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 but the total constraint w as heterogeneous on the individual level and potentially correlated with education and job choice. 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.

1 Le Within

Given a job J, a worker with education e solves

$$\max y(A \circ S, J) \tag{1}$$

$$\text{s.t.} \quad \sum_i S_i \theta_i^e \leq w. \tag{2}$$

The first order condition in dimension i is

$$A_i y_i'(A \circ S^*, J) = \lambda \theta_i^e. \tag{3}$$

Now, assume a CES production function

$$y(A \circ S, J) = \left(\sum_i J_i (A_i S_i)^\sigma \right)^{\frac{1}{\sigma}}. \tag{4}$$

Our first-order condition is now

$$J_i S_i^{\sigma-1} A_i^\sigma \left(\sum_k J_k (A_k S_k)^\sigma \right)^{\frac{1-\sigma}{\sigma}} = \lambda \theta_i^e. \tag{5}$$

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

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

Let's multiply both sides with θ_i^e.

$$\frac{\theta_k^e S_k}{S_i} = \left(\frac{J_k}{J_i}\right)^{-\frac{1}{\sigma-1}} \left(\frac{A_k}{A_i}\right)^{-\frac{\sigma}{\sigma-1}} \left(\frac{\theta_k^e}{\theta_i^e}\right)^{\frac{1}{\sigma-1}} \tag{7}$$

Now, let's sum across is and use the fact that $\sum_i \theta_i^e S_i = w$. Rearranging,

$$S_i = \frac{J_i^{-\frac{1}{\sigma-1}} A_i^{-\frac{\sigma}{\sigma-1}} (\theta_i^e)^{\frac{1}{\sigma-1}}}{\sum_k J_k^{-\frac{1}{\sigma-1}} A_k^{-\frac{\sigma}{\sigma-1}} (\theta_k^e)^{\frac{\sigma}{\sigma-1}}} w \tag{8}$$

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

$$\frac{dS_i}{dA_k} = \frac{\sigma}{\sigma-1} \frac{J_i^{-\frac{1}{\sigma-1}} A_i^{-\frac{\sigma}{\sigma-1}} (\theta_i^e)^{\frac{1}{\sigma-1}}}{\sum_l J_l^{-\frac{1}{\sigma-1}} A_l^{-\frac{\sigma}{\sigma-1}} (\theta_l^e)^{\frac{\sigma}{\sigma-1}}} w \frac{J_k^{-\frac{1}{\sigma-1}} A_k^{-\frac{\sigma}{\sigma-1}-1} (\theta_k^e)^{\frac{\sigma}{\sigma-1}}}{\sum_l J_l^{-\frac{1}{\sigma-1}} A_l^{-\frac{\sigma}{\sigma-1}} (\theta_l^e)^{\frac{\sigma}{\sigma-1}}} \tag{9}$$

which simplifies, rather miraculously, to

$$\frac{dS_i}{d\ln A_k} = \frac{\sigma}{\sigma-1} \frac{S_i S_k \theta_k^e}{w}. \tag{10}$$

Similarly, we have that

$$\frac{dS_i}{d\ln A_i} = \frac{\sigma}{\sigma-1} \frac{S_i^2 \theta_i^e}{w} - \frac{\sigma}{\sigma-1} S_i. \tag{11}$$

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

$$\Delta \ln S_i = \frac{\sigma}{\sigma-1} \left[\sum_k \frac{S_k \theta_k^e}{w} \Delta \ln A_k - \Delta \ln A_i \right] \tag{12}$$

Notice this bakes in a few things we like - the Slutsky symmetry property, for instance. Now, there's just one problem - w is heterogeneous at the individual level, whereas we don't have panel data.

Subtracting across different dimensions, we have

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

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. So we can't actually compute $\Delta \ln S_i$. But, under the assumption that the distribution of w is the same within the occupation-education-job cell before and after the change, we can instead sum across workers within the occ-ed-job cell.

Thus in period t for workers in job J with education e in skill $i \neq$ manual we have

$$\Delta\overline{\ln S^e_{iJetn}} - \Delta\overline{\ln S^e_{\text{manual}eJtn}} = -\frac{\sigma}{\sigma - 1}(\Delta \ln A_{iJ})_t = -\pi_{iJt}. \tag{14}$$

As we now are in a position to produce one estimate of π_{iJt} for each education group, though, we take a simple average.

$$\hat{\pi}_{iJt} = \frac{1}{3}\sum_e \left[\Delta\overline{\ln S^e_{\text{manual}eJtn}} - \Delta\overline{\ln S^e_{iJetn}}\right] \tag{15}$$

Now, we plug (15) into (12) and have

$$\Delta\overline{\ln S^e_{iJetn}} + \hat{\pi}_{iJt} = \sum_k \frac{S_{kJ}\theta^e_k}{w}\hat{\pi}_{kJt}. \tag{16}$$

This must hold at the *individual* level, as the model predicts that $\frac{S_{kn}\theta^e_k}{w_n}$ is constant over individuals in the occupation-education cell. Therefore,

$$\frac{1}{N_{Jet}}\sum_n^{N_{Jet}} S_{kJetn}\theta^e_k = \frac{1}{N_{Jet}}\sum_n^{N_{Jet}} \frac{S_{kJetn}\theta^e_k}{w_n}w_n = \frac{S_k\theta^e_k}{w}\frac{1}{N_{Jet}}\sum_n^{N_{Jet}} w_n \tag{17}$$

so that using the fact that we've assumed $E[w_n] = 1$ in every job,

$$\theta^e_k\overline{S_{kJetn}} = \frac{1}{N_{Jet}}\sum_n^{N_{Jet}} S_{kn}\theta^e_k = \frac{S_k\theta^e_k}{w} \tag{18}$$

finally arriving at our regression equation

$$\Delta\overline{\ln S^e_{iJetn}} + \hat{\pi}_{iJt} = \sum_k \theta^e_k\overline{S_{kJetn}}\hat{\pi}_{kJt} \tag{19}$$

Concern: we're assuming not just that the average ability within each group is $\bar{w} = 1$ but that this is true of the average ability within each occ-educ pair. This is actually falsifiable with our data alone, as we can compute this within-occ average using observed Ss and our estimated θs. There are alternatives. We could iterate, solving for θs and computing individual ws, plugging them back in to (16) and so on. I know iterating thetas left a bad taste in our mouth last time around, but *this time is different...*

2 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 a worker n in job J with wage w_J gets utility

$$\ln w_J + \xi_J + \eta_{Jn} \tag{20}$$

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

$$\frac{w_e(J)e^{\xi_J}}{\sum_{J'} w_e(J')e^{\xi_{J'}}} \tag{21}$$

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

$$\frac{q_e(J)}{q_e(J')} = \frac{P(J)y_e^*(J)e^{\xi_J}}{P(J')y_e^*(J')e^{\xi_{J'}}} \tag{22}$$

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

$$q_e(J) = \frac{P(J)y_e^*(J)e^{\xi_J}}{\sum_{J'} P(J')y_e^*(J')e^{\xi_{J'}}} q_e^{Tot} \tag{23}$$

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

$$\frac{q_e(J)}{q_{e'}(J)} = \frac{y_e^*(J)}{y_{e'}^*(J)} \frac{q_e^{Tot}}{q_{e'}^{Tot}} \frac{\sum_{J'} P(J') y_{e'}^*(J') e^{\xi_{J'}}}{\sum_{J'} P(J') y_e^*(J') e^{\xi_{J'}}} \tag{24}$$

From here, we can exploit changes over time.

$$\Delta \ln \frac{q_e(J)}{q_{e'}(J)} = \sum_i \left(\frac{\partial \ln y_e^*(J)}{\partial \ln A_{iJ}} - \frac{\partial \ln y_{e'}^*(J)}{\partial \ln A_{iJ}} \right) \Delta \ln A_{iJ} \tag{25}$$

$$+\Delta \ln \left[\frac{q_e^{Tot}}{q_{e'}^{Tot}} \frac{\sum_{J'} P(J') y_{e'}^*(J') e^{\xi_{J'}}}{\sum_{J'} P(J') y_e^*(J') e^{\xi_{J'}}} \right] \tag{26}$$

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

$$\Delta \ln \frac{q_e(J)}{q_{e'}(J)} = \sum_i \left(\frac{S_{iJetn}\theta_i^e}{w_{en}} - \frac{S_{iJe'tn}\theta_i^{e'}}{w_{e'n}} \right) \Delta \ln A_{iJ} + const \tag{27}$$

For our purposes, we'll use the mean $S\theta/w$ within the educ-occ cell. Substituting in π,

$$\Delta \left[\ln \frac{q_e(J)}{q_{e'}(J)} \right]_t = \beta_J \sum_k \left(\theta_k^e \overline{S_{kJetn}} - \theta_k^{e'} \overline{S_{kJe'tn}} \right) \hat{\pi}_{kJt} + const_{e,e',t} \tag{28}$$

The interpretation here is that $\beta_J = \frac{\sigma-1}{\sigma}$, 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.