Multidimensional Skills, Sorting, and Human Capital Accumulation

Jeremy Lise Fabien Postel-Vinay

Jordan Peeples

February 15, 2021

Motivation

- A traditional way of studying wage and employment inequality relies on worker skills being one dimensional

Motivation

- A traditional way of studying wage and employment inequality relies on worker skills being one dimensional
- Few quantitative modeling tools exist that use information on multidimensional worker skills and job skill requirements available in the data in a description of labor market equilibrium

Motivation

- A traditional way of studying wage and employment inequality relies on worker skills being one dimensional
- Few quantitative modeling tools exist that use information on multidimensional worker skills and job skill requirements available in the data in a description of labor market equilibrium
    - Sanders & Taber (2012) use a theoretical model of job search and investment in multidimensional skills, but no quantitative analysis

Motivation

- A traditional way of studying wage and employment inequality relies on worker skills being one dimensional
- Few quantitative modeling tools exist that use information on multidimensional worker skills and job skill requirements available in the data in a description of labor market equilibrium
  - Sanders & Taber (2012) use a theoretical model of job search and investment in multidimensional skills, but no quantitative analysis
  - Yamaguchi (2012) uses a Roy-type model of task-specific human capital accumulation and occupation choices to find that higher task complexity results in higher wages and faster growth of relevant skills, but does not use job-shopping as a source of wage growth

Motivation

- A traditional way of studying wage and employment inequality relies on worker skills being one dimensional
- Few quantitative modeling tools exist that use information on multidimensional worker skills and job skill requirements available in the data in a description of labor market equilibrium
  - Sanders & Taber (2012) use a theoretical model of job search and investment in multidimensional skills, but no quantitative analysis
  - Yamaguchi (2012) uses a Roy-type model of task-specific human capital accumulation and occupation choices to find that higher task complexity results in higher wages and faster growth of relevant skills, but does not use job-shopping as a source of wage growth
- No empirical facts – just arguments to fill a gap in the literature and statistical inference used later in the paper

Empirical Question

- What are the origins and costs of mismatch along three dimensions of skills: cognitive, manual, and interpersonal, and the sources of variation in lifetime output?

Model Environment

- To account for the general and specialized skills workers have and how those interact with the technology of a firm, output is represented by a match function:

$$f(\boldsymbol{x}, \boldsymbol{y}) \text{ where } \boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^K \text{ and } \boldsymbol{y} \in \mathcal{Y} \subset \mathbb{R}^L,\ L \leq K$$

- Workers draw initial skills from an exogenous distribution
- Worker's skills gradually adjust to firm's technology:

$$\dot{\boldsymbol{x}} = g(\boldsymbol{x}, \boldsymbol{y})$$

- The market productivity and adjustment of specialized skills depend on the firm's technology, but general skills depend only on experience and have a common effect on output
- Overqualified workers produce more output than qualified workers
- Difference in firm skill requirement and worker skill reduces output

Skills Applied

- Data are from O*NET and NLSY97

Skills Applied

- Data are from O*NET and NLSY97
- Applied to the data, $\boldsymbol{x} = (x_C, x_M, x_I, x_T)$
  - x_C are cognitive skills
  - x_M are manual skills
  - x_I are interpersonal skills
  - x_T are general efficiency skills

Skills Applied

- Data are from O*NET and NLSY97
- Applied to the data, $\boldsymbol{x} = (x_C, x_M, x_I, x_T)$
  - x_C are cognitive skills
  - x_M are manual skills
  - x_I are interpersonal skills
  - x_T are general efficiency skills
- Initial skills determined by education, ASVAB, social skills in NLSY97

Skills Applied

- Data are from O*NET and NLSY97
- Applied to the data, $\boldsymbol{x} = (x_C, x_M, x_I, x_T)$
  - x_C are cognitive skills
  - x_M are manual skills
  - x_I are interpersonal skills
  - x_T are general efficiency skills
- Initial skills determined by education, ASVAB, social skills in NLSY97
- Similarly, $\boldsymbol{y} = (y_C, y_M, y_I)$

Skills Applied

- Data are from O*NET and NLSY97
- Applied to the data, $\boldsymbol{x} = (x_C, x_M, x_I, x_T)$
  - x_C are cognitive skills
  - x_M are manual skills
  - x_I are interpersonal skills
  - x_T are general efficiency skills
- Initial skills determined by education, ASVAB, social skills in NLSY97
- Similarly, $\boldsymbol{y} = (y_C, y_M, y_I)$
- Job skill requirements calculated through O*NET descriptions

Skills Applied

- Data are from O*NET and NLSY97
- Applied to the data, $\boldsymbol{x} = (x_C, x_M, x_I, x_T)$
  - x_C are cognitive skills
  - x_M are manual skills
  - x_I are interpersonal skills
  - x_T are general efficiency skills
- Initial skills determined by education, ASVAB, social skills in NLSY97
- Similarly, $\boldsymbol{y} = (y_C, y_M, y_I)$
- Job skill requirements calculated through O*NET descriptions
- Skills adjust linearly to the job skill requirements

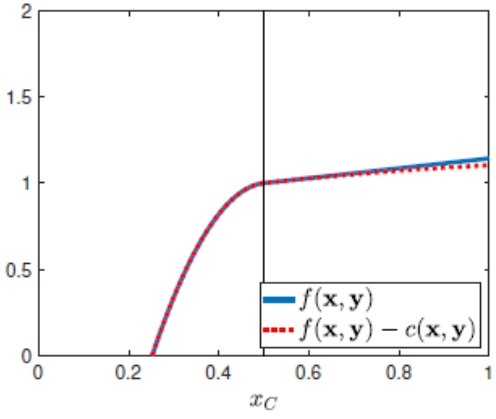
Skills Applied

- Data are from O*NET and NLSY97
- Applied to the data, $\boldsymbol{x} = (x_C, x_M, x_I, x_T)$
  - x_C are cognitive skills
  - x_M are manual skills
  - x_I are interpersonal skills
  - x_T are general efficiency skills
- Initial skills determined by education, ASVAB, social skills in NLSY97
- Similarly, $\boldsymbol{y} = (y_C, y_M, y_I)$
- Job skill requirements calculated through O*NET descriptions
- Skills adjust linearly to the job skill requirements
- Adjustment rates can differ between under and over-qualified

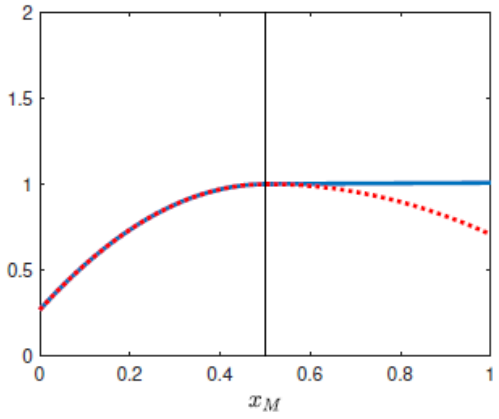
Skills Applied

- Data are from O*NET and NLSY97
- Applied to the data, $\boldsymbol{x} = (x_C, x_M, x_I, x_T)$
  - x_C are cognitive skills
  - x_M are manual skills
  - x_I are interpersonal skills
  - x_T are general efficiency skills
- Initial skills determined by education, ASVAB, social skills in NLSY97
- Similarly, $\boldsymbol{y} = (y_C, y_M, y_I)$
- Job skill requirements calculated through O*NET descriptions
- Skills adjust linearly to the job skill requirements
- Adjustment rates can differ between under and over-qualified
- Worker's specialized skills will adjust to job requirements, but general skills simply grow at constant rate

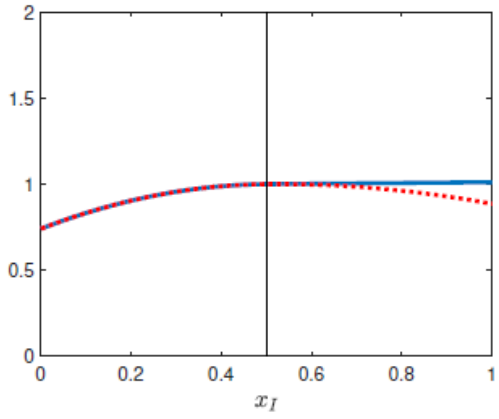
Production Function



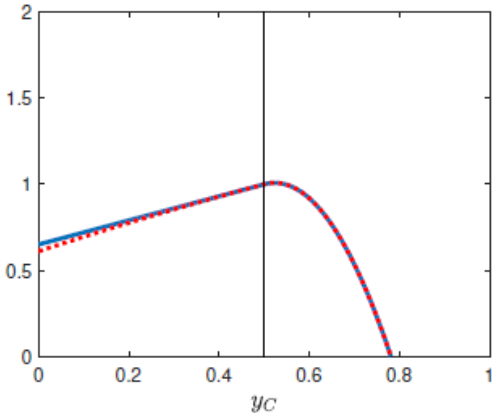
(a) Cognitive skills



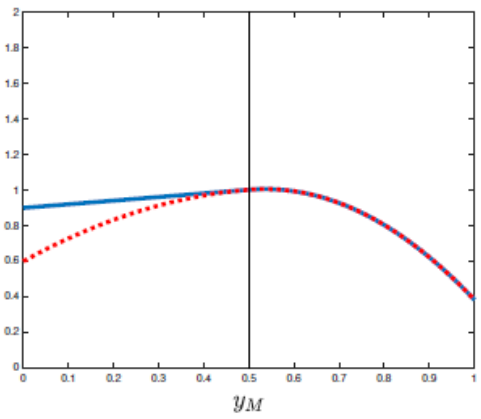
(b) Manual skills



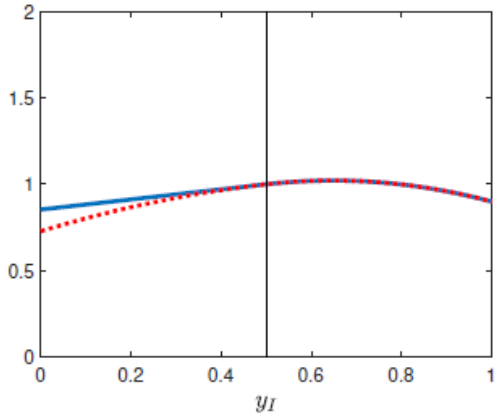
(c) Interpersonal skills



(d) Cognitive requirement



(e) Manual requirement



(f) Interpersonal requirement

Figure 1: The production function

Skill Adjustment

Table 3: Effect of quality and duration of first job on quality of second job

	(1) $\tilde{y}_C^+$	(2) $\tilde{y}_M^+$	(3) $\tilde{y}_I^+$	(4) $\tilde{y}_C^+$	(5) $\tilde{y}_M^+$	(6) $\tilde{y}_I^+$
x_{C0}	0.650	-0.300	0.460	0.659	-0.303	0.472
	(0.062)	(0.074)	(0.061)	(0.062)	(0.074)	(0.061)
x_{M0}	-0.117	0.687	-0.409	-0.124	0.677	-0.401
	(0.062)	(0.074)	(0.061)	(0.063)	(0.075)	(0.062)
x_{I0}	0.054	0.013	0.395	0.062	0.032	0.385
	(0.065)	(0.077)	(0.064)	(0.065)	(0.078)	(0.064)
$\max\{\tilde{y}_C - x_{C0}, 0\}^2$	3.044	0.998	1.102	3.321	0.932	1.379
	(0.694)	(0.827)	(0.686)	(0.696)	(0.836)	(0.690)
$\min\{\tilde{y}_C - x_{C0}, 0\}^2$	-0.677	-0.164	-0.096	-0.678	-0.168	-0.098
	(0.106)	(0.126)	(0.104)	(0.105)	(0.126)	(0.104)
$\max\{\tilde{y}_M - x_{M0}, 0\}^2$	-0.171	0.682	-0.450	-0.230	0.630	-0.484
	(0.227)	(0.270)	(0.224)	(0.228)	(0.274)	(0.226)
$\min\{\tilde{y}_M - x_{M0}, 0\}^2$	0.226	-0.420	0.190	0.213	-0.431	0.178
	(0.123)	(0.146)	(0.121)	(0.123)	(0.148)	(0.122)
$\max\{\tilde{y}_I - x_{I0}, 0\}^2$	-0.049	0.011	0.980	-0.058	0.008	0.981
	(0.312)	(0.371)	(0.308)	(0.312)	(0.375)	(0.309)
$\min\{\tilde{y}_I - x_{I0}, 0\}^2$	0.104	0.026	-0.399	0.121	0.019	-0.381
	(0.109)	(0.129)	(0.107)	(0.109)	(0.130)	(0.108)
duration	0.014	-0.001	0.017	0.016	-0.001	0.018
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
duration$\times(\tilde{y}_C - x_{C0})$	0.050	-0.038	0.035	0.050	-0.036	0.036
	(0.020)	(0.023)	(0.019)	(0.020)	(0.024)	(0.020)
duration$\times(\tilde{y}_M - x_{M0})$	-0.003	0.078	-0.025	-0.006	0.078	-0.028
	(0.016)	(0.019)	(0.016)	(0.016)	(0.019)	(0.016)
duration$\times(\tilde{y}_I - x_{I0})$	0.002	-0.001	0.031	0.002	0.001	0.029
	(0.013)	(0.015)	(0.012)	(0.013)	(0.015)	(0.012)
constant	0.091	0.327	0.161	0.083	0.332	0.159
	(0.040)	(0.047)	(0.039)	(0.043)	(0.052)	(0.043)
controls for occupation-specific wage decile				✓	✓	✓
N	528	528	528	528	528	528
adjusted R^2	0.376	0.276	0.497	0.385	0.274	0.502

Standard errors in parentheses

Matching and Unemployment

- On-the-job search model

Matching and Unemployment

- On-the-job search model
- Workers can be matched to a firm or unemployed

Matching and Unemployment

- On-the-job search model
- Workers can be matched to a firm or unemployed
- If matched, lose job at an exogenous rate

Matching and Unemployment

- On-the-job search model
- Workers can be matched to a firm or unemployed
- If matched, lose job at an exogenous rate
- Both employed and unemployed workers receive job offers from a fixed sampling distribution at different rates

Matching and Unemployment

- On-the-job search model
- Workers can be matched to a firm or unemployed
- If matched, lose job at an exogenous rate
- Both employed and unemployed workers receive job offers from a fixed sampling distribution at different rates
- Workers can exit the market at an exogenous rate

Worker Utility

- Workers have linear utility in wages and disutility of working

$$w - c(\boldsymbol{x}, \boldsymbol{y})$$

- Disutility depends on the type of match and only occurs when the worker is overqualified
- Type $\boldsymbol{x}$ unemployed worker receives flow utility $b(\boldsymbol{x})$ (home production)
- Unemployment income depends on general skills only

Match Values

- Total private value of a match $P(\boldsymbol{x}, \boldsymbol{y})$
- Value of unemployment $U(\boldsymbol{x})$
- Value of wage contract W
- Worker's share of surplus $\frac{W - U(\boldsymbol{x})}{P(\boldsymbol{x},\boldsymbol{y}) - U(\boldsymbol{x})}$

Competition between Employers

- Bertrand competition between current employer and potential employer

Competition between Employers

- Bertrand competition between current employer and potential employer
- New wage contract worth

$$W' = min\{P(\boldsymbol{x}, \boldsymbol{y}), max\{P(\boldsymbol{x}, \boldsymbol{y'}), W\}\}$$

Competition between Employers

- Bertrand competition between current employer and potential employer
- New wage contract worth

$$W' = min\{P(\boldsymbol{x}, \boldsymbol{y}), max\{P(\boldsymbol{x}, \boldsymbol{y'}), W\}\}$$

- Worker's renegotiated share of match surplus:

$$\sigma(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{y}') = \frac{P(\boldsymbol{x},\boldsymbol{y}')-U(\boldsymbol{x})}{P(\boldsymbol{x},\boldsymbol{y})-U(\boldsymbol{x})} \in [0, 1]$$

Competition between Employers

- Bertrand competition between current employer and potential employer
- New wage contract worth

$$W' = min\{P(\boldsymbol{x}, \boldsymbol{y}), max\{P(\boldsymbol{x}, \boldsymbol{y'}), W\}\}$$

- Worker's renegotiated share of match surplus:

$$\sigma(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{y}') = \frac{P(\boldsymbol{x},\boldsymbol{y}')-U(\boldsymbol{x})}{P(\boldsymbol{x},\boldsymbol{y})-U(\boldsymbol{x})} \in [0, 1]$$

- The share of surplus transferred to the worker from a negotiation remains constant between negotiations and only affects time profile of wage payments and timing of renegotiation

Competition between Employers

- Bertrand competition between current employer and potential employer
- New wage contract worth

$$W' = min\{P(\boldsymbol{x}, \boldsymbol{y}), max\{P(\boldsymbol{x}, \boldsymbol{y'}), W\}\}$$

- Worker's renegotiated share of match surplus:

$$\sigma(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{y}') = \frac{P(\boldsymbol{x},\boldsymbol{y}')-U(\boldsymbol{x})}{P(\boldsymbol{x},\boldsymbol{y})-U(\boldsymbol{x})} \in [0, 1]$$

- The share of surplus transferred to the worker from a negotiation remains constant between negotiations and only affects time profile of wage payments and timing of renegotiation
- Implies that the rate at which workers collect offers does not affect the private value of a match

Value Functions and Wage Equations

- The value function of the total private match between a worker and employer depends on:
  - total output
  - disutility from work
  - expected value of being unemployed
  - amount skill adjustments influence private value

Value Functions and Wage Equations

- The value function of the total private match between a worker and employer depends on:
  - total output
  - disutility from work
  - expected value of being unemployed
  - amount skill adjustments influence private value
- The value of unemployment is a stream of income plus the value associated with the speed of losses in certain skills (skill adjustment towards 0)
  - The private value of unemployment is independent of the frequency at which offers arise

Value Functions and Wage Equations

- The value function of the total private match between a worker and employer depends on:
  - total output
  - disutility from work
  - expected value of being unemployed
  - amount skill adjustments influence private value
- The value of unemployment is a stream of income plus the value associated with the speed of losses in certain skills (skill adjustment towards 0)
  - The private value of unemployment is independent of the frequency at which offers arise
- Wage equation is determined by:
  - static sharing of match surplus flow $(\sigma f(\boldsymbol{x}, \boldsymbol{y}) + (1 - \sigma)[b(\boldsymbol{x}) + c(\boldsymbol{x}, \boldsymbol{y})])$
  - value of future outside job offers
  - subtracting off what the worker gains in skills from the job as opposed to the skill diminishing effect of unemployment (as a fraction of the foregone opportunity, i.e. $1 - \sigma$)

Wage Evidence

Table 2: Occupation and Individual Fixed Effects

log wage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
x_{C0}	-0.036	0.567	-0.130	0.449	-0.144					
	(0.153)	(0.116)	(0.127)	(0.105)	(0.121)					
x_{M0}	0.014	-0.153	-0.150	-0.124	-0.065					
	(0.169)	(0.090)	(0.107)	(0.082)	(0.110)					
x_{I0}	0.232	0.311	0.033	0.276	0.105					
	(0.101)	(0.055)	(0.067)	(0.049)	(0.069)					
$\tilde{y}_C$	0.041					-0.532				
	(0.164)					(0.154)				
$\tilde{y}_M$	0.365					0.561				
	(0.171)					(0.154)				
$\tilde{y}_I$	0.395					0.388				
	(0.143)					(0.148)				
$x_{C0} \times \tilde{y}_C$	0.921		1.161		1.114	1.356		0.731		0.752
	(0.221)		(0.102)		(0.123)	(0.228)		(0.117)		(0.116)
$x_{M0} \times \tilde{y}_M$	-0.109		0.202		0.076	-0.279		0.279		0.170
	(0.254)		(0.091)		(0.110)	(0.237)		(0.085)		(0.088)
$x_{I0} \times \tilde{y}_I$	0.095		0.556		0.350	-0.144		0.304		0.183
	(0.233)		(0.101)		(0.124)	(0.257)		(0.109)		(0.112)
tenure	0.234	0.261	0.242	0.232	0.232	0.142	0.121	0.138	0.115	0.134
	(0.025)	(0.025)	(0.024)	(0.023)	(0.023)	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)
experience	0.269	0.289	0.264	0.257	0.244	0.335	0.363	0.334	0.343	0.322
	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
years of education	0.256	0.321	0.294	0.306	0.289					
	(0.081)	(0.085)	(0.080)	(0.075)	(0.073)					
constant	4.603	4.237	4.440	4.579	4.751	5.297	5.303	4.991	5.548	5.332
	(0.148)	(0.194)	(0.200)	(0.221)	(0.248)	(0.058)	(0.173)	(0.185)	(0.130)	(0.151)
occupation FE 1 digit		✓	✓				✓	✓		
occupation FE 3 digit				✓	✓				✓	✓
worker FE						✓	✓	✓	✓	✓
N	232,303	232,303	232,303	232,303	232,303	232,303	232,303	232,303	232,303	232,303
adjusted R^2	0.374	0.347	0.388	0.430	0.448	0.682	0.677	0.684	0.697	0.701

Standard errors clustered at the individual level.

Results: Parameter Estimates

Table 4: Parameter estimates

production function*										disutility of work*			un. inc.
α_T	α_C	α_M	α_I	α_{CC}	α_{MM}	α_{II}	κ_C^u	κ_M^u	κ_I^u	κ_C^o	κ_M^o	κ_I^o	b
137.5	140.3	64.4	92.4	195.6	10.7	15.4	5,165.3	984.5	337.6	54.1	409.6	171.9	137.5
(17.0)	(24.8)	(10.1)	(11.1)	(24.2)	(2.40)	(4.29)	(533.5)	(265.4)	(97.4)	(7.14)	(71.9)	(23.9)	(17.0)
							(45.7)	*(3.9)*	*(4.1)*	*(20.9)*	*(8.8)*	*(5.1)*	

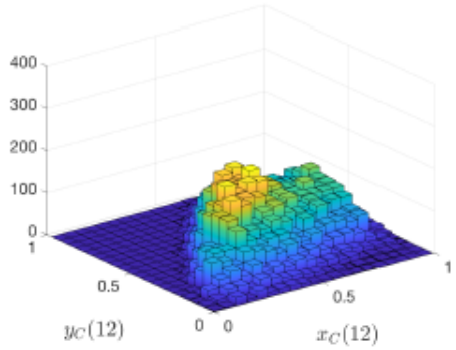
skill accumulation function**							general efficiency			
γ_C^u	γ_C^o	γ_M^u	γ_M^o	γ_I^u	γ_I^o	g	ζ_S	ζ_C	ζ_M	ζ_I
$7.7e-3$	$2.1e-3$	$3.4e-2$	$7.7e-3$	$1.0e-3$	$5.8e-7$	$2.3e-3$	$2.4e-2$	0.18	-0.17	0.20
(8.4e−4)	(5.3e−4)	(3.1e−3)	(6.4e−4)	(3.8e−3)	(1.4e−3)	(5.0e−4)	(.031)	(.501)	(.521)	(.261)
(7.54)	*(27.3)*	*(1.70)*	*(7.51)*	*(55.8)*	*(99,407)*					

sampling distribution***												transition rates		
ξ_C	ξ_M	ξ_I	ρ_{CM}	ρ_{CI}	ρ_{IM}	η_C^1	η_C^2	η_M^1	η_M^2	η_I^1	η_I^2	λ_0	λ_1	δ****
1.21	0.79	0.88	0.14	0.73	-0.44	1.22	2.86	2.15	2.76	0.93	2.96	0.39	0.16	$2.1e-2$
(.038)	(.038)	(.040)	(.019)	(.011)	(.019)	(.066)	(.095)	(.143)	(.117)	(.085)	(.124)	(.011)	(1.2e−3)	(3.3e−7)
			(0.13)	*(0.71)*	*(−0.42)*	*(0.30, 0.20)*		*(0.44, 0.20)*		*(0.24, 0.19)*				

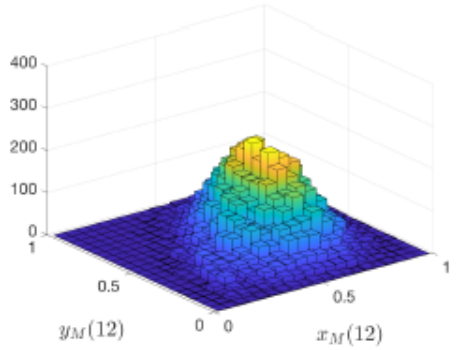
*percent surplus loss caused by deviating from output-maximizing match by 1 SD of Υ at mean x in italics;

** half-life in years in italics ; *** implied correlations and (means, standard deviations) in italics ; **** estimated in first step

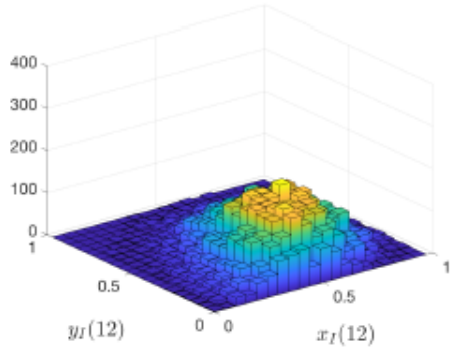
Skill Mismatch



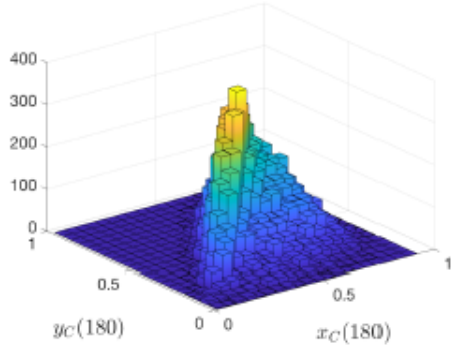
(a) Cognitive, 1 year of exp.



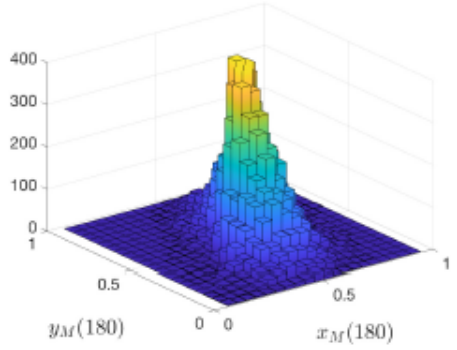
(b) Manual, 1 year of exp.



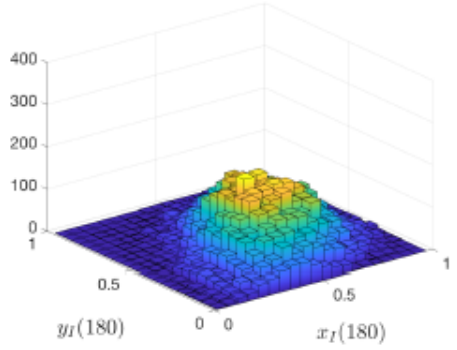
(c) Interpersonal, 1 year of exp.



(d) Cognitive, 15 years of exp.



(e) Manual, 15 years of exp.



(f) Interpersonal, 15 years of exp.

Figure 4: Sorting

Social Output

- Social output is the expected present discounted sum of future output produced by a worker

$$Q_{it} = \mathrm{E}\left[\int_t^{+\infty} \left(\ell_{is}\left[f\left(\mathrm{x}_{is}, \mathrm{y}_{is}\right) - c\left(\mathrm{x}_{is}, \mathrm{y}_{is}\right)\right] + \left(1 - \ell_{is}\right) b\left(\mathrm{x}_{is}\right)\right) e^{-(r+\mu)(s-t)} ds \mid \mathrm{x}_{i0}, \mathrm{ed}_i, \varepsilon_{0i}, \mathrm{x}_{it}, \ell_{it}, \mathrm{y}_{it}\right]$$

Decomposition of Output Variation – Multi-dimensional

Table 5: Decomposition of **Var** $\ln Q_{it}$

	Share of **Var** $\ln Q_{it}$ due to...			
	initial skills $\mathbf{x}_0$ (term 1)	shocks (term 2)	heterogeneity ε_0 (term 3)	education $\mid \mathbf{x}_0$ (term 4)
Whole sample	65.0%	16.4%	18.9%	0.0%
College +	17.2%	48.3%	35.5%	0.0%
Some college	27.5%	34.2%	38.9%	0.0%
Non-college	37.9%	22.4%	40.1%	0.0%

Level of experience: $t = 10$ years.

Decomposition of Output Variation – One-Dimensional

Table 7: Decomposition of **Var** $\ln Q_{it}$: one-dimensional model

	Share of **Var** $\ln Q_{it}$ due to...			
	initial skills x_0 (term 1)	shocks (term 2)	heterogeneity ε_0 (term 3)	education $\mid x_0$ (term 4)
Whole sample	32.5%	3.94%	60.4%	3.16%
College +	10.6%	6.01%	81.3%	2.08%
Some college	28.0%	3.94%	67.6%	0.43%
Non-college	24.6%	3.73%	71.6%	0.13%

Note: Level of experience: $t = 10$ years.

Conclusion

- Manual skills have moderate returns and adjust quickly
- Cognitive skills have much higher returns but are much slower to adjust
- Interpersonal skills have slightly higher returns than manual skills
- Cost of skill mismatch is highest for cognitive skills
  - Employing a worker who is under-qualified in cognitive skills is more than twice as costly in terms of lost surplus as employing an over-qualified worker
- A one dimensional model of skill underestimates the contribution of career shocks in the variation of lifetime output but overestimates the value of unobserved heterogeneity