Response to Editor
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 giving us the chance to resubmit our paper "Welfare and Spending Effects of Consumption Stimulus Policies" to Quantitative Economics. And thank you for your thoughtful comments and suggestions for how to improve our paper. They were all very useful to us in revising it. We believe the paper has improved greatly through the revision and we hope you agree.

In the following, we first summarize the main changes we have made based on your and the referees' suggestions. Thereafter, we go through how we have dealt with the specific requests from you. For each request, we first repeat your comment in italics and then respond how we have dealt with them.

1 Summary of Main Changes

- **The Splurge** We have added section 4.5XX and Appendix AXX 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 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; Ganong and Noel'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 Ganong and Noel 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 Ganong and Noel's results.

2 Comments

1. ***Motivation.*** *I think the paper needs a stronger motivation in the sense that it is a bit unclear why I need to introduce splurge consumption and other features of your model. In particular, incomplete markets/HANK models with liquid and illiquid assets, and/or models with spreads between borrowing and savings rates, can potentially account for both MPCs and for the relevant features of the (liquid) wealth distribution. So why go the way you choose in the paper? Do you have empirical evidence in favor of this alternative model or are there other reasons for exploring this?*

   Let me quote from the feedback I got from another editorial board member:

   "Perhaps the authors want to argue that a HA model + fully constrained agents cannot match their data targets, so that one needs something else, and splurge consumption is a natural solution that also seems intuitive. But then they need to make this argument explicit: they need to compare a version of their model with a fraction of fully constrained agents (which would be the natural first place to go) to their benchmark model and argue how the data rejects this alternative model. They should also discuss why other alternatives to matching iMPCs that have been introduced in the literature are not good enough for what they want to do."

   See below for joint response to this and the next point

2. ***Splurge consumption.*** *A key aspect of your paper is the introduction of splurge consumption. I find this an interesting idea. However, I have issues with your modeling on page 8 contained in equations (1)-(4). Here you assume that splurge consumption is a constant fraction of income which enters the budget constraint, but does not impact on the marginal utility of consumers' optimal choice of consumption. The latter is crucial since a moderate level of splurge otherwise would have no or little impact on the economy. In the other extreme, suppose you had assumed that consumers get utility from splurge goods in exactly the same way as $c_{opt,i,t}$ so preferences are given as $\sum_{t=0}^{\infty} \beta_t^i (1-D)^t \mathbb{E}_0 u(c_{opt,i,t} + c_{sp,i,t})$, which would seem a natural starting point. In this case, splurge would have no effects on total consumption unless $c_{opt,i,t} < 0$ (which you could rule out, I guess).*

   You present your assumptions without defending them, but I think you need to have a convincing story about this as it otherwise looks arbitrary. Your simplest defense, of course would be that preferences are given as $\sum_{t=0}^{\infty} \beta_t^i (1-D)^t \mathbb{E}_0 \left(u_1(c_{opt,i,t}) + u_2(c_{sp,i,t}) \right)$, but in this case you would need to have a good "story" about what type of goods these splurge goods are and why you can treat preferences this way.

   Moreover, as also pointed out by Referee 1, your current set-up is equivalent to a model

in which there is a constant average tax rate, ξ, but you then count tax payments as consumption. Again, this seems inconsistent. Referee 2 points out that your assumptions alternatively can be thought of as each household having some buffer stock members and some hand-to-mouth members. This also seems arbitrary and it is hard to accept that these different branches of the family cannot insure amongst themselves.

In a footnote you mention that splurge might be close to rational in a model with small durables. However, this would seem to me to contradict your calibration (that 30 percent of net income is spent on splurges).

In summary, I think you need a convincing story about splurge consumption, otherwise this seems too arbitrary and also implies that your analysis cannot be used for welfare analyses.

Response. In the consumption literature with which two of the coauthors are quite familiar, there is widespread agreement that a very high immediate MPC out of transitory income shocks is a substantial puzzle that needs explaining, in the sense that it has proven difficult to reconcile such a high MPC with other behaviors we see consumers engage in.

It is true that there is an edge case first clearly articulated by Deaton (1992) in which, with respect to small shocks to income, the MPC out of income shocks can be one. This case requires (1) consumers satisfy an impatience condition that means that, in the absence of a liquidity constraint, they would have perpetually declining spending and perpetually increasing debt and (2) a liquidity constraint to prevent wealth from falling below zero. It is also true that Deaton's model has been augmented in the last decade to permit the existence of an illiquid asset to which consumers can contribute only upon paying a transaction cost. If the rate of return on the illiquid asset is high enough to make them wish to accumulate it, then even wealthy people can cycle through periods in which they are in the Deaton (1992) regime - they oversave for a while to build up enough assets to be worth paying the transaction cost, and then spend a period of recovery during which they are Deaton-like MPC=1 people.

One problem with this is that it doesn't work if there are transitory shocks to income of a magnitude that labor economists have long been measuring, most famously in Abowd and Card's (1989) Econometrica paper. For models like that, the average MPC is much lower than 100 percent. Furthermore, expenditure shocks would have the same effect of inducing agents to hold a buffer stock, and Scott Fulford and David Low at the Consumer Financial Protection Bureau have a new paper (**?**) using a new survey of consumers to document that the size of 'emergency' expenditure shocks

is comparable to the size of transitory income shocks.

Another problem is that it is far from clear that such a model can match what Auclert et al have argued is the key "sufficient statistic" for understanding the effects of income shocks on consumption, the Intertemporal MPC. The Norwegian data that we use is almost the only source in existence that has the crucial information about how spending in years 2, 3, and 4 is affected by a shock in year 1 (that is why the paper is so highly cited).

A final problem is that this model cannot explain the robust finding in recent literature that even wealthy people with *high liquid assets* have a high MPC out of income shocks (Crawley and Kuchler (2023); Graham and McDowall (2024)).

It is also true that some economists still defend the idea that the only kind of heterogeneity needed for macro purposes is the Campbell-Mankiw saver-spender dichotomy. But such models are ill suited to the study, in particular, of unemployment. The paper we cite elsewhere by Ganong and Noel (2019), for example, finds a pattern of behavior among the unemployed that cannot be explained by the saver-spender model, in which the spending of the spenders would drop one-for-one with the drop in their income, while spending of the "savers" would be barely affected by the unemployment spell; neither of these patterns looks like what most unemployed people actually do. And a model that matches what they actually do implies important business cycle dynamics (like strong precautionary reactions to an increase in uncertainty) that cannot be captured in a saver-spender framework.

In the revised version of the paper we have expanded our citations to the rapidly growing literature providing various competing explanations of this "excess initial MPC." **?** has new work that attributes it to a form of "present bias" in which people have time inconsistent preferences, and recent work by **?** offers a similar interpretation. A number of papers, including another by **?** which argue that it's really not an excess MPC, it's a high initial Marginal Propensity to Expend - the difference being that the sensitive component is expenditures on durable goods, not nondurables. **?** have a paper that argues that many expenditures that are classified as nondurable, like vacations or occasional special dinners at fancy restaurants, are "memory goods" on which it is perfectly rational to spend because you are purchasing a memory of the event or service that will last a long time. Another interpretation by Kueng (2018) (following Akerlof and Yellen (1985)) of the phenomenon is that the effort of thinking about how to optimally spread the income surprise over an extended period (like a lifetime) is just not worth it, so busy people just spend the

money now. And of course any of these explanations can combine with each other to provide an even more nuanced story.

The point of our paper was not to wade into this literature and pick a favorite; it was to capture the fact that there is sufficiently strong evidence of an "excess" initial MPC to motivate a burgeoning literature trying to explain it.

Since part of the point of our paper is that the *timing* of the consumption effects of alternative stimulus plans will matter when multipliers are cyclical, we felt it was important to capture in a simple way the fact that motivates the excess MPC literature.

In the end, we now show that we get similar, though somewhat muted, effects in a version of the model without the splurge. But if we had not done the splurge calculations and presented them front and center, we are confident that people in that literature would have criticized us for having gotten the timing wrong.

3. ***Choice setting.*** *You present your analysis as partial equilibrium, but it really is simply a choice setting (with some choices not modelled). I do not have a problem with this as such since such models can be used for many interesting purposes.*

   However, I do think that this means that some of your results may be questioned, and that you need to be very careful with your analysis and perhaps rethink parts of it. Here are my issues:

   - *Clearly, not all of the policies that you consider are equally affected by the lack of (general) equilibrium effects. Tax cuts work mainly through supply side effects which you exclude. I think it is important to point this out and be less dismissive about tax policy. Unemployment insurance is also sensitive to the lack of a supply side modeling as one usually would think of these as potentially hampering job creation. Stimulus checks are more direct demand policies. Hence, in the end, if you extended your analysis to GE, it is unclear to me whether there would be a clear winner.*

     Response. In reponse to these comments, we have introduced a new section of the paper that analyses the three fiscal policies in a canonical HANK and SAM model. Despite this new section, we have a strong preference for our partial equilibrium analysis. As stated in the paper:

     "First, general equilibrium models often struggle to adequately capture the feedback mechanisms between consumption and income, particularly the asymmetric nature of these relationships during recessionary versus expansionary

periods. Additionally, a complete general equilibrium treatment would necessitate the analysis of numerous complex channels including investment dynamics, firm ownership structures and dividend distribution policies, inventory management, and international trade flows—elements that, while important in their own right, would potentially obscure the core mechanisms we aim to investigate."

Nevertheless, we agree that in principle some of the general equilibrium effects could change our results: "fiscal policies can generate labor market responses that our partial equilibrium analysis does not address. Employee tax cuts, for instance, may increase employment through changes in workers' incentives. These supply-side channels can affect both the welfare implications and the fiscal multipliers of different policy interventions."

We demonstrate that the qualitative features of the partial equilibrium model pass through to our general equilibrium approach: the consumption impulse response functions are similar and the difference between the consumption multipliers under the three fiscal policies leads to the same conclusion that the tax cut is significantly less effective at stimulating consumption than either the unemployment insurance policy or the stimulus check policy.

- *I find it very misleading to talk about multipliers in the choice setting. Your model leaves out equilibrium mechanisms that create the potential for such multipliers (the standard Keynesian cross mechanism for example). You do provide the consumption externality feature as an extension, but it was "unclear" to me that this really allows one to interpret results in terms of multipliers. Indeed, since multiplier effects come from general equilibrium effects, one could seriously question why you look at multipliers at all in your paper.*

  Response. We are quite sympathetic to the point that, in the version of our model without the Krueger-Mitman-Perri modification to capture multiplication by *force majeure*, there is no "multiplication" that comes from the pattern of spending; we are instead capturing the intertemporal marginal propensity to consume highlighted by Auclert et al as the key *input* to calculating multipliers.

  In the end we decided that the best way to present our results was in the form that seems to have become standard in the multiplication literature: The cumulative difference in some measure (consumption; GDP) made by the policy being analyzed over some fairly short time interval. When there is some multiplication mechanism present, this measure captures both the impulse and its multiplication. Otherwise, it is just an impulse to consumption.

After struggling a bit to come up with an alternative terminology (besides the iMPC that we do employ), we concluded that we should follow the literature in measuring the quanity of interest in this way, under the assumption that someone reading our paper for this purpose would be comparing our results to those other estimates.

Of course, whatever we call it, the objects we report are measures of the effectivess of the policies that have been explicitly motivated by the goal of stimulating consumption spending in recessions beyond what it would otherwise have been. As we show, even in the absence of any "multiplication," such policies can be grounded in the welfare benefit that they provide.

- *In continuation of these points, showing impulse responses at long forecast horizons and calculating present value multipliers seems a bit odd to me (in the pure choice framework, all of the transfers will be spent sooner or later and the multiplier will go to one; anything you get different from that comes from the ad hoc externality). It is still interesting to see how fast this happens, but why not focus on shorter horizons?*

  Response. Our long horizons were simply motivated just by the goal of presenting our results in a way that would be directly comparable to the results in other papers. We are in agreement that it is somewhat puzzling why the literature has ended up picking such long timeframes for its calculations; perhaps it is precisely because, as you say, over a long enough horizon, in the absence of any general equilibrium phenomena, the full amount of the extra income will be spent. In an old-fashioned perfect foresight model, the timeframe for that extra spending is as long as a full lifetime, so perhaps the literature concluded that 10 years was a good middle ground between the business cycle frequency of 2-3 years and a very long horizon like a lifetime. In the revision to the paper, we have added some further figures showing the full dynamic pattern of spending over time (this seemed, in particular, like the best way to compare our model to the full HANK model).

4. ***Calibration.*** *I found it odd that the calibration mixes up targets/parameters from the US and from Norway. Why not use targets from Norwegian data (or US data) as far as you can and then simply add to this insights from other papers/data (which might be US related of course). Referees 1 and 3 make comments amounting to the same concern.*

   Response. As described above in the "Summary of Main Changes," our main conclusions also hold in an estimated model of the US economy where the splurge is set to zero. These results are derived entirely without use of Norwegian data,

as that data is only used in the estimation of the splurge. As such, the calibration strategy using both US and Norwegian data has no bearing on our ultimate results.

5. ***Welfare Criterion.*** *Referees 2 and 3 complain about the welfare measure that you use. I agree with them, it is somewhat murky. Moreover, how should one actually think about welfare in this choice set-up? I think it is fine to examine some measure of "bang for the buck" but is that really the same as welfare?*

   Given these comments, one option would be to leave the multiplier and welfare analyses to another (general equilibrium) paper and simply focus on how "splurge" consumption can help account for iMPCs and use the model for the analysis of the impact of some selected fiscal policies. If you choose not to go this way, you need to argue your case carefully.

   Response. The ability to calculate welfare gains and compare policies in welfare terms is one of our main motivations for writing this paper and we have a strong preference to keep it. As you point out, in the absence of any multiplication, the case for consumption stimulus policies is weak because they merely change the timing of consumption without resulting in any extra net resources. But some of the rhetoric about the reasons to engage in such policies is best understood in terms of welfare - UI benefit extensions in particular are explained as helping those in need 'through no fault of their own' and 'who need help in tough times' and with similar expressions that have little evident grounding in multiplication. So it is better to have a framework to quantitatively interpret the size and distribution of any welfare effects that might be a motivation for the policies.

   In response to referees 2 and 3, we have overhauled the welfare section as described in the "Summary of Main Changes." Our new analysis provides a "bang-for-the-buck" measure that removes the welfare benefits of redistribution in normal times in a more natural way.

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

References

Akerlof, George A, and Janet L Yellen (1985): "A near-rational model of the business cycle, with wage and price inertia," *The Quarterly Journal of Economics*, 100(Supplement), 823–838.

Crawley, Edmund, and Andreas Kuchler (2023): "Consumption Heterogeneity: Micro Drivers and Macro Implications," *American Economic Journal: Macroeconomics*, 15(1), 314–41.

Ganong, Peter, and Pascal Noel (2019): "Consumer Spending during Unemployment: Positive and Normative Implications," *American Economic Review*, 109(7), 2383–2424.

Graham, James, and Robert McDowall (2024): "Mental Accounts and Consumption Sensitivity Across the Distribution of Liquid Assets," *Available at SSRN 4793885*.

Kueng, Lorenz (2018): "Excess Sensitivity of High-Income Consumers*," *The Quarterly Journal of Economics*, 133(4), 1693–1751.

Ravn, Morten O., and Vincent Sterk (2017): "Job uncertainty and deep recessions," *Journal of Monetary Economics*, 90, 125–141.