Educational Mismatch, Gender and Racial Wage Inequality:

Evidences on the persistence of high-skilled blacks and women in low-skilled jobs in Brazil

Ana Maria Hermeto

Drawing on Brazilian evidences spanning an interval of the last three decades, regularities and disruptions in the incidence of over- and under-education are outlined, as well as consequences for individual earnings. This paper studies the impact of increased schooling in the labor market between 1982 and 2012, from the perspective of occupational allocation, and uses the data from large Brazilian. We compute indicators of mismatch, trying to interpret whether this reflects over-education or increased skills requirements. Hierarchical regression models are estimated to capture the effects of mismatch on individual wage returns by gender and race, controlling for occupational characteristics. Preliminary results indicate both an increase of mismatch over the period and suggest increased skills requirements. And it was found that the wage returns are lower among the overeducated and higher for undereducated when compared to similar individuals in the same occupations, over the entire period.

Extended Abstract

Over the last decades important changes have been observed in the Brazilian labor market, including changes in the myriad of occupations, either through extinction, loss of weight or appearance of new jobs, increase and decrease in demand for certain occupations, impacts of technology on occupations, increasing number of women entering the labor market, etc. This picture of changes has been based, among other factors, on the constant restructuring of firms in the labor market, through the impact of technological change in the required educational pattern of the labor force. This change, in turn, has generated an increased demand for workers with higher educational levels, required to carry out the tasks related to the occupation.

This article is in the line of research that seeks to determine such changes in the Brazilian labor market, from the analysis of occupational data in a comparison of wage returns attributed to the various occupational categories in the early 1980's to the 2000's. The analysis will be based on PME large datasets from 1982 to 2012, conducted by IBGE. This database allows a representative historical overview of changes in occupations of the Brazilian labor market.

We investigate the relationship between educational mismatch and wage returns. That is, whether the impact of over and undereducation on the determination of wages, as well as their trend over time. Using the classification of underqualified, overqualified and appropriately skilled, incorporated in mincerian equations, a set of variables were controlled: individual, occupational, sectoral, regional and cyclical. Thus, it was possible to arrive at annual estimates of the impact of mismatch on wages, controlled by a number of variables which also influence the composition of wages.

The data source is the microdata from PNAD, IBGE for the years 1986 to 2011. The construction of databases and their analysis depended on the compatibility of occupational classifications. PNAD allows a study of characteristics and changes in the labor market, first, because it has a database of a few decades and, second, to always have in your research structure, one or more topics aimed to capture information from the Brazilian labor force. The concepts and classifications have changed over the period. Among the 1980, 1990 and 2000 changes occurred, trying to adapt the research

standards used by the ILO. Moreover, during this period, there were changes in the existing range of occupations, with the appearance, weight loss or fusion of occupations previously considered distinct. This was also reflected on the relationship of occupations adopted by IBGE.

We only considered individuals above 24 years and below 60 years of age. Due to the large difference between rural and urban labor markets, the sample is restricted to those living in urban areas and whose work is non-agricultural.

Indicators of Educational Mismatch

We applied the methodology proposed by Clogg and Shockey (1984) for the calculation of an indicator of educational mismatch, which allows analyzing over time the trend on under and overeducation of the workforce in Brazil between 1986 and 2011. Unlike other methods, it uses only two variables - formal education and occupation, not depending on data on sex, age, race, sector etc. This fact allows this indicator to be calculated for different databases and different periods.

For the 120 occupational titles we calculated the mean and standard deviation of years of schooling of individuals employed in 1987 (baseline). Years of schooling are considered proxy for the qualification acquired by individuals. Adding (subtracting) to the mean the standard deviation creates the upper (lower) bound. Between the two limits the individual is classified as matched. Below the lower limit, the individual is classified as undereducated. If the individual has more years of schooling than the upper limit, it defines the overeducation. Since the goal is to calculate the trend over the years, the range calculated in 1987 for the groups is maintained in subsequent years. This means that the measure can be arbitrary for the absolute prevalence, but is satisfactory when used for relative comparisons. After computing this measure of educational mismatch at each year, for each occupational title, it is possible to analyze the evolution of mismatch over the period. The evaluation will be done by occupational groups, gender, race, macro-regions, industries, and cohorts, trying to understand the evolution of mismatch and its main determinants. Since the distribution of occupational groups is not the same over the years, data was standardized before analysis of the results to avoid composition effects.

Effects on Wages

In order to assess the impact of mismatch on wages, we run the regressions of the logarithm of wages for a set of variables, which includes the traditional variables of the mincerian specification, plus dummies of over and undereducation, a set of sectoral and occupational variables, and set of control variables: gender, working hours, formal job, etc.

Hierarchical regression models were estimated, including individuals at the basic level, occupations as the second level and earnings as the dependent variable, in specific estimates by race, gender and period. The method allows that second level observations have different random intercept and elasticity coefficients. The ANOVA model with random effects is important because it allows to decompose the variance into two separate components, namely, σ^2 representing the variance at the individual level, and τ_{00}, the variance at the occupational level. They allow the computation of the coefficient of correlation (ρ), which indicates the proportion of the variability of the wages between the second level and the total sample, i.e., how variation of the whole model is due to between-occupations wage variation. With the aim of deepening the analysis of the reduction of the wage gap between men and women over time, we estimate for each year of analysis, hierarchical models controlled for sex and race of individuals. At the individual level, it is estimated the logarithm of deflated earnings as a function of individual human capital, age and an error with random

distribution. The second level reflects the sensitivity of the parameters that characterize a group of occupation to remunerate men and women, blacks and whites differently, in order to investigate the increase in the remuneration of women and blacks in more sophisticated occupations. The technological and educational requirements necessary for the performance of the work required in these occupations specifies the second level. Two variables are dummies for technology groups that can be exchanged for dummies related to the nature of the task; FF ranges from 1 to 4 according to the amount of physical effort required for performance of the task, REGR, REGM and REGL range from 1 to 6 indicating, respectively, grammatical requirements, mathematical and logical; AES indicates the minimum education accepted by employers for full performance of work; AEX expresses the average experience of workers required in number of years.

Preliminary Results

Mismatch rose in Brazil over time, the decline in the undereducation is accentuated. On the other hand, overeducation shows an upward trend. Regarding age, there is a direct relationship between age and overeducation and the increase in this phenomenon occurs over time in all age structure.

Analysis by gender shows a tendency for men to be more overeducated than women. In all the years women had lower rates than males. This fact is surprising, because historically women had higher average education to male. In more recent years, however mismatch rates by gender converged. Among workers, those in the formal sector are more overeducated, in the whole period. By industry, the overeducated are concentrated in the modern manufacturing sector and in the distributive services sector. The discussion can be further illustrated by the occupational titles themselves and be thorough in terms of the discussion of the differences by race and other occupational attributes.

The most important result concerns the average change in wages between the beginning and end of time interval. While the mean in the universe of workers varied slightly, the fastest growing occupational groups showed greater growth and groups whose mismatch rates grew more had a huge in their average hourly wages.

Within the mismatch debate and increased educational requirements, the phenomenon of overeducation is linked to the maintenance of average salaries, because excess skilled labor force imposes to some highly skilled workers, who could not find a job in occupations matched to their schooling, to assume functions historically performed by less qualified people. And as it comes to excess supply, we would see only schooling increasing in these occupations, without a counterpart in their real wages. But if a particular occupation, when deploying technological innovations or change work routines, demand more skilled workers than historically employed, ceteris paribus, will need to raise wages to get paid to hire professionals to meet the new demand. That is, the increased educational requirements (or required skills) lead not only to increase the average education of professionals in a specific occupation, but also to wage increases (above average labor market).

The fact that the group with the highest increase in the mismatch rate had an adjustment as above average in the period strongly suggests that such occupational groupings have not gone through a process of overeducation itself, but the occupations included in these groups began to demand greater skills over the period and rewarded - well above the market - skills that more educated individuals had.

All annual cross-section regressions of log hourly wage - both in the basic model, as in the full model - have significant coefficients for the mismatch variables. There are two conspicuous changes between models. The first is the inclusion of the gender variable - which captures part of the individual effects -

on the composition and reinforces the effect of mismatch on wages, especially between in the 1980´s and 1990´s. The other is the reduction of the effect after the inclusion of occupational variables, especially in the 2000´s. That is, controlling the heterogeneity of the sample in terms of individual attributes increases the effect of mismatch. When we control the sample heterogeneity also for occupational, industry and regional characteristics, the impact of occupational mismatch is attenuated, suggesting that there are not just inequalities between workers in the same occupation, but also between-occupations.

Other variables show important trends of the Brazilian population. The growth of mismatch among women, which, ceteris paribus, tends to reduce wages for individuals with the same education is simultaneous to the reduction of reducing occupational segregation by gender. This segregation, by itself, would have negative effects on wages. It remains to be explored the racial differentials and many other occupational characteristics in the hierarchical models.