Narrow Bracketing Design October 2019

Francesco Fallucchi and Marc Kaufmann

December 1, 2019

Rationale for new design

Our last design as registered in the AEA registry (https://doi.org/10.1257/rct.3412-3.0) used a design where different individuals had different amounts of tasks and we would elicit the willingness to do additional work (WTW) by asking how many extra tasks subjects were willing to do for various piece-rates, such as "How many extra tasks are you willing to do for $0.15 per task?". We ran a pilot of this in November 2019, with 30 to 40 subjects per treatment group, for Narrow10, Narrow20, Broad10, and Broad20. We find that subjects in the group with 20 initial tasks to 2.16 more *total* tasks than in the group doing 10 initial tasks (p-value 0.06, clustered at the individual level, pooling Narrow10 with Broad10 and Narrow20 with Broad20, which is not the regression we'd run for the full study).

There are two issues that this highlighted. First, in order to compare the results of Narrow10 and Narrow20, we need to account for the fact that subjects in the latter group must at least do 20, while the former group can choose 10 to 20. Similarly, since we kept the number of extra tasks subjects can choose constant, the group starting with 20 tasks can do up to 50 tasks (if we allow 30 extra tasks), while the Narrow10 group maxes out at 40. In order to compare across groups, we therefore imputed a choice of 20 to subjects in Narrow10 who chose less than 20 tasks, and we imputed a choice of 40 to people in Narrow20 who chose 40 or more.

After doing this, we ended up with 80% of all the choices being corner solutions: either choosing to do 0 extra tasks or 30. This severely limits the power of the design. A further pilot where we allowed up to 50 extra tasks and elicited more piece-rates in a middle range led to a drop in corner solutions, but it is still a concern.

However, more problematic than this is the fact that the current design is also not clean from the perspective of keeping choice sets exactly equal. A person in Narrow10 who chooses to do 20 extra tasks and a person in Narrow20 who chooses to do 10 extra tasks do the same amount of work, and receive the same *marginal* incentives for a given piece-rate. However the person in Narrow10 receives 20 times the piece-rate, while the person in Narrow20 only receives it 10 times. While we could equalize it by making our design such that the piece-rate also applies to the initial baseline, we figured out an alternative design that is cleaner and should get rid of all of these concerns.

Design 3

The main change is in the extra tasks subjects can do and how we elicit this. The baseline pay is M which everyone receives conditional on completing the study. The questions are phrased roughly as follows:

```
Option A: Decode 20 additional sequences for an extra $4.00

Option B: Decode 30 additional sequences for an extra $[INPUT FIELD] or more (between $4.00 and $8.00)
```

In the broad treatment, the phrasing is

```
Option A: Decode 20 total sequences for a total $4.00

Option B: Decode 30 total sequences for an total of $[INPUT FIELD] or more (between $4.00 and $8.00)
```

Thus in the broad treatment, subjects are comparing the total tasks for the study and the total money.

In the description below I shorten this to "What is the minimal payment for (20,$4.00) vs (30, X).

We have three treatments:

Broad: "What is the minimal payment for (T + 16, $X) vs (S + 16, $(5 + M))" Narrow0: "What is the minimal payment for (T + 16, $X) vs (S + 16, $5)" Narrow16: "What is the minimal payment for (T, $X) vs (S + 16, $5)"

Narrow0 and Narrow16 means that subjects are told that they have a baseline of 8 and 16 required tasks to do. However, they may have to do additional tasks no matter which option they choose, since we do not always offer 0 extra tasks.

The difference between broad and narrow0 is that the money is not framed broadly in the narrow treatment, and that the wording is as 'extra' vs 'total' tasks. However since narrow0 has 0 baseline tasks, the extra tasks are total tasks.

As should be clear from the choices, Narrow0 has the same choice as Narrow16, except that they number of *extra* tasks is 16 higher for all options. That way, their *total* tasks are identical. The total money is also kept constant, but it is framed broadly in Broad. This means that subjects are told that they choose between 'extra money' in the narrow treatments and between 'total money' for the broad treatment.

The broad treatment is equivalent to "Narrow0", which is as if they had 0 baseline tasks, except that money is framed broadly too.

Therefore this new design keeps the choice sets *exactly* equal across all treatments, both in terms of total money and in terms of total tasks. This alleviates any concern that differences in choices could be driven by some broadly framed preference. One remaining concern is that the baseline number of tasks serves as a reference point, in which case reference-dependent loss aversion would lead subjects with a lower reference point to consider more tasks a loss, and thus want to do fewer tasks. To avoid this, we tell subjects what the number of baseline tasks is only on the page where they are asked to make their decisions.[1]

We will have several values for T, S, and the payment amount for each person, but always such that they lead to identical broadly framed outcomes across treatments. We have two main hypotheses about the following regression, where i is the individual, j one of the 5 (or so) choices faced by that individual, and t the treatment they are in:

$$X_{i,j,t} = \alpha_j + \beta_8 \cdot \mathbb{1}(Narrow8) + \beta_1 6 \cdot \mathbb{1}(Narrow16) + \varepsilon_{i,j,t}$$

Our main parameter of interest are β_t's: They should be 0 for broad bracketing, since the choices j are identical when interpreted as broad choices. Thus the primary test is whether β_8 and β_{16} are statistically significant from 0 and from each other. Furthermore, we expect that Broad will be less willing to do more work, which implies that the amount of money asked by the other groups should be lower (they are more willing to do the work). I.e. the type of narrow bracketing we expect has $\beta_{16} < \beta_8 < 0$. (This assumes that subjects are always giving their willingness for doing *more* work, i.e. that $T > S$. If this sign is flipped, we'd have to flip the sign on β too, but right now our design always has $T > S$.)

Our secondary prediction is that people narrowly bracket by acting more like people who entirely ignore their baseline. In the extreme, this means that people should make the same choice when choosing to do an extra of 8 tasks rather than 0 (say), whether or not they have 0 or 16 baseline tasks. Thus the prediction is that Narrow16 + 8 should lie between Narrow0 + 8 and Broad0 + 24 - and the degree to which it does so indicates the degree of narrow bracketing in the population. This secondary result will have less power than the former, since we cannot make all choices line up with another one.

Finally, we will test whether the medians and means of the treatments are equal or not.

[1] One might still be worried about reference points being set to this baseline, but that itself would be a form of narrow bracketing. Nonetheless it might be interesting to explore mechanisms further. One way to do this would be to give subjects two choices, and two choices only, both of which will be implemented and therefore should be bracketed together. One can then split the same broad choice in two choices in several different ways. That design would get around reference dependence, since it would allow us to give the same baseline choices to all subjects.