

# Response to Referee 3

## 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<sup>XX</sup> and Appendix A<sup>XX</sup> 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 inverse of 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 ?. 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; ?'s found that the drop was much steeper than could be explained by a consumption model with standard features including a time preference rate calibrated to match, as best it can, the entire path of consumption pre- and post-expiry. If we were to add the splurge to that baseline model, it would indeed make the drop in spending at UI expiration substantially steeper. However, our model is one in which agents have heterogeneous time preference factors. In a little-noticed exercise later in the ? paper, the authors examine a model with heterogeneous discount factors, and find that when the distribution of time preference rates is calibrated to match the consumption paths, the model goes a long way to explaining the sharp consumption drop. The difference reflects the fact that in a model with a single homogeneous

time preference rate, the model must match a pattern of consumption that implies that the great majority of consumers will have a substantial buffer stock of savings at the time when UI benefits expire. But a model with heterogeneous time preference rates has an additional degree of freedom to have heterogeneous MPCs at the point of expiry. In the HA-prefs model, the more impatient consumers will have a larger drop in consumption upon becoming unemployed, but the more patient ones will have a smaller initial drop. The total pattern of consumption decline however can match the observed total decline. The difference is that by the end of the UI-covered period, some households have run down their wealth more drastically than others, and those impatient households will have a much bigger drop in consumption than in the identical-prefs model. In other words, there is much less of a puzzle in their robustness test than in their headline model - and our model is already like that robustness variant. 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 directly compare this to ?'s results.

## 2 Essential points

- **Results without the splurge factor.** *The first half of the paper spends some time on motivating the splurge factor in consumption, but it is then taken for granted in the policy evaluations. Thus, the two sections appear somewhat disconnected. It would help me to see the policy results for a model calibrated without the splurge factor, at least in an appendix. Is matching the initially very high MPC out of transitory shocks essential for the policy conclusions drawn in the paper, or would the same conclusions be obtained from a more standard heterogeneous agent model? I imagine a removal of the splurge factor will make stimulus checks and income tax cuts much less attractive, but the strength of the channel should be reported.*

**Response.** As described above in the "Summary of Main Changes," we now include Appendix AXX that discusses the relevance of the splurge for our results. We show that with a wider distribution of discount factors even a model without the splurge is able to account for the relevant empirical evidence. Also our main results, including the ranking of the policies in terms of welfare impact, are robust to removing the splurge from the model.

- **Education specific transition rates, E to U vs. U to E.** *In the model, transition rates from employment to unemployment (E to U) are calibrated to different values across education groups, while job-finding rates out of unemployment (U to E) are constant across groups. The probability of transitioning out of unemployment also appears to not change over the cycle, while the separation rate does (this seems odd, as the DMP literature tends to assume that separation rates are constant and job-finding rates move with the cycle (see, e.g., ?). This focus on E to U rates is not motivated and the reader is left to wonder whether this is a crucial assumption or not. It appears that the model could be calibrated to match the same moments with heterogeneous U to E probabilities and constant separation rates, instead. I presume this would make recessions much worse for unemployed dropouts, boosting the welfare effects of a UI extension. Given that the welfare results of the paper strongly favor UI benefit extensions, a discussion of the heterogeneous E to U rates assumption seems warranted.*

**Response.** In response to your comment and some questions/comments from another referee, we have updated our presentation of the calibration of the transition probabilities between employment and unemployment in section 3.3.1XX in the paper.

As you point out, our strategy is to calibrate the same probability of transition-

ing from unemployment to employment for each education group and education-specific probabilities of transitioning from employment to unemployment. This enables us to fit our calibration targets which are the average duration of unemployment and the education-specific unemployment rates in normal times. As we are using the 2004 wave of the SCF, we use numbers for 2004 from the Bureau of Labor Statistics for the average duration of unemployment of 1.5 quarters and for the education-specific unemployment rates.

Note that this calibration strategy is consistent with the results in ? who finds that the main difference between education groups is in the incidence of unemployment, and not its duration. He states that “the reduction of the incidence of unemployment [at higher education levels] is found to be far more important than the reduced duration of unemployment in creating the educational differentials in unemployment rates” (page 1).

? reports probabilities of transitioning into unemployment as the product of the probability of job separation times the conditional probability of unemployment given a job separation, and both of these are lower for higher education levels. For our calibration, this means that a higher job finding rate *within* the quarter of the job separation for more educated workers translates into a lower probability of transitioning from employment to unemployment during a quarter. In that sense, our calibration is consistent with short-term job-finding rates being higher for more educated workers.

In our model of a recession, unemployment increases because the probability of transitioning from employment to unemployment increases for each group. This would be consistent with a model where the job separation rate is constant, but there is an increase in the conditional probability of being unemployed at the end of the quarter when the separation happens. However, our model is not detailed enough to separately identify these two probabilities, so it is only the overall probability of an E to U transition which increases in a recession.

? refers to data from 1979 and 1980, and there may have been substantial changes in the labor market since then. However, more recent work by ? include data up to 2009 and echo ?’s findings (see their Figure 8).

- ***The welfare criterion.*** *I had trouble understanding the welfare criterion used to evaluate the three different policies. My own intuition for comparing the policies would have been to simulate, for each policy, two recessions, one with the policy and one without, and then compare aggregate welfare between the two; potentially also to an economy without a*

*recession, to quantify the cost of business cycles. This approach is deemed not viable at the top of page 27, because it would give the social planner incentive to redistribute outside of recessions. I think this point warrants additional explanation. Is the problem that recessions end stochastically, making it difficult to exclude welfare gains from, e.g., a UI extension after the recession is over? In this case, one could compute this measure for each recession length, not including the additional welfare gains from continuing policies after the recessions, and weight them by their probabilities. I'm sure the explanation is perfectly obvious to the authors, but a reader will strongly benefit from an additional discussion. It was also unclear to me why the recessionary policies investigated had fiscal costs in non-recessionary periods, as suggested in equation (12). Relatedly, page 9 states that all policies have the same fiscal cost and are financed through taxes in the far future, but I could not find how  $PV(\text{policy}, \text{Rec})$  or  $PV(\text{policy}, 0)$  are constructed. How can they be included in equation (12) without further assumptions about the timing and implementation of taxation schemes? This also matters for the results presented on page 25—is  $NPV(\infty, \Delta G)$  roughly constant across policies?*

**Response.** We have overhauled the welfare section as described in the “Summary of Main Changes.” In the paper we explain why a standard aggregation of household welfare gives rise to welfare benefits and costs to redistribution policies even in normal times:

“In our model, some households consume much less than other households, and a social planner with equal weights on each household could significantly increase welfare through redistribution across households even in normal times. We are interested in the benefit of carrying out fiscal policies in a recession, so we do not want our results to reflect the benefits of redistribution inherent in our model in normal times.” (see section 4.3XX)

As a consequence, our new welfare measure weights are chosen such that, in normal times, there is no welfare benefit to *any* redistribution policy. We believe this new measure best captures the welfare benefits of each policy in a recession without biasing policies (such as extended UI) that redistribute wealth even in normal times.

In the description of the model, we are now clear that: “To keep our analysis as simple as possible, we do not model the debt repayment.” The exact fiscal rule that is chosen to pay for the policies will not alter our conclusions about which policies are more effective as this will be the same for all three policies. Given that, we think the exposition is clearer if we do not model the debt repayment. As such, the net present value of each policy is equal to the total cost of policy at time zero. For example, the NPV of the stimulus check policy is simply the total value of all the

stimulus checks that are sent out.

In equation (9)XX,  $NPV(\infty, \Delta G)$  is not equal across policies. For the example the total cost of stimulus checks by far exceeds the total cost of the UI extension, since the latter applies only to a subset of the population. Dividing by  $NPV(\infty, \Delta G)$  in the definition of the multiplier serves the purpose to assess the amount of induced consumption *relative* to the total cost of the policy. We have added a sentence to further clarify this when the multiplier is first introduced.

### 3 Suggestions

- **Robustness.** *In my view, the robustness section can be relegated to an appendix. None of the exercises come close to changing the welfare ranking across policies. Further, in three of the four exercises (risk aversion, interest rates and recession properties), the quantitative welfare results change only marginally.*

**Response.** Thank you for this suggestion. We have removed the robustness section as and we are happy to add it as an appendix if that is thought to be useful.

- **Liquid wealth distribution in splurge estimation.** *When estimating the splurge factor, the authors calibrate a model to the Norwegian economy to match marginal propensities to spend out of lottery winnings and the liquid wealth distribution. Unfortunately, the latter does not exist for Norway, hence the authors instead target the US liquid wealth distribution from the 2004 SCF. I think it would be prudent to target a liquid wealth distribution for a country whose labor market institutions more closely match Norway's, e.g., Germany or France. The European Union's Household Finance and Consumption Survey provides such a measure, the code to compute the relevant metrics are provided with the Brookings Paper by ?.*

**Response.** As mentioned above, we now include Appendix AXX that discusses the relevance of the splurge for our results. There we show that our conclusions about the ranking of the different stimulus policies hold up even in a model without the splurge, though the effects are somewhat muted without it. The approach you suggest may provide an improved estimate of the splurge, but the new appendix indicates that our overall results are unlikely to be sensitive to such an improved estimate.

- **Targeted stimulus checks.** *In the final paragraph of the introduction, the authors discuss the possibility of targeted stimulus checks (as opposed to the case discussed in the paper:*



*means tested, but universal distribution). I think estimating this case would help readers benchmark the welfare results by providing an upper bound. As of now, it is clear that UI extensions in the model do most to increase welfare; but are they far away from this hypothetical upper bound?*

**Response.** Thanks for this suggestion. Our motivation for writing this paper was to provide practical real-world advice to policymakers considering different fiscal policies during a recession. As such, we limited ourselves to three policies that have been used regularly in such situations and are likely to be used again.

The practicalities of getting stimulus checks out in a timely manor, and the limitations of government data and IT systems, means that the actual choice set of policymakers is likely limited to simple means-testing.

Having said that, there are many households in our model that are at their borrowing constraint and therefore have an MPC equal to one. If the government were able to target them directly, that would be optimal for stimulating the greatest amount of consumption during the recession. We have added footnote to the section on Comparing the Policies, “Theoretically, stimulus checks could be targeted to the highest-MPC households which, for small-sized policies, would mean households with an MPC of one. However, data limitations and other practicalities make means-testing stimulus checks by income the extent of targeting in practice.”

- **Effect of UI extensions on job-finding rates.** *The model presented in the paper does not feature endogenous search, which precludes the UI extensions from having negative effects on job-finding, as proposed in, e.g., ?. In this sense, the authors, in the name of tractability, stack the deck in favor of the UI extensions. I’m strongly sympathetic to the assumption, but it would be interesting to know how bad the negative incentive effects would need to be for the UI extension to lose its effectiveness. For example, assuming that job-finding probabilities decrease linearly with benefit size, how negative would the function’s slope need to be for the paper’s conclusions to revert? A back of the envelope calculation could provide an answer. My prior would be that the number is too high to be realistic, strengthening the conclusions in the paper.*

**Response.** We have undertaken the following back-of-the-envelope calculation.

1. We define the cost associated with a reduced job finding rate as consisting of two components: a) loss in economic activity due to additionally unemployed agents and b) fiscal cost due to UI benefits to additionally unemployed agents.

- The number of additionally unemployed agents is calculated as  $AU = 4 \times N \times JFR \times (1-x/100)$ , where  $N$  is the number of unemployed agents affected by the reduction in the job-finding rate,  $JFR$  the job-finding rate during a recession (calibrated to 0.25 in our model) and  $x$  is the %-reduction in the job-finding rate due to the UI extension. In our model results,  $N \times JFR$  many agents find a job in each period. In the back-of-the-envelope calculation we assume that number reduces by a factor of  $x/100$ . As the UI extension policy is in place over 4 quarters, we furthermore multiply the number by 4.
  - The loss in economic activity is assumed to be given by the reduction in labor income due to the lower job finding rates, i.e. the product of the number of additionally unemployed agents and their average permanent income.
  - The fiscal cost is assumed to be given by the replaced labor income.
  - Hence, the total cost is given by  $AU \times \text{perm. income} \times 1.2$  (as the replacement rate provided by the government is 0.2)
2. Next, we determine the additional cost, which the UI extension policy would need to cause, such that its welfare benefit equals that of the second-ranked policy, the check stimulus. We then determine  $x$  (the reduction in the job-finding rate) such that additional costs as defined above equal the amount necessary to equate welfare benefits between the UI extension and the check.
  3. Finally, we translate the increase in the job finding rate to an increase in job search time to allow for a literature comparison.

We find that the job search time in response to the UI extension would need to increase by more than 1.3-2.3 quarters (depending on whether AD effects are switched on or off) in order for our welfare ranking to flip. The empirical literature surveyed in ? finds that “a one week increase in potential benefit duration increases the average duration of the unemployment spells of UI recipients by 0.1 to 0.2 weeks.” Hence, based on that empirical literature, our extension of UI benefits for 2 quarters should yield an increase in the job search time by around 0.2 to 0.4 quarters, well below the 1.3-2.3 quarters needed to overturn our results.

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