## **Responses to Anonymous Referee Report 2**

## We thank very much to the anonymous referee for his/her valuable comments and suggestions.

The paper aims to measure the extent of pay discrimination between male and female regular wage earners in Turkey at two points in time. The authors use the 2003 and 2010 Household Budget Surveys of the Turkish Statistical Institute. Using the Oaxaca-Blinder and Juhn-Murphy-Pierce (JMP) decomposition methods, they arrive at the conclusion that 57-61% of the gender wage gap is due to discrimination and that discrimination is higher among the higher income groups.

This paper is one of the many that look at the gender wage gap. It is clearly written – though a bit repetitive at times – offering evidence on the gender wage gap from a developing country where female participation in the labor market is rather low. I remunerate below some concerns and ways to improve the paper:

1. We do not find out how big the gender wage gap is until much later in the paper. The authors need to motivate the paper by presenting fact and figures about the gender wage gap. It is not sufficient to provide the extent of the wage discrimination, the reader needs to know the size of the wage gap itself. We will discuss the gender wage gap in more detail in the introduction of the paper in order to improve the motivation of our study.

2. The authors state that they contribute to the (national) literature by providing evidence on the gender-wage discrimination at two points in time (2003 and 2010). What is so special about these two dates? Why not use some later year in place of 2010? Do we expect the wage gap and the extent of discrimination to fall or rise over time? Since there is a time component, there is a need to provide a sense of why the authors compare the two time periods. Why not look at the evolution of the wage gap over time? **2003 is the first year for which healthy data is available. In order not to further lengthen an already long paper, we have chosen one more year that belongs to the post-crisis period. However, following the suggestion of an anonymous reader, we will reestimate all models for each year between 2004 and 2014 by using a different data source, the Labor Force Survey.** 

3. The review of the empirical evidence fails to provide a sense of why the estimates vary so quite widely and the way in which these results can be reconciled. We will try to reconcile the findings of empirical studies in Section 3.

4. The theoretical background is unnecessarily long. Much of what is described in this section would be well known to anyone even mildly interested in this topic. I suggest cutting much of this section and concentrating on the advantages and drawbacks of using Oaxaca–Blinder and Juhn-Murphy-Pierce decomposition methods in measuring wage discrimination. How is selection handled in the Juhn-Murphy-Pierce decomposition? Especially when selection changes over time? We will revise and shorten the section on theoretical background. Selection is handled in the Juhn-Murphy-Pierce decomposition in a similar way (see for example, Datta Gupta, N., R. L. Oaxaca and N. Smith (2001), Swimming Upstream, Floating Downstream: Trends in the U.S. and Danish Gender Wage Gaps, Centre for Labour Market and Social Research, WP 01-06). We will mention it in the revised version of our paper.

5. The dependent variable used is monthly wages. The human capital model dictates that this should be hourly wages. Given the substantial difference in the hours of work between men and women, and hours of work being a choice variable – at least to some extent - using it as a RHS variable does not seem appropriate. Workers in Turkey are generaly paid monthly. That is why we used monthly wages as our dependent variable. However, we will change it to hourly wages in the new estimations.

6. Why disregard the causal workers? The sample should consist of all wage earners. Assigning zero wages to causal workers is not correct. **Wage workers in our analyses include casual workers. The** 

information given in the paper will be corrected accordingly. However, casual workers are not assigned zero values. What we mean in the last sentence of the fourth paragraph of Section 6 is that the minimum value of zero belongs to casual workers. This sentence will be rephrased to avoid any confusion. It should be noted, however, that our paper does not include wage workers in agriculture because the agriculture sector has different dynamics.

7. Much of the RHS variables in the wage equation are endogenous, which must be acknowledged. Given that the paper falls short of establishing any causal relationship, the claim on how union status affects wages is unfounded (p. 21). In general, I would shy away from making strong claims based on the wage function presented in the paper. The endogenous nature of some of the explanatory variables will be acknowledged. We will be careful to avoid suggesting a casual relationship between wages and union status (and the other independent variables).

8. Are there controls for regions? Urban/rural settlements? Data does not include information on region and urban/rural residence. However, we will include an urban/rural dummy in our models in the new estimations.

9. Footnote 6: Why not express wages in real terms? One would like to see how real wages evolve and compare across time. **We will use real wages in the new estimations.** 

10. The selection correction must be related to what is chosen as the dependent variable in the second stage. Since the dependent variable is being wage/salary worker, this is the status that must be estimated in the first stage and not lfp. The referee is right; what we mean by lfp in the first stage is whether a worker is a wage/salary worker or not. That is, we estimate this status in the first stage. We will replace the term lfp with an appropriate term throughtout the text. (We could express it explicitly as 'being a wage/salary worker', or it could be called just 'employment' instead of lfp.)

11. There are some puzzling results – possibly typos – in Table 5. The intercept terms in 2010 look wrong. Since this is a probit estimation, along with the coefficients, one would like to see marginal effects and standard errors not t-values. The diagnostics must be reported as well. Yes, there are typos regarding the intercept term in 2010. We will be more careful in reporting the results in the new estimations. Also, in the new estimations we will report marginal effects and standard errors as well as the diagnostics.

12. Table 6: Intercept terms in 2010 are funny. Apart from this some results are strange. Is it possible that women's monthly earning are maximized at 25.83 years in 2003 but at 45 years in 2010? For men, the drop from 46.3 years to 33.8 years from 2003 to 2010 also looks strange. Yes, the years of experience for which earnings is maximized indeed look strange. If we find similar findings in the new estimations, we will provide an explanation of it.

13. The interpretation of the inverse mills ratio is not correct (p.22). We will rephrase the interpretion of the coefficient estimate for  $\lambda$  to avoid any confusion.  $\lambda$  is inversely related to the probability of employment. A negative coefficient on  $\lambda$  thus indicates that workers with higher probabilities of being employed will earn higher wages (see: Neuman & Oaxaca (2003), Estimating Labor Market Discrimination with Selectivity-Corrected Wage Equations: Methodological Considerations and An Illustration from Israel, The Pinhas Sapir Center for Development Tel-Aviv University, Discussion Paper No. 2-2003, p.15)

14. The results are interpreted as if one cross-section is used. There is very little attempt in the paper to discuss changes over time in the estimated coefficients. There is a lot of repetition in section 7. Points discussed earlier are raised again in this section with very little value added. We will discuss changes over time in the estimated coefficients (at least for important variables). We will refrain from repetitions in Section 7.

15. In reduced form estimation (the results not given) tenure is endogenous. If the point is to estimate wage equations based on exogenous human capital variables, one needs to omit the tenure variable. How is experience defined? The point in estimating the reduced form equations is to assess discrimination in the presence of only human capital variables. The definition of experience is given in Section 5.

16. Reduce or eliminate the emphasis on cohorts since no robust finding emerges. **The emphasis on cohorts will be reduced.** 

17. Why should discrimination increase over time? If it is indeed increasing, this is a point that must be emphasized and further explored in the paper. If we find a similar finding in the new estimations, we will explore it further.

18. The other interesting finding is higher discrimination faced by women in higher income groups. This is a point that would increase the value added of the paper. Why is this the case? Go beyond conjectures, try to provide answers both at a point in time and across time. If the new estimations yield a similar finding, we will try to provide a sound explanation for it.