# Pre-analysis plan: Task Autonomy, Motivation & Performance

Jonas Radbruch Timo Freyer Sebastian Schaube

August 26, 2020

Regarding the analysis of the pooled sample: Is there a reason against standardizing outcomes individually for the two sets of hypotheses (1-3, 4-6)? That is, standardize by task across treatments SelfSelected and AssignedPreferred for H1-3, and by task across treatments AssignedPreferred and AssignedNonPreferred for H4-6?

## Introduction

This project investigates how task autonomy affects intrinsic motivation and performance by means of a real-effort online experiment. More precisely, we study how autonomy affects performance indirectly through the selection of a suitable task, as well as directly through a change in the intrinsic motivation due to the opportunity to actively select a task in itself. Therefore, we decompose the overall effect of task autonomy into an indirect effect (selection into tasks) and a direct effect (change in intrinsic motivation).

## Experimental Design

For the online experiment we aim to recruit 500 subjects from the BonnEconLab subject pool to work on one out of two tasks for a flat payment of 5 €. To create an environment conducive to intrinsic motivation, we picked two real-effort tasks which are arguably productive. The first one is a classic data entry tasks where subjects have to transfer student numbers by university and gender from German statistical yearbooks. The data subjects have to digitize are not available in a digitalized form. Within the scope of the second task, subjects have to assess and evaluate tweets on the topic of the German fiscal debt brake with respect to different categories. Notably, the data generated could be used to study the respective theme, and this is the way the tasks will be presented to the subjects to ensure that they perceive their effort as valuable.

Data collection will proceed in two steps. In a first session, subjects will be presented the two different tasks, before they are asked to indicate their preference for one of them. In a second session, on the following day, subjects are asked to work on one of the tasks for at least

10 minutes. Throughout this working stage, they have the opportunity to quit work by clicking on a button.

After the first session we will randomize subjects into treatments, stratified by task preference. Treatments differ with regard to whether subjects can self-select their task (Self-Selected), are assigned their preferred task (Assigned-Preferred), or are assigned their non-preferred task (Assigned-Not-Preferred).

A simple treatment comparison between Assigned-Preferred and Self-Selected yields the direct effect of task autonomy. Subjects work on the same kind of task and they both preferred that kind. The only difference between both groups is that the task was assigned in Assigned-Preferred vs. self-selected in Self-Selected. In addition, knowledge of subjects' task preferences allows to check for the selection into tasks due to autonomy. For this purpose, we can compare individuals in Assigned-Preferred, i.e., who work on their preferred task, with others who have been assigned to the task they had not preferred (Assigned-Not-Preferred).

As the main interest of this project is the estimation of the direct effect of autonomy, we oversample the groups Self-Selected and Assigned-Preferred. We aim at 200 observations for each of those groups (100 subjects for each task-treatment cell). Additionally we aim at 100 observations for the Assigned-Not-Preferred control group (50 subjects for each task cell).

## Main Outcome Measures

Our main outcome measures capture three different facets of motivation and performance. First, we are interested in how the treatment manipulation affects total working time. In particular, we define total working time as the amount of time until subjects reach the last subtask[1] they processed as our primary measure of working time.[2] Our second outcome measure is a subjects' output, measured as the total number of tweets assessed and the total number of data rows transcribed, respectively. Third, for those subjects working on the data entry task we will investigate how the treatment manipulation affects the share of errors committed, defined as the number of erroneous entries divided by the total number of entries to be made.[3]

[1] A subtask is defined as either a single tweet they have to assess or the data from one particular statistical yearbook they are supposed to enter.

[2] This is due to the fact that subjects can only exit the work stage after 10 minutes and might remain on some task without actually working. As a robustness check, we will use other definitions of working time in the subsidiary analysis.

[3] In case a subject does not start working on a specific subtask at all and does not enter data, the subtask is not considered for the calculation.

### Control Variables

Our primary set of independent variables only consists of treatment and task indicators. In addition, we elicit a set of control variables including subjects' age, gender, high school grade and interest in either task, measured on individual seven-point Likert scales.

## Hypotheses

### The Direct Effect of Task Autonomy

Our main hypotheses concern the potential direct effect of task autonomy. To isolate the direct effect of autonomy on intrinsic motivation and performance, we compare subjects who self-selected their preferred task (Self-Selected) with those subjects who got assigned their preferred task (Assigned-Preferred). We hypothesize that making an autonomous decision between the two tasks itself increases subjects' intrinsic motivation. As a consequence, we should observe a treatment effect on subjects' working time (hypothesis 1). Increased intrinsic motivation should translate into a better performance. However, a priori it is not clear that increased motivation leads to a better performance in terms of quantity: It could be the case that subjects assess more tweets or transcribe more data if they work longer; yet, it could also be that they take more time to assess a given tweet or transcribe a given set of data, improving the quality of their work. In the latter case, increased motivation is revealed by subjects working more diligently and making fewer errors. We test both of these possibilities by exploring whether the treatment affects subjects' output (hypotheses 2), or, for subjects working on the data entry task, the share of errors commited (hypothesis 3).

**Hypothesis 1.** *Subjects who self-selected their preferred task (*Self-Selected*) work longer than subjects who got assigned their preferred task (*Assigned-Preferred*).*

**Hypothesis 2.** *Subjects who self-selected their preferred task (*Self-Selected*) produce more output than subjects who got assigned their preferred task (*Assigned-Preferred*).*

**Hypothesis 3.** *Subjects who self-selected their preferred task (*Self-Selected*) commit less errors than subjects who got assigned their preferred task (*Assigned-Preferred*) (data entry task only)*

### The Selection Effect

We exploit our control treatment Assigned-Not-Preferred to check if selection into tasks based on subjects' preferences changes their performance. Because in this context the interaction is one-shot and subjects receive a fixed wage, we expect that they simply prefer the task they find more interesting and thus are more motivated to work. In addition, the tasks subjects prefer might be those they are more able at, reinforcing the selection effect.

We investigate this effect by testing whether working on the preferred task (keeping the exogenous assignment constant) affects working time and output (hypotheses 4 and 5). Additionally, for subjects working on the data entry task, we scrutinize whether task autonomy makes them work more diligently (hypothesis 6).

**Hypothesis 4.** *Subjects who got assigned their preferred task (* Assigned-Preferred*) work longer than subjects who got assigned their non-preferred task (* Assigned-Not-Preferred*).*

**Hypothesis 5.** *Subjects who got assigned their preferred task (* Assigned-Preferred*) produce more output than subjects who got assigned their non-preferred task (* Assigned-Not-Preferred*).*

**Hypothesis 6.** *Subjects who got assigned their preferred task (* Assigned-Preferred*) commit less errors than subjects who got assigned their non-preferred task (* Assigned-Not-Preferred*) (data entry task only).*

# Statistical Analysis

For each hypothesis, we first conduct nonparametric Wilcoxon rank-sum tests. We complement the non-parametric analysis with OLS regressions, both with and without controls.

## The Direct Effect of Task Autonomy

In order to investigate the direct effect of task autonomy, it is important to thoroughly control for differences in outcomes which originate from subjects in Self-Selected being able to chose their preferred task, which they may find more interesting or be more able at. Therefore, we condition on the set of subjects working on their preferred task, i.e. the treatments Assigned-Preferred and Self-Selected. In order to test hypotheses 1 - 3, we conduct a Wilcoxon rank-sum tests for treatment differences for each of the two tasks individually. To pool the data from the two tasks, we complement the non-parametric tests with a regression analysis. For this purpose we estimate the following model.

$$y_i = \alpha + \beta SelfSelected_i + \gamma X_i + \tau_t + \epsilon_i \tag{1}$$

where $y_i$ is the respective outcome of subject i, $\mathit{SelfSelected}_i$ indicates whether i was in treatment SELF-SELECTED or not, $X_i$ is a vector of control variables and $\tau_t$ is a task fixed effect. Parameter $\beta$ measures how the outcomes differ between the two treatments and thus yields the overall effect of task autonomy. We estimate this model on the pooled sample and for each task individually, with the exception of hypothesis 3. In a second step we add the control variables (described above).

## The Selection Effect

In order to investigate the selection effect, i.e. hypotheses 4-6, we compare the subjects in ASSIGNED-NOT-PREFERRED with those in ASSIGNED-PREFERRED. We use the same statistical analysis as above.

First, we conduct Wilcoxon rank-sum tests for each task separately.

Thereafter, we estimate a slightly adapted version of equation 1. In particular, we estimate the OLS model

$$y_i = \delta + \lambda \mathit{Preferred}_i + \rho X_i + \eta_t + \kappa_i \tag{2}$$

where $y_i$ is the respective outcome of subject i, $\mathit{Preferred}_i$ indicates whether i worked on her preferred task or not, $X_i$ is a vector of control variables and $\eta_t$ a task fixed effect. $\lambda$ estimates the effect of selection due to getting assigned the preferred task.

Again, we estimate this model on the pooled sample and for each task individually, with the exception of hypothesis 6. In a second step we add the vector of control variables.