{ijt}$, with $\varepsilon_{ijt} \sim N(0, \varsigma_j^2)$
- Under **piece-rate contracts** and perfect competition:

$$W_{ijt} = P_{jt} \exp(\varepsilon_{ijt})\, Y_{ijt}$$

- **Log-wage equation:**

$$w_{ijt} = p_{jt} + \boldsymbol{\lambda}_j' \boldsymbol{\xi}_i + h_j(\boldsymbol{\tau}_{it}) + \varepsilon_{ijt}$$

  where $\boldsymbol{\xi}_i = \log \boldsymbol{\theta}_i$ (latent log-skills), $p_{jt} = \log P_{jt}$

▸ Back

Appendix: Most Informative Occupations at Entry

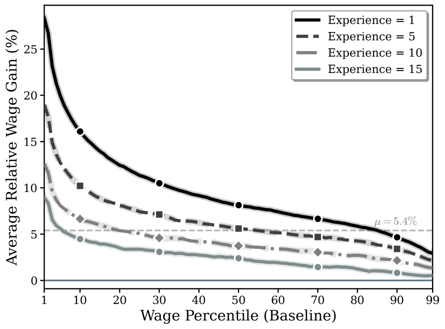
Which occupations maximize learning?

- **High-Educated & Medium-Educated Men:**
  $\Rightarrow$ Assigned to **Manual-intensive** occupations (Occ_6)
- **Low-Educated (Men & Women):**
  $\Rightarrow$ Assigned to **Purely Cognitive** occupations (Occ_7)
- **Medium-Educated Women:**
  $\Rightarrow$ Assigned to **Cognitive-intensive** occupations (Occ_8)

- **Key Takeaway:** Informative occupations are typically **highly specialized**. They load strongly on the skill dimension where the demographic group has the highest prior uncertainty.

Back

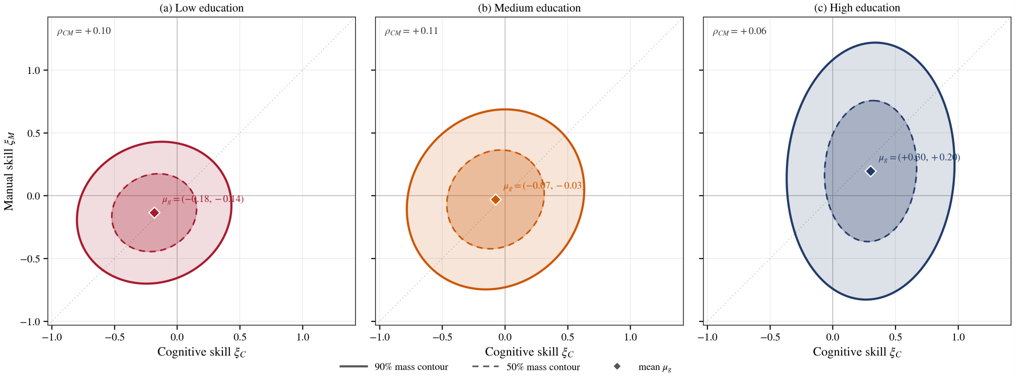
Appendix: Wage Gains across the Wage Distribution



- **Low-wage workers gain most**: ∼13% at 10th percentile vs ∼5% at 90th
- Removing information frictions **reduces wage inequality**

▸ Back

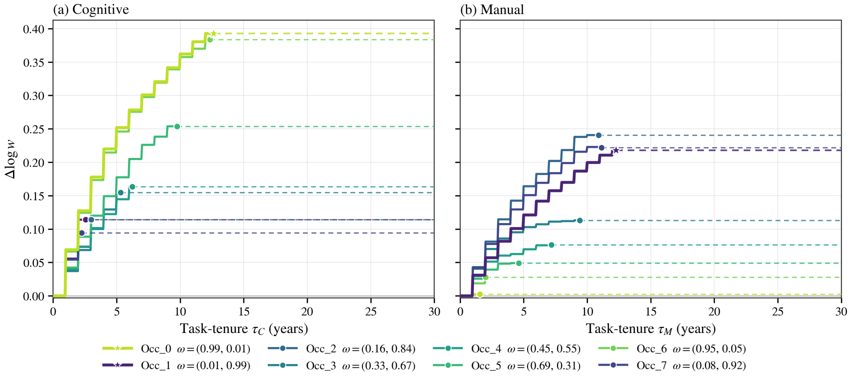
Appendix: Estimated Skill Distribution by Education Group



- **Higher education** $\Rightarrow$ **higher skills** in both dimensions
- Cognitive and manual skills **weakly correlated** within each group

▸ Back

Appendix: Estimated Returns to Task Tenure



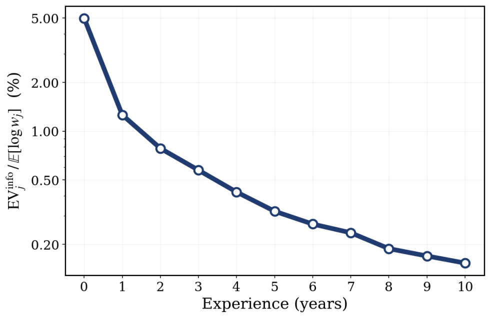
- Returns to task tenure are **higher and steeper in cognitive than manual tasks**

Back

Appendix: The Ex-Ante Value of Learning over the Lifecycle

Value of information in the chosen occupation $j^\star$, per unit of its expected log wage:

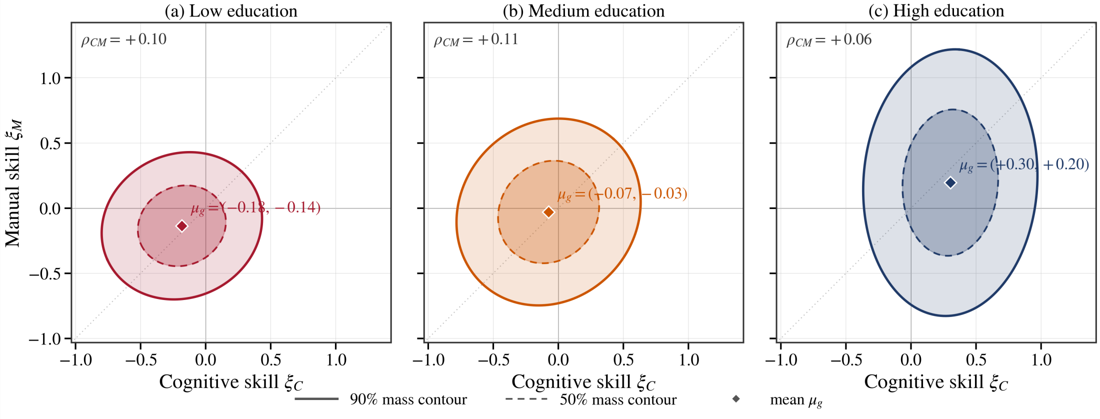
$$\underbrace{\mathrm{EV}_j^{\text{info}}(\boldsymbol{x}) \;=\; \beta\,\mathbb{E}_{s,\epsilon}\big[V_j(\tilde{\xi}', \tilde{\Sigma}'; \tau+\omega_j) \mid \boldsymbol{x}\big] \;-\; \beta\,\mathbb{E}_{\epsilon}\big[V_j(\tilde{\xi}, \tilde{\Sigma}; \tau+\omega_j) \mid \boldsymbol{x}\big]}_{\text{gain from updating beliefs on the wage signal vs. holding them fixed}}$$



- Learning is most valuable **early in the career** and **decays** as beliefs sharpen with experience

▸ Back

Appendix: Initial Skill Distribution/Beliefs by Education Group



- The **higher the education level, the higher the initial skill level**
- The **higher the education level, the higher the comparative advantage in cognitive skills**
- More educated individuals start higher in the occupational job ladder and immediately into more

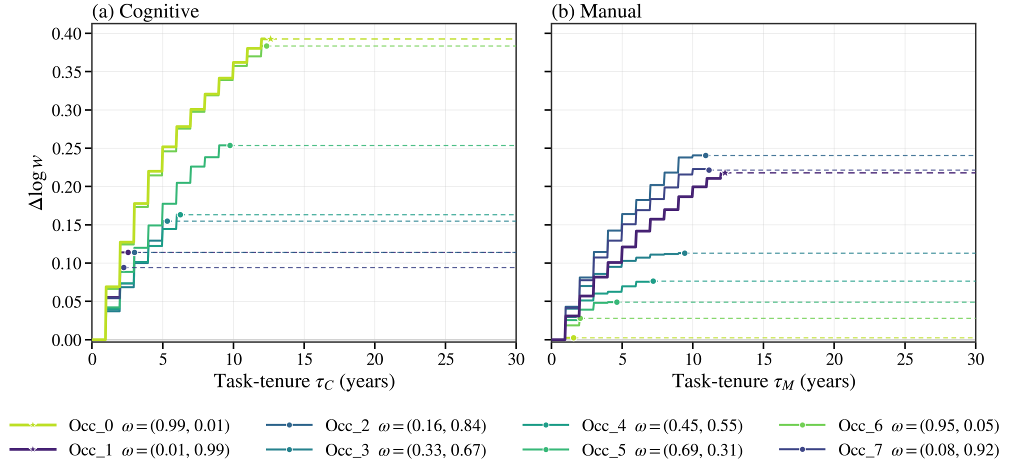
Appendix: Estimated Returns to Skills by Occupation

Occupation	$\omega_{j,C}$	$\omega_{j,M}$	$\lambda_{j,C}$	$\lambda_{j,M}$	Type
Occ_0	0.936	0.064	0.726	0.584	**C**
Occ_1	0.201	0.799	0.446	0.800	**M**
Occ_2	0.566	0.434	0.495	0.875	**B**
Occ_3	0.015	0.985	0.000	1.000	**M**
Occ_4	0.322	0.678	0.773	0.238	**B**
Occ_5	0.712	0.288	0.227	1.005	**C**
Occ_6	0.117	0.883	0.000	1.085	**M**
Occ_7	0.998	0.002	1.000	0.000	**C**
Occ_8	0.846	0.154	0.885	0.207	**C**
Occ_9	0.432	0.568	0.427	0.796	**B**

- **Large heterogeneity** in task importance and skill elasticities across occupations
- **C** = Cognitive-intensive, **M** = Manual-intensive, **B** = Balanced

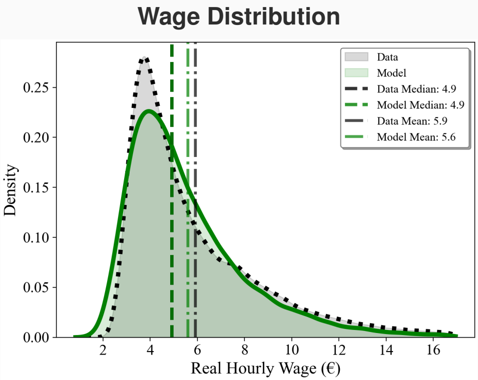
Back

Appendix: Estimated Returns to Task Experience by Occupation



- Experience in tasks has **different returns across occupations**
- Returns to experience **higher and steeper in cognitive tasks than manual**

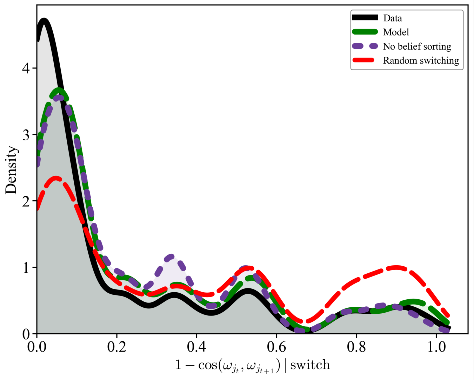
Appendix: Model vs Data: Wage Distribution and Wage Experience-Profile



- Model captures **shape of the wage distribution** and **increasing wage profile**

▸ Back

Appendix: Occupational Switches Are Closer in Task Space

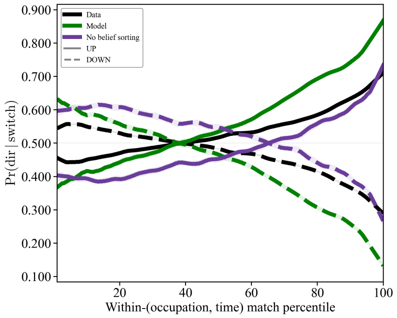


- Model reproduces the empirical density of **task distance conditional on switching**; far from random reallocation

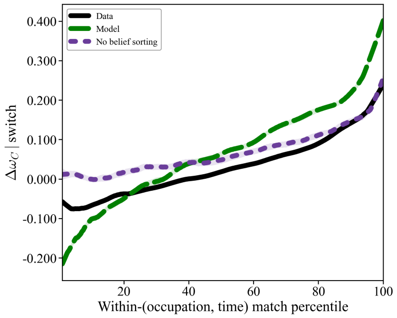
▸ Back

Appendix: Directionality and Cognitive Content Conditional on Switching

Direction of the Move (Up/Down)



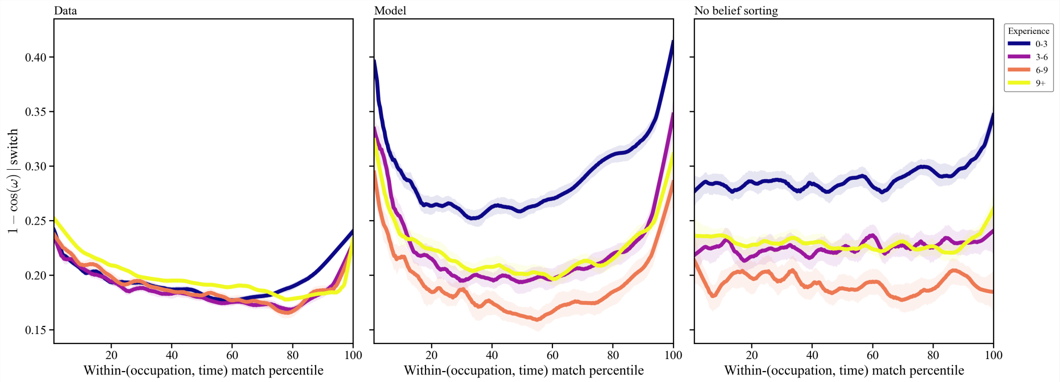
Change in Cognitive Task Content



- **Well-matched** switchers move **up** the ladder and toward **more cognitive** task content
- **Low-match** quality moves downward to **more manual** occupation

▸ Back

Appendix: Model vs Data: Task Distance of Switches by Position and Experience

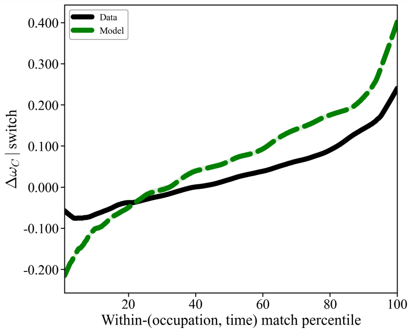


- Task distance of moves **declines with experience**: younger workers make larger task jumps

▸ Back

Appendix: Directionality of Task-Content Change

Data vs Model

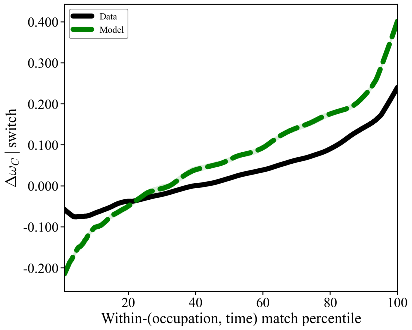


- Bad matches move to **less cognitive** occupations ($\Delta\omega_C < 0$), good matches to more

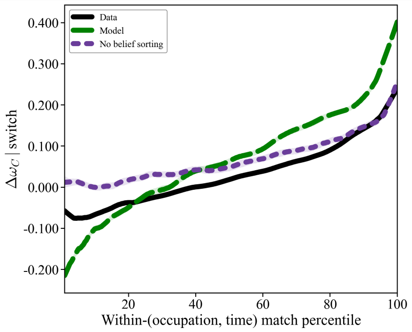
Back Task Distance Density Directionality By Experience

Appendix: Directionality of Task-Content Change: the Role of Belief Sorting

Data vs Model



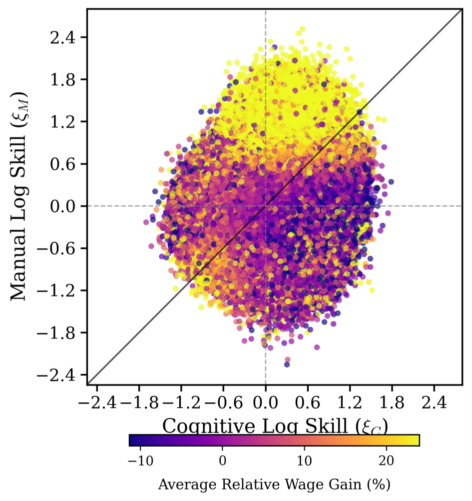
No Belief Sorting



- Bad matches move to **less cognitive** occupations ($\Delta\omega_C < 0$), good matches to more
- Without **belief sorting** the gradient **flattens** and never turns negative

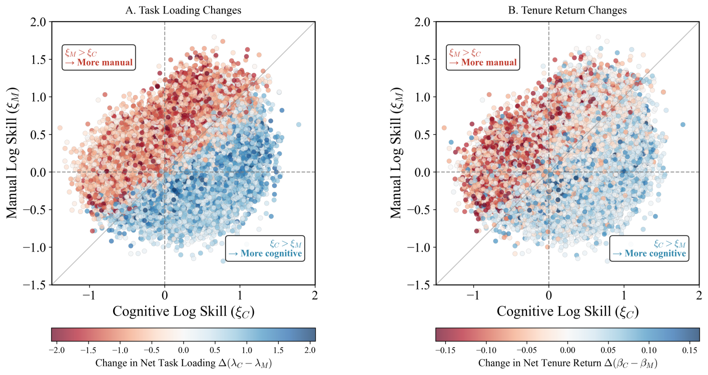
▸ Back ▸▸ Task Distance Density ▸▸ Directionality ▸▸ By Experience

Appendix: AI Wage Gains by Skill Type



- **Cognitive specialists** (below 45°): harmed by AI (primary skill exposed)
- **Manual specialists** (above 45°): gain as manual skills become more valuable
- **Generalists**: benefit by sorting into more manual-intensive occupations

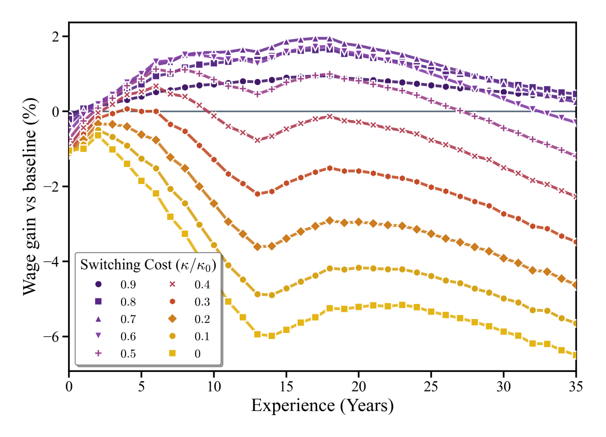
Appendix: Tenure & Skill Returns under Perfect Information



▸ Back

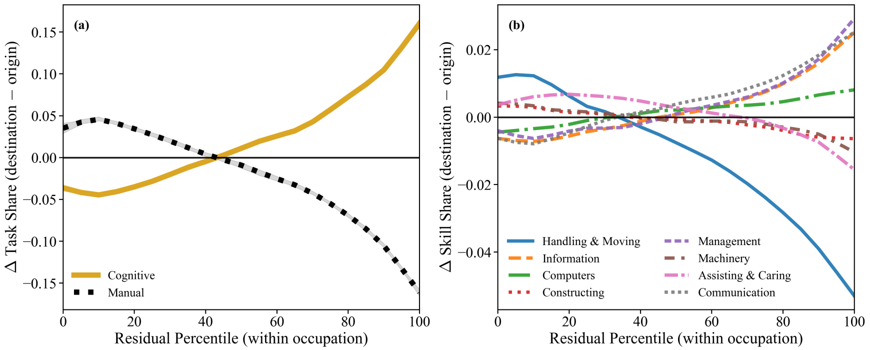
Appendix: Lifecycle Wage Gains from Reducing Switching Costs

Direct effect: Wage gain relative to original baseline (imperfect beliefs, full switching costs)



- Workers earn higher wages even without better information
- Lower switching costs allow faster reallocation toward better-matched occupations

Appendix: Empirical Task Directionality



▸ Back

Appendix: LLM Classification of Skills – Examples

E0 – No Exposure

```
{
 "esco_skill": "moving and lifting",
 "esco_code": "S5.5.1",
 "dimension": "Manual",
 "exposure": "E0",
 "rationale": "Physical task"
}
{
 "esco_skill": "negotiating",
 "esco_code": "S4.7.2",
 "dimension": "Cognitive",
 "exposure": "E0",
 "rationale": "Human interaction required"
}
```

E1 – Direct LLM

```
{
 "esco_skill": "translating and interpreting",
 "esco_code": "S4.8.3",
 "dimension": "Cognitive",
 "exposure": "E1",
 "rationale": "Core LLM capability"
}
```

E2 – LLM + Tools

```
{
 "esco_skill": "programming computer systems",
 "esco_code": "S3.4.1",
 "dimension": "Cognitive",
 "exposure": "E2",
 "rationale": "Needs IDE integration"
}
```

E3 – Vision AI

```
{
 "esco_skill": "monitoring quality of products",
 "esco_code": "S1.4.1",
 "dimension": "Cognitive",
 "exposure": "E3",
 "rationale": "Visual inspection required"
}
```

296 ESCO skills classified: 138 manual (all E0), 158 cognitive (37 E0, 41 E1, 66 E2, 14 E3)

▸▸ Back

Appendix: Manual Skills Revaluation under AI

AI reduces $\omega_{jC} \Rightarrow$ after renormalization, ω_{jM} rises

Manual skills gaining weight:

- *using precision hand tools*
- *assembling mechanical products*
- *operating machinery*
- *measuring dimensions*
- *testing mechanical systems*

All classified (M, E0) $\Rightarrow$ immune to AI

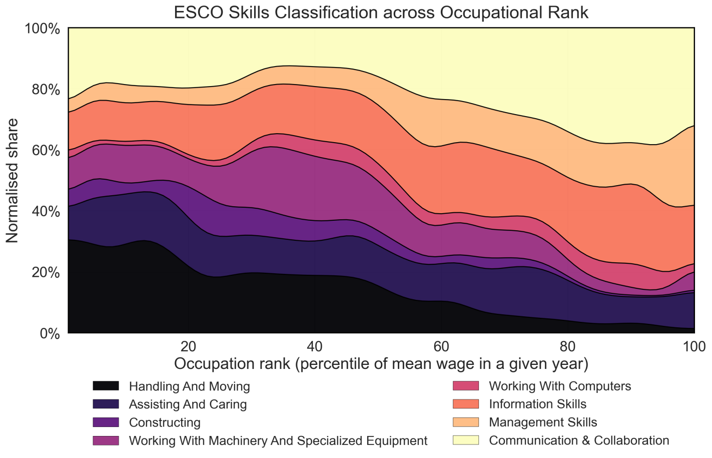
▶▶ Back

Example: Balanced occupation

	ω_{jC}	ω_{jM}
Pre-AI	0.57	0.43
Post-AI ($\alpha_{jC} = 0.4$)	0.57×0.6	0.43
Renormalized	0.44	**0.56**

Manual tasks: 43% $\rightarrow$ **56%** of output weight

The Skill Content of the Occupational Job Ladder



- Occupations where abstract (manual) skills are important at the top (bottom) of job ladder

Back

Sample Selection Details

1. exclude part-time workers: monthly hours < 130/weekly hours < 30
2. exclude monthly wage less than minimum wage
3. exclude primary sector/public administration/military (survey optional)
4. keep observations for workers observed at least two consecutive years and age $\in [20, 55]$

Back

Solving the Dynamic Program with Multidimensional Skills

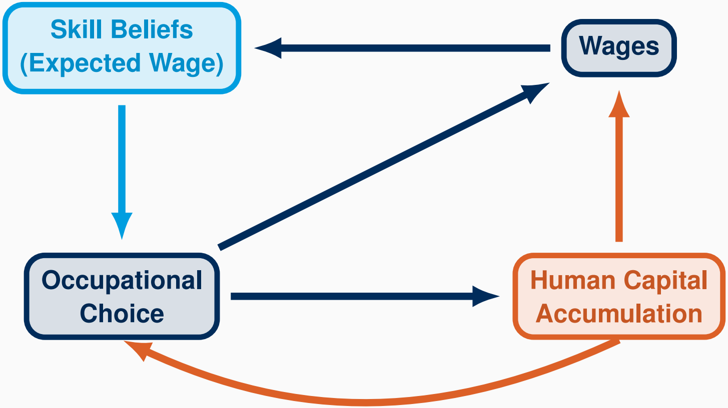
$$\widehat{EV}(\mathbf{x}_t) = \varrho \ln\left(\sum_{j=1}^{O} \exp\left\{\frac{1}{\varrho}\left[u\left(\mathbb{E}[w_{ijt} \mid \mathbf{x}_t]\right) + \beta\mathbb{E}\left[\widehat{EV}(\mathbf{x}_{t+1})\right]\right]\right\}\right)$$

$$= \varrho \log\left(\sum_{j=1}^{O} \exp\left(u\left(\mathbb{E}[w_{ijt} \mid \mathbf{x}_t]\right) + \beta\, \pi^{-K/2} \sum_{k_1}^{N_1} \cdots \sum_{1}^{N_K} \omega_1 \ldots \omega_K\, \widehat{EV}\left(\mathbf{x}_{t+1}\left(\nu_1, \ldots, \nu_K\right); \phi\right)\right)^{\frac{1}{\varrho}}\right)$$

- **Smolyak Approximation**: the value function $\widehat{EV}(\cdot)$ is projected on a sparse grid
- **GH Quadrature**: Gaussian–Hermite quadrature for correlation across skills

Back

Appendix: Dynamics in the Occupational Choice Model



▸ Back

EV is a Contraction: Blackwell Sufficiency Conditions

The Operator: let $\bar{u}_j(\boldsymbol{x}_t)$ be the deterministic utility and $\beta < 1$ the discount factor. The integrated Bellman is:

$$\Gamma(EV)(\boldsymbol{x}_t) = \varrho \ln \left(\sum_{j \in O} \exp \left\{ \frac{\bar{u}_j(\boldsymbol{x}_t) + \beta \mathbb{E}[EV(\boldsymbol{x}_{t+1}) \mid \boldsymbol{x}_t, j]}{\varrho} \right\} \right)$$

2. Monotonicity

- Assume $EV_1(\cdot) \leq EV_2(\cdot)$. Since $\exp(\cdot)$ and $\ln(\sum \cdot)$ are strictly increasing:
- $\Rightarrow \Gamma(EV_1)(\boldsymbol{x}_t) \leq \Gamma(EV_2)(\boldsymbol{x}_t)$ **(Satisfied)**

3. Discounting

- Consider $EV(\cdot) + c$ with $c \geq 0$. The constant shifts the exponent:

$$\Gamma(EV + c) = \varrho \ln \left(\sum \exp \left\{ \frac{\dots + \beta(\mathbb{E}[EV] + c)}{\varrho} \right\} \right)$$
$$= \varrho \ln \left(e^{\frac{\beta c}{\varrho}} \sum \exp \{ \dots \} \right) = \Gamma(EV) + \beta c$$

- Since $\beta < 1$, the discounting condition holds. **(Satisfied)**

Back

Appendix: Optimal Time Allocation Under Symmetric Beliefs

Firms maximize **expected output** under **symmetric beliefs** $\tilde{\xi}_{it}$, subject to $\sum_k \ell_{ijkt} = 1$:

$$\max_{\{\ell_{ijkt}\}} \mathbb{E}_{it}\left[Y_{ijt}\right], \qquad Y_{ijt} = e^{\varepsilon_{ijt}} \prod_{k=1}^{K} \left(\frac{Y_{ijkt}}{\omega_{jk}}\right)^{\omega_{jk}}$$

Plugging in the task technology $Y_{ijkt} = \ell_{ijkt}\,\theta_{ik}^{\lambda_{jk}^0} H_{jk}(\tau_{ikt})$ and collecting terms in ℓ:

$$\mathbb{E}_{it}\left[Y_{ijt}\right] = \underbrace{\mathbb{E}_{it}\Big[e^{\varepsilon_{ijt}} \textstyle\prod_k \left(\theta_{ik}^{\lambda_{jk}^0} H_{jk}(\tau_{ikt})\right)^{\omega_{jk}}\Big]}_{\text{independent of } \ell_{ijkt}} \cdot \prod_{k=1}^{K} \left(\frac{\ell_{ijkt}}{\omega_{jk}}\right)^{\omega_{jk}}$$

FOC $\omega_{jk}/\ell_{ijkt} = \mu$ with $\sum_k \omega_{jk} = \sum_k \ell_{ijkt} = 1 \Rightarrow$ allocation **independent of worker skills**:

$$\ell^*_{ijkt} = \omega_{jk}$$

Back

Appendix: Estimation Details

Stage 1: EM Algorithm

Wage likelihood integrates out latent skills ξ_i:

$$L_i^w(w_i; \boldsymbol{\Gamma}) = \int \left[\prod_t f_{\varepsilon j}\Big(w_{ijt} - p_j - \boldsymbol{\lambda}_j' \boldsymbol{\xi} - h_j(\tau_{it}) \Big)^{d_{ijt}} \right] \phi(\boldsymbol{\xi}; \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g)\, d\boldsymbol{\xi}$$

- **E-step:** $Q_i(\boldsymbol{\Gamma} \mid \boldsymbol{\Gamma}^{(q)}) = \mathbb{E}\Big[\log p(w_i, \boldsymbol{\xi}_i; \boldsymbol{\Gamma}) \,\Big|\, w_i; \boldsymbol{\Gamma}^{(q)} \Big]$
- **M-step:** $\boldsymbol{\Gamma}^{(q+1)} = \arg\max_{\boldsymbol{\Gamma}} \sum_i Q_i(\boldsymbol{\Gamma} \mid \boldsymbol{\Gamma}^{(q)})$ — closed-form updates

Stage 2: Nested Fixed-Point (NFXP)

- **Inner loop:** Solve for integrated value function $W(\boldsymbol{x}_t)$:

$$W(\boldsymbol{x}_t) = \varrho \log \sum_{j \in \mathcal{O}} \exp\left\{ \frac{\nu_j(\boldsymbol{x}_t)}{\varrho} \right\}, \quad \nu_j = r_j + \eta_{jg} - \mathcal{C}_j + \beta\, \mathbb{E}_t[W(\boldsymbol{x}_{t+1})]$$

- **Outer loop:** Update $\boldsymbol{\Delta}$ via MLE with multinomial logit CCPs: $\mathbb{P}(j \mid \boldsymbol{x}_t) = \exp\{\nu_j / \varrho\} / \sum_m \exp\{\nu_m / \varrho\}$

Back

Appendix: Separable Likelihood

Starting from the joint likelihood, apply the chain rule:

$$p(w_i, d_i, \xi \mid g) = \phi(\xi; \mu_g, \Sigma_g) \prod_t p(w_{it} \mid d_{it}, \xi, \tau_{it}; \Gamma) \prod_t p(d_{it} \mid \mathcal{H}_{it}; \Gamma, \Delta)$$

Key: choices depend on history only through beliefs $(\tilde{\xi}_{it}, \tilde{\Sigma}_{it}, \tau_{it}, d_{it-1})$, which are *deterministic* functions of observables — so $\prod_t p(d_{it} \mid \mathcal{H}_{it})$ does not depend on ξ and factors out of the integral:

$$\begin{aligned} L_i(w_i, d_i; \Gamma, \Delta) &= \int \phi(\xi) \prod_t p(w_{it} \mid \xi; \Gamma) \prod_t p(d_{it} \mid \mathcal{H}_{it}; \Gamma, \Delta)\, d\xi \\ &= \underbrace{\prod_t p(d_{it} \mid \mathcal{H}_{it}; \Gamma, \Delta)}_{L_i^c(d_i; \Gamma, \Delta)} \cdot \underbrace{\int \phi(\xi) \prod_t p(w_{it} \mid \xi; \Gamma)\, d\xi}_{L_i^w(w_i; \Gamma)} \end{aligned}$$

$\Rightarrow$ Wages and choices can be estimated **sequentially**

Back

Appendix: Identification of the Wage Equation under Learning

Potential wage in occupation j: $W_{it}(j) = p_j + h_j(\boldsymbol{\tau}_{it}) + \boldsymbol{\lambda}_j'\boldsymbol{\xi}_i + \varepsilon_{ijt}$

Pure learning (Bunting 2024; Arcidiacono et al. 2025): workers sort on beliefs formed from the observed history $\mathcal{H}_{it} = (g_i,\, D_i^{t-1},\, W_i^{t-1},\, \boldsymbol{\tau}_i^t)$:

$$D_{it} \perp\!\!\!\perp \boldsymbol{\xi}_i \mid \mathcal{H}_{it} \qquad \Longrightarrow \qquad \text{selection on } \underline{\text{observables}}$$

Control function — history is a sufficient control:

$$\mathbb{E}[W_{it}(j) \mid \mathcal{H}_{it}, g_i, D_{it} = j] = p_j + h_j(\boldsymbol{\tau}_{it}) + \boldsymbol{\lambda}_j' \underbrace{\mathbb{E}[\boldsymbol{\xi}_i \mid \mathcal{H}_{it}, g_i]}_{\text{posterior mean}}$$

Normalizations fix the latent space:

- anchor loadings $\boldsymbol{\lambda}_{a_k} = \boldsymbol{e}_k$ — **rotation**
- weighted zero mean $\sum_g \pi_g \boldsymbol{\mu}_g = \mathbf{0}$ — **location**
- base experience $h_j(\boldsymbol{\tau}^0) = 0$ — **price vs. return** split

$\Rightarrow \boldsymbol{\Gamma} = \{(p_j, \boldsymbol{\lambda}_j, \boldsymbol{\delta}_j, \varsigma^2_{\varepsilon j})_j, (\boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g)_g\}$ **point identified**: CCPs reweight (IPW) the Gaussian wage path — means $\rightarrow$ prices $+$ group means; covariances $\rightarrow \boldsymbol{\Sigma}_g$ and loadings

▸ Back ⏩ Dynamic Choice Parameters

Appendix: Identification of the Dynamic Choice Parameters

Stage 1 $\Rightarrow$ expected wages $\tilde{w}_j(x)$ and the belief law of motion. **Rationality** (entry beliefs $=$ true skill distribution) $\Rightarrow$ exact belief sequence $\Rightarrow$ **CCPs** $P_j(x)$ identified from choices and states.

Hotz–Miller (1993) inversion — log-odds recover value-index differences, and the ex-ante value is expressed through <u>any</u> single choice:

$$\varrho \log \frac{P_j(x)}{P_m(x)} = \nu_j(x) - \nu_m(x), \qquad W(x) = \nu_j(x) - \varrho \log P_j(x)$$

Arcidiacono–Miller (2011) — iterating the inversion along the path that keeps choosing j, the value index is affine in $(\boldsymbol{\eta}, \kappa_0)$, with the continuation term known:

$$\nu_j(x) = \underbrace{\tilde{w}_j(x) + \beta \Psi_j(x)}_{\text{known: Stage 1 \& CCPs}} + \frac{\eta_{jg}}{1-\beta} - \kappa_0 \mathbf{1}_{\{j \neq d\}}$$

$$\Psi_j(x) = \mathbb{E}\Big[\textstyle\sum_{s \geq 1} \beta^{s-1} \big(\tilde{w}_j(x_{t+s}) - \varrho \log P_j(x_{t+s}) \big) \;\Big|\; x_t = x,\ d_{t+s-1} = j \ \ \forall s \geq 1 \Big]$$

Two comparisons separated by whether the choice is the origin ($\eta_{0g} = 0$; $\varrho = 0.271$ calibrated):

$$\eta_{jg} - \eta_{mg} = (1-\beta)\Big[\varrho \log \frac{P_j}{P_m} - (\tilde{w}_j - \tilde{w}_m) - \beta(\Psi_j - \Psi_m) \Big] \qquad (j, m \neq d)$$

$$\kappa_0 = (\tilde{w}_j - \tilde{w}_d) + \beta(\Psi_j - \Psi_d) + \frac{\eta_{jg} - \eta_{dg}}{1-\beta} - \varrho \log \frac{P_j}{P_d} \qquad (j \text{ vs. stay } d)$$

▸ Back ▸▸ Wage Equation

Appendix: Two-Stage Estimation Strategy

Likelihood is separable into wages and choices Why?:

$$L_i(w_i, d_i; \Gamma, \Delta) = \underbrace{L_i^w(w_i; \Gamma)}_{\text{wages}} \cdot \underbrace{L_i^c(d_i; \Gamma, \Delta)}_{\text{choices}}$$

▶▶ Estimation Details

Appendix: Two-Stage Estimation Strategy

Likelihood is separable into wages and choices Why?:

$$L_i(w_i, d_i; \boldsymbol{\Gamma}, \boldsymbol{\Delta}) = \underbrace{L_i^w(w_i; \boldsymbol{\Gamma})}_{\text{wages}} \cdot \underbrace{L_i^c(d_i; \boldsymbol{\Gamma}, \boldsymbol{\Delta})}_{\text{choices}}$$

Stage 1 — EM Algorithm: estimate wage parameters $\boldsymbol{\Gamma} = \{(\boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g)_{g \in \mathcal{G}}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \varsigma, p_j\}$

$$\widehat{\boldsymbol{\Gamma}} = \arg\max_{\boldsymbol{\Gamma}} \sum_{i=1}^{N} \log L_i^w(w_i; \boldsymbol{\Gamma})$$

- EM iterates E-step (posterior over ξ_i) and M-step (closed-form updates)

Estimation Details

Appendix: Two-Stage Estimation Strategy

Likelihood is separable into wages and choices Why?:

$$L_i(w_i, d_i; \boldsymbol{\Gamma}, \boldsymbol{\Delta}) = \underbrace{L_i^w(w_i; \boldsymbol{\Gamma})}_{\text{wages}} \cdot \underbrace{L_i^c(d_i; \boldsymbol{\Gamma}, \boldsymbol{\Delta})}_{\text{choices}}$$

Stage 1 — EM Algorithm: estimate wage parameters $\boldsymbol{\Gamma} = \{(\boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g)_{g \in \mathcal{G}}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \varsigma, p_j\}$

$$\widehat{\boldsymbol{\Gamma}} = \arg\max_{\boldsymbol{\Gamma}} \sum_{i=1}^{N} \log L_i^w(w_i; \boldsymbol{\Gamma})$$

- EM iterates E-step (posterior over ξ_i) and M-step (closed-form updates)

Stage 2 — NFXP: estimate $\boldsymbol{\Delta} = \{\boldsymbol{\eta}, \kappa_0, \kappa_1\}$ given $\widehat{\boldsymbol{\Gamma}}$ (ϱ calibrated)

$$\widehat{\boldsymbol{\Delta}} = \arg\max_{\boldsymbol{\Delta}} \sum_{i=1}^{N} \log L_i^c\left(d_i; \widehat{\boldsymbol{\Gamma}}, \boldsymbol{\Delta}\right)$$

- Inner loop solves the DP; outer loop maximizes choice likelihood

▶▶ Estimation Details

Back