Response letter to the Reader's Comments for the article entitled as "Gendered Economic Policy Making: The Case of Public Expenditures on Family Allowances"

I am thankful to the anonymous reader for the comments. Please see below the relevant replies for each comment. Invited reader's comments are provided in bold characters.

Primary Comments:

1. Because the author aims to investigate the effect of a critical mass of female parliamentarians on public spending for family allowance, it might be useful to spend some time motivating either why this policy is important to look at in and of itself, or (more importantly) why this policy might be particularly well supported by women -- over and above other policies that women favour. Consider: if the dependent variable is public spending on family allowance (as a fraction of either GDP or of total government expenditure), this variable may actually decrease if funds are spent on other policy initiatives that are strongly supported by women (social welfare, health, maternal/paternal benefits, education).

Indeed, all above-mentioned policies such as total social welfare, health, maternal/paternal benefits, education are strongly supported by women as previous literature suggests and the related studies are deeply investigated in the paper. The reason that my paper focuses on the case of ''family allowances'' is the inexistence of a study that so far analysed this spending category. Instead, all others are already analysed by different studies, but none of them point out those categories as a reason of a possible crowding out effect. The only spending category that is shown as a reason for a possible crowding out effect is old age benefits, which is already controlled for the entire analysis in the paper. Moreover, there is a clear explanation in the paper on ''why family allowances as a policy might be particularly well supported by women and as well as why old-age benefits can create crowding out effects on women and children-friendly policies. Among women friendly policies (e.g. between maternal benefits and family benefits) I am unaware of the existence of a literature which points out a crowding out effect.

2. If the author's goal is to provide evidence for one theory over another (Citizen-Candidate versus Median Voter) I would have liked to see a clearer treatment of the model comparison. Specifically, are there any cases (demographic structures) under which both models would generate similar predictions? If this comparison is not a primary concern of the paper, the discussion could be effectively cut back.

The aim of the paper is not making a comparison between those two models but to give some information about the ''background'' that the paper lies behind. The theoretical background for this study is based on a model supporting the fact that the identity of a politician matters for policy-making. In other words, it is assumed that the representation of female politicians will support issues that are women's preferences as long as women can vote in the elections. I agree that some phrases in the paper, which lead to some misunderstanding, should be removed from the paper (e.g. the first sentence of Section 2 which talks about testing those two theories). However I believe that it is useful to explain both theories to show which theory lies behind the main idea of the paper.

3.The author might consider a one year lag in the measure of female parliamentary representation, since policy changes can take some time to implement. That said, an increase in spending on a pre-existing policy might be quicker than the introduction of a new family allowance.

Once the female political representation with one year lag is included into analyses, the corresponding regression and the results from the replication of analyses are as following;

$$y_{it} = a_i + \beta_1 wompar_{it} + \beta_2 wompar_{it-1} + \phi' \mathrm{X}_{it} + country + year$$
(1)

Table 1. Estimation of (1)

	First Sample	*Second Sample*	*Third Sample*
Coefficient	**Fixed Effects**	**Fixed Effects**	**Fixed Effects**
β_1	(0.0105)	0.0086	0.0251
	(0.0069)	(0.0072)	(0.0188)
β_2	(-0.0017)	0.0011	0.0071
	(0.0029)	(0.0020)	(0.0052)
	GMM	GMM	GMM
β_1	(0.0062*)	0.0143	0.0132
	(0.0032)	(0.0091)	(0.0179)
β_2	(-0.0010)	0.0014	0.0084
	(0.0020)	(0.0045)	(0.0078)

Robust standard errors within brackets, * indicates significance 10% level. Regressions include all control variables that are mentioned in the paper.

The inclusion of the lagged female representation do not make any change in the main results of the paper and the coefficient of one year lag variable is not significant in any type of specification.

Did any (many) of the OECD countries ever have zero spending on family allowance? If so, it might be interesting to explore whether a critical mass of female parliamentarians was associated with the introduction of such a policy and/or with the type of allowance provided (i.e. do some countries have means tested cash benefits while others have tax rebates?)

Unfortunately there is no data on family allowances earlier than 1980 allowing us to test whether female politicians played a role for the introduction of these policies. The latest data belong to ''OECD Social Expenditures Database'' starting from 1980.

It would also be helpful to have a more detailed description of the family allowance variable (what policies it encompasses) and how it may vary across country.

I am thankful to the reader for this useful comment. There is a description of family allowances in the paper but there is not a deep discussion how it may vary across countries. I would add this discussion to the final version of the paper.

4. The paper would benefit from a greater exploration of the key explanatory variable (female parliamentary representation). I was curious as to why the author stopped at the 30% threshold? Does the theory suggest that women have no effect prior to the threshold and then a constant effect thereafter? Are results similar if the threshold is 40%? What proportion of the sample is represented in each threshold? (This would be helpful to see in table 1).

As the first referee also suggested, the percentage level of female representation for every year (and for each country) will be represented in the final version of the paper.

Unfortunately, the average female political representation in OECD Parliaments has been always lower than some certain levels and there are not many elections in countries resulted with high level of female representation like 40 per cent or over. For instance, the number of observations for the female representation at 40 per cent (and over) is only 18 out of 551 observations from a dataset of 19 countries in 38 years (1980-2008). <u>More importantly, the proportion of the sample for each threshold is not equal among countries which causes sampling bias</u>. For instance, those 15 out of 18 observations belong only to Sweden. Sweden was the only country where female parliamentary seats passed 40 per cent in more than one election (15 times), Finland and Netherlands follow her with 1 and 2 times respectively in 38 years history. Therefore, testing the relevant hypothesis for the higher level of representation might be incorrect due to the sampling bias problem and insufficient number of observations.

It might also be interesting to test an interaction of the fractionalization variable with your threshold dummy (or with the fraction of female parliamentarians). Your results suggest that greater fractionalization, while it may theoretically lead to higher public spending, does not increase spending on family allowance. However, one might expect that the impact of fractionalization on women favoured spending would be seen only in fractionalization parliaments with a critical female mass.

As the reader emphasized as well, the entire analysis using any of the three samples do not support the theory on the positive relationship between fractionalization and public spending. Once I replicate the applications, analysis that includes the interaction of the fractionalization variable with the percentage of female parliamentarians give the same insignificancy for the interaction variables. However an analysis for each specific threshold level gives positive significance for some threshold levels.

For instance, a sample regression for the electoral fractionalization and related results are as following:

$$y_{it} = a_i + \beta_1 wompar_{it} + \beta_2 rae\_ele_{it} + \beta_3 threshZ_{it} \times rae\_ele_{it} + \varphi' X_{it} + country + year \quad (2)$$

Table 2. Estimation of (2)

Threshold	Coefficient	Fixed Effects	GMM
20	β_3	-0.0002	-0.0004
		(0.0010)	(0.0003)
25	β_3	0.0005	5.1e-0.4
		(0.0008)	(4.5e-0.4)
28	β_3	0.0016	0.0005
		(0.0010)	(0.0004)
29	β_3	0.0025**	0.0011***
		(0.0009)	(0.0003)
30	β_3	0.0027***	0.0012***
		(0.0008)	(0.0005)
31	β_3	0.0030***	0.0015**
		(0.0010)	(0.0007)
32	β_3	0.0034***	0.0017***
		(0.0010)	(0.0006)
33	β_3	0.0024**	0.0009**
		(0.0010)	(0.0004)
34	β_3	-0.0008	0.0004
		(0.0062)	(0.0003)
35	β_3	-0.0008	0.0003
		(0.0006)	(0.0003)

Robust standard errors within brackets, ***, ** and * indicate significance at 1%, 5% and 10% level.

Only the interaction term of each threshold level between 29 and 33 with electoral fractionalization gives significant results supporting the main findings of the paper and also the presumes that explained for the 4th comment. Similar results are obtained for the interactions with legislative fractionalization as well.

Another consideration is that if strategic alliances are more important in a fractionalized parliament, such parliaments may have a lower critical level at which minority groups can influence policy. It would be interesting to explore this hypothesis.

Thanks a lot to the reader for the idea of minority groups. I will work on this idea for another article that focuses on the minority groups and relevant public spending categories.

Minor Comments:

1.The description of the empirical model and techniques is a bit long, and occasionally contradicts itself. For example, on page 9, equation (1) suggests that the author will estimate a model with country and year fixed

effects (consistent with the notes below Table 2). However, the author then states that country-specific time trends are used. Perhaps these were included in an alternative specification as robustness check which the author then omitted?

Except Pooled-OLS specifications, all estimation techniques (FE, GMM, PCSE) include country specific time trends (in addition to country and year fixed effects) for the entire analyses in the paper. Note only in page 9, this note exists below the all Tables. Pooled-OLS estimations represented in all Tables is the only one does not control for country specific time trends and country fixed effects, as they should not include them.

2.The author might consider reducing other parts of the estimation methods discussion, some of which are straight forward and don't require elaboration/presentation (inconsistency of pooled OLS p.13), and some of which are never presented in the paper anyway (fixed effects with lagged dependent variable p.10).

I agree about the unnecessary presentation of some results such as OLS. Relevant results of Pooled-OLS will be removed in the final version of the paper.

Moreover, the reason that I did not present the results of FE estimations with lagged variable was because I did not find any change in the results. FE estimations with lagged or without lagged variable give so similar results. Angrist and Pischke (2008) also suggest presenting them both if the results are really different than from each other. Otherwise they are more positive to present one of them alone, rather than both, as it is indicated with the following sentence in their book *''Applied researchers may face a choice between fixed effects and lagged dependent variables models; one solution is to include both, but the conditions for consistent estimation are much more demanding than for either alone''* (Angrist and Pischke, Ch.5 for more discussion).

3.Finally, though it reads well in general, the paper could be improved with the correction of grammatical errors, and with a bit more care over word choice, for example the cases of "prove"(p.2) or "unique" (p.9).

I am thankful to anonymous reader for this suggestion as well. As the first referee also suggested, the paper will be thoroughly proof-read especially for the correction of grammatical errors.

REFERENCES

Angrist, J. D. and Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.