

## Abstract:

Are migrants positively or negatively self-selected from within their populations of origin? We study this question by comparing the heights of Italian immigrants to the United States between 1907 and 1925 to the height distributions of their respective birth cohorts in their provinces of origin. We created a novel data set based on 3.2 million Italian passengers entering Ellis Island whose last place of residence was geo-located, and on further transcription of stature and other personal information of a random subsample of 88,000 individuals. Interpreting stature as a measure of migrants' "quality," we find that the Italian migration was (a) negatively selected at the national level; but (b) positively selected at the local level; and (c) the selection varied systematically within the country, with more positive selection from shorter and poorer provinces. We show that the systematic variation in selection was not spuriously generated by differential geo-location probabilities or by non-classical measurement error caused by random errors in geo-location. We find that the patterns of selection were partly, but not entirely, driven by the post-1917 literacy requirement; they also cannot be explained by variations in rural-urban or occupational composition, though we cannot rule out that to some extent they are affected by selection across destinations. Our findings are consistent with theories of migration that highlight the importance of liquidity constraints that are partly solved by the support of networks of friends and relatives. They also highlight the significance of the distinction between selection at the local and the national levels.