Appendix A

The following appendix briefly summarizes the basic estimators and relates them to the theoretical models presented in Section 2. Suppose the goal is to estimate the effect of leaving the training firm on wages for $t = 1,3,5$ periods of potential labor market experience of young workers after the end of the apprenticeship. For notational simplicity, dependence on observable characteristics other than a constant is suppressed. The estimation procedure itself is discussed in Appendix B. The equations estimated for OLS and OLSFE are

$$w_{it} = \mu + \delta_t D_{i0} + \varepsilon_{it}$$
$$w_{it} = \mu + \delta_t D_{i0} + \phi_j + \varepsilon_{it}$$

The resulting estimates are calculated as

$$\Rightarrow \hat{\delta}_t^{OLS} = \frac{\operatorname{cov}(w_{it}, D_{i0})}{\operatorname{var}(D_{i0})}$$
$$\Rightarrow \hat{\delta}_t^{OLSFE} = \frac{\sum (w_{it} - \overline{w}_{j(i)})(D_{i0} - \overline{D}_{j(i)0})}{\sum (D_{i0} - \overline{D}_{j(i)0})^2}.$$

If D_{i0} is instrumented by the fraction of other apprentices moving out of the training firm, the corresponding IV estimates are calculated from

$$\Rightarrow \hat{\delta}_t^{IV} = \frac{\operatorname{cov}(w_{it}, z_{ij(i)})}{\operatorname{cov}(D_{i0}, z_{ij(i)})}$$
$$\Rightarrow \hat{\delta}_t^{IVFE} = \frac{\sum (w_{it} - \overline{w}_{j(i)})(z_{ij(i)} - \overline{z}_{j(i)})}{\sum (D_{i0} - \overline{D}_{j(i)0})(z_{ij(i)} - \overline{z}_{j(i)})},$$

where the necessary assumptions are $\operatorname{cov}(a_i - \overline{a}_{j(i)}, z_{ij(i)}) = 0$ and $\operatorname{cov}(V_{i0}, z_{ij(i)}) = 0$ for IV and $\operatorname{cov}(a_i - \overline{a}_{j(i)}, z_{ij(i)} - \overline{z}_{j(i)}) = 0$ and $\operatorname{cov}(V_{i0} - \overline{V}_{j(i)}, z_{ij(i)} - \overline{z}_{j(i)}) = 0$ for IVFE.

The core mechanisms of the five theoretical approaches to job and wage mobility discussed in the text imply the following stylized wage equations:

Adverse Selection

$$\Rightarrow w_{it} = \gamma + a_i + \varepsilon_{it}$$

Initial Assignment

$$\Rightarrow w_{it} = \gamma + (a_i - \overline{a}_{j(i)}) + \overline{a}_{j(i)} + \varepsilon_{it}$$

Job Search

$$\Rightarrow w_{it} = \gamma + \delta_{It} D_{i0} + (\delta_{Vt} - \delta_{It}) V_{i0} + \varepsilon_{it}$$

Sequential Sorting

$$\Rightarrow w_{it} = \gamma + (a_i - \overline{a}_{j(i)}) + \overline{a}_{j(i)} + \varepsilon_{it}$$

Institutions/Rosen

$$\Rightarrow w_{it} = \gamma + (a_i - \overline{a}_{j(i)}) + \overline{a}_{j(i)} + \delta_{It} D_{i0} + \varepsilon_{it}.$$

Combining these data generating processes with the various statistical procedures, for each theory in turn one obtains the following estimates.

Adverse Selection

$$\Rightarrow \hat{\delta}_{It}^{OLS} = \frac{\operatorname{cov}(a_i, D_{i0})}{\operatorname{var}(D_{i0})} < 0$$

$$\Rightarrow \hat{\delta}_{It}^{OLSFE} = \frac{\operatorname{cov}(a_i - \overline{a}_{j(i)}, D_{i0} - \overline{D}_{j(i)})}{\operatorname{var}(D_{i0} - \overline{D}_{j(i)})} < 0$$

$$\Rightarrow \hat{\delta}_{It}^{IV} = 0$$

$$\Rightarrow \hat{\delta}_{It}^{IVFE} = 0$$

As mentioned in the text, if there is some initial sorting based on firms' retention rates, one would expect that IV is still negative (potentially more than OLS), but that IVFE is zero.

Initial Assignment

$$\Rightarrow \hat{\delta}_{It}^{OLS} = \frac{\operatorname{cov}(\overline{a}_{j(i)}, D_{i0})}{\operatorname{var}(D_{i0})} < 0$$

$$\Rightarrow \hat{\delta}_{It}^{OLSFE} = 0$$

$$\Rightarrow \hat{\delta}_{It}^{IV} = \frac{\operatorname{cov}(\overline{a}_{j(i)}, z_{ij(i)})}{\operatorname{cov}(D_{i0}, z_{ij(i)})} < 0$$

$$\Rightarrow \hat{\delta}_{It}^{IVFE} = 0$$

Under initial assignment, IV would be more negative than OLS since $\operatorname{cov}(\overline{a}_{j(i)}, z_{ij(i)}) = \operatorname{cov}(\overline{a}_{j(i)}, D_{i0})$ but $\operatorname{cov}(D_{i0}, z_{ij(i)}) < \operatorname{var}(D_{i0})$. The weaker the instrument, the larger will the difference be.

Job Search

$$\Rightarrow \hat{\delta}_{It}^{OLS} = \delta_{It} + (\delta_{Vt} - \delta_{It})\frac{\operatorname{cov}(V_{i0}, D_{i0})}{\operatorname{var}(D_{i0})}$$

$$\Rightarrow \hat{\delta}_{It}^{OLSFE} = \delta_{It} + (\delta_{Vt} - \delta_{It})\frac{\operatorname{cov}(V_{i0} - \overline{V}_{j(i)}, D_{i0} - \overline{D}_{j(i)})}{\operatorname{var}(D_{i0} - \overline{D}_{j(i)})}$$

$$\Rightarrow \hat{\delta}_{It}^{IV} = \delta_{It}$$

$$\Rightarrow \hat{\delta}_{It}^{IVFE} = \delta_{It}$$

The basic search model assumes individuals are equal but jobs differ. Thus, neither selection based on unobserved ability nor initial sorting matters. As discussed in the text, the model predicts that losses from an involuntary displacement are temporary.

Sequential Sorting

$$\Rightarrow \hat{\delta}_{It}^{OLS} = \frac{\text{cov}(a_i - \bar{a}_{j(i)}, D_{i0})}{\text{var}(D_{i0})} + \frac{\text{cov}(\bar{a}_{j(i)}, D_{i0})}{\text{var}(D_{i0})}$$

$$\Rightarrow \hat{\delta}_{It}^{OLSFE} = \frac{\text{cov}(a_i - \bar{a}_{j(i)}, D_{i0})}{\text{var}(D_{i0})}$$

$$\Rightarrow \hat{\delta}_{It}^{IV} = \frac{\text{cov}(a_i - \bar{a}_{j(i)}, z_{ij(i)})}{\text{cov}(D_{i0}, z_{ij(i)})} + \frac{\text{cov}(\bar{a}_{j(i)}, z_{ij(i)})}{\text{cov}(D_{i0}, z_{ij(i)})}$$

$$\Rightarrow \hat{\delta}_{It}^{IVFE} = 0$$

In the case of sequential sorting, the bias of OLS coming from within firm selection is ambiguous; on the one hand, good firms will separate from their least able workers, such that $\text{cov}(a_i - \bar{a}_{j(i)}, D_{0i}) < 0$. On the other hand, less desirable firms will separate from their better workers, implying $\text{cov}(a_i - \bar{a}_{j(i)}, D_{0i}) > 0$. The sign of the bias is indicated by OLSFE, which is net of the effect of initial assignment (and less negative than OLS). Since good workers are sorted out of firms with high average turnover rates (low average retention rates), $\text{cov}(a_i - \bar{a}_{j(i)}, z_{ij(i)}) > 0$. This implies that IV is less negative than OLS unless initial assignment is very strong (i.e., unless $\text{cov}(\bar{a}_{j(i)}, z_{ij(i)}) / \text{cov}(D_{i0}, z_{ij(i)})$ is very negative).

Institutions/Rosen

$$\Rightarrow \hat{\delta}_{It}^{OLS} = \delta_{It} + \frac{\text{cov}(a_i - \bar{a}_{j(i)}, D_{i0})}{\text{var}(D_{i0})} + \frac{\text{cov}(\bar{a}_{j(i)}, D_{i0})}{\text{var}(D_{i0})} < 0$$

$$\Rightarrow \hat{\delta}_{It}^{OLSFE} = \delta_{It} + \frac{\text{cov}(a_i - \bar{a}_{j(i)}, D_{i0})}{\text{var}(D_{i0})} < 0$$

$$\Rightarrow \hat{\delta}_{It}^{IV} = \delta_{It} + \frac{\text{cov}(\bar{a}_{j(i)}, z_{ij(i)})}{\text{cov}(D_{i0}, z_{ij(i)})}$$

$$\Rightarrow \hat{\delta}_{It}^{IVFE} = \delta_{It}$$

In the case that some larger firms have internal labor markets and provide sheltered and well-defined career paths, the loss of such an opportunity could imply permanent negative effects. However, if wages are rigid in the internal labor market these firms might fire the worst workers instead of lowering their wages, thus $\text{cov}(a_i - \bar{a}_{j(i)}, D_{i0}) < 0$. Similarly, these firms are more desirable, and they might attract the best workers, suggesting that $\text{cov}(\bar{a}_{j(i)}, D_{i0}) < 0$ and $\text{cov}(\bar{a}_{j(i)}, z_{ij(i)}) < 0$. Thus, OLSFE and IV are predicted to be both less negative than OLS (unless $\text{cov}(D_{i0}, z_{ij(i)})$ is very small). Under the simple sequential human capital accumulation, wages are assumed to be fully flexible and thus $\text{cov}(a_i - \bar{a}_{j(i)}, D_{i0}) = 0$. Similarly, the model has no predictions regarding the supply of career jobs between different firms, and thus $\text{cov}(\bar{a}_{j(i)}, D_{i0}) = 0$.