

Response to Referee 2

Quantitative Economics MS 2442

“Welfare and Spending Effects of Consumption Stimulus Policies”

Christopher D. Carroll, Edmund Crawley, William Du,
Ivan Frankovic, and Håkon Tretvoll

Thank you for your thoughtful comments and suggestions on our paper “Welfare and Spending Effects of Consumption Stimulus Policies”. They were all very useful to us in revising the paper. We hope you agree that the paper has improved. In the following, we summarize the main changes we have made based on your, the other referees’, and the editor’s suggestions. Thereafter, we state each of your comments in italics and provide point-by-point responses to them.

1 Summary of Main Changes

- **The Splurge** We have added section 4.5~~XX~~ and Appendix A~~XX~~ to discuss the implications of the splurge for our model.

To do so we have reestimated the US model without imposing any splurge consumption. We show that with a wider distribution of discount factors, the model is still able to match the empirically observed liquid wealth distribution in the US. It also matches untargeted moments on the consumption dynamics in response to transitory income shocks fairly well and still generates a substantial drop in consumption when unemployment benefits expire. The empirical fit is not as good as for the model that includes a splurge, however.

We then simulate the three fiscal policies in the reestimated models, and we show that our multiplier and welfare metrics for the policies differ only marginally between the model with and without the splurge. Importantly, the ranking of the policies

remains unaffected. While the splurge is thus helpful in matching available empirical evidence, it does not affect the main conclusions of our paper. In fact, including the splurge in the model, increases the cumulative multipliers of both the stimulus check and the tax policy, while leading to a slight reduction in the multipliers for the UI extension.

The key question is whether it is necessary for our purposes to include splurge consumption in the model. Including the splurge provides a better fit to the dynamics of spending after a temporary income shock, and when ranking the policies we discuss the timing of spending that they induce as an important distinguishing characteristic. For this reason, we have opted to keep it in our baseline version of the model. However, the empirical fit without the splurge is not so much worse that it substantially affects our results. The main moment the model cannot generate without the splurge is the high MPC for the wealthiest quartile, which is not so important for our evaluation of consumption stimulus policies.

- **Welfare Measure** We have completely overhauled our section on welfare and have introduced a new measure that we think best captures the idea that we want to measure welfare gains from carrying out each policy during a recession, but give no benefit in normal times to policies that in our model would increase welfare through redistribution.

Our welfare measure weights the felicity of a household at time t by the marginal utility of the same household in a counterfactual simulation in which neither the recession occurred nor the fiscal policy was implemented. This weighting scheme means that in normal times the marginal benefit to a social planner of moving a dollar of consumption from one household at one time period to another household at the same or different time period is zero. Hence, in normal times, any re-distributive policy has zero marginal benefit. However, in a recession when the average marginal utility is higher than in normal times, there can be welfare benefits to government borrowing to allow households to consume more during the recession.

Our new welfare measure leads to the same qualitative conclusions as in the previous version of the paper. It improves on the previous measure in several ways:

- The new measure does not scale with the size of the fiscal stimulus. We divide by the net present value of the fiscal policy so the measure is a ‘bang-bang-for-the-buck’ measure. As pointed out by referee 2, our previous measure was biased by the change in the size of the UI extension policy in a recession relative to normal times.

- Our new measure more naturally removes the bias to policies that redistribute from low to high-marginal utility households in normal times. Previously, we took away the welfare benefit in normal times of each policy. In our new measure, ANY marginal redistributive policy has no welfare benefit in normal times.

For more details on how we treat welfare, see section 4.3XX.

- **Robustness in a General Equilibrium Model** The main results of this paper are presented in a partial equilibrium setup with aggregate demand effects that do not arise from general equilibrium effects. We think there are many advantages to studying the welfare and multiplier effects in this setting without embedding the model in general equilibrium.

However, we now complement our analysis with a general equilibrium HANK and SAM model similar to Ravn and Sterk (2017). This model is as standard as possible, but able to capture supply-side effects that are absent from the partial equilibrium setup. In this model we also introduce a fiscal rule to balance the government budget. We find that the consumption multipliers across horizons follow the same qualitative pattern as we have in our partial equilibrium analysis.

The results from this HANK and SAM model are presented in section 5XX, and the details of the model are in Appendix BXX.

- **Consumption drop upon expiry of unemployment benefits.** One referee suggested that we report the size of the drop in consumption for households who remain unemployed for long enough for unemployment benefits to expire and compare this to the result reported by Ganong and Noel (2019). We now plot the path of income and spending around the expiry of unemployment benefits in our model in section 3.3.3XX (and for the version of the model without splurge consumption in Appendix A.2XX), and compare this to Ganong and Noel's results.

2 Essential points

1. *How badly does a model without the splurge factor actually match the data?* As far as I can tell, the need for the splurge factor is asserted (p. 14), but not shown. I believe the authors but would like them to show the best fit of the model without the splurge, for example the counterpart of Figure 1.

Response. See joint response below.

2. *Does the splurge factor matter?* A heuristic, right or wrong, is that the distribution of (intertemporal) MPCs is all that matters for consumption dynamics.¹ If we calibrate the model to evidence on marginal propensities to consume but allow ourselves to miss the liquid-wealth distribution, how does that affect the results?

To be clear, this is not a “robustness check” since the result is interesting and informative regardless of the outcome. Either (i) the splurge factor is useful for matching the liquid-wealth distribution, but does not matter for policy evaluation or (ii) the splurge factor changes the quantitative results quite a bit. Since the splurge factor is a novelty of the paper and the paper’s ambition is quantitative, it behooves the authors to explore and report how the splurge factor affects the results.

Response to 1. and 2. As described above in the “Summary of Main Changes,” we now include a section that discusses the relevance of the splurge for our results. We show that with a wider distribution of discount factors a model without the splurge is able to account fairly well for the empirical evidence on the dynamics of spending after an temporary income shock and the distribution of liquid wealth. (A chart illustrating the fit in that case, i.e. a counterpart to Figure 1 as you have suggested, is now included in the new appendix.) The main moment the model cannot generate without the splurge is the high MPC for the wealthiest quartile, which it turns out is not so important for our evaluation of consumption stimulus policies. Our main results, including the ranking of the policies in terms of welfare impact, are thus unchanged when removing the splurge from the model. In terms of the multipliers, the main effect of removing the splurge is to reduce the short-term multipliers for the stimulus check.

While the splurge does not matter for our ultimate policy conclusions, we agree that this is an interesting result in itself. We have opted to keep the splurge in our baseline version of the model since this provides a better fit to the dynamics of spending after

¹This heuristic is clearly not exactly right since it ignores precautionary saving responses, but it is a useful starting point.

a temporary income shock and since we view the timing of spending induced by the different policies as an important distinguishing characteristic in our comparison.

3. **Motivate the introduction of the aggregate-demand externality.** *In the model with an aggregate-demand externality, an increase in spending leads to an increase in everyone's income. One can easily imagine a world where aggregate spending instead reduces unemployment, primarily benefitting the unemployed (with higher MPCs etc.). More broadly, one of the main purported goals of fiscal stimulus is to shorten the duration of a recession, something which the model setup does not allow for. Qualitative, these distinctions may not matter but the ambition of the paper is quantitative.*

Response. We have expanded on the motivation for the aggregate demand effects in section 2.3:

"The motivation for this specification comes from the idea that spending in an economy with substantial slack and in which the central bank is unable to prevent a recession will result in higher utilization rates and greater output. By contrast, government spending in an economy running at potential with an active monetary policy will not succeed in increasing output. The recent inflation experience of the U.S. provides some evidence that output responds highly non-linearly to aggregate demand. This idea is explored in a recent revival of research into non-linear Phillips curves, such as Benigno and Eggertsson (2023) and Blanco, Boar, Jones, and Midrigan (2024)."

We agree that there are different ways of modeling these effects and that it might be the case that one channel could be through lower unemployment. Our modeling choice is primarily for simplicity. While our model is quantitative, there are a broad range of estimates for the size of aggregate demand effects and this is partly why we have chosen to present results both in a model with and without these effects.

Further, the aggregate-demand externality they have introduced implies, by construction, that promised fiscal stimulus in the future has no effect on output today. This is stark, in my opinion too stark to be introduced without further motivation! It cannot be dismissed out-of-hand that promised stimulus in the future spurs investment (and hiring etc.) already today. Further, by analogy with forward guidance (with respect to monetary policy), I conjecture that a new-Keynesian model would have an effect of promised fiscal stimulus in the future on output today (perhaps even a "too strong" effect!). The aggregate demand externality introduced in the paper, and the sharp distinction between stimulus in and out of a recession, is thus not consistent with the workhorse model.

Response. In our model, fiscal spending promised in the future causes increased spending today through the usual channels—households are forward looking and many are on an interior solution of their Euler equation. As such, future fiscal spending will activate the aggregate-demand effects today.

The sharp dichotomy behind spending in a recession and outside of a recession lies behind their support of stimulus checks (since the spending associated with them arrives during the recession).² The way they model the aggregate-demand externality therefore matters for their bottom-line results.

Given the partial-equilibrium focus of the paper, it is to me unclear what role the aggregate-demand externality serves. The particular aggregate-demand externality they introduced needs to be motivated and contrasted with alternatives such as a new-Keynesian aggregate demand externality. If the authors simply want to distinguish consumption during vs. after the recession, they can do so without this reduced-form way of introducing an aggregate-demand externality. The authors should then, with or without a theoretical framework, make a plausibility argument for why spending during recessions is likely to be more robustly stimulative than promised spending after recessions.

Response. We have now included a section in which we embed our households in a simple general equilibrium HANK and SAM model. In this model, fiscal spending increases output through the intertemporal Keynesian cross. This HANK model is a linear model and, as such, the aggregate demand effect endogenous to the model is in effect at all times. This is part of the reason we prefer our partial equilibrium approach—we think it is better able to capture differences in the effects of fiscal stimulus in a recession. Nevertheless, the impulse response and consumption multipliers are qualitatively similar in this HANK model, even though the tax cuts—which are slower acting than the other policies—have a relatively greater stimulative effect over a long horizon.

3 Suggestions

4. *Why not report stimulated consumption during recession?* In Table 3, “Share of policy expenditure during recession” is reported for the three policies. The policy-relevant object is,

²In footnote 3, they argue that a conventional new-Keynesian aggregate-demand externality would “bring in too many other confounding and confusing elements what would be likelier to obscure than to illuminate our points.” Confounding and confusing elements aside, a main qualitative difference between the two conceptualizations of aggregate demand is whether future stimulus affects output today.

however, “share of consumption increase during recession” and it would be informative to report this number as well.

Response. Indeed, the share of the consumption induced by the policy is a relevant metric when explaining the multiplier when aggregate demand effects are present. We have therefore added a fourth line to Table 3 (which is now Table 5) showing that share. Interestingly, the UI extension features the highest share of induced consumption during the recession among the three policies, slightly higher than the share of the check stimulus. This is because while the check occurs early it also is paid out to agents with low MPCs who spend the additional income very slowly. In contrast, the UI extension is paid out slightly later, but at that point its target population has high MPCs and spend the additional income more quickly. Given those insights, we have adjusted the language slightly when explaining the cumulative multipliers. Our main conclusions are, however, unaffected.

5. **Welfare criterion – weak inequality or equality?** When introducing the desired properties of the welfare criterion, the authors write (p. 26): “There is no social benefit to implementing any of the policies outside of the recession.” This way of phrasing it invites the interpretation of a weak inequality when the correct interpretation is an equality. The phrasing should be changed to something along the lines of “There is no social benefit or cost to implementing any of the policies outside of the recession.”

Response. Thanks for this suggestion - in our overhaul of the welfare section we have taken care to describe the welfare measure as having “no social benefit or cost” in normal times.

6. **Welfare criterion – measure in “bang for the buck”.** There is a mechanical reason why the UI extensions program will yield a larger welfare gain in recessions than in expansions: since twice as many households are unemployed, twice as many receive unemployment benefits. The programs are thus comparable in an expansion but not in a recession (since they are of different fiscal size in recessions). For a fair comparison of the programs, it would be helpful to compute “bang for the buck” numbers (i.e., dollars of consumption equivalent per dollars spent). In particular, I would suggest

$$\underbrace{\left(\frac{W(\text{policy, recession}) - W(\text{none, recession})}{W^c} \right) \bigg/ \frac{PV(\text{policy, recession})}{P^c}}_{\text{“bang for the buck” in recession}}$$

$$\underbrace{\left(\frac{W(\text{policy, expansion}) - W(\text{none, expansion})}{W^c} \right) \bigg/ \frac{PV(\text{policy, expansion})}{P^c}}_{\text{"bang for the buck" in expansion}}$$

Response. Thanks for this suggestion and in particular the observation that our previous measure gave a mechanical advantage to the extended UI benefits policy because of the size change during a recession. Our new welfare measure briefly described above and in more detail in section 4.3XX is a “bang-for-the-buck” measure in that we divide by the net present value of the policy when calculating the welfare measure. Our new welfare measure also has the property that it more naturally accounts for redistribution in normal times and thus we think it is the more appropriate measure to use for our analysis.

7. *Explaining the nuances of the welfare measure.* Relatedly, I would work through the text and make sure that the explanations match the fairly nuanced welfare measure used. Concretely, on p. 29 it says “The extended UI payments are well targeted to consumers with high MPCs and high marginal utility, giving rise to large multipliers and welfare improvements”. This is a true statement but this is also a true statement in expansions, and the welfare measure is about the difference between the two. One could instead write (for the welfare calculations) that “the extended UI payments are well targeted to the households who are particularly severely hit by the recession, giving rise to large welfare improvements relative to pursuing the policy in an expansion.”

Response. Thanks for highlighting the importance of getting the nuance right in this section. In our overhaul of this section, we have taken care to make sure we accurately describe why the extended UI insurance policy has the greatest welfare benefits.

8. *Robustness: changing real interest rate changes the discounting of the social planner.* In the robustness section, the authors (i) change the real interest rate and (ii) change the risk-aversion parameter. As far as I can tell, when they do this, they change the parameters entering the welfare calculations (i.e., the social planner’s preferences) and the behavior of households. These are two conceptually distinct robustness exercises and it would be clearer to change them one by one. Report the effect of changing household preferences, without changing the social planner’s preferences, and vice versa. The authors suggest (p. 31) that

changing the interest rate primarily affects the results by changing the preferences of the social planner, I think it would be clearer to spell this out.

Response. Following the suggestion of another referee, we have removed the robustness section. We would be happy to add it as an appendix if that is thought to be useful.

Finally, we would like to thank you again for your careful advice on our paper. We hope you find our revision satisfactory.

References

- BENIGNO, P., AND G. B. EGGERTSSON (2023): “It’s Baaack: The Surge in Inflation in the 2020s and the Return of the Non-Linear Phillips Curve,” Working Paper 31197, National Bureau of Economic Research.
- BLANCO, A., C. BOAR, C. J. JONES, AND V. MIDRIGAN (2024): “Non-Linear Inflation Dynamics in Menu Cost Economies,” Working Paper 32094, National Bureau of Economic Research.
- GANONG, P., AND P. NOEL (2019): “Consumer Spending during Unemployment: Positive and Normative Implications,” *American Economic Review*, 109(7), 2383–2424.
- RAVN, M. O., AND V. STERK (2017): “Job uncertainty and deep recessions,” *Journal of Monetary Economics*, 90, 125–141.