Referee Report

This paper estimates the gender pay gap in Turkey between 2003-2010 using the Turkish Household Budget Survey. The authors use different decomposition methods and focus on the change in the wage gap over time (from 2003 to 2010). They find that the differences in the observed part of the gap have declined, while the part attributable to discrimination has increased over time.

The research question is a very standard and old one. The methods used in the paper are old fashioned and over-used. In its current form, the paper does not offer a significant contribution to the literature. Below I provide some suggestions to improve the quality of the paper.

1- I don't understand why the authors limit themselves to the period 2003-2010. As far as I know, it is possible to get the micro-level datasets for 2014 (for the Labor Force Survey). Given that the authors do not make a significant contribution to the literature, it might be a good idea to use the most up to date data. New data can offer new stories, which can improve the quality of the paper.

2- 2003 and 2010 may not be the ideal years to perform a comparison. 2003 is right after the 2001 crisis in Turkey and 2010 is right after the 2008 global crisis. The nature and depth of these two crises are quite different from each other. These differences have a potential to affect the results. The authors should use the micro level Labor Force Survey dataset from 2004 to 2014 (for every year) to capture the time series evolution of the gender pay gap over time rather than naively focusing on the change between two years. Maybe the differences in the nature of those two crises are driving the results.

3- Both Oaxaca-Blinder and JMP are "very old" techniques. The authors should not try to sell the JMP method as if it is a new one.

4- The Heckman correction exercise is not performed correctly and the interpretation of the coefficients is also not correct. It is not clear what is the first step, what is the second step, which variables are used as instruments (i.e., exclusion restrictions), why they are used as instruments, why this is a plausible assumption.

5- Rather than using hours as an explanatory variable, the authors should take the more conventional way and use hourly pay as the dependent variable. At least hourly pay should be used in the analysis as a robustness check exercise.

6- The "Introduction" is useless. The authors should summarize the findings in the introduction.

7- The theoretical part is trivial. It is absolutely redundant to write the equations for OB and JMP decompositions, as they are extremely well known in the literature.

8- The authors do not explain the main story behind their findings. Why did the gender pay gap increase over time?

9- It is well-documented in the modern literature that interpreting the "unexplained" part in the old-fashioned methods (such as OB) as "discrimination" is problematic. This is one of the reasons why I did not like the paper. The authors should provide a satisfactory explanation about the limitations of the methods used in the paper.