To the editors,

Thank you again for giving us the opportunity to revise our manuscript "Can Citizens Set City Policy? Evidence From A Decentralized Welfare State" for consideration for publication in the Urban Affairs Review.

We greatly appreciate the time and effort taken by yourself and the three reviewers to review the manuscript. Your comments have been extremely helpful and we have made our best effort to accommodate these suggestions in the revised version of the manuscript. We believe the revisions, made because of your feedback, have made for a stronger, clearer, and ultimately more persuasive paper. We hope that you agree with us and deem the revisions satisfactory. However, needless to say, we would be willing to make further revisions if desired.

Below we first address the points raised by the editor in the decision letter, which also summarizes the main changes made to the manuscript. We then respond to the specific comments made by each of the reviewers on a point-by-point basis. For ease of reading, we reproduce your and the reviewers’ comments before each of our responses.

Sincerely,

The authors

**Editor comments**

*1. Reviewer #1 has a few minor comments regarding data-related choices (comment #1) and the substantive interpretation of the results (comment #3). The most important comment by this reviewer relates to the use of local vote share to assess ideological preferences. An abundant literature, both in the US and European contexts, suggests that ideological preferences are more pronounced in national level elections than in local level ones. This is partly due to the fact that many local government systems feature nonpartisan elections (with local lists/candidates independent from national parties). Even though this argument may actually support the authors’ assertions, we agree with the reviewer that a better explanation should be provided to ground your choice.*

*2. Reviewer #2 recommends working on the theoretical development of the manuscript. Even though the paper was submitted as a Research Note and should remain as such, some modest level of theoretical grounding and/or advancement is always desirable. The reviewer suggests two possible ways to accomplish this, either by providing a better understanding of the time structure of policy responses to local preferences or by exploring the theoretical mechanism behind the link policy preferences-policy responses.*

*3. Reviewer #3 is primarily concerned with the clarification of methodological procedures. The first three comments by this reviewer are related to the lack of transparency/clarity in the way data, methods, and results are presented. We encourage you to improve these aspects of your work and see some benefit in moving some information from the appendices into the main text. For example, Appendix F should be added to the main document to allow a better understanding of the results.*

*4. Both Reviewer #2 and #3 require a better justification for case selection. While data availability is a good reason to replicate previous studies, generalizability of results is an important concern expressed by the reviewers and shared by the editors.*

*5. Please provide complete references to OECD (2016) (p.1), Fiva et al. (2016) (p.1), Sances (2017b) (p.2), and Ansolabehere and Schaffner (2015) (p.5).*

**Response to Reviewer 1 (R1)**

We are happy to note that the reviewer thinks the “manuscript makes an important contribution to the literature on urban policy responsiveness, and that it is an excellent fit for the Urban Affairs Review.” Naturally, reviewer 1 also has some concerns and raises a number of good point.

*1. I would like more details about why these fiscal policy indicators are more “fine-grained” than those used in previous research. Most previous studies centered in American politics (which comprise the bulk of this literature) rely on the U.S. Census of Governments. The authors should make more explicit why their indicators are more precise than those of the Census of Governments (COGs) specifically. The authors certainly have better longitudinal coverage, which they note (and could make an even bigger deal about). But, are the policy items measured in Danish data more precise than the items in the COGs? This is certainly plausible, as the COGs data are quite coarse—but, the authors should show this.*

*Sætte os ind I COG. Hvorfor er det her bedre? Researche COG !  
  
2. The authors should offer more justification for why local vote share is actually the right independent variable measure. Previous studies used national vote share not because it was the only data available, but because they believed it would better correlate with local ideological preferences than local election returns. Consider the case of Massachusetts, which as a long history of electing Republican state governors. Public opinion surveys and national election returns indicate a far more liberal electorate than gubernatorial Republican vote share would indicate. I would anticipate that local policies in Massachusetts would on average reflect the liberal leanings of its mass public, which would be more accurately measured using national vote return data. I’m curious what the author(s) results look like if they rely on national election returns as their measure of ideology instead?*

Problem med nat: valg falder ikke samtidig. Potentielt undervurderer vi effect. Hvis det er en ren selektionseffekt, så er lokalvalg bedre, men hvis det ere n signaleffekt er det mere uklart.

*3. I liked the authors’ efforts at providing substantive significance, but it would be helpful to have more details, since I don’t know Danish cities: “With an effect of this magnitude, moving the voters from the Social Democratic stronghold Albertslund to the highly conservative Solrød would transform the fiscal policy in Gentofte to roughly that of Stenløse. This would move Gentofte down by more than 20 positions (out of 271 municipalities) in our ranking of fiscal conservatism.” A more intuitive approach for a broader readership might be to tell us what kind of policies we might expect a city to pursue “with an effect of this magnitude.”*

We agree with R1 that this interpretation is unlikely to be helpful to an audience without knowledge of Danish municipalities. On page 13 of the revised manuscript, we provide an interpretation of the results, which we believe is more meaningful to a broader audience. We rely on the correlation between the overall measure of fiscal conservatism and a number of its component parts to gauge what kind of change in real-world policy we would expect. This allows us to elicit expected reductions in various taxes and forms of spending under a change in the electoral support for right-wing parties.

**Response to Reviewer 2 (R2)**

We are glad that R2 thinks the paper in the data are “terrific”. R2 mains concern concerns the theoretical ambitions of the paper. In particular, R2 recommends *“that the authors refocus the paper. One option would be to highlight the dynamic results that they report toward the end of the paper. Thus, the paper might be reframed to analyze the timing of municipal policy change. They find that responsiveness occurs between four and eight years after local preferences change. This is interesting and could be explored much more thoroughly. Why does the timing take this structure? What does this mean for accountability? Another possibility would be for the authors to try to develop the mechanism by which policy responsiveness occurs. They note that it does not appear to be through the avenue of election of partisan officials. This seems hard to make sense of given that their independent variable is vote share. That aside – if the mechanism isn’t elections…what is it? They also say that it isn’t demographic change (although this was not directly tested). Determining the mechanism at work could help us to better understand the functioning of local democracy.”*

BLABLA

In addition to this, R2 has some other suggestions.

1. The literature review should motivate the authors’ theory rather than simply review what people have argued. The authors should explain what new insights they bring with their new data.

2. The authors include a lengthy appendix. However, the additional tests are disjointed and not well motivated in the paper. The text needs to better detail the problems that the various choices are seeking to address.

3. Some of the policies included in the fiscal policy index were not well explained. Co-payment for public services does not seem to be a direct measure of city policy.

4. The paper needs to include a longer discussion of generalizability of Danish cities. It is possible that we want to know about Danish cities just for the sake of knowing about Danish cities, but I was not convinced.

**Response to Reviewer 3 (R3)**

We are grateful for the reviewer’s assessment that the paper “will contribute to the discussion on citizen preferences and municipal response’’, and for noting that they “strongly support this manuscript.” R3 also has some different concerns which we lay out and respond to below.

*1.  First, despite the detailed robustness checks and nice visualization of results, I was surprised with the lack of basic statistical information in the manuscript. I do not see any basic information, which is commonly seen in academic papers, such as descriptive statistics, correlation matrix, VIF. Figure 2 only shows results of independent variables. However, the readers do not know what control variables included in the models and coefficients of the control variables. Also, if I understand the manuscript correctly, the author(s) created a fiscal conservativism index from various fiscal policy indicators presented in table B1. The paper should present how each indicator is correlated to the index and show the validity of index showing Cronbach's alpha or other measures. Table F1 does not indicate information on significance level for each star. Does \*\*\* indicate statistical significance at which level? Also, table F1 omits many details. Readers like myself may want to see full results of analysis.*

We agree that the presentation of our data and results should be made clearer, and we would like to thank R3 for pointing this out. To increase the transparency of the results in the main text, we have:

1. Moved the visualization of fiscal conservatism over time (what was previously Figure D2) into the main text as Figure 1, on page 8. This should give the reader a better sense of the distribution of the index and how it changes over time.
2. Added a new Table 1 on page 10 with summary statistics on the main variables as well as some other socio-economic features of the municipalities. The table allows for comparison between the pooled versions of the variables as well as the within-municipality changes, which allows the reader to better assess differences in the coefficients we present later.
3. In the revised manuscript, we present the main results in Table 2 on page 12 instead of plotting the coefficients (Figure 2 in the original manuscript). This makes it easier to inspect the model specification. It also allows the reader to assess coefficient sizes directly against the summary statistics presented in Table 1. Consequently, we have deleted Appendix F.

As suggested by R3, we have added a number of additional analyses to Appendix C. The aim is to allow for better interpretation of what the index measures as well as inspection of its validity and reliability:

1. To assess the reliability of the index, in Table C1, we present the overall Chronbach’s alpha as well as measures of how the alpha would change if we were to leave out each single item.
2. Figure C1 shows the pairwise correlation between each item and the overall measure of fiscal conservatism. This allows us to discuss more specifically what it is our index captures and, relatedly, to assess which factors are most important in defining the measure.
3. We have added summary statistics (Table B2) on the items as well. While all items are mean centered and variance standardized before being used to estimate fiscal conservatism, this allows the reader to inspect key characteristics of municipal fiscal policy.

*2. Table B1 shows that availability of number of data years differ depending on indicator. That’s why the authors used imputation to cover the missing data. However, as the author(s) may have noticed, imputing data sometimes create bias. In my opinion, one reason why many components of the index fail to reach statistical significance as showed in figure J1 is imputation bias, while the index itself achieved statistical significance results. I’d like to see more authors’ elaboration and explanation for this in the manuscript.*

We agree that the use of multiple imputation is not without risk, and we do take steps to reassure ourselves and the reader that our results are not driven by bias induced through the imputation procedure. In the original manuscript, we were not sufficiently clear about this, and we have sought to clarify in the revised version.

1. As noted by the reviewer, many of the items in the full index of fiscal conservatism have missing values after 1993. To make sure that our results are not driven by bias induced through imputation, we reproduce the results from the full index with a smaller index comprised only of the three variables for which we have complete data for the period. These are property tax, income tax and spending per capita. On page 10 of the revised manuscript, we have added a clarification of why we conduct this robustness check.
2. Because we use a Bayesian estimation technique to construct the measure of fiscal conservatism, missing values are imputed as an integral part of the estimation itself. The benefit of this approach over, e.g., frequentist factor analysis is that we can make use of indicators with missing values during different years in the same measure. The benefit over first imputing the data and then constructing the index is that the Bayesian technique puts less weight on items during the years, where data on them are missing. In this way, items like the cost of daycare, which is largely missing before 1993, adds to our knowledge of fiscal conservatism after 1993, but only to a limited degree before.  
   On page 7 of the revised manuscript, we have sought to bring more clarity to this part of the procedure.
3. This also means that, in Appendix J, when we estimate the relationship between support for right-wing parties and each component of the overall index, we do not impute missing values. We have added a few lines to notify the reader of this. Like the reviewer, we have been considering why some of the individual components do not reach statistical significance, and we believe that it is, because some of the items only have few non-missing years. For instance, the price of staying at a nursing home only has seven years of data, which explains why it can achieve a very large coefficient, yet is estimated with quite a lot of noise and therefore does not reach standard levels of statistical significance it itself.   
   In that sense, the combined measure of fiscal conservatism becomes more than the sum of its parts. This is in part, because the Bayesian estimation technique allows us to use information from variables that – like the price of nursing homes – would not allow us to draw any firm conclusions, but still are valuable in conjunction with other information on municipal fiscal policy.

*3.  Table B1 shows meaning of lower/higher values of each indicator in terms of fiscal conservatism. I do not understand why higher values of cost of day care, price of relief stay, food delivery, and stay in nursing home are considered as fiscal conservatism? I’d like to see some explanations for this.*

We very much agree with the reviewer that this feature of the original manuscript was not sufficiently clear, and we have sought to shed more light on these expectations in Appendix B for the revised manuscript. As we describe there, it is an interesting feature of the Danish welfare state that all municipalities are deeply involved in providing day care, relief stay, food delivery for the elderly and stays in nursing homes. The extent of municipal involvement and subsidies, however, vary between local governments. Therefore, in the Danish case, how much the municipality chooses to spend on keeping the prices of these services down, is a very central aspect of fiscal conservatism. We would like to stress that while this is an important feature of fiscal policy in Scandinavian countries, it might is not a central aspect in most other countries. This could be a potential threat to generalizability, however, we obtain similar results, when only using the smaller conservatism index with items that are not tailored to the Scandinavian experience.

4.  On a related note, I’d like the author(s) to justify the case selection of Danish municipalities in a more meaningful way, focusing on other than the data availability. I agree with the author(s) that the majority of municipal level studies focus on US municipalities. However, please don’t’ get me wrong, but in my opinion, the Danish municipalities, are also becoming “over-studied” because of the data availability and convenience. In fact, the majority of recent studies on municipal amalgamation/merger focuses on Danish municipalities. What can readers learn from this study of Danish municipalities other than utilizing dynamic models with the available rich data? How are Danish municipalities different from municipalities in other advanced democratic countries? In particular, what does fiscal conservatism mean in the Danish context? How is it different from one in other countries such as US, Germany, Japan, France? My main point is that the authors should justify why they select Danish municipalities from non-methodological perspectives. If the data availability is a main reason, it is difficult to see future studies of countries where data availability at the local level is limited. That is not what we want.

Diskussion af genraliserbarhed? Sig mere om konteksten. Hvad er fiscal conservatism I Danmark. At der er potentiel ”more at stake for voters” because so much spending is decideded on at the local level. This might hurt generalizability, but it also means, that to the extent that local governments are tasked with big responsibilities voters will have a say in how thise responsibilities are carried out. Parralell til city government I USA: stor diskrtion oer ting som ikke er spending relateret. Sances Hye Young You.

*5. Regarding voter preferences, can we assume that voters always have certain choices of electoral candidates with diverse ideologies? The independent variable is a difference between left-wing support and right-wing support.  But, do voters usually have choice? Are there enough candidates at the local election? For example, some studies such as Suzuki & Han (2018) show that walkover election (no-electoral contestation) is associated with fiscal autonomy.  I understand that the microdata of electoral campaigns may not be available. But, I encourage the author(s) to at least refer to the potential bias that may have been created by this issue.*

We agree with the reviewer that this could be a source of bias. However, part of the answer to this concern can be found in the specificities of the Danish case. Besides the relatively high turnover in Danish local elections, they are generally also strongly contested. For instance, in the latest municipal election in 2017, which was not extraordinary, there were in the excess of 9.000 candidates – i.e. approximately 1/400 if the citizens, who legally can run, choose to run for local office. In combination with the multiparty, proportional electoral system, this ensures that there generally are no walkover elections in the Danish municipalities. On page 6, we have added a footnote explaining the concern as well as outlining these features of Danish municipal elections.

*6. On a related note, what about the effects of electoral budget cycle and municipal merger? Results of previous studies suggest fiscal behavior of municipalities are somewhat different right before elections and merger reforms. Such data might not be available or it may be time-consuming to get such data, but again I’d like the author(s) to be at least aware of the potential biases without considering these effects. The author may want to cite relevant literature such as follows.*

We agree with the reviewer that both electoral budget cycles and municipal merger could be causes for concern. However, we do not believe that these factors would induce major biases in our results.

1. We believe that the results in Figure 3 alleviate concerns regarding electoral budget cycles. The results do show that an effect of a change in the electoral support for right-wing parties can be detected the one year prior to the election. However, they also indicate that the relationship between electoral support for right-wing parties and fiscal conservatism remain stable in the period between four years and 9 years into the future. Thus, the effect is stronger in the period between the two up-coming elections than it is during the year before the election takes place – i.e. the period, where we would expect an electoral budget cycle. On page 15 of the revised manuscript, we have added these thoughts.
2. … Municipal mergers …??