Study Outline:

The goal of this study is to estimate the shape of short-term time preferences using incentive-compatible decisions over a real-effort task. The basic idea of the study is fairly simple: after practicing the task in the first week, study participants are asked the number of tasks they wish to complete for different wages on a given date as that date approaches over the next three weeks. For example, a participant might be asked questions (via text message) on Monday morning, Wednesday afternoon, and Thursday afternoon about work preferences for that Friday afternoon. Broadly, the study design and analysis plan largely follows that of Augenblick and Rabin (2016), using the same task and potential wages. The main differences are that the choices are all within one week in order to determine a more precise shape for short-term time preferences and there is no focus on sophistication or projection bias.

Outcome Data:

The main data of the study is fairly simple - each decision represents the number of tasks chosen by a specific participant at a given time for a given wage for work that occurs for a given time window.

Sample:

As this is an experiment that takes place over multiple weeks, significant attrition is expected (my expectation is 15-20%). Some participants will likely quit prior to making a decision. Consequently, in an attempt to gather 100 participants to reach the point of making decisions, 110 participants will be initially recruited into the study. In addition to this pre-decision attrition, I expect that other participants will quit after making some decisions. My expectation is not to include any analysis of the drop-out decision in the main text as I don't believe there will be enough power to make meaningful statements.

Non-Parametric Analysis:

[All of the analysis will cluster the standard errors at the participant-level]

The non parametric analysis will summarize the raw data in a variety of ways (there are some important degrees of freedom in the analysis, which I discuss below):

- Show the aggregate level of tasks chosen for different wages.

- Raw comparisons and difference tests of aggregate wage choices as the work time approaches and/or smoothed comparison of wage choices as the work approaches.

- Raw comparisons and difference tests that loosely test the predictions of different models on the aggregate data (exponential, quasi-hyperbolic, hyperbolic, fixed-cost, etc.).

- Discussion of heterogeneity in individual responses.

Parametric Analysis:

[All of the analysis will cluster the standard errors at the participant-level]:

The functional form of the non-parametric analysis will largely follow Augenblick-Rabin 2016 (there are some important degrees of freedom in the analysis, which I discuss below):

- The aggregate analysis will structurally estimate parameters from the main time discounting models: exponential, hyperbolic, quasi-hyperbolic, and (secondarily) fixed-cost (Benhabib, Bisin, and Schotter 2010) and potentially others.

- In the structural estimation, the disutility of work will be modeled as a power function (with exponent parameter gamma and scaling parameter 1/psi) with the utility of wages modeled as a linear function of money earned.

- For some models (such as fixed-cost and quasi-hyperbolic), it is possible to estimate multiple models with different parameters. For example, a one-parameter quasi-hyperbolic model specifies that the discount function is 1 for t=0 and β for all t>0. One two-parameter model would allow the discount function to be to allow $\beta\delta^t$ for all t>0. However, both of these models do not precisely specify the time-period length and therefore it is likely necessary to define an additional parameter which determines when the discounting discontinuity occurs.

- The fit of the models will be compared. Here, there is an issue which I am unfortunately not sure of the appropriate solution. Usually, measures of goodness-of-fit penalize additional parameters. However, in this case, although the fully specified quasi-hyperbolic model includes three parameters and the hyperbolic model includes one parameter, my intuition is that the quasi-hyperbolic model is still effectively as restrictive as the hyperbolic model. I require more work to determine the effective way to compare the fit of these models.

- One issue is that averaging over quasi-hyperbolic discounters might create patterns that appear hyperbolic in the aggregate. It will should be possible to estimate individual parameters and potentially categorize individual choices to deal with this issue.

Main degrees of freedom:

There are a few degrees of freedom in the analysis. Broadly, my goal is to present the most natural or representative analysis in the main paper and then include other choices in the Appendix.

- Bins/Smoothing over the week: The goal of the non-parametric analysis is to show how decisions change over the week. This presentation requires either a choice of bins or a smoothing method (and smoothing parameter). Unfortunately, based on simulated data, I do not believe it is appropriate to pre-specify the exact bins or smoothing parameters. That is, in the simulated data, different assumed discount functions and parameters create different patterns that only appear visually given different smoothing parameters. My plan is to use multiple smoothing techniques/bins and present the ones that I feel best represent the data in the paper. I will then include a representative set of other possibilities in the Appendix.

- Bins over wages: Either in the paper or the Appendix, I expect to show a graph that demonstrates raw changes in work decisions given wage changes. Based on simulations and data from other papers, I will aggregate the data into 10 wages bins for visual ease. I will provide a similar graph which shows decisions for each wage in the Appendix.

- Participant Sample: It will likely be difficult to properly identify the structural parameters for some participants, either due to lack of decisions due to attrition, lack of variation in decisions, or decisions that consistently violate monotonicity. The analysis will be run on the entire sample as well as the subsample for which it is possible to identify parameters. I expect to focus on the subsample in the paper and only include the results for the entire sample in the Appendix.

- Decision Sample: There are two ways in which I foresee restricting or breaking up the sample of decisions. First, it might be argued that participants will be potentially confused in the first week of decisions - therefore, I will provide analysis which either removes the first week or separates the results week-by-week. Second, there is a concern that boredom or a desire for consistency leads people to choose similarly after making many decisions. Therefore, I will potentially restrict the analysis to earlier decisions, exploiting the fact that I assign some participants in some weeks to only make decisions closer to the work time.

- Controls: When presenting the non-parametric results in the text, I will using controls for wage/individual fixed effects, although I presume that these should not affect the qualitative conclusions. If the results do change substantially given the controls, I will discuss the differences in the text. When possible, I will present results both with and without controls in the text. Given that the graphs will likely only have one set of controls, I will include graphs with different sets of controls in the Appendix if the results appear to differ.

- Demographic information: I do not have strong ex-ante hypotheses concerning the interaction of demographic variables (such as gender, intelligence, or major) with the results. However, I have collected these data mainly for use by future researchers who might have an interest in this data. Therefore, I will not use demographic information in the main analysis. I will present all heterogeneity results in the Appendix regardless of the outcome and make a small reference to the results in the paper.

- Location of error term: For the structural estimation, I believe there are two reasonable locations for the error term. The first location, which I expect to be in the main text, is an additive zero-mean normal-shaped error term on the work decision. The second, which I expect to include in the Appendix, is an error term on the effort-cost parameter gamma.

- Estimation of individual parameters: hopefully, it will be possible to estimate individual parameters for a large set of the participants. However, for some of the discounting models with multiple parameters, it might be that there are not enough individual observations to achieve

this goal. In this case, a second-best solution is to estimate certain individual parameters, but constrain other parameters to be constant across participants.

- Definition of when work occurs: In the experiment, participants choose a three-hour window to complete their work (Friday, February 26 10am-1pm) rather than specifying an exact time (Friday, February 26 10:30am). This is done to reduce the possibility that workers have small events arise that are unrelated to time discounting but make it difficult for them to complete work at an exact time. However, this creates endogeneity in the exact timing of work and complicates the task of tightly identifying very short-term (hourly) effects. An alternative analysis is not completely obvious, but it might be possible to use the average work time within the time window or the pre-stated expected work time to test for these effects.