
December 2 , 2024

Report prepared for the Journal of Economics, Race, and Policy

Title: The Impact of Hispanic Last Names on Educational and Labor Market Outcomes

This paper studies the adult labour market outcomes for a demographic group composed of one parent born in a Spanish-speaking country of birth ('Hispanic') and one parent born in the United States ('White'). The paper distinguishes between individuals with an Hispanic father and White mother (HW) and individuals with a White father and Hispanic mother (WH). The identifying assumption is that individuals with HW parents exhibit comparable characteristics, which are important determinants of labor market outcomes, to individuals with WH parents. The paper is essentially a replication of Rubinstein & Brenner (2014) to the United States. The research question is clearly outlined and well motivated.

My first concern relates to the threats to identification. It is feasible that children born to HW couples may differ systematically to children born to WH couples, particularly with respect to unobservable characteristics important for labor market outcomes. For example, HW pairs may exhibit differing parental characteristics related to parenting styles, parental preferences for education, gender norms and beliefs. The paper would be improved by including an expanded discussion around the threat to identification. In light of these issues, the paper would be improved by providing a more thorough discussion of the relative advantages of the current approach, compared to the conventional Oaxaca-Blinder decomposition.

The analysis sample is restricted to individuals who self-identify as 'White'. It is likely that this self-identification of ethnicity is endogenous to labor market outcomes. Importantly, the unobservable determinants of racial identification are potentially correlated with labor market outcomes. The paper would be improved by including a more thorough discussion of this (sample) selection issue. Related to this is the depiction of parents born in the United States as 'White'. This seems particularly over-simplifying and ignores the long history of immigration. A more accurate description would be to refer to this group as US born.

The discussion on the construction of the 'synthetic parents' is somewhat brief. The paper suggests that the potential parents are matched using the birth year of the child and the parent's place of birth to the children's information collected in the CPS sample. While not explicitly stated in the paper, I assume that this sample of synthetic parents is used to construct mean parental education and family income at the time of the birth of the child. The process for matching the potential parents to the children is somewhat aggregated and there is likely considerable heterogeneity in the mean educational attainment of the potential parents within the set of children with the same birth year with the same parent's country of birth. The paper could be improved by exploring the possibility of improving the quality of the match of potential parents by including further characteristics to match the potential parents to their children.

The description of the estimated model is somewhat brief. However, there are two

main issues associated with statistical inference on the estimated parameters in model (1). First, there is a ‘generated regressor’ issue associated with using estimated group-level parental education. A failure to account for this sampling variation will lead to misleadingly small standard errors. Second, given this group structure for parental education, it seems reasonable to assume that the model errors are uncorrelated across clusters but correlated within (potential parents) clusters. It is well understood that failing to account for this ‘clustering’ problem can lead to misleadingly small standard errors, narrow confidence intervals, and low p-values.

The paper would be improved through an expanded discussion of the impact of measurement error on the reported estimates. While there is a well understood result in the measurement error literature that measures of group mean parental education will provide estimates that are more robust to the presence of measurement error in individual level measures of parental education, there is still an issue with the non-random attrition of potential parents. Specifically, not all potential parents have children in the CPS sample, nor do all children in the CPS sample have parents in the sample of potential parents. Moreover, the non-random attrition of potential parents implies that mean parental education may be systematically higher or lower than the actual parental education of the children in the CPS sample. The paper would be improved by providing an expanded discussion of the likely impacts of measurement error on the reported estimates.