1 Introduction

[The current plan has introducing Workers Do Jobs differently in the intro. We might also show triangle graphs.]

2 Education and Skill Acquisition

2.1 A model of technical change, education, and skills

2.2 Using the [Data] Data to estimate the model

2.3 Results: how education facilitates skill acquisition

3 The Educational Makeup of Occupations

3.1 Jobs' compositions may change in opposite directions

[This subsection will only exist to show triangles if we do not show triangles in the intro.]

3.2 A model of job choice by those workers from earlier

3.3

Each worker

Let's start with our usual skill choice problem. One dimension we want to leave open is individual heterogeneity in the skill constraint. That is, we think of relative weights θ^e as given by the worker's education group. The total budget constraint is normalized to 1, so we won't be normalizing the cost of any skill. Also, we're letting A vary by job.

For reference, J is a job, i and j are skill dimensions, n and m are workers, and various es are education levels.

4 Le Within

Given a job J, a worker with education e solves

$$
\begin{aligned}
&\max y(A \odot S, J) \\
&\text{s.t.} \quad \sum_i S_i \theta_i^e \leq \omega.
\end{aligned} \tag{1}
$$

The first order condition in dimension i is

$$
A_i y_i'(A \odot S^*, J) = \lambda \theta_i^e. \tag{2}
$$

Now, assume a CES production function

$$
y(A \odot S, J) = \left(\sum_i J_i (A_i S_i)^{\sigma_J} \right)^{\frac{1}{\sigma_J}}. \tag{3}
$$

Our first-order condition is now

$$
J_i S_i^{\sigma_J - 1} A_i^{\sigma_J} \left(\sum_k J_k (A_k S_k)^{\sigma_J} \right)^{\frac{1-\sigma_J}{\sigma_J}} = \lambda \theta_i^e. \tag{4}
$$

Dividing across different skills and raising to the power $\frac{1}{\sigma_J - 1}$, we have

$$
\frac{S_k}{S_i} = \left(\frac{J_k}{J_i} \right)^{-\frac{1}{\sigma_J - 1}} \left(\frac{A_k}{A_i} \right)^{-\frac{\sigma_J}{\sigma_J - 1}} \left(\frac{\theta_k^e}{\theta_i^e} \right)^{\frac{1}{\sigma_J - 1}} \tag{5}
$$

Let's multiply both sides with θ_k^e.

$$
\frac{\theta_k^e S_k}{S_i} = \left(\frac{J_k}{J_i} \right)^{-\frac{1}{\sigma_J - 1}} \left(\frac{A_k}{A_i} \right)^{-\frac{\sigma_J}{\sigma_J - 1}} \frac{(\theta_k^e)^{\frac{\sigma_J}{\sigma_J - 1}}}{(\theta_i^e)^{\frac{1}{\sigma_J - 1}}} \tag{6}
$$

Now, let's sum across is and use the fact that $\sum_i \theta_k^e S_k = 1$. Rearranging,

$$
S_i = \frac{J_i^{-\frac{1}{\sigma_J - 1}} A_i^{-\frac{\sigma_J}{\sigma_J - 1}} (\theta_i^e)^{\frac{1}{\sigma_J - 1}}}{\sum_k J_k^{-\frac{1}{\sigma_J - 1}} A_k^{-\frac{\sigma_J}{\sigma_J - 1}} (\theta_k^e)^{\frac{\sigma_J}{\sigma_J - 1}}} \tag{7}
$$

...which looks terrible. Instead of losing hope, take a derivative with respect to A_k where $k \neq i$.

$$
\frac{dS_i}{dA_k} = \frac{\sigma_J}{\sigma_J - 1} \frac{J_i^{-\frac{1}{\sigma_J - 1}} A_i^{-\frac{\sigma_J}{\sigma_J - 1}} (\theta_i^e)^{\frac{1}{\sigma_J - 1}}}{\sum_l J_l^{-\frac{1}{\sigma_J - 1}} A_l^{-\frac{\sigma_J}{\sigma_J - 1}} (\theta_l^e)^{\frac{\sigma_J}{\sigma_J - 1}}} \frac{J_k^{-\frac{1}{\sigma_J - 1}} A_k^{-\frac{\sigma_J}{\sigma_J - 1} - 1} (\theta_k^e)^{\frac{\sigma_J}{\sigma_J - 1}}}{\sum_l J_l^{-\frac{1}{\sigma_J - 1}} A_l^{-\frac{\sigma_J}{\sigma_J - 1}} (\theta_l^e)^{\frac{\sigma_J}{\sigma_J - 1}}} \tag{8}
$$

which simplifies, rather miraculously, to

$$\frac{dS_i}{d\ln A_k} = \frac{\sigma_J}{\sigma_J - 1} S_i S_k \theta_k^e. \tag{9}$$

Similarly, we have that

$$\frac{dS_i}{d\ln A_i} = \frac{\sigma_J}{\sigma_J - 1} S_i^2 \theta_i^e - \frac{\sigma_J}{\sigma_J - 1} S_i. \tag{10}$$

Thus, we have for each worker and each skill dimension i a first-order expansion of the form

$$\Delta \ln S_i = \frac{\sigma_J}{\sigma_J - 1}\left[\sum_k S_k \theta_k^e \Delta \ln A_k - \Delta \ln A_i\right] \tag{11}$$

Using the Slutsky equation on demand for 'effective' skills $A \odot S$, we can break this down into summed income and substitution effects:

$$\Delta \ln S_i = \underbrace{\sum_k S_k \theta_k^e \Delta \ln A_k}_{\text{Income Effects}} + \underbrace{\frac{1}{\sigma_J - 1}\sum_k S_k \theta_k^e \Delta \ln A_k - \frac{\sigma_J}{\sigma_J - 1}\Delta \ln A_i}_{\text{Substitution Effects}}. \tag{12}$$

Subtracting across different dimensions, we have

$$\Delta \ln S_i - \Delta \ln S_k = \frac{\sigma_J}{\sigma_J - 1}\left[\Delta \ln A_k - \Delta \ln A_i\right] \tag{13}$$

This time around, we're allowing A to vary at the job level; so we'll be writing A_{iJ} for the technological coefficient of skill i in job J. Of course, these equations have quite a bit of redundancy, as everything is HOD0. So, let's pick a skill - manual - and set $\Delta \ln A_{\text{manual}J} = 0$ in every job.

Unfortunately, we don't observe individual-level skill changes. But, we've assumed that workers within the occ-ed cell start off identical, so we can deal in averages. Thus in period t for workers in job J with education e in skill $i \neq$ manual we have

$$\Delta \overline{(\ln S_i)}_{Je} = \frac{\sigma_J}{\sigma_J - 1}\left[\sum_k \overline{(S_k)}_{Je} \theta_k^e \Delta \ln A_{kJ} - \Delta \ln A_{iJ}\right] \tag{14}$$

Using our normalization $\Delta \ln A_{manualJ} = 0$, we have

$$\Delta\overline{(\ln S_i)}_{Je} - \Delta\overline{(\ln S_{manual})}_{Je} = -\frac{\sigma_J}{\sigma_J - 1}\Delta \ln A_{iJ} =: -\pi_{iJ}. \qquad (15)$$

We are now in a position to produce one estimate of π_{iJt} for each education group. As these quantities are not theoretically distinct, though, we take a simple average.

$$\hat{\pi}_{iJ} := \frac{\sigma}{\sigma - 1}\Delta \ln A_{iJ} = \frac{1}{3}\sum_e \left[\Delta\overline{(\ln S_{manual})}_{Je} - \Delta\overline{(\ln S_i)}_{Je}\right] \qquad (16)$$

Now, we plug (16) into (11) and have

$$\Delta\overline{(\ln S_i)}_{Je} = \sum_k \overline{(S_k)}_{Je}\theta_k^e \hat{\pi}_{kJ} - \hat{\pi}_{iJ}. \qquad (17)$$

5 Le Between: Discrete Choice of the Proletariat

Unlike in the DOT paper, we want a theory of sorting across jobs other than indifference. To that end, we'll employ a discrete choice framework in which worker n in job J with wage w gets utility

$$\ln w + \xi_J + \eta_{Jn} \qquad (18)$$

where ξ_J is a parameter expressing a common preference for job J, and η_{Jn} expresses a combination of idiosyncratic preference for and productivity in J and is i.i.d Gumbel-distributed.

Thus, the probability job J is chosen by a worker of type e is equal to

$$\frac{w_e(J)e^{\xi_J}}{\sum_{J'} w_e(J')e^{\xi_{J'}}} \qquad (19)$$

Where do wages w_e come from? As workers solve the problem in (1), and each intermediate output J is priced by the market at some $P(J)$ somehow, $w_e(J) = y_e^*(J)P(J)$. Therefore the ratio of e workers in job J to those in J' is

$$\frac{q_e(J)}{q_e(J')} = \frac{P(J)y_e^*(J)e^{\xi_J}}{P(J')y_e^*(J')e^{\xi_{J'}}} \qquad (20)$$

Summing across jobs, if the total amount of workers of type e is q_e^{Total}, we get

$$q_e(J) = \frac{P(J)y_e^*(J)e^{\xi_J}}{\sum_{J'} P(J')y_e^*(J')e^{\xi_{J'}}} q_e^{Tot} \tag{21}$$

Instead of imposing assumptions and estimating the discrete choice model at hand, we take a different tack: comparing across types.

$$\frac{q_e(J)}{q_{e'}(J)} = \frac{y_e^*(J)}{y_{e'}^*(J)} \frac{q_e^{Tot}}{q_{e'}^{Tot}} \frac{\sum_{J'} P(J')y_{e'}^*(J')e^{\xi_{J'}}}{\sum_{J'} P(J')y_e^*(J')e^{\xi_{J'}}} \tag{22}$$

From here, we can exploit changes over time.

$$\Delta \ln \frac{q_e(J)}{q_{e'}(J)} = \sum_i \left(\frac{\partial \ln y_e^*(J)}{\partial \ln A_{iJ}} - \frac{\partial \ln y_{e'}^*(J)}{\partial \ln A_{iJ}} \right) \Delta \ln A_{iJ} \tag{23}$$

$$+\Delta \ln \left[\frac{q_e^{Tot}}{q_{e'}^{Tot}} \frac{\sum_{J'} P(J')y_{e'}^*(J')e^{\xi_{J'}}}{\sum_{J'} P(J')y_e^*(J')e^{\xi_{J'}}} \right] \tag{24}$$

Using the envelope theorem, and noticing that the last term is a constant across J, we arrive at

$$\Delta \ln \frac{q_e(J)}{q_{e'}(J)} = \sum_i \left(S_{iJe}\theta_i^e - S_{iJe'}\theta_i^{e'} \right) \Delta \ln A_{iJ} + const_{e,e',t} \tag{25}$$

For our purposes, we'll use the mean S_{iJe} within the educ-occ cell. Also substituting in π,

$$\Delta \ln \frac{q_e(J)}{q_{e'}(J)} = \beta_J \sum_i \left(\overline{(S_i)}_{Je}\theta_e^i - \overline{(S_i)}_{Je'}\theta_i^{e'} \right) \hat{\pi}_{iJ} + const_{e,e',t} \tag{26}$$

The interpretation here is that $\beta_J = \frac{\sigma_J - 1}{\sigma_J}$, which may or may not be interesting to know. A test of our model is that each $\sigma_J < 1$. More importantly, this will allow us to test our model's goodness of fit. Or something.