Pre-analysis plan: Using similar workers' labor market outcomes in different countries to decompose cross-country wage gaps

Michael A. Clemens ~ March 4, 2018

1 Rationale and goals

A recent research literature has challenged prior findings that workers experience very large gains in productivity when they migrate from poor to rich countries. It concludes that the large majority of cross-country income differences in fact arise from differences in portable human capital, but relies on methods that can produce a downward bias in the estimated gain to migration. The goal of this research is to produce superior estimates of the fraction of intercountry real income differences that arise from portable human capital—and thereby superior estimates of the gain to migration—using alternative methods less subject to bias.

2 Outcomes of interest and analysis methods

The study will test what portion of the occupation-specific wage gap between rich and poor countries can be explained by the returns to observed and unobserved human capital. It will do this using measures of cognitive and other skills—both those typically observed at the individual level and typically unobserved—required for each occupation.

Straightforward parametric empirical tests would resemble the linear regressions $\frac{w_{\text{des}}-w_{\text{ori}}}{y_{\text{des}}-y_{\text{ori}}} = \beta_1 + \beta_2 s + \varepsilon$ for country dyads $\left(\text{alternatively } \frac{w_{\text{des}}/w_{\text{ori}}}{y_{\text{des}}/y_{\text{ori}}} = \beta_1 + \beta_2 s + \varepsilon\right)$, where 'des' is a high-income migrant-destination country such as the US and 'ori' is a lower-income non-US migrant-origin country; or $\ln w = \alpha + \beta_0 s + \beta_1 \ln y + \beta_2 (s \times \ln y) + \varepsilon$ across all countries, for occupation-by-country data, where w is the wage, y is national income per capita, s is occupational skill requirement, ε is an error term, and α and vector β are to be estimated. But nonparametric versions of these tests would be preferable if statistical power allows, given that available occupational skill indices are qualitatively scored and linearity is a strong assumption. Theory also suggests that occupational wage gaps should be larger for occupations with non-traded outputs so s could be recast as a measure of output tradability. The estimates will be closer to population parameter estimates if weighted by estimates of the country population in each occupation-country cell.

3 Hypotheses

The primary hypothesis is that in regressions of the broad type outlined above, the estimate $\hat{\beta}_1 > 0.5$. Failing to reject this hypothesis would imply failing to reject the hypothesis that the majority of cross-country gaps in real earnings are caused by cross-country differences in income determinants other than the individual human capital of workers. Important secondary hypotheses are that $\hat{\beta}_2 < 0$ when s represents skill requirements, that is, cross-country real wage gaps are narrower for occupations with greater skill requirements; and that $\hat{\beta}_2 < 0$ when s represents tradability of output.

4 Timeline and Public Posting

Research planned during 2018, exact timing uncertain. This plan was irreversibly published on March 4, 2018, at https://dataverse.harvard.edu/dataverse/mac-pap.