# Flexible Work Arrangements and Precautionary Behavior: Theory and Experimental Evidence*

Andreas Orland†
Davud Rostam-Afschar‡

January 13, 2019

**Abstract**

We examine the relationship of saving and work time flexibility, i.e. workers can choose the relative length of work-shifts. We generalize the standard model to allow for endogenous non-separability over periods and show that allocation of work time and saving are perfect substitutes. Results from a laboratory experiment show that most of the predictions derived from the extended model are correct but that substantial heterogeneity between subjects exists. In particular, we show that (not predicted) static-model-behavior and predicted precautionary behavior coexist and that our most flexible model can explain a much larger part of observed behavior than the standard model.

**Keywords** precautionary saving, labor supply, intertemporal substitution, experiment

**JEL Classification** D14 · E21 · J22 · C91 · D81

*We thank Alan Auerbach, Peter Blair, Pedro Brinca, Lisa Bruttel, David Card, Christopher Carroll, Gary Charness, Stefano DellaVigna, Nadja Dwenger, Frank Fossen, Ulrich Glogowsky, Marcus Giamattei, Shachar Kariv, Amir Kermani, Patrick Kline, Dmitri Koustas, Peter Kuhn, Thomas Meissner, Georg Nöldeke, Luigi Pistaferri, Emmanuel Saez, Johannes Spinnewijn, Eric Swanson, and seminar and conference participants at University of Basel, University of California–Berkeley, University of Hohenheim, University of Potsdam, University of Nevada–Reno, San Francisco State University, Federal Reserve Banks of Minneapolis and Chicago, the 4th Maastricht Behavioral Economic Policy Symposium, the 1st Behavioral Macroeconomics Workshop in Bamberg, the 2018 ESA International Meeting, the SABE/IAREP 2018 Conference, and the 2018 BEAM–ABEE Workshop for valuable comments. The usual disclaimer applies. Declarations of interest: none.

†Department of Economics, University of Potsdam, August-Bebel-Strasse 89, 14482 Potsdam, Germany (e-mail: aorland@uni-potsdam.de).

‡Department of Economics, University of Hohenheim, 70593 Stuttgart, Germany (e-mail: davud.rostam-afschar@uni-hohenheim.de).

# 1 Introduction

Flexible work arrangements are increasingly common. The so-called gig economy, for instance, changed the environment for both labor supply and consumption decisions of many professionals from ride-sharing drivers, bicycle couriers, and craftspersons to data scientists, engineers, architects, accountants, consultants or designers. These professionals increasingly supply labor on online platforms like Lyft, Uber, Delivery, HomeAdvisor or Upwork without submitting to a specific work time schedule. In contrast to traditional work arrangements, working as a freelancer allows determining the (expected) wage of a shift and its risk endogenously even if the wage rate varies over time exogenously by choosing how to allocate a fixed amount of work time (in the following referred to as shifting). Therefore, by shifting work time such that the (expected) wage adjusts in a way that it smooths marginal utility over time, the same intertemporal substitution may be achieved as by saving through adjustment of consumption, leisure, or effort. If saving and shifting are substitutes, the increasing prevalence of variable work arrangements may reduce savings. At the same time, wages in the gig economy may be very volatile, giving rise to precautionary behavior. However, the extant literature (reviewed in the next section) is ambiguous on how important precautionary behavior is or— more importantly—whether it even exists in the first place. To know about these facts is crucial for the design of fiscal policies (e.g., Jappelli and Pistaferri 2014).

In this study, we provide theoretical and experimental evidence on behavior underlying many important questions raised by changes in work arrangements. First, we break new ground by showing in which sense labor supply and savings may be substitutes and discuss the mechanics of this alternative strategy for intertemporal substitution (that works through endogenous non-separability over periods) in an extension to the standard model for consumption and labor supply decisions. This channel may be very costly in practice because it can be viewed as a link between savings and the extent of tax avoidance in a progressive tax system.[1] Second, we provide clear evidence for the existence of precautionary saving from laboratory experiments, where we are able to control wage risk and to make sure that the only reason to save is precaution, which is not possible using survey or administrative data.[2] In fact, we deliberately chose a laboratory experiment to avoid the often

[1] Our model rationalizes the recent evidence that self-employed and workers are able to shift income to take advantage of a Swiss tax holiday (Martinez et al. 2018).

[2] Our study provides important evidence that complements and speaks to research with data from outside the lab. Many attempts to quantify causal effects in real-world data are plagued by problems like absence of credible mea-

epistemological discussion whether preferences or constraints or measurement errors drive the results. Instead, we induce preferences and only vary constraints on saving and work arrangements to see how choices are affected under ideal circumstances for theoretical predictions to hold.[3] This allows to identify heterogeneity due to different decision rules and to assess, e.g., whether the static or the intertemporal model predicts behavior better. As a third contribution, we show that not only precautionary saving but also precautionary shifting exists in our experimental data, and fourth, that these two channels are indeed substitutes, albeit not perfect substitutes, as theoretically predicted. In particular, we can classify two behavioral strategies: those who follow the model of intertemporal substitution as predicted and those who simplify the intertemporal decision problem and behave like hand-to-mouth consumers. This means that the true extent of precautionary behavior would be understated if these two groups are not considered separately. The absence of information on shift-specific wages in non-experimental settings until recently made it difficult to take shifting behavior into account and is thus a further possible explanation for why standard saving regressions using survey data have provided only mixed evidence on the importance of precautionary saving.

We set up a two-period model in which the wage (i.e., the piece-rate) is certain during the first period and perturbed by a mean-zero shock that either takes a high or a low realization with equal probability in the second. We induce preferences via incentives and compare decisions in a treatment where only effort but neither saving nor the length of work-shifts can be chosen to treatments where either a saving or a work-shift choice is allowed. While precautionary leisure cuts are not possible in any treatment, income is determined in a real-effort task and thus subjects could provide extra effort in the anticipation of risk. Independently of the existence of extra effort, a part of the income may be shifted from one shift to another via savings or work time allocation for the induced precautionary reasons. We formulate four hypotheses to show whether the hand-to-mouth model, the standard precautionary saving model, and the precautionary shifting model sustain the test with our experimental data and whether saving and shifting are indeed substitutes.

First, we examine whether subjects behave according to Jensen's inequality, which implies that they provide more effort in the certain first period than in the risky second. Theory could

sures of wage risk, effort, wealth or shift-specific information, the presence of many different saving motives or incentives provided, e.g., by saving subventions or tax policy, and the presence of heterogeneous decision rules. Given this, it is surprising that there are only very few studies that investigate heterogeneity of behavior in labor market experiments (see, e.g., Dickinson 1999) and of precautionary behavior.

[3] See Falk and Heckman (2009) for a more general discussion on the advantages and limitations of lab experiments.

already fail at this stage if subjects would not adjust effort to risk but we find that subjects follow this prediction and reduce effort in our setting by 14-24 percent. Second, does precautionary saving exist? The results resoundingly reject the non-existence of precautionary saving with at least 85 percent of the subjects engaging in precautionary saving. Is precautionary effort absent as predicted by the model? Non-parametric tests show that medians and distributions of effort are identical. Thus, there is no evidence for precautionary effort. Third, does precautionary shifting exist? At least 52 percent of all subjects engaged in precautionary shifting. Fourth, are saving and shifting perfect substitutes? While we find that around 30 percent of all subjects can be classified as substituters since their absolute deviations from optimal saving conditional on shifting choice are minimal, we reject that saving and shifting are *perfect* substitutes on aggregate. Another way to test this relies on the hypothesis that if they were perfect substitutes, expected payoffs would not significantly differ across treatments. However, expected euro earnings differ at all usual levels of significance.

Finally, we demonstrate that the most flexible model with saving and shifting predicts behavior reasonably well on the aggregate level and few subjects disregard effort costs. In particular, predicted and actual savings are not statistically different. However, the distribution of actual savings is widely dispersed, such that around 7% of subjects who follow hand-to-mouth behavior and others who oversave generate this result. Similarly, while shifting is observed to be important, the model systematically overpredicts its extent. This results from the fact that a non-negligible part of subjects follows hand-to-mouth behavior. We show that a combination of the static and our extended intertemporal model may predict behavior best and advocate using estimation procedures, like finite mixture models applied in this study, that take this into account.

These results provide grounds to study behavior under flexible work arrangements further. In practice, shifting may be more relevant for decisions on how much time to allocate to high versus low risk projects, particularly in professions with frequently changing or highly volatile wages. Also, for tax avoidance or if saving is not possible, e.g., due to use of a large part of income to serve tuition fees or very high cost of living in his or her area of residence, shifting may be relevant. However, our results also have more broad implications from a policy perspective beyond the freelancers or self-employed: If governments or labor unions decide to promote variable work arrangements (flexible hours or days) as an alternative to the traditional fixed, 40-hour work week, saving and thus economic growth may be reduced.[4]

[4] According to Koustas (2018) households would be willing to pay on average around 1,800 dollars per year for

The paper is structured in the following way: After providing an overview of the literature in the next section, Section 3 presents our extension of the standard model of consumption and labor supply and our behavioral hypotheses, Section 4 describes our experimental design and procedures, and Section 5 presents and discusses our findings. Finally, Section 6 summarizes and concludes.

## 2 Review of the Literature

To understand how shifting relates to saving (i.e., either to forgo utility by cutting consumption or incur extra disutility by cutting leisure or by providing more effort) in the context of precautionary behavior, a brief review of the literature is instructive. Precautionary saving is usually defined as the difference between consumption under certainty and in the presence of risk (see, e.g., Kimball 1990, p. 55). Some empirical evidence from survey data shows that this kind of precautionary behavior may be economically important (see Section A in the Appendix for an overview of results on the existence and importance of precautionary behavior complementing the excellent surveys in Jappelli and Pistaferri 2017 and Lugilde et al. forthcoming). Gourinchas and Parker (2002) attribute 60-70 percent of wealth to precautionary saving in early life. Kazarosian (1997) and Carroll and Samwick (1998) estimate the precautionary component of wealth to be in the range of 20-50 percent. However, the evidence is not unambiguous. With subjective earnings uncertainty, Guiso et al. (1992) estimate the precautionary component of wealth at only a few percentage points. Lusardi (1998, 1997) and Engen and Gruber (2001) find small precautionary wealth as well. Hurst et al. (2010) and Fossen and Rostam-Afschar (2013) argue that estimates are sensitive to whether business owners (who often have more flexible work arrangements than regular employees) are included in the dataset.

Some of the problems in survey data may be avoided in experimental settings, but the experimental literature on precautionary saving is relatively small. Fuchs-Schündeln and Schündeln (2005) show that, in accordance with a model that includes substantial precautionary effects, saving rates of most East Germans increased sharply after the natural experiment of the German unification, but saving rates of civil servants did not. By contrast, West Germans—who would have been subject to more selection into jobs based on risk preferences—exhibited little difference

flexible labor supply. Chen et al. (2017) report that UberX drivers, who have almost no restrictions on their work time flexibility, would reduce the hours they supply by more than two-thirds if required to supply labor inflexibly at prevailing wages.

in saving rates between civil servants and others with riskier jobs, either before or after reunification. Meissner and Rostam-Afschar (2017) show in a controlled laboratory environment that more than 50 percent of subjects simplify consumption decisions by ignoring incentives for precautionary behavior. Ballinger et al. (2003) study precautionary saving and social learning. They find that subjects save too little early in the life cycle. Still, the qualitative features of behavior are corroborated in their experiment, even if it misses the point predictions of the standard model with precautionary motives. Brown et al. (2009) test two explanations for apparent undersaving in life-cycle models: bounded rationality and a preference for immediacy, and find evidence for both.

With flexible hours of work a second channel emerges through which individuals may react in a forward-looking way: They may take into account their expectation about future wage risk when deciding how much to work in a given period (see Carroll and Kimball 2008, Flodén 2006 or Low 2005). Individuals with higher risk, e.g., self-employed, would work longer *before* the realization of shocks to accumulate precautionary wealth. Precautionary labor supply is then defined as the difference between work hours supplied in the presence of risk compared to certainty. As an alternative to precautionary consumption cuts, precautionary labor supply always increases savings.

On the empirical side, very little research has been devoted to precautionary labor supply. As reported by Mulligan (1998), "there is no empirical evidence that precautionary motives for delaying leisure are important" (p. 1034). Pistaferri (2003) finds that the effect of wage risk on labor supply is in line with theory, but in practice negligible. Jessen et al. (2017) find that the self-employed would work 4.5 percent less if they faced the same wage uncertainty as the median civil servant. Their estimates show that inflexible work arrangements constrain blue and white collar workers who can adjust hours only after about four and two years, respectively, and thus act less in anticipation of risk. Because of measurement issues, various important studies treat labor supply as synonymous with effort at work and time spent working (Heckman 1993, p. 116). However, already Marshall (1920) notes that "for even if the number of hours [of work] in the year were rigidly fixed, which it is not, the intensity of work would remain elastic" (p. 438). Therefore, we redefine labor supply as a function of two endogenous choices: effort and work-shift allocation (shifting). Accordingly, we define precautionary shifting as the difference in the relative length of a work-shift under risk and without risk when leisure is held constant.

While we focus on shifting, Eeckhoudt et al. (2012) and Wang and Li (2015) recognized that effort may also serve to accumulate precautionary savings. Empirically, precautionary effort has only been indirectly addressed by Huck et al. (2018). In their experiments, subjects work on a task and information about the two possible realizations of a piece-rate is either (i) resolved; (ii) unknown; or (iii) chosen to be learned by the subjects. About a third of the subjects are information avoiders in the last treatment who subsequently outperform the information seekers. Similarly, we define precautionary effort in this study as the difference between effort costs under (unresolved) risk and under certainty. Although it is straightforward to formulate a model that results in positive precautionary effort, we use a version that predicts absence of precautionary effort. Accordingly, we find only little evidence for it in our experimental data. While we stress that precautionary effort—like precautionary labor supply—*increases* savings, precautionary shifting is an *alternative* to saving. In other words, while labor supply (extra effort or less leisure) can be transformed into consumption and thus savings *intra*temporally, we show that savings and labor supply (i.e., the distribution of fixed work time) are substitutes *inter*temporally.

Our study is related to experimental research on labor markets, surveyed, e.g., in Charness and Kuhn (2011). Dickinson (1999) sets up a related model where workers can substitute on- and off-the-job leisure and tests it in experiments. In one of the conducted treatments, the subjects are only allowed to choose their effort (while work hours are fixed), and in another, they are also allowed to leave the experiment early. In both treatments, the piece-rate for the real effort-task is varied within-subject. The results confirm the predictions of the model: subjects in the lab experiments substituted leisure on- and off-the-job, which explains the negative substitution effects. Our model provides an alternative explanation for negative wage elasticities, since a wage increase may change the optimal allocation of work-shifts.

# 3 A Simple Two-Period Model

We propose a model of intertemporal choice that deviates in three aspects from the standard model of consumption and labor supply (e.g., Flodén 2006; Low 2005): the margin of intertemporal choice (relative length of work time), heterogeneity in productivity, and the specification of the valuation (work-shifts do not have to coincide with the time in which a certain wage is paid).[5] A fourth semantic but innocuous difference results from adoption of a slightly different terminology: we speak of a progressive tax function instead of a utility function returning consumption payoff.

[5] We build on insights on intertemporal choices from standard labor supply models in Heckman and MaCurdy (1980), MaCurdy (1981), and Blundell and Walker (1986) and in particular on the precautionary labor supply model analyzed in Flodén (2006). See Card (1994), Blundell and MaCurdy (1999), Keane (2011), and Jappelli and Pistaferri (2017) for surveys.

What usually is consumption, in turn, is referred to as income net of expenses for effort and savings. When we speak of income cuts is therefore equivalent to what is referred to by consumption cuts in the standard model. Income is then revenues from effort before effort costs and savings are deducted. This framing allows inducing preferences in a way that is easy to understand for subjects.

Labor supply is a function of two endogenous variables, since we distinguish between the provision of effort and the provision of work time. Supply of effort is total cost incurred during a given time and a work-shift is defined as calendar time spent working with continuous effort. Notably, a work-shift ends with the valuation of total work net of total effort costs accumulated during the work-shift.

Most importantly, our model generalizes the standard model by allowing the time span underlying intertemporal choices to differ from the periodicity of the decision environment. This keeps the decision environment exogenous but provides a choice for non-separable valuation. We only introduce uncertainty about the second-period wage rate in order to derive clear predictions that we subsequently test in our experiments.[6]

## 3.1 A General Model of Intertemporal Substitution

Consider a simple two-period model with two work-shifts, where ex-ante total consumption $C$ is the sum of consumption in work-shift 1, $c_1$, and expected consumption in work-shift 2, $E[c_2]$. In both work-shifts $i = 1,2$, consumption is a concave function of income $c(y_i)$ which can be interpreted as a progressive tax. The individual's problem is

$$\max_{y_1,y_2} C = c(y_1) + E_\varepsilon[c(y_2)]. \tag{1}$$

Work-shift specific income, in turn, depends on exogenously given, period-specific wage rates $w_j$ with periods $j = 1,2$ and three kinds of shift-$i$-specific choices: effort $e_i$, savings $s$, and choice of the relative length of the first work-shift $t \in [0,1]$.[7] Both periods' joint absolute length is exogenously fixed and lasts $T$ units of time. For simplicity and without loss of generality, we assume

[6] Our model also applies to cases where the wage changes occur not through time but concurrently, as long as the certain and the uncertain components can be separately related to work-shift allocations. For instance, consider bonus payment schemes, where on top of a fixed wage, a payment may be realized that varies with uncertain outcomes like annual revenues or the current stock price of a company. It is fairly straightforward to extend our model to a case in which the probability of receiving a bonus payment depends on one's productivity.

[7] Accordingly, the second shift's relative length is then $1-t$.

that each of the two periods takes $0.5 \times T$ units of time. At the beginning of the second period, the period-specific wage rate $w_{j=1}$ changes exogenously to $w_2$. The first-period wage rate is certain, $w_1 = w$, while the second-period wage rate $w_2$ is uncertain. In the second period, a mean-zero wage shock $\varepsilon$ shifts $w_2 = w \pm \varepsilon$ either up or down.

It is important to emphasize that the occurrence of the wage shock is only revealed after all decisions have been made. This renders it possible to isolate the effects of uncertainty. Furthermore, this represents a direct implementation of the models described in Flodén (2006), Parker et al. (2005), Hartwick (2000), and Eaton and Rosen (1980). This design feature is often an element of real-world settings. For example, to obtain a bonus payment, it might be necessary to allocate effort before the particular amount of payment is known (e.g., because it depends on the business cycle and is out of control of the individual worker). E.g., UberX drivers accept rides without neither knowing the distance of the ride nor the fee they receive (which depends in part on supply and demand in the local market they serve).

The choice of $t$ causes income $y_i$ in each shift to be determined by wage of period $j = i$ or wages of both $j = i$ and $j \neq i$. In particular, income in shift 1 is

$$y_1 = \begin{cases} y_1(t, w_1, e_1, s) & \text{if } t < 0.5 \\ y_1(0.5, w_1, e_1, s) & \text{if } t = 0.5 \\ y_1(t, w_1, e_1, w_2, e_2, s) & \text{if } t > 0.5 \end{cases} \tag{2}$$

and income in shift 2 is

$$y_2 = \begin{cases} y_2(t, w_1, e_1, w_2, e_2, s) & \text{if } t < 0.5 \\ y_2(0.5, w_2, e_2, s) & \text{if } t = 0.5 \\ y_2(t, w_2, e_2, s) & \text{if } t > 0.5. \end{cases} \tag{3}$$

This shows that by different choices of $t$, different period wage rates determine the shift-specific income in the generalized model and that $t = 0.5$ nests the standard model as a special case. While we assume shift-separable consumption, by choosing $t$ different from 0.5 implies non-separability over periods. Figures 1 and 2 illustrate this. In the standard model in Figure 1, labor supply is chosen according to only one (expected) wage prevalent in a given period. However, this setup precludes the possibility of influencing the expected wage by choosing the length of a work-shift.

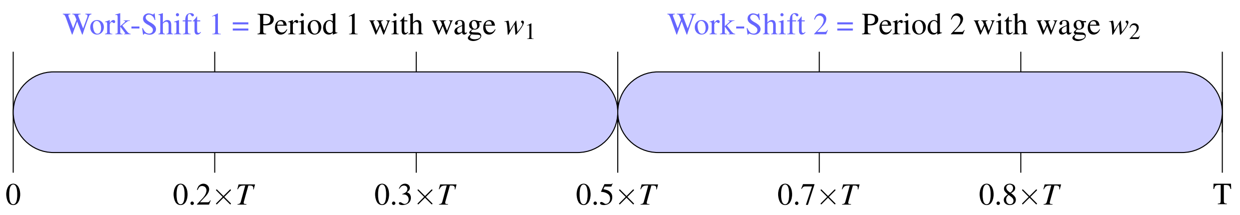



Figure 1: Labor Supply and Wage Changes in the Standard Model

*Source:* Authors' presentation.

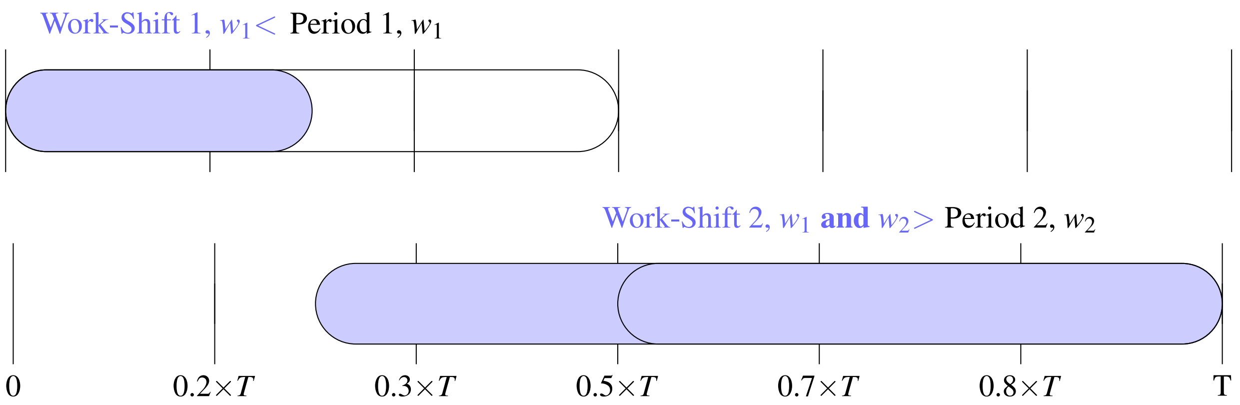



Figure 2: Labor Supply and Wage Changes in the General Model

*Source:* Authors' presentation.

Figure 2 shows an example where a worker decides to end the first work-shift early. Hence, he or she earns wage rate $w_1$ in this first work-shift. In the second work-shift, he or she earns both wage rates: $w_1$ for each unit produced in the time until $0.5 \times T$ and $w_2$ for each unit produced until $T$. In this way, we generalize the standard model, which does not take into account that the choice of labor supply may determine the (expected) hourly wage and that income may be valuated at the end of a shift—not at the end of a period.

**Precautionary motive** The choice of $t$ does not only change input factors to income but also determines which amount is evaluated in the concave consumption function $c(y_i)$. At the end of the first work-shift, first savings are chosen and then a tax resulting in after-tax consumption $c_i$ is imposed.

In each work-shift, after-tax consumption is related to income by a scaled and shifted isoelastic function

$$c(y_i) = \left( [1/(1-\tau)(y_i^{1-\tau} - 1)] - \eta \right) \zeta. \tag{4}$$

In this specification, parameter $\tau$ determines the degree of risk aversion (the concavity of the consumption function) and prudence (the convexity of marginal consumption), and its reciprocal $1/\tau$ determines the intertemporal elasticity of substitution induced by the tax system.[8] Most importantly for our analysis, this tax schedule leads to a positive third derivative implying prudence and thus risk will affect optimal choices.[9] Prudence is measured by the parameter of relative prudence (Flodén 2006; Kimball 1990). In our case, this parameter is $-y_i \frac{c'''}{c''} = 1 + \tau$. Accordingly, marginal after-tax consumption is higher when before-tax income is low, and the *rate* at which marginal after-tax consumption rises when before-tax income falls is greater when before-tax income is low than when it is high.

Given the precautionary motive, there are two margins of choice that reflect precautionary behavior, both of which we will analyze separately and jointly. First, as in the standard model, prudent individuals have an incentive to save in anticipation of wage risk. Insurance against wage risk is the only reason for saving in our experiment since the expected wage is identical in periods 1 and 2. The possibility to end a work-shift before or after a change in wage risk creates another route to engage in precautionary behavior: prudent individuals have an incentive to sacrifice some payoff in shift 1 and end it before the wage becomes uncertain in order to ensure that some income from the certain wage rate enters the consumption function in the second shift.

**Budget constraint** The second major difference to standard labor supply models is that we specify a function $c(y_i)$ that valuates benefits net of costs instead of the additive separable valuation of benefits and costs of work. Instead of describing the situation of employees, where disutility of work accrues privately (and is valuated in an additive separable way), in our model the costs of work can immediately be deducted as business expenses before valuation. This resembles the situation of self-employed or freelancers more closely. The main reason for this design feature is that it disincentivizes precautionary effort, i.e., higher effort in the first work-shift under uncertainty than under certainty. Since precautionary effort works via savings, its existence does not confound our results. A positive side effect is that the self-employment specification also requires fewer non-linear functions, which makes the experimental setup simpler to explain.

---

[8] The standard definition of a progressive tax function is that the marginal tax rate is larger than the average tax rate for every level of gross consumption.

[9] In order to maximize expected total consumption payoff, the worker must reduce—in our wording—income in the first shift and increase it in the second, so as to reduce the expected marginal consumption payoff of income in the second shift and increase it in the first. See Jappelli and Pistaferri (2017, p. 97) for details.

In particular (we present the other cases below), the shift-specific budget constraint $y_i$ for the case $t = 0.5$ is given by

$$y_1 = w_1 \times q(e_1) - v(e_1) - s, \tag{5}$$

$$y_2 = w_2 \times q(e_2) - v(e_2) + s. \tag{6}$$

Shift-specific effort $e_i$ translates into a production quantity according to a production function $q(e_i)$ from which costs of effort $v(e_i)$ are deducted. Income at the end of a work-shift, before saving, is then the product of wage times production minus effort costs. From this income, savings are deducted in the first work-shift and added in the second work-shift (without bearing interest).

**Production function** We do not impose a production function in the experimental design as we suspect productivity to be highly heterogeneous. In line with Gächter et al. (2016), who introduced the experimental real effort task, our aim is to estimate the production function from the experimental data. As Gächter et al. (2016), we will estimate a production function of the form

$$q(e_j) = \beta_1(e_j)^{0.5} + \beta_2(e_j)^2 + \gamma, \tag{7}$$

where $\gamma$ is fixed production, i.e., effortless output. As our structural predictions regarding the *level* of savings and shiftings depend on the production function, we will later estimate the production function to obtain point predictions. However, we will also investigate deviations from the optimal *share* of income saved or shifted as this measure does not depend on individual productivity.

**Cost function** We specify a cost function for effort that limits the optimal level of effort:

$$v(e_i) = \sum_{k=0}^{e_i} \varphi \times (k)^2. \tag{8}$$

At the beginning of each work-shift, the cost function is reset to zero. This and the quadratic form of the cost function resembles fatigue effects with increasing effort frequently specified in the labor supply literature (cf. Moffitt 1984). Note that we defined this cost function with our experiments in mind, where effort levels are discrete.

## 3.2 Treatments

Here, we describe the four scenarios that arise due to the two precautionary channels that workers can or cannot use. We will speak of treatments in order to use the same terminology as in the empirical part of this paper. Our objective is to test if and how subjects use the different precautionary channels. To do this, we simplify the setup dramatically for the first treatment group and allow more choices in subsequent treatments, as shown in Table 1.

Table 1: Treatments and Choices

| | Static | Intertemporal | | Choices | | |
|---|---|---|---|---|---|---|
| Treatments | Effort | Saving | Shifting | | | |
| TI | Allowed | Not Allowed | Not Allowed | $e_1, e_2$ | | |
| TII | Allowed | Allowed | Not Allowed | $e_1, e_2$ | s | |
| TIII | Allowed | Not Allowed | Allowed | $e_1, e_2$ | | t |
| TIV | Allowed | Allowed | Allowed | $e_1, e_2$ | s | t |

*Source:* Authors' presentation.

**Treatment I: No Intertemporal Choice** In reality, many consumers lead a 'hand-to-mouth' existence: they simply consume their net income and do not save (Kaplan and Violante 2014). This may be due to unsophisticated behavior (non-optimizing, or 'rule-of-thumb' consumers), or due to the inability to trade in asset markets because of infinitely high transactions costs.[10] In our experiment, we restrict subjects in Treatment I to be hand-to-mouth consumers, i.e., $t = 0.5$ and $s = 0$. In this way, we generate a 'control' treatment where intertemporal consumption smoothing is not possible. In this treatment, the only effect of wage risk is that the second period optimal level of effort is smaller than under certainty (because of Jensen's inequality induced by the concave tax function).

To find the optimal effort, a Lagrange function $\mathscr{L}^I$ for Treatment I can be written for each work-shift $i$ as

$$\mathscr{L}_i^I = E_\varepsilon[c(y_i, e_i)] + \mu^I(E_\varepsilon[w_i \times q(e_i) - v(e_i) - y_i]),$$

where the expectation operator $E_\varepsilon$ is only relevant for work-shift $i = 2$. As the two work-shifts are not connected via savings or time-allocation, each optimization can be considered separately. With fixed work arrangements, the only choice variable in this setting is effort $e_i$ in each work-shift.

[10]This corresponds to the static model discussed in section 3.1 in Keane (2011).

The first order conditions are (with partial derivatives denoted, e.g., in the form of $c_{y_i} = \frac{\partial c(y_i,e_i)}{\partial y_i}$, and the Lagrange multiplier denoted for treatment $k$ as $\mu^k$)

$$\frac{\partial \mathscr{L}_i^I}{\partial y_i} = E_\varepsilon[c_{y_i}] - \mu^I = 0,$$
$$\frac{\partial \mathscr{L}_i^I}{\partial e_i} = E_\varepsilon[c_{e_i}] + \mu^I(E_\varepsilon[w_i q_{e_i} - v_{e_i}]) = 0.$$

Hence, combining the two first order conditions, income and effort can be traded at a rate equal to the difference between valued marginal production and marginal costs.

$$E_\varepsilon[c_{y_i}(w_i q_{e_i} - v_{e_i})] = -E_\varepsilon[c_{e_i}]. \tag{9}$$

With the first order condition and the production function (as mentioned before, we will estimate the production function when we present our results), we can derive optimal levels of effort. With our experimental setting, we can test the following hypothesis:

**Hypothesis 1 – Instantaneous reduction of effort by risk:** *With valuation of income as a concave function, optimal effort will be smaller under uncertainty than under certainty because Jensen's inequality implies* $E_\varepsilon[c(w_2,\bar{e})] = E_\varepsilon[c(w_1+\varepsilon,\bar{e})] \leq c(w_1+E[\varepsilon],\bar{e}) = c(w_1,\bar{e})$*, i.e., effort in the second work-shift will be smaller than in the first work-shift.*

**Treatment II: Precautionary Saving** While hand-to-mouth behavior can be observed in many situations, another important behavioral tendency is to 'save for a rainy day' (a term originally coined by Campbell to describe the savings effect due to an expected change of income, but now used more widely, see Jappelli and Pistaferri 2017, p. 79). This understanding of precautionary behavior has received much attention in the literature, although, as mentioned in Section 2, the empirical evidence for it is mixed. The main purpose of Treatment II is to shed light on whether precautionary saving occurs at all. In each work-shift, the choice variable is effort $e_i$ and at the end of the first work-shift, the amount of savings $s$ is chosen.

By allowing that subjects save, the Lagrange function $\mathscr{L}^{II}$ for determining optimal effort changes only slightly compared to Treatment I: The intertemporal budget, not the intratemporal, constrains choices (see Equations (5) and (6)). Ex-ante, the sum of the net payoff from both periods is relevant. Abstracting from the borrowing constraint, $\mathscr{L}^{II}$ can be written as

$$\mathscr{L}^{II} = c(y_1,e_1) + E_\varepsilon[c(y_2,e_2)] + \mu^{II}(E_\varepsilon[w_1 \times q(e_2) + w_2 \times q(e_2) - v(e_1) - v(e_2) - y_1 - y_2]).$$

Compared to Treatment I, optimal behavior is subject to one more necessary condition, namely the net income Euler equation:

$$c_{y_1} = E_\varepsilon[c_{y_2}], \tag{10}$$

$$E_\varepsilon[c_{y_i}(w_i q_{e_i} - v_{e_i})] = -E_\varepsilon[c_{e_i}]. \tag{11}$$

In the absence of risk, the intertemporal condition ensures that the net payoff in both work-shifts will be smoothed. In our experiment, the expected wage in the second period is identical to the certain wage in the first period. Therefore, given productivity and effort, there is no reason for a difference in net payoffs in both work-shifts except for precautionary cuts of first-period income, i.e., $y_1(s > 0) < y_1(s = 0)$.

There are two things to note: First, under certainty, savings will be zero, but under uncertainty a strictly positive amount of savings is optimal. This shows that of the many reasons to save enumerated by Browning and Lusardi (1996), the precautionary motive is the only determinant of all savings. Second, if the third derivative of $c(y_i)$ is not positive, (precautionary) saving will be zero: $c_{y_1} = E_\varepsilon[c_{y_2}] = c_{E_\varepsilon[y_2]}$, such that $e_1 = e_2$ and $c(y_1, e_1) = c(y_2, e_2)$.[11]

**Hypothesis 2 – Precautionary saving and effort:** *i. Absence of precautionary effort: Savings do not result from increased effort in the first work-shift, which will be the same as under certainty and as in Treatment I, but from a deduction in the payoff-relevant income in the first work-shift.*

*ii. Existence of precautionary motive: In anticipation of risk in the second period, a strictly positive fraction of income will be saved.*

**Treatment III: Precautionary Shifting** While precautionary behavior through savings is a well-known concept, another option for shifting income intertemporally has been missed to date: precautionary time allocation. To the best of our knowledge, we are the first to consider this alternative mechanism, e.g., due to flexible work arrangements.

To study this mechanism, we change two particular features in the setup for Treatment III compared to Treatment II. First, we do not allow saving anymore, i.e., $s = 0$. Second, we introduce the possibility to choose when work-shift 1 ends. This changes the budget constraints for each work-shift such that each has now three cases as shown in Equations (2) and (3).

[11] Jappelli and Pistaferri (2017, chapters 4 and 6) explain the different model predictions with and without a positive third derivative of the consumption function in great detail.

The choice of $t$ determines which of the three cases is relevant. If $t = 0.5$, the periodicity of wage change coincides with the work-shift change such that income in work-shift 1 depends on the effort and wage in period 1. Income in work-shift 2 depends on the effort and wage in period 2. If $t < 0.5$, work-shift 1 is shorter than work-shift 2. All income in work-shift 1 is determined by the effort and wage in period 1. However, income in work-shift 2 is determined by the effort and wage of both period 1 and 2. If $t > 0.5$ income in work-shift 1 is determined by the effort and wage of both periods, while income in work-shift 2 depends only on the effort and wage in period 2. In this case, shifting is equivalent to borrowing, since income is shifted from period 2 to 1.

The task is to choose the length of work-shifts $t$ and, in each work-shift, effort $e_i$. The Lagrange function is

$$
\begin{aligned}
\mathscr{L}^{III/IV} &= c(y_1,e_1) + E_\varepsilon[c(y_2,e_2)] + \mu^{III/IV}\Bigg\{ \\
+\mathbb{1}_{\{t=0.5\}} &\times \Big[2\times t[w_1 \times q(e_1) - v(e_1)] - y_1 \\
&+ 2\times(1-t)E_\varepsilon[w_2\times q(e_2) - v(e_2)] - y_2\Big] \\
+\left(1-\mathbb{1}_{\{t=0.5\}}\right)\mathbb{1}_{\{t<0.5\}} &\times \Big[2\times t[w_1\times q(e_1) - v(e_1)] - y_1 \\
&+ 2\times(0.5-t)[w_1\times q(e_1) - v(e_1)] \\
&+ 2\times 0.5E_\varepsilon[w_2\times q(e_2) - v(e_2)] - y_2\Big] \\
+\left(1-\mathbb{1}_{\{t=0.5\}}\right)\left(1-\mathbb{1}_{\{t<0.5\}}\right) &\times \Big[2\times 0.5[w_1\times q(e_1) - v(e_1)] \\
&+ 2\times(t-0.5)E_\varepsilon[w_2\times q(e_2) - v(e_2)] - y_1 \\
&+ 2\times(1-t)E_\varepsilon[w_2\times q(e_2) - v(e_2)] - y_2\Big]\Bigg\},
\end{aligned}
\tag{12}
$$

where $\mathbb{1}_{\{\text{condition}\}}$ is an indicator that equals 1 if the condition is true. To keep this Lagrangian comparable to the two previous ones, we need to account for the two periods in our model, therefore we write *two* times the time allocation term. Compared to Treatment II, only the budget constraint is different: instead of subtracting a specific savings amount at the end of a shift, the choice of the length of a work-shift determines how much is subtracted from income. Thus, the intertemporal optimality condition is identical in Treatments II and III. For an optimizing decision-maker, both options are substitutes for intertemporal substitution. In practice, however, behavior may vary depending on whether saving or time allocation is allowed. Here, the first order conditions are given by:

$$c_{y_1} = E_\varepsilon[c_{y_2}], \tag{13}$$

$$E_\varepsilon[c_{y_i}(w_i q_{e_i} - v_{e_i})] = -E_\varepsilon[c_{e_i}]. \tag{14}$$

In the absence of risk, there is perfect smoothing, i.e., effort in both periods will be identical, and $t$ will be 0.5. Since intertemporal shifts can be achieved with the same costs as in Treatment II, under risk, both effort levels and the net payoff will be identical to that of Treatment II. Prudent decision-makers will find it optimal to finish work-shift 1 early in anticipation of the higher risk in the second period. The rate at which marginal after-tax payoff falls when before-tax income rises is greater when before-tax income is low than when it is high due to the convexity of marginal after-tax payoff. Therefore, prudent decision-makers will use the certain wage to build up a level of wealth before working under the uncertain piece-rate begins.

**Hypothesis 3 – Precautionary shifting:** *Existence of precautionary shifting in Treatment III: Work-shift 1 is shorter than work-shift 2.*

**Treatment IV: Precautionary Shifting and Saving** Under flexible work-arrangements, for instance, if self-employed, workers may have the option of both determining when to finish a work-shift and whether to store value over time. Therefore, we consider the case where both saving and work-shift allocation are allowed. The budget constraints change slightly in comparison to Treatment III because savings $s$ are not restricted to zero anymore and may be non-negative. The budget constraints are given in Equations (2) and (3). The Lagrange function and the optimality conditions, however, are identical to that of Treatment III (hence $\mathscr{L}^{III} = \mathscr{L}^{IV}$). This is because while savings enter the work-shift budget constraints, they cancel out in the intertemporal budget constraint. Therefore, the resources on which the decisions are based remain unchanged.

In the general setup of this analysis, shiftings and savings are in fact perfect substitutes. The comparison of the data from Treatments II and III shows whether subjects can actually achieve the same expected payoff by choosing the predicted work-shift allocation and the predicted amount of savings, respectively. Treatment IV goes one step further to analyze whether subjects do not only substitute the extreme cases but also choose combinations of work-allocation and saving. We can test the following hypothesis implied by the redundancy of saving.

**Hypothesis 4 – Precautionary saving and shifting:** *i. Either work-shift 1 is shorter than work-shift 2, or there are positive savings, or both, since choosing savings after the first work-shift and the time allocation between work-shifts are perfect substitutes.*

*ii. Identical choices and outcomes with intertemporal substitution: Effort levels are identical in Treatments II, III, and IV. Expected payoff is identical in Treatments II, III, and IV.*

# 4 Experimental Design and Procedures

Our experimental design follows our previously described model. Because we expect a high level of heterogeneity of productivity between the subjects that we want to account for, we use a within-subject design for our individual decision-making experiments. Before we describe the different stages of the experiment, we explain the real-effort task that resembles work in our stylized labor market situations.

**The task** We use the ball-catching task proposed by Gächter et al. (2016). Figure 3 shows an example screen of this task. In the ball-catching task, subjects are presented a rectangular box. There are balls hanging at the top of the box in four columns and a tray is positioned at the bottom of the box. As soon as subjects click the start button, balls fall down the screen in either one of the four columns at constant speed (probabilities are equal for the next ball to fall down in any column). Subjects earn the piece-rate $w_j$ within period $j$ by catching balls with the tray (hence, the expected work-shift revenue is $Er_i = q_{i1} \times w_1 + E(q_{i2} \times w_2)$ with $q_{ij}$ the number of caught balls in period $j$ and shift $i$). In order to catch the balls, the subjects can move the tray from one column to the other by clicking two buttons under the rectangular box labeled LEFT and RIGHT.

Moving the tray is costly in monetary terms; this can be interpreted as the labor effort employed in a shift. $e_i$ designates the number of movements in a shift.[12] To implement an increasing marginal cost of effort, we use the following unit cost function in each shift $i$: $\kappa(e_i + 1) = 0.1 \times (e_i)^2$ with $e_i + 1$ being the next movement and $e_i$ the number of movements so far.[13] At the beginning of each

[12] In contrast to many other real-effort tasks that are designed to be tedious for the subjects in order to "bring the task more in line with what people consider labor" (Charness and Kuhn 2011, pp. 243-244), the ball-catching task explicitly quantifies the cost of effort in monetary terms. Hence, we consider the ball-catching task ideal for our research questions. Even if subjects enjoy the task, the cost of effort should keep them from exercising more effort than necessary.

[13] We round the unit costs up to one integer in order not to confuse subjects with the decimals.

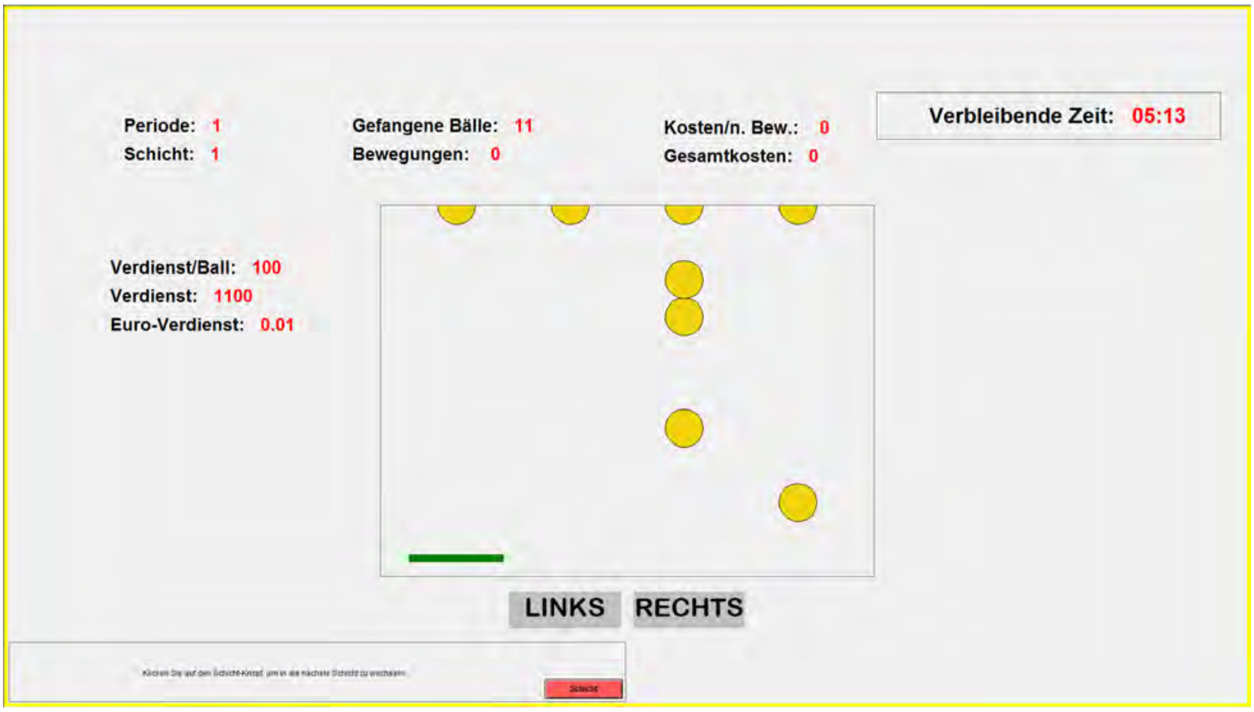



Figure 3: Example screenshot of the ball catching task (with 'shift change' button at the bottom)
*Source:* Authors' presentation.

work-shift, the unit cost function is reset, $e_i(0) = 0$. The total cost per shift is given by the sum of unit costs, $v_i(e_i) = \sum_{k=0}^{e_i} \kappa(k)$. Therefore, this task generates a tradeoff between the revenue from catching balls, $r_i$, and the total cost of effort, $v_i(e_i)$. The point earnings in any one of the two shifts are then given by revenue minus cost, $y_i = r_i - v_i - s$, where $s$ denote savings. The euro earnings in each shift are calculated by $\text{Euro}_i = 4 \times [\ln(y_i) - 7]$.[14] This payoff function implies a coefficient of relative risk aversion of $\tau = 1$ and of relative prudence of $\tau = 2$. This is in line with estimates in Fagereng et al. (2017, p. 396) using Norwegian administrative data. The variables number of caught balls, unit cost of the next movement, total cost, point earnings per ball, and total point and euro earnings in the current work-shift are continuously updated on-screen during the task. Once the task is started by pressing the start button, it cannot be paused. When the work-shift ends, a feedback screen with the statistics mentioned before is shown.

**Part 1 of the experiment: Trial periods** During this phase, we let subjects play three incentivized trial periods so they can familiarize themselves with the user interface and mechanics of the task. Only one of the three trial periods is chosen randomly for payoff and feedback about the

[14]This logarithmic utility function is a special case of the shifted isoelastic utility function in Section 3.1: $\lim_{\tau \to 1} \left( [1/(1-\tau)(y_i^{1-\tau} - 1)] - \eta \right) \zeta = (\ln(y_i) - \eta)\zeta$.

chosen period is only shown at the very end of the experiment.[15] In a first trial period, we deviate from the costly effort-incentive structure and make movements costless. We also abstract from the concave consumption function. Subjects are given 180 seconds to catch balls, with each caught ball generating earnings of 1 euro cent. There is no tradeoff between the returns from catching balls and the monetary costs of effort in this trial period. In the following two trial periods (and for the rest of the experiment), subjects work with the concave consumption function, the convex cost function, and the point earnings outlined before. In the second trial period, subjects work on the task and earn the certain wage, $w = 100$ for 180 seconds. In the third trial period, subjects work under uncertainty and either earn the low or the high wage, $w = 20$ or $w = 180$, with equal probability, for 180 seconds.

**Part 2 of the experiment: Main treatments** In Part 2, we conduct the four main treatments in the order in which we described them in Section 3.[16] In the instructions and on-screen we talk about four rounds, not about treatments. Here we will stick to the term 'treatments'. Each of the four treatments consists of two periods of 180 seconds each (the first one with the certain wage, the second with the uncertain wage) and two work-shifts (which are defined as the time where subjects work without a break on the task). Only one of the four treatments is chosen randomly for payoff. Furthermore, feedback about the chosen treatment is only presented at the very end of the experiment.

**Treatment I** This is the simplest treatment as subjects have neither the savings nor the time allocation option at their disposal. Subjects work on the task for two work-shifts that coincide with the periods (à 180 seconds). In the first work-shift, subjects earn the certain piece-rate $w_1 = 100$. In the second work-shift, they work under uncertainty and earn either the high rate, $w_2 = 180$, or the low rate, $w_2 = 20$. The instructions stress that the probability for the low or high rate is equal and independently drawn in each of the treatments.

[15]This is a common technique in order to avoid portfolio effects when subjects make multiple decisions. It also helps us to keep each decision salient by paying a relatively high amount per decision. See Charness et al. (2016) for a discussion of paying one or few decisions vs. paying all.

[16]We thought about randomizing the order of the four treatments between subjects. But by sticking to the order in Section 3, we gradually increase the level of difficulty. Thereby, we intend to limit our subjects' confusion that could occur due to, e.g., taking both a shifting decision and then a saving decision in the first round.

**Treatment II** This treatment differs from Treatment I only in the savings decision. If subjects earned a positive euro amount in the first work-shift, they enter a screen where they can calculate the consequences of hypothetical saving decisions with a slider. They then enter the number of points they would like to save in a separate box. (The savings amount has to be non-negative, $s \geq 0$, and the highest amount that subjects can save is limited so that the euro earnings in work-shift 1 cannot become negative). After that, they press the OK button and proceed to the second work-shift. The amount of points saved is then deducted from the point earnings of the first work-shift and added to the point earnings in the second work-shift. See Figure E.1 in Section E of the Appendix for a screenshot of the savings screen.

**Treatment III** This treatment differs from Treatment I in the time allocation between the two work-shifts. Subjects can freely divide the total amount of time, $T = 360$ seconds, between the two work-shifts. This is implemented in the following way: In work-shift 1 subjects are shown a button that allows them to immediately switch to work-shift 2 at any point of time (see Figure 3 for a screen-shot of a first work-shift with the switch button at the lower left corner of the task screen). The time remaining of the initial 360 seconds is then spent in the second work-shift. As soon as subjects enter the second period, the low wage's point revenue and euro earnings are displayed on the left-hand side of the task box, and the high wage's on the right-hand side of the task box.

**Treatment IV** In this treatment, both the savings decision of Treatment II and the time allocation of Treatment III are available to the subjects. First, subjects have to decide when to end work-shift 1. After being shown feedback on their outcomes in work-shift 1, subjects enter the savings screen where they can enter their savings decision.

**Part 3 of the experiment: Elicitation of risk aversion and prudence** In order to elicit the risk aversion and prudence of the subjects, we consecutively presented them with 12 binary choices between lotteries, as suggested by Noussair et al. (2014). Figure E.2 in Section E of the Appendix shows an example screenshot. Due to the potentially very high payoff of up to 165 euros, each subject only had a 1 in 20 chance of being randomly selected to receive a monetary payment from Part 3 of the experiment (which lasted about five minutes).

**Post-experimental questionnaire** In the post-experimental questionnaire, we asked the subjects for their gender, age, field of study, their number of semesters at university (including undergraduate studies), and how strenuous they perceived the experiment to be (on a scale from 1 (not at all strenuous) to 6 (very strenuous)). We also asked for a subjective self-assessment of their general level of risk aversion (the wording of this question is identical that used by the German Socio-economic Panel (SOEP), with answers ranging from 0 (not at all willing to take risks) to 10 (very willing to take risks)). Moreover, we asked whether subjects knew anybody who previously participated in this experiment and whether they tended to pay attention to the low or to the high piece-rate in the periods where the rate was uncertain.[17]

**Procedures and subjects** Upon arrival at the laboratory, the subjects were seated in separate booths. Then, the subjects received printed instructions which included tables with selected values and graphs of the cost and consumption functions. After reading the instructions, the subjects had to answer a set of control questions correctly in order to proceed.[18] The experiment was computerized. Only after the subjects completed the three parts of the experiment and answered the questionnaire they did receive feedback about the outcomes of the experiment and their euro earnings. Finally, the payoff took place privately in a room separate from the other subjects.

All experiments were conducted in PLEx, the Potsdam Laboratory for Economic Experiments at Universität Potsdam, in November and December 2017. All 192 subjects were students of Universität Potsdam and other nearby universities (Freie Universität Berlin, Filmuniversität Potsdam, and University of Applied Sciences Potsdam). Subjects were invited using ORSEE (Greiner 2015). The experiments were run on z-Tree (Fischbacher 2007), in 19 sessions of 4 to 14 subjects (depending on enrollment to the experimental sessions and attendance of subjects). The laboratory sessions took about 90 minutes. On average, subjects earned about 15 euros (with a minimum of 0 euros and a maximum of 66.20 euros).[19] Section F in the Appendix presents a summary of our sample and discusses it briefly.

[17] We use both items from the post-experimental questionnaire and the elicitation of risk aversion and prudence from Part 3 of the experiment to explain behavior in the main part of the experiment in Section 5.2.

[18] You can find a translation of the instructions, the supplied tables, and the control questions in Sections B, C, and D of the Appendix.

[19] One might criticize the low euro earnings in this experiment. As mentioned in the introduction, some subjects disregarded effort costs. In fact, 23 of the subjects earned zero euros. If we calculate the average without them, each subject earned about 17 euros. However, we do not exclude them from our analysis.

# 5 Results

Before we turn to the results, we illustrate the most important data in one figure. Figure 4 shows the expected payoff across and within treatments conditional on observed behavior. One point represents one subject's performance in one treatment. All subjects in Treatment I are piled up in one dimension at exactly half of the work-shift axis since by design each of the work-shifts is restricted to 180 seconds. Depending on effort choices in each work-shift, earnings range from very low (blue circles) to very high (red triangles).

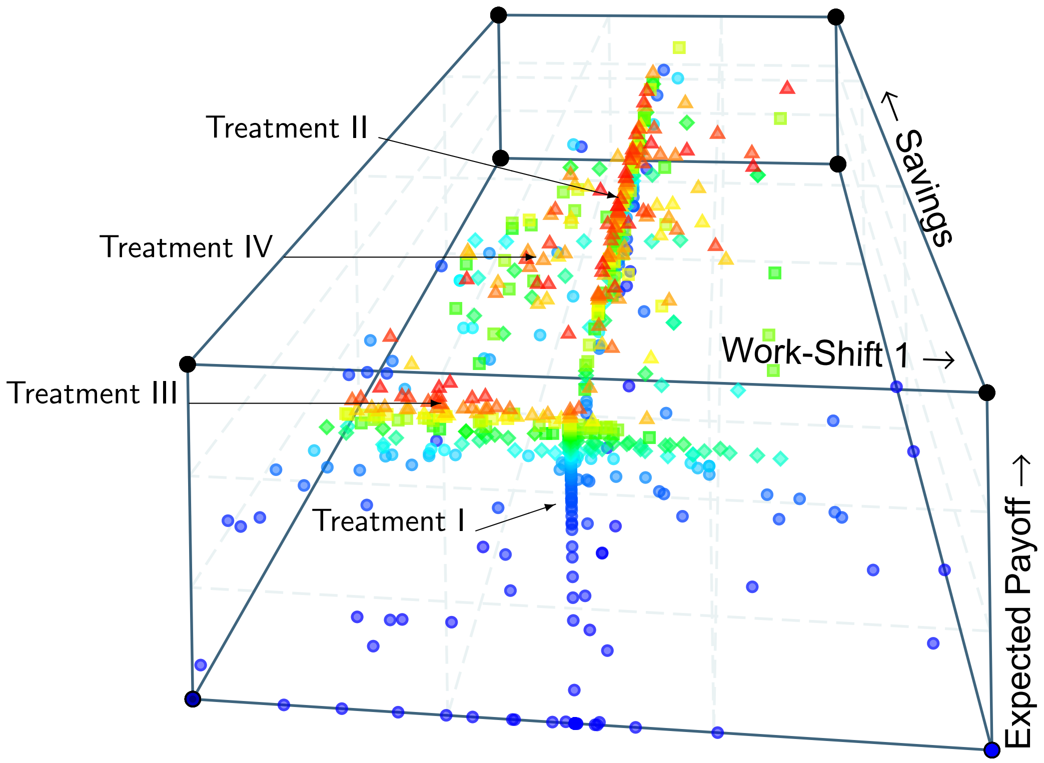



Figure 4: Work-Shift-Savings-Payoff Space

*Legend:* Payoffs range from very low (blue circles) to very high (red triangles). The color changes with the distribution and each of the four markers covers a quartile of expected payoffs from 0 euros to 14.89 euros. The fraction allocated to work-shift 1 ranges from 0 to 100 percent, savings from 0 points to 6,473 points.

*Source:* Own presentation using GRAPH3D.

Treatment II allows subjects to save, therefore subjects are scattered in the two dimensions spanned by the savings-axis and the vertical expected earnings axis. Clearly, those who do not save sufficiently or save too much cannot achieve high expected earnings (blue circles or green squares); at about 1/3 of the length of the saving axis (at about 2,000 points saved), there is a cluster of red triangles, indicating that expected earnings are high. Treatment III does not allow saving but does allow shifting (equivalent to saving if work-shift 1 ends before half of the total time and borrowing otherwise).

Thus, a two-dimensional space is once again spanned, this time by the horizontal work-shift axis and the expected earnings axis. Those who chose to end work-shift 1 at about 1/3 of the work-shift axis (at around 130 seconds) expect the highest earnings, as illustrated by the red triangles while choosing to end shift 1 earlier or later reduces expected earnings substantially. Finally, Treatment IV allows subjects to scatter in a three-dimensional space because both saving and shifting are allowed. The figure shows that subjects on or close to the theoretical line of substitution between shifting and saving—spanned between the expected income peaks of Treatment II and Treatment III (and a little beyond that point)—expect highest earnings. Figure G.1 in the Appendix shows the same graph with colors and markers indicating treatments, not expected payoff.

## 5.1 Aggregate-Level Results

### 5.1.1 Tests of Hypotheses 1 to 3

In this section, we discuss our results regarding Hypotheses 1 to 3. Table 2 reports the main results. First, we turn to Hypothesis 1. The results of t-tests suggest that there is the theoretically predicted direct effect of risk: in Treatment I we observe significantly fewer movements in work-shift 2 than in work-shift 1 (p-value$<0.01$). This difference is also reflected in the logarithm of effort costs; again, equality is clearly rejected (p-value$<0.01$). This provides strong evidence for Hypothesis 1. Figure 5a illustrates this. The distributions of log effort costs and movements are clearly different; the median values of each distribution, as indicated by the blue and red vertical bars, differ substantially (Wilcoxon rank-sum test has p-value$<0.01$ and Pearson $\chi^2$ test of the equality of the medians has p-value$<0.01$).

**Result 1:** *In Treatment I, effort in the second work-shift is smaller than in the first work-shift. Subjects behave according to Jensen's inequality and reduce effort under uncertainty by 14-24 percent.*

Table 2: Tests of Hypotheses 1-3

| | **H1: Effort Smaller in Second Work-Shift than in First Work-Shift of TI** | | **H2(i): Absence of Precautionary Effort (Higher First Shift Effort in TII)** | |
|---|---|---|---|---|
| | Movements | Log Effort Cost | Movements | Log Effort Cost |
| Mean | TI Shift 1: 32.71 | TI Shift 1: 6.66 | TI Shift 1: 32.71 | TI Shift 1: 6.66 |
| Mean | TI Shift 2: 26.54 | TI Shift 2: 5.99 | TII Shift 1: 30.73 | TII Shift 1: 6.46 |
| Δ 95% Conf. | 4.61 to 7.75 | 0.52 to 0.83 | 0.37 to 3.59 | 0.05 to 0.35 |
| t-test p-val. | <0.001 | <0.001 | 0.016 | 0.010 |
| $\chi^2$ p-val. | <0.001 | <0.001 | 0.540 | 0.540 |
| Rank-sum p-val. | <0.001 | <0.001 | 0.358 | 0.358 |

| | **H2(ii): Proportion With Savings Higher than 100 Points Greater Zero** | | | | **H3: Proportion With Work-Shift 1 Shorter than 180 Seconds Greater Zero** | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std. Err. | 95% Conf. | p-val.= 5% | Mean | Std. Err. | 95% Conf. | p-val.= 5% |
| TII | 89.58 | (2.20) | 85.26 to 93.90 | $H_0: 93.16$ | — | — | — | — |
| TIII | — | — | — | — | 58.85 | (3.55) | 51.89 to 65.81 | $H_0: 65.57$ |
| TIV | 86.98 | (2.43) | 82.22 to 91.74 | $H_0: 91.02$ | 47.40 | (3.60) | 40.33 to 54.46 | $H_0: 54.44$ |

*Notes:* P-values are from t-tests with the hypothesis that the average difference of observations Δ is zero, from Pearson $\chi^2$ tests with the hypothesis that the two independent samples were drawn from populations with the same median, and Wilcoxon rank-sum test with the hypothesis that the samples are from populations with the same distribution. All proportions are given in percent. Proportions in column labeled "p-val.= 5%" are obtained from tests of proportion with p-value equal to 5%.
*Source:* Own calculations.

Next, we discuss Hypotheses 2(i) and 2(ii). Hypothesis 2(i) states our expectation that precautionary effort does not exist. Again, we use t-tests to test our hypothesis. The results show that movements in both work-shifts are statistically different (p-value=0.016). Although this supports the notion that some subjects might have tried to exercise precautionary effort, the difference is not economically relevant.[20] Similarly, the logarithm of effort costs is statistically different (p-value= 0.01). The confidence interval shows that this difference is not in the order of magnitude of the direct effect of the wage risk (Hypothesis 1). This provides some evidence for Hypothesis 2(i). Figure 5b illustrates this. The medians lie exactly on top of each other (Pearson $\chi^2$ test of the equality of the medians has p-value= 0.54) and also the distributions do not differ much (Wilcoxon rank-sum test has p-value= 0.36).

Table 2 also presents the evidence for precautionary saving. It shows that Hypothesis 2(ii)—that states that precautionary saving exists—can not be rejected in a test of proportions ($H_0$: 93.16

[20]In fact, as we show later subjects caught slightly fewer balls per move in all first shifts. If anything, this is evidence against precautionary effort.

(91.02) percent of the subjects saved more than 100 points[21] in Treatments II (IV), p-value$= 0.05$). In fact, at least 85.26 percent (82.22 percent) behaved in this way in Treatment II (Treatment IV) according to the 95 percent confidence interval. This provides strong evidence for Hypothesis 2(ii). Both the upper and lower bound of the confidence interval are smaller in Treatment IV compared to Treatment II. This suggests that the possibility to end the first work-shift earlier allowed some subjects to substitute shiftings entirely for savings, and save less than 100 points in Treatment IV. Figure 5c displays the saving amounts in Treatments II and IV. It shows that 20 (25) subjects out of the 192 chose not to save in Treatment II (IV). These observations could have followed a different behavioral strategy than implied by the model. We denote the optimal choice under the strategy where not to save or to shift income is optimal as hand-to-mouth behavior $\vartheta(s_{sr}/y_{sr})_{\text{hand-to-mouth}} = 0$, where $s_{sr}$ is the amount saved or shifted and $y_{sr}$ is income. Moreover, the medians of the amounts saved (blue and red vertical lines) are of similar magnitude as the means reported in Table 4, later in the paper. To allow for comparisons across subjects conditional on their individual income and to the optimal share $\vartheta(s_{sr}/y_{sr})_{\text{optimal}}$ of 23 percent,[22] we define the share of income saved as $s_{sr}/y_{sr} = \frac{\text{savings}}{\text{income in shift 1}} \times 100$. Figure 5d presents this variable, indicating with the left gray bar the optimal share of income saved and with the right gray bar the share that to be chosen under the assumption that subjects are risk pessimists, i.e. focus only on the low wage. Then $\vartheta(s_{sr}/y_{sr})_{\text{risk pessimist}} = 35$ percent.

**Result 2:** *i. As predicted, precautionary effort is virtually absent, since median effort and thus effort costs are identical whether or not saving is allowed.*

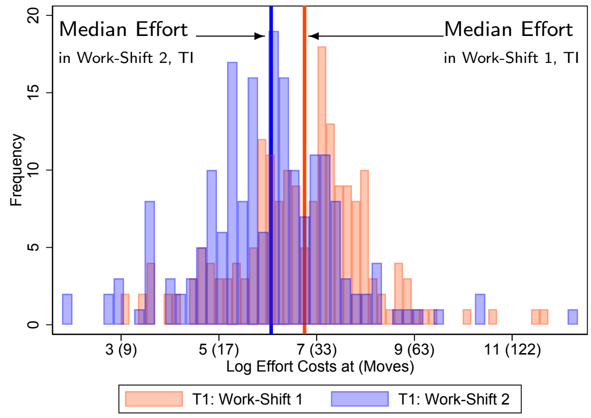
*ii. In anticipation of risk in the second period, savings are strictly positive for at least 85 percent of subjects in Treatment II (and 82 percent in Treatment IV).*

Next, we test Hypothesis 3. A test of proportions strongly rejects the hypothesis that precautionary shifting does not exist. Even the hypothesis that 66 (54) percent of the subjects in Treatment III (IV) shift work time (in order to insure against risk) can only be rejected at a relatively high level of significance (p-value$= 0.05$). According to the 95 percent confidence interval,

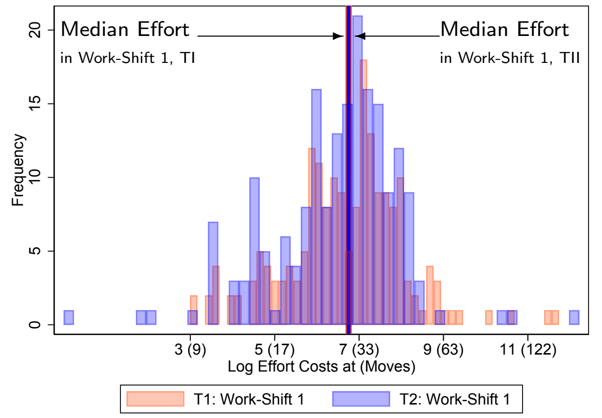
---

[21] We chose 100 points as the threshold in order to be far enough away from the zero-lower bound but the results change very little if we use $> 0$ points or $> 1$ point as criterion.

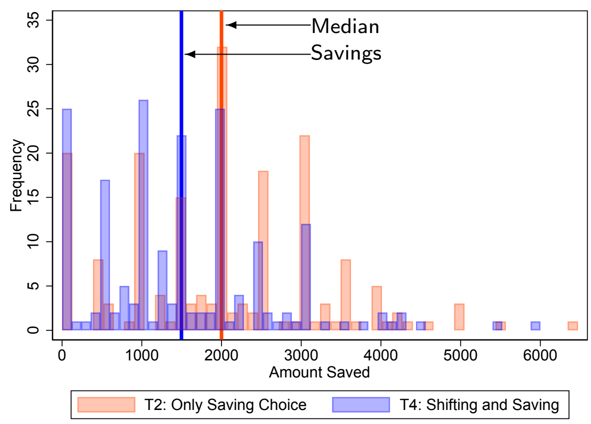
[22] The optimal share of income saved or shifted results from the theoretical model; it is independent of income and thus of the ability to catch balls.



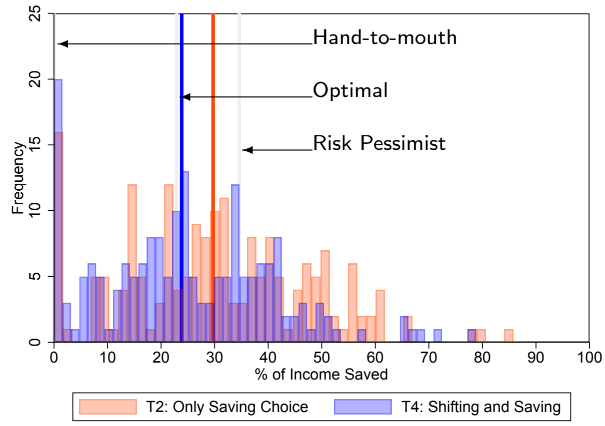

(a) H1: Effort Smaller in Second Work-Shift



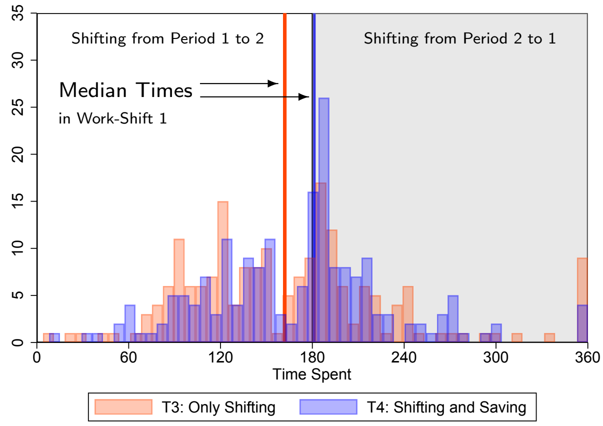

(b) H2(i): Absence of Higher First Shift Effort



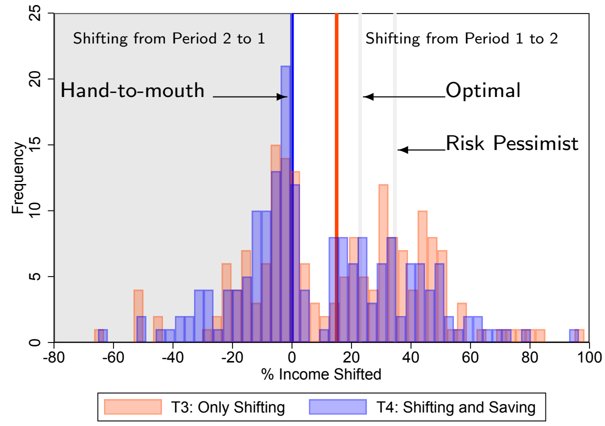

(c) H2(ii): Precautionary Saving




(d) H2(ii): Savings Share of Income




(e) H3: Precautionary Shifting




(f) H3: Shifting Share of Income

Figure 5: Tests of Hypotheses 1-3

*Notes:* Blue and red vertical lines represent medians, gray bars show optimal values, and gray background indicates borrowing.

*Source:* Authors' presentation.

at least 51.89 percent (40.33 percent) of the subjects ended their first work-shift before 180 seconds in Treatment III (Treatment IV). This provides clear evidence in favor of Hypothesis 3. We can also see that a smaller fraction of subjects chose to shift in Treatment III than in Treatment IV. Once the possibility of intertemporal substitution via savings was given, many subjects substituted finishing work-shift 1 early with savings — this forestalls Hypothesis 4.

Interestingly, the difference of the average proportions across treatments is smaller for saving than for shifting ($89.58\% - 86.98\% < 58.85\% - 47.40\%$). This suggests that some subjects treat these two choices as equivalent or that they find it harder to determine the optimal end of the work-shift than to chose the optimal amount of savings. Figure 5e shows the time spent in work-shift 1. The median work-shift choice indicated by the solid red bar is less than 180 seconds in Treatment III, where only shifting is possible. However, the median in Treatment IV is exactly at 180 seconds. Both distributions have two local peaks, one at 180 seconds and one at 120 seconds in both treatments. To be able to compare behavior across treatments, we convert time into the share of income in work-shift 1 that was shifted by calculating the share of income shifted as $s_{sr}/y_{sr} = \frac{\text{income in period 1} - \text{income in shift 1}}{\text{income in period 1}} \times 100$ for positive incomes. Figure 5f presents this variable. In this figure, we can see that the median share of shifted income in Treatment IV is exactly zero percent and less than the optimal share of 23 percent in Treatment III, where only shifting is possible. Abstracting from the gray area, where subjects could borrow, allows comparing the distributions resulting from the shifting and saving treatments. In all cases, distributions have two local peaks, one at zero percent and one near the optimal share of income cuts. This suggests that subjects choose at least two distinct strategies.

**Result 3:** *Work-shift 1 is shorter than work-shift 2 for at least 52 percent of subjects in Treatment II (and 40 percent in Treatment IV).*

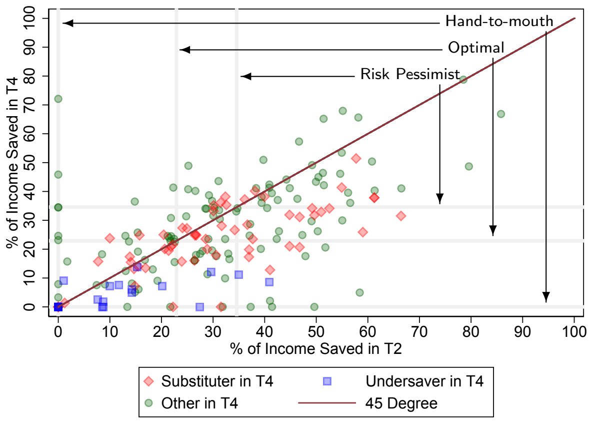
### 5.1.2 Test of Hypothesis 4: Is Shifting a Substitute for Saving?

In this section, we test whether and how subjects substitute savings for shiftings. Figure 6 presents visual evidence for how subjects substitute the two methods of intertemporal substitution. Figure 6a plots the share of income saved in Treatment II, where only saving is allowed, on the horizontal axis and the same figure from Treatment IV, where both saving and shifting are allowed, on the vertical axis. The solid red line shows all points where the shares in both treatments are identical. Strikingly, 64 percent of all saving decisions are below the 45-degree line (though many observations appear on and above the 45-degree line, 4 and 32 percent, respectively). This is consistent with the substitution of saving and shifting, which we will discuss in further detail with reference to different groups indicated in the legends of Figure 6.

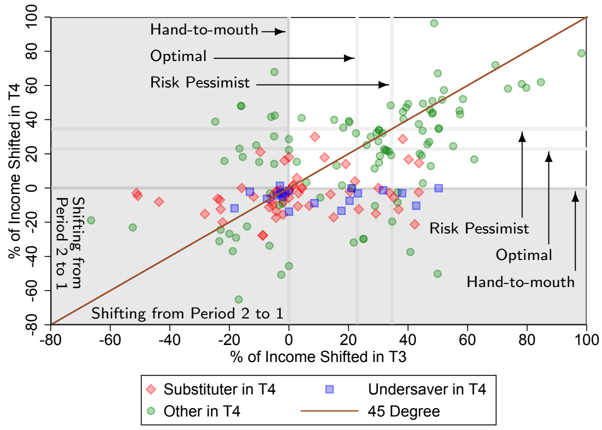
Figure 6b shows similar evidence for shifting in Treatments III and IV. The gray area shades all observations of subjects who either took a credit by shifting income from shift 2 to shift 1 or who did not shift at all. The observations outside of the gray area imply positive income cuts and are thus directly comparable to Figure 6a. Here the share of income shifted in Treatment III is presented on the horizontal axis and the same figure in Treatment IV on the vertical axis. While 56 percent of all subjects do not save via shifting of work time, 44 percent do so. Forty-six percent of the data points lie on or below the 45-degree line. This shows that most subjects *substitute* shifting. We identify three cases: either, (i), subjects prefer shifting over saving and choose to be on or close to the 45-degree line or, (ii), they prefer saving over shifting (but use shifting when saving is not possible) and hence end work-shift 1 at 180 seconds in Treatment IV while performing shifting in Treatment III or, (iii), they do not substitute at all but just replicate the restriction of Treatment I in both Treatment III and IV, resulting in the cluster in the lower left corner of the non-shaded area.[23] The last group seems to abstract from risk to simplify the intertemporal model and behave according to the predictions of $\vartheta(s_{sr}/y_{sr})_{\text{hand-to-mouth}}$. The willingness to sacrifice payoff in order to reduce complexity is apparent since the average (and the median) subject in this group chose slightly more than 180 seconds in Treatment III.

Finally, Figure 6c focuses on Treatment IV, where both saving and shifting are allowed. This figure shows combinations of shifting share (horizontal axis) and saving share (vertical axis). The solid blue line is the connection between the theoretical predictions, marked by the gray lines, which we derive via the structural model. Points on the blue line yield the highest payoffs. At first glance, some subjects are very close to the solid blue line, which shows that they substituted almost perfectly.

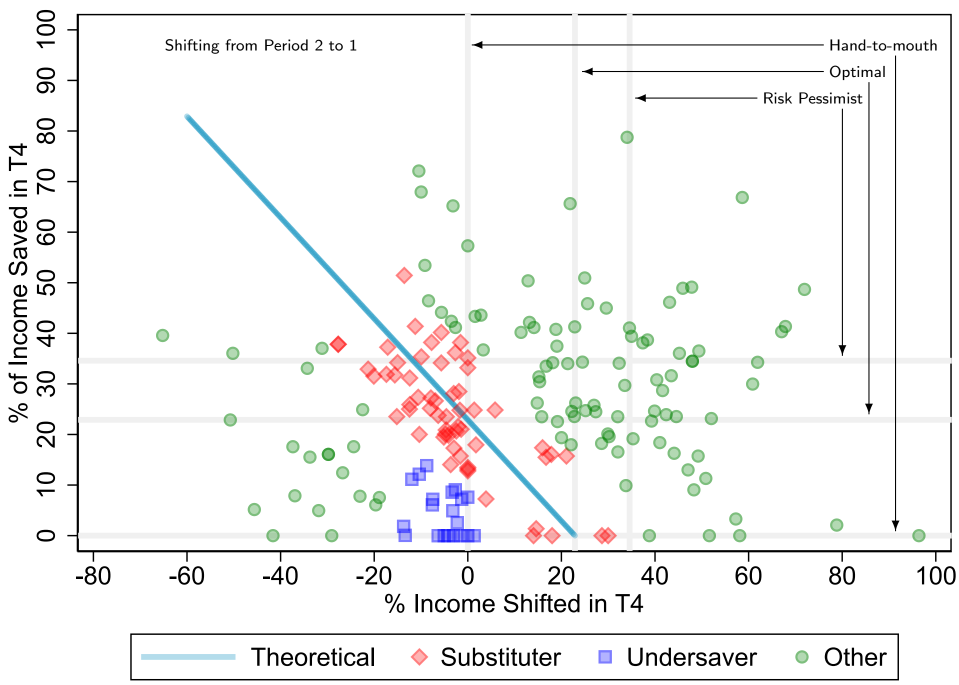
[23] A description of how we categorized the subjects' behavior is in order. Six percent of all subjects can be classified into group (i) having saved only between -5 and 5 percent of income in Treatment IV and having shifted more in Treatment IV, 11 percent of all subjects in group (ii) having finished work-shift 1 with $<$ 180 seconds in Treatment III and chosen to shift between -5 and +5 percent of income but saved more in Treatment IV, and (iii) nine percent of all subjects chose to shift only between -5 and +5 percent of income in both treatments. Thirty-one percent of all subjects cut income by more than 5 percent by shifting and saving, respectively, in Treatment IV. Thirty-nine percent of all subjects borrow or do not shift income in Treatment III.




(a) H4(i): Less Savings if Shifting is Allowed




(b) H4(i): Longer First Work-Shift if Saving Allowed




(c) Shifting as a Substitute for Saving

Figure 6: Test of Hypothesis 4

*Notes:* Gray bars indicate optimal values and gray background indicates borrowing.

*Source:* Authors' presentation.

Table 3: OLS regression of euro earnings on treatment dummies

| | Full Sample | | | Income Cut≥0 | | |
|---|---|---|---|---|---|---|
| Euro earnings | Expected | Low | High | Expected | Low | High |
| Treatment I | (baseline) | (baseline) | (baseline) | (baseline) | (baseline) | (baseline) |
| Treatment II | 2.43***,b | 5.01***,b | -0.14 | 2.43*** | 5.01*** | -0.14 |
| | (0.41) | (0.58) | (0.37) | (0.41) | (0.58) | (0.37) |
| Treatment III | 1.09**,a,c | 2.79***,a,c | -0.61 | 1.48** | 3.43***,c | -0.46 |
| | (0.52) | (0.68) | (0.52) | (0.63) | (0.84) | (0.61) |
| Treatment IV | 2.09***,b | 4.69***,b | -0.51 | 2.40*** | 5.06***,b | -0.25 |
| | (0.54) | (0.68) | (0.53) | (0.61) | (0.74) | (0.61) |
| Constant | 8.76*** | 2.38*** | 15.14*** | 8.76*** | 2.38*** | 15.14*** |
| | (0.71) | (0.84) | (0.67) | (0.71) | (0.84) | (0.67) |
| Adjusted $R^2$ | 0.010 | 0.039 | -0.003 | 0.012 | 0.043 | -0.004 |
| Observations | 768 | 768 | 768 | 665 | 665 | 665 |

Robust standard errors clustered at subject level.
Significantly different from zero at the 1%-level: ***, 5%-level: **.
Significantly different from Treatment II's coefficient at the 1%-level: $^a$, from Treatment III's: $^b$, from Treatment IV's: $^c$.
*Source:* Own calculations.

However, 60 percent of all subjects are above this line, which means that they over-saved given their shifting choice.[24] The remaining 40 percent of all subjects lie below the solid blue line, which means that they under-saved. Again, apparently many subjects tried to replicate the Standard-Model 180-second restriction, indicating a status quo bias. Those who did not save or only very little (4 percent of subjects cut income by between -5 and +5 percent in Treatment IV) seem to abstract from risk to simplify the decision problem.

Taken together, our results from this and the previous section provide evidence that shifting and saving are indeed substitutes, though not for all subjects. If they were *perfect* substitutes, the average earnings would be statistically identical across Treatments II, III, and IV. Do expected euro earnings differ depending on which choices are available? To answer this question, we conduct OLS regressions of expected euro earnings (euro earnings for low and high wages, weighted with equal probability), euro earnings if the low wage is realized, and euro earnings if the high wage is realized on treatment dummies. Table 3 shows the results.

[24]Eighteen percent of all subjects borrow more than 5 percent of their income in shift 1 via ending work-shift 1 late but (over)compensate this with savings.

First, we consider all observations. In the first column, we observe that subjects in Treatments II, III, and IV earn significantly more than in Treatment I (before knowing which of the two possible states of the world occurs). When we compare the earning differences in Treatments II, III, and IV, we see that earnings in Treatments II and IV are not significantly different from one another and that earnings in Treatment III are significantly lower than in Treatments II and IV. As reported later, the saving behavior in Treatment II, which is indistinguishable from optimum, leads to higher euro earnings than both the shifting behavior in Treatment III (where we observed too little shifting) and the mix of saving and shifting in Treatment IV (where we observed that 60 percent of subjects over-saved).

The second and third column show how euro earnings are affected ex-post. In case the 'bad' state of the world occurs, in column 2 the same pattern as under uncertainty emerges (only with higher magnitudes of the treatment dummies' coefficients). This means that subjects use saving and shifting as precautionary measures against a 'rainy day' but not as a perfect substitute for one another. In column 3 we can see how much income the subjects give up in case they do not need to insure, i.e., the 'good' state of the world occurs. None of the coefficients is significantly different from zero: hence, the price subjects pay for their precautionary behavior is rather low thought the benefits are high. Our results show that more flexibility does not necessarily lead to better outcomes. In fact, subjects attained the highest payoffs in Treatment II, where they did not need to make an impulsive decision but instead had time to contemplate several possible choices. Now we consider the last three columns where we restrict the sample to subjects who did not borrow. Here, the estimates are identical for Treatment II and the constant. The point estimates, particularly in Column 5, are larger.

To distinguish different behavioral strategies, we classify individuals as substituters (marked with red diamonds in Figure 6) whose absolute difference between the actual share of savings and the share of savings implied by the observed share of income shifted does not exceed 15 percentage points in Treatment IV (30.2 percent of all subjects).[25] Similarly, we define another group according to whether the absolute difference from optimal income cuts in work-shift 1 exceeds 15 percentage points in Treatment IV provided that they have not been classified as substituters (10.4 percent of all subjects). We label this group (marked with blue squares in Figure 6) undersavers because they replicate the exogenously given work-shift distribution of Treatments I and II. In fact, in the group of undersavers there were more subjects that saved too little or nothing at all, which

[25]Concretely, we define a subject to be a substituter if $|\text{Share of income saved} - (23 - \text{Share of income shifted})| \leq 15$.

is true by definition for Treatment IV but strikingly also observed in Treatment II. In contrast, substituters save most frequently amounts close to the optimum even in Treatment II and 10.3 percent of them shifted between 15 and 30 percent of income in Treatment III. Our results for Hypothesis 4 may be summarized as follows:

**Result 4:** *i. We reject that saving and shifting are perfect substitutes on aggregate. In our experiments, we observe considerable heterogeneity between the subjects in respect with substituting saving and shifting. 30 percent of subjects do not exceed a given threshold for deviations from actual savings to those implied by shifting choices. 10 percent do not exceed a given threshold for deviations from choosing work-shifts of equal length. 10 percent of those who were classified as substituters (30 percent) shifted close to optimally in Treatment III.*

*ii. Expected payoff is significantly higher if either shifting or saving is allowed than when neither is allowed. It is significantly higher if shifting and saving is allowed compared to the case where only shifting is allowed. It is not significantly different from the case where only saving is allowed. Thus, while saving and shifting are substitutes, they are not perfect substitutes.*

#### 5.1.3 Differences across treatment averages

So far, we tested whether the model predicts the direction of the behavioral responses of the subjects to constraints in the different treatments correctly (and found that the model fares well). Now, we compare the averages of four key variables across the different treatments. The comparison of averages still does not require that we estimate the production function since we control for subject-specific fixed effects.

The variables we consider important are savings, time spent in shift 1, (hypothetical) income cuts, and the number of balls caught per movement. The upper panel of Table 4 shows the differences of the mentioned variables in Treatments II, III, and IV from Treatment I using subject fixed effects regressions (Treatment I's mean of the respective variable is equal to the constant in the regressions). The subject fixed effects assist in controlling for unobserved heterogeneity due to a subject's ability to catch balls or other time-invariant characteristics. The lower panel uses the same approach but compares Treatments II and III with Treatment IV. The first column considers savings. In the upper panel, we observe significant precautionary savings in Treatment II. In Treatment IV, the coefficient is smaller than in Treatment II, which suggests that some savings might have been substituted by shifting.

Table 4: Differences Across Treatments

| | Savings | Time Shift 1 | Time Shift 1 ≤180 | Income Cut | Income Cut≥0 | Balls per Move S1 | Balls per Move S2 |
|---|---|---|---|---|---|---|---|
| Treatment II-I | 2011.64*** | | | 2011.88*** | 2011.88*** | 0.21*** | 0.17 |
| | (89.98) | | | (90.01) | (90.03) | (0.08) | (0.13) |
| Treatment III-I | | -14.34*** | -58.94*** | 934.84*** | 2094.02*** | 0.17* | 0.24** |
| | | (5.10) | (3.25) | (146.90) | (124.37) | (0.09) | (0.11) |
| Treatment IV-I | 1511.16*** | -9.47** | -54.68*** | 2117.63*** | 2593.48*** | 0.22*** | 0.31*** |
| | (80.65) | (4.41) | (3.46) | (158.77) | (132.70) | (0.08) | (0.10) |
| Constant (I) | 0.00 | 180.00*** | 179.13*** | 0.28 | 142.34 | 2.84*** | 3.36*** |
| | (49.94) | (2.77) | (1.51) | (75.71) | (119.07) | (0.05) | (0.07) |
| Subject Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Adjusted $R^2$ | 0.617 | 0.021 | 0.683 | 0.327 | 0.544 | 0.013 | 0.009 |
| Observations | 576 | 576 | 397 | 768 | 665 | 767 | 755 |
| Treatment II-IV | 500.48*** | | | -105.75 | -519.35*** | | |
| | (82.17) | | | (153.68) | (142.07) | | |
| Treatment III-IV | | -4.88 | -7.85** | -1182.79*** | -501.14*** | | |
| | | (4.70) | (3.52) | (153.83) | (146.52) | | |
| Constant (IV) | 1511.16*** | 170.53*** | 125.61*** | 2117.63*** | 2617.62*** | | |
| | (41.09) | (2.35) | (1.94) | (87.50) | (82.86) | | |
| Subject Fixed Effects | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| Adjusted $R^2$ | 0.161 | 0.003 | 0.060 | 0.153 | 0.056 | | |
| Observations | 384 | 384 | 205 | 576 | 473 | | |

*Estimation Equation:* Differences across treatments estimated using individual-specific fixed effects.
*Inference:* Cluster robust (individual level) standard errors are in parentheses, significance levels are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
*Source:* Authors' calculations.

The lower panel of the table shows that this difference of about 25 percent in savings is significantly different from zero between Treatment II and Treatment IV. The results in the second column confirm this. In the upper panel, we can see that our subjects spent on average significantly less time in Treatment III (-14 seconds) and Treatment IV (-9 seconds) than in Treatment I (where time in shift 1 was fixed). The lower part of this column shows that we cannot reject that the amount of shifting is the same in Treatments III and IV, though. The point estimate for Treatment IV is smaller than for Treatment III, which again points to substitution, but is not statistically different from zero. In the third column, we restrict the sample to subjects who decided to end work-shift 1 on or before 180 seconds. Of course, the coefficients in the upper panel become larger in comparison to the second column.[26] In the lower panel, we see that the difference between Treatment III and IV is significant in the restricted sample. Our interpretation of this is that a substantial fraction of subjects understands saving and shifting as substitutes and that they choose to combine these two ways of intertemporal substitution.

How can we compare precautionary shifting (which is measured in seconds) with precautionary savings (which is measured in points)? The fourth column shows the results of a thought experiment: We know from theory that savings should be zero in the absence of wage risk and that both work-shifts should be equally long. Therefore, the observed number of earned points in period 1, during the first 180 seconds, serves as the benchmark level of income $y_1$ under certainty.

To obtain a comparable measure of precautionary behavior, we compute the expected income at the end of shift 1 minus savings and subtract it from $y_1$. That means that income cuts are defined as income in period 1 − (expected income in shift 1 − savings). If shifts and periods coincide, the only difference is income cuts due to savings; if savings are zero and shift 1 is shorter than period 1, this difference measures income cuts due to shifting (we assume for this calculation that point earnings are uniformly distributed over time). Since income cuts are measured in points, we can simply compare this measure across treatments. The fourth column shows that income cuts just equal savings from the first column in this case. In the third treatment (second row), shifting is allowed, but not saving. Therefore, we calculate the difference of income in period 1 to income obtained in shift 1. This is the measure for income cuts in Treatment III. Similarly, in Treatment IV, this difference minus the amount of savings gives total income cuts.

Using the entire sample, average income cuts in Treatment III seem to be too low to achieve the optimal intertemporal substitution implied by the theoretical model calibrated with the estimated ability to catch balls, while in Treatment IV the average amount is very close to the optimal. Statistically, the former value is different from that of Treatment II but the latter is not. In the next column, we exclude all subjects from the sample that did not save, i.e., who have negative income cuts. This shows that while in Treatments II and III savings and shiftings were virtually perfectly substituted, in Treatment IV significant excess income cuts take place.

Finally, the last two columns show that productivity is on average economically not significantly different across treatments, although some statistically significant differences are detected. The constant shows that on average, the subjects caught three balls.

[26]The point estimate of -59 seconds for Treatment III suggests that subjects on average shift too much since 179-59=120 seconds is smaller than the prediction of 131 seconds that we report in the next section.

### 5.1.4 Fit of structural predictions

As discussed in Section 3 (p. 10f), we need to estimate a production function as supplied in Equation 7 in order to calculate the point predictions for our variables of interest (savings and time spent in the two work-shifts) using the structural model. In a first step, we pool Treatments I to IV to estimate the production function from the real-effort task data.[27] The following coefficients are estimated from the number of movements and caught balls:

$$\text{balls}(\text{moves}) = 63.337 + 12.491 \times \sqrt{\text{moves}} - 0.001 \times \text{moves}^2.$$

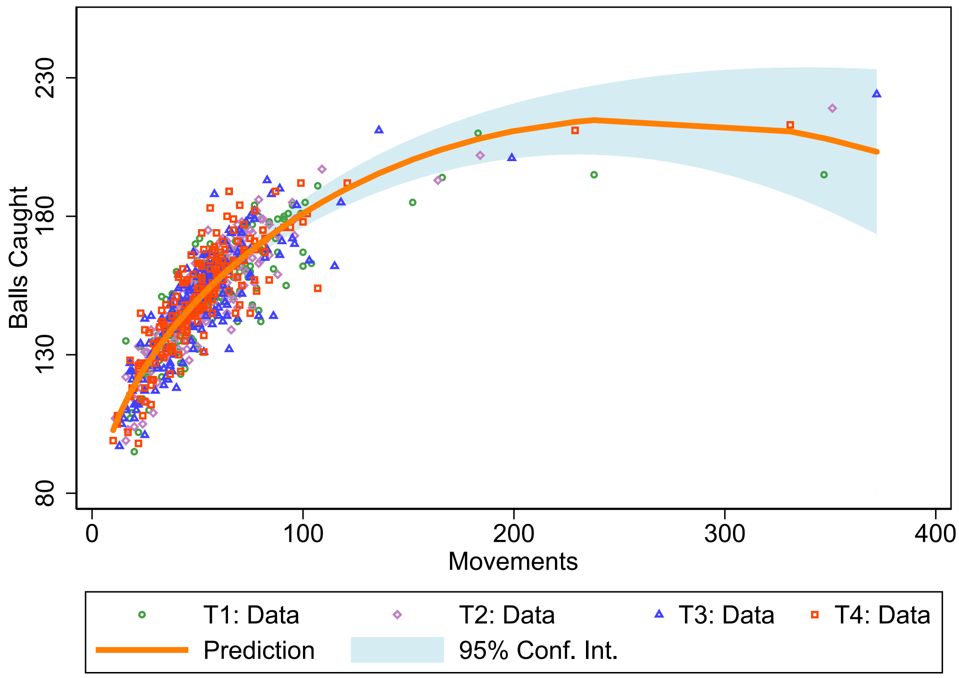



Figure 7: Overall estimates of the production functions and observations in all four treatments
*Source:* Authors' presentation.

Figure 7 displays the estimated production function and the observations in all four treatments (similar to Figure 3 in Gächter et al. 2016, p. 696). Despite the heterogeneity between the subjects' ability, the fit of this simple polynomial regression is quite high with $R^2 = 0.77$. One problem with using this function is that it does not cover the values of movements and balls needed in one single

[27] In Section H in the Appendix, we also discuss the stability of performance of subjects over time and supply graphical evidence of their heterogeneity of productivity.

period or shift. Hence, we repeat the exercise separately for the two periods[28] with individual-specific fixed effects:

For period 1 (overall $R^2 = 0.65$):

$$\text{balls}(\text{moves}) = 43.8091 + 6.3099 \times \sqrt{\text{moves}} - 0.0001 \times \text{moves}^2.$$

For period 2 (overall $R^2 = 0.73$):

$$\text{balls}(\text{moves}) = 40.8174 + 6.9724 \times \sqrt{\text{moves}} - 0.0010 \times \text{moves}^2.$$

Figure H.2 in the Appendix shows that these two production functions have a similar form and fit as the one from the pooled data. Table 5 displays aggregate predictions based on these two production functions, the realized means and their standard deviations in all treatments. This table shows how well the point estimates of the production function fit the average number of balls caught per period and second how well our model predicts movements per period, average savings, and the average shifting choice. The first prediction exercise performs very well; the average number of balls caught is predicted quite precisely for all treatments. No t-test indicates a significant deviation of the mean from the prediction. This is not obvious, because we jointly fit the production function for all treatments as an econometrician would when being unable to identify under which restriction choices were made.

Still, the model prediction exercise is much harder, since deviations from optimal behavior of only a few subjects could lead to a rejection of the model. The model predictions are conditional on the estimates of the production function. Of course, these estimates are measured with an error that we do not take into account in the model prediction explicitly. Instead, we report the results from using the point estimates in Table 5.

On average, subjects systematically moved about six to eight moves more often than predicted in all treatments. Of course, this is partly due to a few subjects making up to six-times the predicted number of moves (those disregarding effort costs). This is similar for period 2 but only in Treatment I and II. In Treatments III and IV, where shifting is allowed, it cannot be rejected that the average number of movements in period 2 is equal to the prediction.

[28]Using periods comes with the advantages that the two periods are equally long and that the wage uncertainty only affects period 2. That is not the case when we use shifts.

Table 5: Predictions of Extended Model and Data

| | TI | TII | TIII | TIV |
|---|---|---|---|---|
| | **Prediction** Mean Std. Dev. | **Prediction** Mean Std. Dev. | **Prediction** Mean Std. Dev. | **Prediction** Mean Std. Dev. |
| Production function predictions | | | | |
| Balls Caught in Period 1 | **78.77** | **77.72** | **78.80** | **78.07** |
| | 78.47 | 78.32 | 78.98 | 77.97 |
| | (10.82) | (10.53) | (11.64) | (11.43) |
| Balls Caught in Period 2 | **74.59** | **73.95** | **70.58** | **71.22** |
| | 73.92 | 73.48 | 70.76 | 72.61 |
| | (10.42) | (11.05) | (12.28) | (12.23) |
| Model predictions | | | | |
| Movements in Period 1 | **25.46** | **25.46** | **25.46** | **25.46** |
| | 32.71*** | 30.73*** | 32.99*** | 31.50*** |
| | (18.44) | (17.37) | (19.13) | (17.82) |
| Movements in Period 2 | **17.20** | **19.54** | **19.54** | **19.54** |
| | 26.54*** | 25.20*** | 21.17 | 21.69* |
| | (17.53) | (14.93) | (16.60) | (15.26) |
| Savings | **0.00** | **1917.16** | **0.00** | **Substitutes?** |
| | 0.00 | 2011.64 | 0.00 | 1511.16 |
| | (0.00) | (1244.67) | (0.00) | (1115.59) |
| Time Spent in Shift 1 | **180.00** | **180.00** | **131.00** | **Substitutes?** |
| | 180.00 | 180.00 | 165.66*** | 170.53 |
| | (0.00) | (0.00) | (70.54) | (61.02) |
| Observations | 192 | 192 | 192 | 192 |

*Notes:* Predicted means from period-specific individual-specific fixed effects estimations (Balls Caught) and model predictions based on point estimates of these estimates. Below predictions are sample means and standard deviations. Significance levels of one-sample t-tests against predicted means are * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
*Source:* Authors' calculations.

How well does the model predict precautionary savings and shiftings? The table shows that the average savings in Treatment II are not statistically different from the prediction. This shows that the model captures the savings decision extremely well. However, although shifting does occur (time spent in shift 1 is on average 166 seconds), the model predicts with 131 seconds in shift 1 even more shifting. Thus, a t-test rejects equality of average time spent in shift 1 and the model prediction. Regarding Treatment IV, there is no single prediction for savings and shiftings. Instead, we can predict optimal combinations of these two choices and compare them to actual combinations chosen by the subjects. We did this to classify substituters (as described in Section 5.1.2) and found that about 30 percent of all subjects follow the theoretical prediction that saving and shifting are perfect substitutes closely.

## 5.2 Heterogeneous Behavioral Strategies

One of the main objectives of this study is to investigate whether the intertemporal model predicts behavior under ideal laboratory conditions or whether other behavioral strategies are (also) important. First, we analyze deviations from the optimum that characterize different behavioral strategies and propose two measures for precautionary behavior in general and for shifting behavior in particular that may be used in survey or administrative data. Second, to account for the possibly many heterogeneous behavioral strategies that our subjects may have used, we view the data through the lens of a random behavior model. This model is based on the assumption that the actual choice is the optimal solution plus a random error term with a continuous distribution (see Moffatt 2016, chapter 8). Alternative structural models are based on modeling of the utility function or assumptions on random (preference) parameters. Since we deliberately induced preferences, the random behavior approach allows us to identify different behavioral strategies and optimization errors as the source of non-optimal behavior. To this end, we formally estimate a finite mixture model (FMM) of random behavior.[29]

### 5.2.1 Deviations from Optimum and Measures of Shifting

We derive theoretical predictions for three possible behavioral strategies—one of them static and two intertemporal—and use our subjects' deviations from these theoretical predictions to assign an individual-round observation to a strategy group. Then, we calculate the shares of subjects in a given group among all subjects by treatment. Table 6 provides the classification of observations in our data. A few subjects disregarded effort costs and ended up with negative income in the first work-shift (last column in Table 6). Neither the static nor the intertemporal model can predict this behavior because it stems from a violation of the intratemporal optimality condition in Equation (9). The classification rules assign the remaining observations to one of four behavioral strategies. If savings are zero and/or the distance of time in work-shift 1 to 180 seconds is

[29] This approach has been described in detail in an overview of recent developments in Compiani and Kitamura (2016) and been applied in Bruhin et al. (2010) or Cappelen et al. (2007).

less than or equal to 10 seconds, observations are classified as group (1), *Hand-to-Mouth*. If the absolute distance to the optimal savings-to-income ratio is less than 10 percentage points, we classify behavior as consistent with group (2), *Saving*. If the distance to optimal shifting is less than 10 percentage points, observations are assigned to group (3), *Shifting*. Finally, the sum of shares of observations classified as saving or shifting or using a combination of saving and shifting are classified as an group (4), *Extended*. The group *Combined* (1)+(4) includes the same observations as *Extended* and in addition the share of *Hand-to-Mouth* observations. For now, we label the remainder of the observations as *Unexplained*, a group that we investigate in more detail later.

To see how subjects' classifications correspond to the main mechanisms of the extended model, we propose the following two measures that can be calculated in any data with both shift- and period-specific information: first, the average wage in the second work-shift in case the bad state occurs in the second period and, second, the average reduction of income risk in work-shift 2. Since shifting allows to endogenously determine the ex-post wage in a given shift, the intuitive way to shield against the rainy day wage is to allocate time in such a way that the average wage in the bad case increases, i.e., $\Delta\tilde{\bar{w}}_2^{\text{Bad}} > 0$. This first measure for each subject $s$ in round $r$ is a weighted average of two components:

$$
\begin{aligned}
\Delta\tilde{\bar{w}}_{2,sr}^{\text{Bad}} &= \frac{\text{average bad wage in shift } 2_{sr}}{\text{average bad wage in period } 2_{sr}} - 100 \\
&\approx \{\omega_{1,sr} \times w_1^{\text{Certain}} + \omega_{2,sr} \times (w_2^{\text{Bad}} + \text{savings}_{sr}\text{per ball}_{sr})\}/w_2^{\text{Bad}} - 100,
\end{aligned} \tag{15}
$$

where the weights are

$$
\begin{aligned}
\omega_{1,sr} &= \max[T(1/2 - t_{sr}), 0]/[T(1 - t_{sr})], \\
\omega_{2,sr} &= 1 - \omega_{1,sr}.
\end{aligned} \tag{16}
$$

The intuition of this measure is that the average wage in the second work-shift in the bad case increases, starting from the low wage of the second period, linearly with the weight $\omega_1 = 1 - \omega_2$, which depends on the relative time allocation. More specifically, $\omega_1$ measures how much of the total time of the second work-shift has been spent under the certain wage. If a subject switches immediately at the beginning of a round into work-shift 2, $\omega_1$ equals $1/2$ (one half of the time in work-shift 2 has been worked under the certain wage and the other half under the uncertain). Ending work-shift 1 later means that both the time spent under the certain wage becomes shorter and the total time spent in work-shift 2. Since the former diminishes faster than the latter, $\omega_1$ declines non-linearly from $1/2$ to 0, while $\omega_2$ increases non-linearly from $1/2$ to 1.

Saving also increases the average wage in work-shift 2. Since savings are a component of income, we need to convert them into a measure that is comparable to the wages. A straightforward approximation is to use savings per ball. This can be understood as a premium to the low wage in the second period. Since $t_{sr} = 0.5$ in Treatment II, we obtain the relationship between the distance to optimal income cuts from Treatment III to calculate the weight.

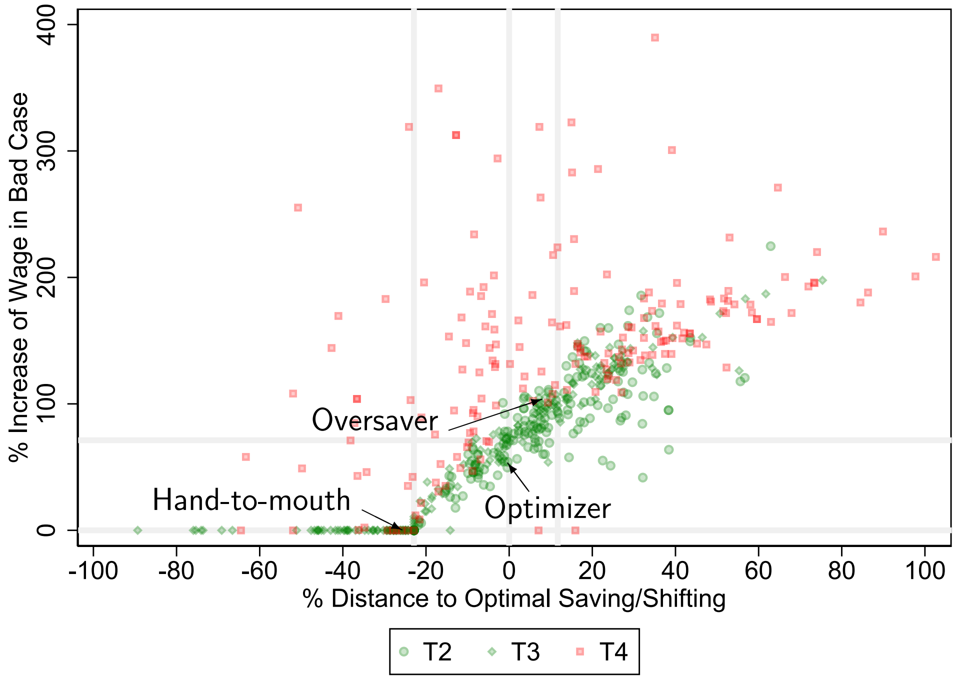
Figure 8a shows that the optimal wage increase is almost twice as high as the low wage (in TII and TIII: 71.2 percent) in the second work-shift (horizontal gray line). Of course, subjects could have found the optimal share of income cuts but had not enough income to increase the low wage of the second period optimally. This is the reason that some of the observations are located below the bulk of the data that follows the non-linear curve. Moreover, this figure clearly shows that there seem to be at least three types of subjects: those who do not increase the low wage in the second shift (hand-to-mouth), those who optimize, and those who exceed the optimal wage increase (oversaver). Finally, this figure shows that a combination of borrowing and saving allows increasing the low wage in the second shift much more (Treatment IV). The two observations of this treatment that are on the horizontal line represent subjects who chose not to save and had lower income from the first work-shift than during the first period though the first work-shift exceeded its length.

The second measure is very useful to detect how subjects use shifting behavior. It is defined for each subject $s$ in round $r$ as

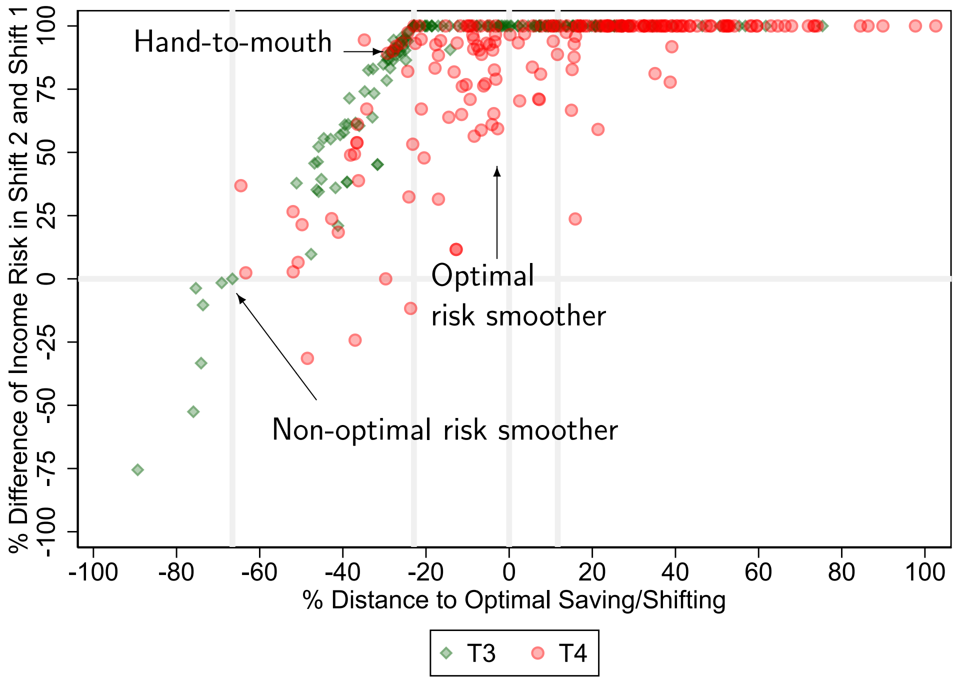
$$\begin{aligned}\Delta\tilde{\sigma}_{y_2,sr} &= 100 - \frac{\text{standard deviation of income in shift } 2_{sr}}{\text{standard deviation of income in period } 2_{sr}} \times 100 \\ &= 100 - \frac{\sqrt{(y^{\text{Good}}_{\text{shift }2,sr} - \bar{y}_{\text{shift }2,sr})^2 + (y^{\text{Bad}}_{\text{shift }2,sr} - \bar{y}_{\text{shift }2,sr})^2}}{\sqrt{(y^{\text{Good}}_{\text{period }2,sr} - \bar{y}_{\text{period }2,sr})^2 + (y^{\text{Bad}}_{\text{period }2,sr} - \bar{y}_{\text{period }2,sr})^2}} \times 100,\end{aligned} \tag{17}$$

where $\bar{y}_{\text{shift }2,ir} = (y^{\text{Good}}_{\text{shift }2,sr} + y^{\text{Bad}}_{\text{shift }2,sr})/2$ and $\bar{y}_{\text{period }2,sr} = (y^{\text{Good}}_{\text{period }2,sr} + y^{\text{Bad}}_{\text{period }2,sr})/2$.

Optimally, $\Delta\tilde{\sigma}_{y_2}$ should be zero in Treatment III, since shifting of *safe* income from the first to the second shift does not change the standard deviation of income. In contrast, shifting uncertain income from the second to the first shift (i.e., borrowing) reduces income risk measured as the standard deviation of income in the second shift at the expense of the standard deviation of income in the first shift. In fact, as shown in Figure 8b, subjects can smooth income risk across shifts (horizontal gray line).




(a) Second Period Wage in the Bad Case Across Treatments




(b) Income Risk Across Treatments

Figure 8: Optimality of Behavior versus Second Period Wage in the Bad Case and Income Risk

*Notes:* Gray lines show optimal behavior (distance to optimum is 0 percent), risk pessimist behavior (distance to optimum is 12 percent), and hand-to-mouth behavior (distance to optimum is -23 percent).

*Source:* Authors' presentation.

Table 6: Behavioral Strategies

| | Static Models | Intertemporal Models | | | Combined | | Negative |
|---|---|---|---|---|---|---|---|
| | (1) Hand-to-Mouth | (2) Saving | (3) Shifting | (4) Extended | (1)+(4) | Unexplained | Income |
| TI | 96.9% | — | — | — | 96.9% | — | 3.1% |
| TII | 6.8% | 43.8% | — | 43.8% | 50.6% | 47.9% | 1.6% |
| TIII | 17.2% | — | 20.3% | 20.3% | 37.5% | 56.8% | 5.7% |
| TIV | 3.1% | 40.1% | 21.4% | 81.8% | 84.9% | 11.5% | 3.6% |
| Average rainy day wage increase (optimal in TII/TIII: 71.2%) | | | | | | | |
| TII | 0% | 71.0% | — | 71.0% | 61.5% | 101.1% | 0% |
| TIII | 3.3% | — | 77.7% | 77.7% | 43.3% | 66.4% | 95.5% |
| TIV | 1.4% | 137.0% | 138.0% | 144.7% | 134.4% | 170.4% | 55.4% |
| Average risk reduction (optimal in TIII: 0%) | | | | | | | |
| TIII | 2.3% | — | 0% | 0% | 1% | 10.8% | 0% |
| TIV | 2.3% | 6.2% | 0.4% | 5.6% | 5.4% | 11.4% | 1.8% |

*Notes:* Classification rules assign observations to a behavioral strategy as follows: (1) *Hand-to-Mouth* if savings are zero and/or distance to 180 seconds is less or equal to 10 seconds. (2) *Saving* if absolute distance to optimal savings to income ratio is less than 10 percent. (3) *Shifting* if absolute distance to optimal shifting is less than 10 percent. (4) *Extended* sum of shares of observations classified as saving or shifting. *Combined* (1)+(4) like *Extended* but including share of *Hand-to-Mouth* observations. Rainy day wage increase is calculated as $\frac{\text{average bad wage in shift 2}}{\text{average bad wage in period 2}} - 100$. Risk reduction is $100 - \frac{\text{standard deviation of income in shift 2}}{\text{standard deviation of income in period 2}} \times 100$.
*Source:* Authors' calculations.

However, only borrowing to reduce risk is not optimal, since this alone does not provide any insurance against the realization of the low wage: it implies no increase of the average wage in the bad case of the second shift ($\Delta\tilde{\bar{w}}_2^{\text{Bad}} = 0$). Therefore, another behavioral strategy that subjects could follow is to i) smooth income risk via borrowing through shifting income and to simultaneously ii) use savings to ameliorate the rainy wage. This observation is stated in Result 5 and could explain the behavior of some subjects who have been classified previously as *Unexplained*.

**Result 5:** *An alternative behavioral strategy is to smooth income risk in both shifts. This strategy can be combined with precautionary saving. Few (2.1 percent) smooth income risk in both shifts alone. None of the subjects combines optimal precautionary behavior with smoothing income risk but 15.1 percent follow this strategy partly.*

Coming back to Table 6, we describe the share of the subjects classified to a specific behavioral strategy and compare the two measures of precautionary saving and shifting by treatment and classification of subjects. While in Treatment II observations predicted by the Standard-Saving model seem to be easy to detect, there is a substantial group that follows the static model instead, which the theoretical conditions pin down at around 7 percent (see the first column of the second row). Compared to the previous treatments, in Treatment III we again find strong heterogeneity. While these subjects do not reduce income risk substantially, there is a marked difference in the increase of the average low wage of the second shift. In Treatment IV only somewhat fewer observations can be classified to group (2) and the shares in group (3) remain around 20 percent. A very large share of observed behavior can be rationalized with the extended model, which is the result of allowing for more choices. The extended model allows to combine saving and shifting, for instance, one may borrow a share of income in shift 1 by ending work-shift 1 late and then compensate this with savings. Since borrowing by ending work-shift 1 late reduces the time available to earn income in shift 2, this reduces income risk in shift 2 (at the expense of income risk in shift 1) as shown in Figure 8b. As mentioned, the vertical difference to observations lined up at the top of the graph at the point where the distance to the optimum is 0 percent in Treatment IV shows that some subjects manage to reduce risk by substituting optimally. Because of this additional strategy, the share assigned to the extended model exceeds the sum of the individual saving and shifting shares. Even more important, however, is the result that a combination of the static and the extended model can explain behavior for virtually all subjects in Treatment IV. This shows that there is a role for both the static and the intertemporal model because also outside the laboratory a substantial part of the population might rely on heuristics that simplify intertemporal decision problems by abstracting from risk or entirely from intertemporal substitution. The group of unexplained make much more use of intertemporal substitution than the other groups. The average rainy day wage increase indicates that they save more and borrow more. Also, the average risk reduction measure indicates more borrowing than in any other group. Therefore, this group comprises those subjects who are non-optimizing risk smoothers and those who oversave.

### 5.2.2 Classification with a Finite Mixture Model

The previous results are now corroborated from the econometric perspective of a random behavior model. We assume that individuals are of different types, cannot change type during a round but possibly across rounds. If an individual's type determines uniquely the behavioral strategy for a given round, actual behavior may or may not coincide with optimal behavior due to optimization errors.

In our panel data set, we define $d_{sr} = s_{sr}/y_{sr} - \vartheta_2(s_{sr}/y_{sr})$ as the distance of individual $s =$

$1,...,192$ in round $r = 1,...,4$ to the optimal share of income saved, $\vartheta_2(s_{sr}/y_{sr})$, and to the optimal share of income shifted, also denoted as $\vartheta_2(s_{sr}/y_{sr})$. This will serve as our dependent variable. Inspecting the distribution of this variable across treatments (see Figure 9) shows that, while random deviations from the optimum may explain some of the data, much of the probability mass appears to pile up exactly at a deviation of -23 percent to the optimum which implies neither savings nor shiftings, i.e. hand-to-mouth behavior. Moreover, we know from the theoretical solution $\vartheta$ that the share of income is 35 percent for risk pessimists who only focus on the low realization of the wage $w_2^{\text{Bad}} = w - \varepsilon$. Therefore, theory gives us the following three groups $g = 1,2,3$ of strategy-types implied by their share saved or shifted $(s_{sr}/y_{sr})$:

**Hand-to-mouth** $(g = 1)$: The subjective expectation $E[w_2] = w$ implies $\vartheta_1(s_{sr}/y_{sr}) = 0\%$,

**Optimizer** $(g = 2)$: $E[w_2] = p \times (w+\varepsilon) + (1-p) \times (w-\varepsilon)$ implies $\vartheta_2(s_{sr}/y_{sr}) = 23\%$,

**Over-saver** $(g = 3)$: $E[w_2] = w - \varepsilon$ implies $\vartheta_3(s_{sr}/y_{sr}) = 35\%$.

We assume that variables included in the matrix $X_{sr}$ may deterministically explain part of the deviations from each strategy. $X_{sr}$ includes indicators for the subjective effort of participating in the experiment, an indicator of a coefficient of relative risk aversion that is greater than one, and another indicator for a coefficient of relative prudence greater than two, as well as indicators of the degree of prudence (see Appendix F).

The random behavior model specifies the allocation by subject $s$ of type $g$ as

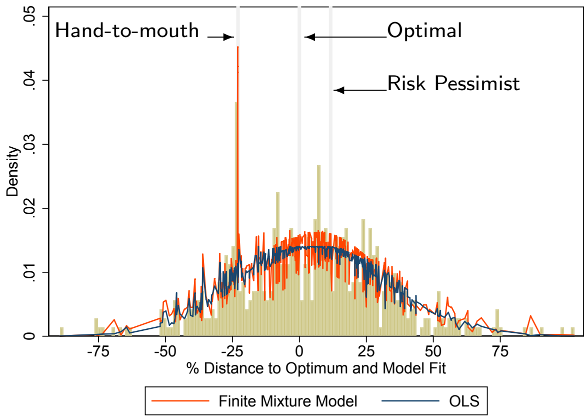
$$d_{sr} = \vartheta_g(s_{sr}/y_{sr}) - \vartheta_2(s_{sr}/y_{sr}) + X'_{sr}\beta_{gr} + \varepsilon_{sr}, \tag{18}$$

$$\varepsilon_{sr} \sim N(0, \sigma^2_{\varepsilon,g}), \tag{19}$$

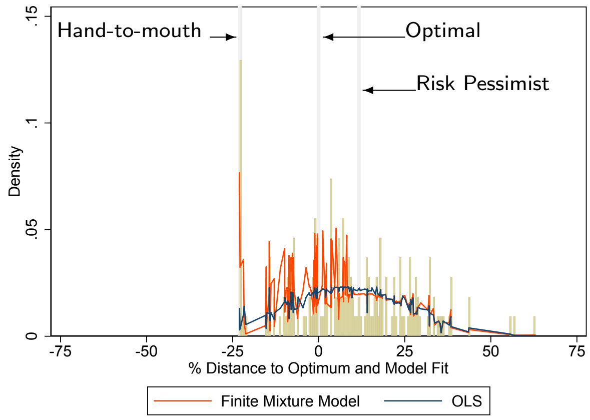
The non-deterministic variation in behavior is explained by the mean-zero error $\varepsilon_{sr}$ for which a new realization is independently drawn in each round. Using this specification with a finite mixture model allows for consistent estimates of the proportions $\pi_1, \pi_2, \pi_3$ of the population who belong to a given group. Since this model is probabilistic, groups may overlap. It combines conditional probability density functions $f_1, f_2, f_3$ for the response in the respective group. The density of a three-component mixture model is

$$f(d_{sr}|\hat{\pi}_g, \hat{\beta}_g, \hat{\sigma}^2_{\varepsilon,g}) = \sum_{g=1}^{3} \hat{\pi}_g f_g(d_{sr}|\hat{\beta}_g, \hat{\sigma}^2_{\varepsilon,g}),$$

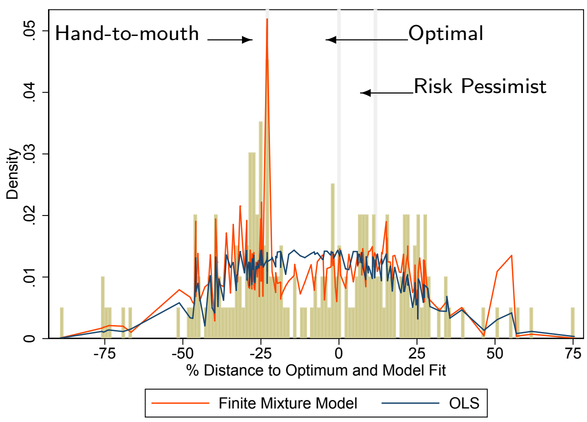
where $d_{sr}$ is the observed distance to optimum in percent, $0 \leq \hat{\pi}_k \leq 1$ and $\sum_{g=1}^{3} \hat{\pi}_g = 1$. The sum of the probability-weighted conditional likelihoods from each group is optimized with maximum likelihood methods.



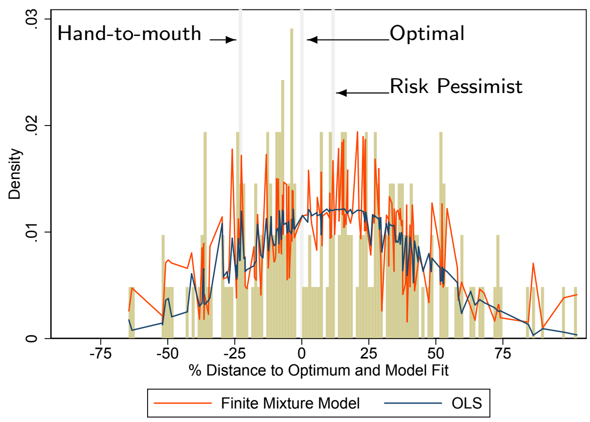

(a) Full Sample: TII through TIV




(b) TII




(c) TIII




(d) TIV

Figure 9: Fit of Finite Mixture Models and OLS

*Notes:* The histogram of the distance to optimal income cuts is shown in yellow bars, blue lines represent the fit of an OLS model and red lines of finite mixture models. Gray lines show optimal behavior (distance to optimum is 0 percent), risk pessimist behavior (distance to optimum is 12 percent), and hand-to-mouth behavior (distance to optimum is -23 percent).

*Source:* Authors' presentation.

Since, as discussed, much of the probability mass is *exactly* at a deviation of -23 percent to the optimum, we specify $\hat{\pi}_1 = \hat{q}$ as a degenerate distribution such that

$$\beta_{gr} = \beta_g \text{ with probability } (1-q),$$

$$\beta_{gr} = 0 \text{ and } \varepsilon_{sr} = 0 \text{ with probability } q.$$

The individual likelihood function is

$$L_s(\hat{q}, \hat{\pi}_2, \hat{\pi}_3, \hat{\beta}_g, \hat{\sigma}^2_{\varepsilon,g}) = \tag{20}$$
$$\hat{q}\mathbb{1}_{\{d_{sr}=\vartheta_{1,sr}\}} + (1-\hat{q})\mathbb{1}_{\{d_{sr}\neq\vartheta_{1,sr}\}}\left\{\sum_{g=2}^{3}\hat{\pi}_g\prod_{r=1}^{4}\left[\frac{1}{\hat{\sigma}_{\varepsilon,g}}\phi\left(\frac{d_{sr}-\vartheta_g(s_{sr}/y_{sr})+\vartheta_2(s_{sr}/y_{sr})-X'_{sr}\hat{\beta}_g}{\hat{\sigma}_{\varepsilon,g}}\right)\right]\right\},$$

where $\phi$ is the probability density function of the normal distribution and the predicted probability for the group $k$ is given by the multinomial logistic function

$$\hat{\pi}_k = Pr(k=1|X_{sr}) = \frac{\exp(\hat{\gamma}_k)}{\sum_{g=2}^{3}\exp(\hat{\gamma}_g)}.$$

Intuitively, we use this method to estimate several OLS models jointly, one for each of two classes and a degenerate distribution for the hand-to-mouth class. Then we predict the density of an OLS model with only one class and a mixture density from the finite mixture model. Figure 9a shows the resulting fit when the sample includes Treatments II through IV. Clearly, the finite mixture model provides a better fit, since the simple OLS model misses the hand-to-mouth behavior, which is located at -23 percent. Though oversavers are predicted better with the finite mixture model, this model does not result in the predicted bimodal distribution. Figure 9b is based on data of Treatment II only and shows that the plain OLS model underpredicts a large portion of observed behavior. In particular, the finite mixture model can account for the hand-to-mouth subjects and for the optimizers better. Similarly, from Figures 9c and 9d it becomes apparent that behavior is much better fitted by a mixture of densities since at least for Treatment III the predicted distribution is bimodal (though not clearly trimodal as our behavioral theory suggests). To quantitatively assess the model fit, we compare the respective log-likelihoods and the Akaike and Bayes information criteria (AIC and BIC). While the log-likelihood is always higher and the AIC always smaller for the finite mixture model, the BIC is only smaller when we use the sample of Treatment II. The main conclusion from this evidence is that both empirical and theoretical economists need to explicitly take into account heterogeneity in behavioral strategies. Otherwise, e.g., estimates of the willingness to pay for flexible work arrangements will be biased.

A further criterion of model success is how well it assigns subjects to types. To see this, we calculate the posterior probabilities of each subject in the sample belonging to a specific class using Bayes' Rule.

The posterior probabilities for the model are

$$\tilde{\pi}_k = \frac{Pr(g=k|s) \times Pr(s|g=k)}{Pr(s)} = \frac{\hat{\pi}_k \times L_{s,k}}{\sum_{g=1}^{3} \hat{\pi}_g L_{s,k}}. \tag{21}$$

where $L_{s,k}$ represents the component of the likelihood function corresponding to type $k$. Figure I.1 in the Appendix presents the posterior probabilities in four rows. The figures for Treatment II in the first row indicate successful classification if there is mass in the upper left and lower right corners of a respective figure. In all three figures, some individuals can be classified with high probabilities to either class. Only in the second figure, the downward sloping scatter indicates uncertainty over types. Rows two to four present the probabilities that a subject followed a specific behavioral strategy in more than one treatment or if the subject switched strategies. Probability mass in an upper right corner indicates that a subject followed the same strategy in two treatments, mass in the lower right or upper left corner indicates a change of strategy. According to the figures, many subjects stuck to their strategy in both Treatments II and III. Comparing Treatment III and IV, many subjects changed behavioral strategies.

**Result 6:** *i. A random behavior model estimated with OLS under-predicts a substantial part of the probability mass. In particular, it cannot capture hand-to-mouth behavior. Estimates of structural models to assess for instance the willingness to pay for flexible work arrangements will be biased if hand-to-mouth behavior is ignored.*
*ii. Subjects stick to the same strategy in Treatments II and III but switch strategies in Treatment IV.*

# 6 Conclusions

This paper presents an increasingly important strategy for intertemporal substitution. If the wage rate varies over time, the choice of how much time to work allows one to determine the expected wage. This is not possible if work-shifts and periods coincide, as in the standard model. Therefore, the predictions of the standard model might be misleading when work-shift allocation is possible. This shifting behavior allows intertemporal substitution in the same way as saving behavior and may in practice be important for self-employed who shift income, e.g., between fiscal years to avoid taxes. Since it is difficult to impose saving restrictions and control for differences in preferences and wage risk, we conduct laboratory experiments that show that agents do in fact regard these channels as substitutes, albeit not as perfect substitutes, as was theoretically predicted.

In the experiment, the only reason for engaging in intertemporal substitution was future wage risk. Accordingly, a second important contribution of the paper is that we shed light on an empirical puzzle contained in the precautionary saving literature. In this literature, different methods yield very different results on the importance of precautionary saving ranging from a few percents to 100 percent of overall savings. Hence, we contribute to the debate on whether the model of intertemporal substitution or a static hand-to-mouth model is more useful.

We show in favor of the intertemporal model that (i) a potentially large part of precautionary behavior occurs through shifting and does not affect savings, (ii) both channels for intertemporal substitution are important for a fraction of subjects, but that (iii) both hand-to-mouth consumers and intertemporal substituters coexist such that even a single level of wage risk may lead to a very broad distribution of savings. This implies that there is a role for the hand-to-mouth model even if intertemporal substitution would be optimal. Therefore, these behavioral strategies need to be jointly modeled to allow making statements on the willingness to pay for flexible work arrangements. To distinguish different behavioral strategies, surveys should, therefore, gather data on shift-specific wages, wage risk and the flexibility of the work arrangement.

However, our findings come with a caveat. The introduced model and the experiments based on it deliberately ignore how workers in the field also might cope with a potentially boring and exhausting work situation. The two most obvious strategies are trading on- and off-the-job leisure[30] (the absolute work time is not a decision in our model and we did not observe subjects refraining from work on the real-effort task during our experiments) and working overtime hours (which we did neither model nor allow in the experiments in order to avoid that our subjects' behavior is driven by confounds outside the laboratory). We designed our model jointly with the laboratory experiment and we intended to make use of the major advantage of laboratory experiments: control over unobservable variables. Field experiments in labor economics (see Heinz et al. 2017 for a recent example) provide a more realistic alternative to laboratory experiments. Hence, our approach is a pioneering study that can also lead to (more costly) field experiments.

[30]Which is in the focus in the experiments by Dickinson (1999).

# References


BALLINGER, T. P., M. G. PALUMBO, AND N. T. WILCOX (2003): "Precautionary saving and social learning across generations: an experiment," *Economic Journal*, 113(490), 920–947. Cited on pages 6 and A.1.

BARTZSCH, N. (2008): "Precautionary saving and income uncertainty in Germany: new evidence from microdata," *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)*, 228(1), 5–24. Cited on page A.1.

BENITO, A. (2006): "Does job insecurity affect household consumption?," *Oxford Economic Papers*, 58(1), 157–181. Cited on page A.2.

BLUNDELL, R., AND T. MACURDY (1999): "Labor supply: a review of alternative approaches," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3 of *Handbook of Labor Economics*, chap. 27, 1559–1695. Elsevier. Cited on page 7.

BLUNDELL, R., AND I. WALKER (1986): "A life-cycle consistent empirical model of family labour supply using cross-section data," *Review of Economic Studies*, 53(4), 539–558. Cited on page 7.

BROADWAY, B., AND J. P. HAISKEN-DENEW (2018): "Keep calm and consume? Subjective uncertainty and precautionary savings," *Journal of Economics and Finance*. Cited on page A.2.

BROWN, A. L., Z. E. CHUA, AND C. F. CAMERER (2009): "Learning and visceral temptation in dynamic saving experiments," *Quarterly Journal of Economics*, 124(1), 197–231. Cited on pages 6 and A.1.

BROWNING, M., AND A. LUSARDI (1996): "Household saving: micro theories and micro facts," *Journal of Economic Literature*, 34(4), 1797–1855. Cited on page 15.

BRUHIN, A., H. FEHR-DUDA, AND T. EPPER (2010): "Risk and rationality: Uncovering heterogeneity in probability distortion," *Econometrica*, 78(4), 1375–1412. Cited on page 39.

CABALLERO, R. J. (1991): "Earnings uncertainty and aggregate wealth accumulation," *American Economic Review*, 81(4), 859–871. Cited on page A.2.

CAGETTI, M. (2003): "Wealth accumulation over the life cycle and precautionary savings," *Journal of Business and Economic Statistics*, 21(3), 339–353. Cited on page A.2.

CAMPBELL, J. Y. (1987): "Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis," *Econometrica*, 55(6), 1249–1273. Cited on page 14.

CAPPELEN, A. W., A. D. HOLE, E. . SØRENSEN, AND B. TUNGODDEN (2007): "The pluralism of fairness ideals: an experimental approach," *American Economic Review*, 97(3), 818–827. Cited on page 39.

CARD, D. (1994): "Intertemporal labor supply: an assessment," in *Advances in Econometrics: Sixth World Congress*, ed. by C. A. Sims, vol. 2. Cambridge University Press. Cited on page 7.

CARROLL, C. D., AND M. S. KIMBALL (2008): "Precautionary saving and precautionary wealth," in *The New Palgrave Dictionary of Economics*, ed. by N. Durlauf, and L. Blume. MacMillan, 2nd edn. Cited on page 6.

CARROLL, C. D., AND A. A. SAMWICK (1998): "How important is precautionary saving?," *Review of Economics and Statistics*, 80(3), 410–419. Cited on pages 5 and A.1.

CHARNESS, G., U. GNEEZY, AND B. HALLADAY (2016): "Experimental methods: pay one or pay all," *Journal of Economic Behavior & Organization*, 131, 141–150. Cited on page 20.

CHARNESS, G., AND P. KUHN (2011): "Lab labor: what can labor economists learn from the lab?," *Handbook of Labor Economics*, 4, 229–330. Cited on pages 7 and 18.

CHEN, M. K., J. A. CHEVALIER, P. E. ROSSI, AND E. OEHLSEN (2017): "The Value of Flexible Work: Evidence from Uber Drivers," Working Paper 23296, National Bureau of Economic Research. Cited on page 5.

COMPIANI, G., AND Y. KITAMURA (2016): "Using mixtures in econometric models: a brief review and some new results," *The Econometrics Journal*, 19(3), C95–C127. Cited on page 39.

DARDANONI, V. (1991): "Precautionary savings under income uncertainty: a cross-sectional analysis," *Applied Economics*, 23(1), 153–160. Cited on page A.1.

DICKINSON, D. L. (1999): "An experimental examination of labor supply and work intensities," *Journal of Labor Economics*, 17(4), 638–670. Cited on pages 3, 7, and 49.

DYNAN, K. E. (1993): "How prudent are consumers?," *Journal of Political Economy*, 101(6), 1104–1113. Cited on page A.2.

Eaton, J., and H. S. Rosen (1980): "Labor supply, uncertainty, and efficient taxation," *Journal of Public Economics*, 14(3), 365 – 374. Cited on page 9.

Eeckhoudt, L., R. J. Huang, and L. Y. Tzeng (2012): "Precautionary effort: a new look," *Journal of Risk and Insurance*, 79(2), 585–590. Cited on page 6.

Engen, E. M., and J. Gruber (2001): "Unemployment insurance and precautionary saving," *Journal of Monetary Economics*, 47(3), 545 – 579. Cited on page 5.

Fagereng, A., L. Guiso, and L. Pistaferri (2017): "Firm-related risk and precautionary saving response," *American Economic Review*, 107(5), 393–397. Cited on page 19.

Falk, A., and J. J. Heckman (2009): "Lab experiments are a major source of knowledge in the social sciences," *Science*, 326(5952), 535–538. Cited on page 3.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2), 171–178. Cited on page 22.

Flodén, M. (2006): "Labour supply and saving under uncertainty," *Economic Journal*, 116(513), 721–737. Cited on pages 6, 7, 9, and 11.

Fossen, F. M., and D. Rostam-Afschar (2013): "Precautionary and entrepreneurial savings: new evidence from German households," *Oxford Bulletin of Economics and Statistics*, 75(4), 528–555. Cited on pages 5 and A.1.

Fuchs-Schündeln, N., and M. Schündeln (2005): "Precautionary savings and self-selection: evidence from the German reunification "experiment"," *Quarterly Journal of Economics*, 120(3), 1085–1120. Cited on pages 5 and A.1.

Gächter, S., L. Huang, and M. Sefton (2016): "Combining "real effort" with induced effort costs: the ball-catching task," *Experimental Economics*, 19, 687–712. Cited on pages 12, 18, and 36.

Gourinchas, P.-O., and J. A. Parker (2002): "Consumption over the life cycle," *Econometrica*, 70(1), 47–89. Cited on pages 5 and A.2.

Greiner, B. (2015): "Subject pool recruitment procedures: Organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1(1), 114–125. Cited on page 22.

GUISO, L., T. JAPPELLI, AND D. TERLIZZESE (1992): "Earnings uncertainty and precautionary saving," *Journal of Monetary Economics*, 30(2), 307 – 337. Cited on pages 5 and A.1.

HARTWICK, J. M. (2000): "Labor supply under wage uncertainty," *Economics Letters*, 68(3), 319–325. Cited on page 9.

HECKMAN, J. J. (1993): "What has been learned about labor supply in the past twenty years?," *American Economic Review*, 83(2), 116–121. Cited on page 6.

HECKMAN, J. J., AND T. E. MACURDY (1980): "A life cycle model of female labour supply," *Review of Economic Studies*, 47(1), 47–74. Cited on page 7.

HEINZ, M., S. JEWORREK, V. MERTINS, H. SCHUMACHER, AND M. SUTTER (2017): "Measuring indirect effects of unfair employer behavior on worker productivity: a field experiment," *CESifo Working Paper Series*, 6760. Cited on page 49.

HEY, J. D., AND V. DARDANONI (1988): "Optimal consumption under uncertainty: an experimental investigation," *Economic Journal*, 98(390), 105–116. Cited on page A.1.

HUCK, S., N. SZECH, AND L. M. WENNER (2018): "More effort with less pay: on information avoidance, optimistic beliefs, and performance," Discussion paper. Cited on page 6.

HURST, E., A. LUSARDI, A. KENNICKELL, AND F. TORRALBA (2010): "The importance of business owners in assessing the size of precautionary savings," *Review of Economics and Statistics*, 92(1), 61–69. Cited on pages 5 and A.1.

JAPPELLI, T., AND L. PISTAFERRI (2014): "Fiscal policy and MPC heterogeneity," *American Economic Journal: Macroeconomics*, 6(4), 107–136. Cited on page 2.

JAPPELLI, T., AND L. PISTAFERRI (2017): *The economics of consumption: theory and evidence*. Oxford University Press. Cited on pages 5, 7, 11, 14, 15, and A.2.

JESSEN, R., AND D. ROSTAM-AFSCHAR (2014): "GRAPH3D: Stata module to draw colored, scalable, rotatable 3D plots," Statistical Software Components, Boston College Department of Economics. Cited on pages 23 and G.1.

JESSEN, R., D. ROSTAM-AFSCHAR, AND S. SCHMITZ (2017): "How important is precautionary labour supply?," *Oxford Economic Papers*, 70(3), 868–891. Cited on pages 6 and A.2.

Kaplan, G., and G. L. Violante (2014): "A model of the consumption response to fiscal stimulus payments," *Econometrica*, 82(4), 1199–1239. Cited on page 13.

Kazarosian, M. (1997): "Precautionary savings: a panel study," *Review of Economics and Statistics*, 79(2), 241–247. Cited on pages 5 and A.1.

Keane, M. P. (2011): "Labor supply and taxes: a survey," *Journal of Economic Literature*, 49(4), 961–1075. Cited on pages 7 and 13.

Kimball, M. S. (1990): "Precautionary saving in the small and in the large," *Econometrica*, 58(1), 53–73. Cited on pages 5 and 11.

Koustas, D. K. (2018): "Consumption insurance and multiple jobs: evidence from rideshare drivers," mimeo. Cited on page 4.

Low, H. (2005): "Self-insurance in a life-cycle model of labor supply and savings," *Review of Economic Dynamics*, 8(4), 945–975. Cited on pages 6 and 7.

Lugilde, A., R. Bande, and D. Riveiro (forthcoming): "Precautionary saving: a review of the empirical literature," *Journal of Economic Surveys*. Cited on pages 5 and A.2.

Lusardi, A. (1997): "Precautionary saving and subjective earnings variance," *Economics Letters*, 57(3), 319 – 326. Cited on pages 5 and A.1.

——— (1998): "On the importance of the precautionary saving motive," *American Economic Review*, 88(2), 449–453. Cited on pages 5 and A.1.

MaCurdy, T. E. (1981): "An empirical model of labor supply in a life-cycle setting," *Journal of Political Economy*, 89(6), 1059–85. Cited on page 7.

Marshall, A. (1920): *Principles of economics*. Macmillan, London, 8 edn. Cited on page 6.

Martinez, I. Z., E. Saez, and M. Siegenthaler (2018): "Intertemporal labor supply substitution? Evidence from the Swiss income tax holidays," Working Paper 24634, National Bureau of Economic Research. Cited on page 2.

Mastrogiacomo, M., and R. Alessie (2014): "The precautionary savings motive and household savings," *Oxford Economic Papers*, 66(1), 164–187. Cited on page A.1.

Meissner, T., and D. Rostam-Afschar (2017): "Learning Ricardian equivalence," *Journal of Economic Dynamics and Control*, 82(Supplement C), 273–288. Cited on pages 6 and A.1.

Moffatt, P. G. (2016): *Experimetrics: econometrics for experimental economics*. Palgrave. Cited on page 39.

Moffitt, R. (1984): "The estimation of a joint wage-hours labor supply model," *Journal of Labor Economics*, 2(4), 550–566. Cited on page 12.

Mulligan, C. B. (1998): "Pecuniary incentives to work in the United States during World War II," *Journal of Political Economy*, 106(5), 1033–1077. Cited on page 6.

Noussair, C. N., S. T. Trautmann, and G. van de Kuilen (2014): "Higher order risk attitudes, demographics, and financial decisions," *Review of Economic Studies*, 81(1), 325–355. Cited on page 21.

Parker, S. C., Y. Belghitar, and T. Barmby (2005): "Wage uncertainty and the labour supply of self-employed workers," *Economic Journal*, 115(502), C190–C207. Cited on pages 9 and A.2.

Pijoan-Mas, J. (2006): "Precautionary savings or working longer hours?," *Review of Economic Dynamics*, 9(2), 326 – 352. Cited on page A.2.

Pistaferri, L. (2003): "Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply," *Journal of Labor Economics*, 21(3), 729–754. Cited on pages 6 and A.2.

Skinner, J. (1988): "Risky income, life cycle consumption, and precautionary savings," *Journal of Monetary Economics*, 22(2), 237 – 255. Cited on page A.2.

Ventura, L., and J. G. Eisenhauer (2006): "Prudence and precautionary saving," *Journal of Economics and Finance*, 30(2), 155–168. Cited on page A.2.

Wang, J., and J. Li (2015): "Precautionary effort: another trait for prudence," *Journal of Risk and Insurance*, 82(4), 977–983. Cited on page 6.

Zeldes, S. P. (1989): "Optimal consumption with stochastic income: deviations from certainty equivalence," *Quarterly Journal of Economics*, 104(2), 275–298. Cited on page A.2.

# Appendix

# A Existence and Importance of Precautionary Saving in Extant Literature

Table A.1: Literature on Precautionary Saving

| Study | Data Set | Data Period | Measures of Risk | Precautionary Saving |
|---|---|---|---|---|
| Lab experiment | | | | |
| Meissner and Rostam-Afschar (2017) | Students at TU-Berlin | Eight life cycles à 25 periods | 35% of expected value with probability 0.5 | No evidence |
| Brown et al. (2009) | Students at National University of Singapore and California Institute of Technology | Seven life cycles à 30 periods | Log-normally distributed | Undersaving |
| Ballinger et al. (2003) | Students at Univeristy of Huston and Stephen F. Austin State University | One life cycle à 60 periods | Two treatments: 3 francs (5%) or 5 francs (5%); otherwise, 4 francs, 50% 8 francs and 50% 0 francs | $>0\%$, but undersaving |
| Hey and Dardanoni (1988) | Students at University of York | between 5 and 15 periods | normally distributed | — |
| Wealth regression | | | | |
| Mastrogiacomo and Alessie (2014) | DHS | 1993-2008 | Subjective earnings variance, second income earner | 30% |
| Fossen and Rostam-Afschar (2013) | SOEP | 2002, 2007, 1984–2007 | Heteroskedasticity function | 0-20% |
| Hurst et al. (2010) | PSID | 1984, 1994, 1981-1987, 1991-1997 | Permanent and transitory components of earnings regression | $<10\%$ |
| Bartzsch (2008) | SOEP | 2002, 1980-2003 | Variance of income | 0-20% |
| Fuchs-Schündeln and Schündeln (2005) | SOEP | 1992-2000 | Civil servant indicator | 12.9-22.1% |
| Carroll and Samwick (1998) | PSID | 1984, 1981–1987 | Variance of income | 32-50% |
| Lusardi (1998) | HRS | 1992 | Self-reported | 1-3.5% |
| Lusardi (1997) | SHIW | 1989 | Self-reported | 2.8% |
| Kazarosian (1997) | NLS | 1966-1981 | Permanent and transitory components of earnings regression | 29% |
| Guiso et al. (1992) | SHIW | 1989 | Self-reported | 2% |
| Dardanoni (1991) | UK Family Expenditure Survey | 1984 | Variance of labor income | $>60\%$ |

*Table continued on next page.*

| Study | Data Set | Data Period | Measures of Risk | Precautionary Saving |
|---|---|---|---|---|
| Hours of work regression | | | | |
| Jessen et al. (2017) | SOEP | 2001-2012 | Standard deviation of past detrended log wages | 1.16 hours per week |
| Benito (2006) | BHPS | 1991-2007 | Difference between actual and expected financial situation | $<$ 1.4 hours per week |
| Parker et al. (2005) | PSID | 1968-1993 | Standard deviation of past wages | 1.68 hours per week |
| Pistaferri (2003) | SHIW | 1989, 1991, and 1993 | Subjective information on future income | negligible |
| Saving regression | | | | |
| Broadway and Haisken-DeNew (2018) | HILDA, CASiE | 2002, 2006 and 2010 | Subjective and objective uncertainty | 0.35% |
| Ventura and Eisenhauer (2006) | SHIW | 1993;1995 | Average income variance | 15-36% |
| Skinner (1988) | CEX | 1972-1973 | Occupation indicators | 0% |
| Estimation of Consumption Euler Equation | | | | |
| Dynan (1993) | CEX | Four quarters of 1985 | Consumption variability | 0% |
| Skinner (1988) | CEX | | | 56% |
| Method of Simulated Moments | | | | |
| Cagetti (2003) | SCF, PSID | 1989, 1992, 1995; 1984, 1989,1994 | Permanent and transitory components of earnings regression | 50-100% |
| Gourinchas and Parker (2002) | CEX, PSID | 1980-1993 | Permanent and transitory components of earnings regression, prob of zero earnings | 60-70% |
| Numerically Simulated Consumption Function | | | | |
| Pijoan-Mas (2006) | PSID | | | 18.0% |
| Zeldes (1989) | from other studies | | | 1.6-10% |
| Skinner (1988) | CEX | | | 56% |
| Calibrated Closed Form Consumption Function | | | | |
| Caballero (1991) | | | | $>$ 60% |

*Notes:* Importance figure is sometimes calculated from several sources in the respective paper, please read the paper for details. Datasets are De Nederlandsche Bank household survey (DHS), German Socio-Economic Panel (SOEP), Italian Survey of Household Income and Wealth (SHIW), Household, Income and Labour Dynamics in Australia (HILDA), Consumer Attitudes, Sentiments and Expectations (CASiE), British Household Panel Survey (BHPS), National Longitudinal Survey (NLS), Health and Retirement Study (HRS), Consumer Expenditure Survey (CEX), Survey of Consumer Finances (SCF), Panel Study of Income Dynamics (PSID). Surveys of further related studies can be found in Jappelli and Pistaferri (2017) or Lugilde et al. (forthcoming).

# B Translation of the Instructions

INSTRUCTIONS

Welcome to this experiment!

In this experiment, you can earn a considerable amount of money. Your earnings in this experiment depend *only* on the choices *you* make during the experiment. Please read the printed instructions and those shown on-screen carefully.

During the experiment, you are not allowed to use electronic devices other than your PC or to talk to other participants. Please only use the computer programs and functions designated for the experiment. Should you have any questions, please raise your hand. We will then quietly answer your question. If the question is of relevance for all participants, we will loudly repeat and answer it.

**Outline**

Please read the instructions carefully. Afterward, you will answer a few **quiz questions** to make sure you understand everything. Overall, the experiment will take about 1.5 hours.

The experiment is made up of **three parts**. The payoff you are able to receive in each separate part does not depend on your behavior in the other parts.

**Part 1**

Part 1 is made up of three test periods, which gives you the opportunity to practice the **assignment** you will work on in the second part (the assignment will be explained further down in Part 2). One of the three test periods is randomly chosen for payoff. Only at the end of the experiment, you will be informed about which period was chosen. Further information will show up on your screen.

**Part 2**

The second part is made up of **four different rounds**, which consist of **two shifts** each. *The time during which you are working on your assignment without an interruption is referred to as a* ***shift****.* Only one of the four rounds is relevant for your payoff. Which of the four round earnings will be paid out will be chosen at random. Only at the very end of the experiment, you will be informed about the round that was chosen for payoff.

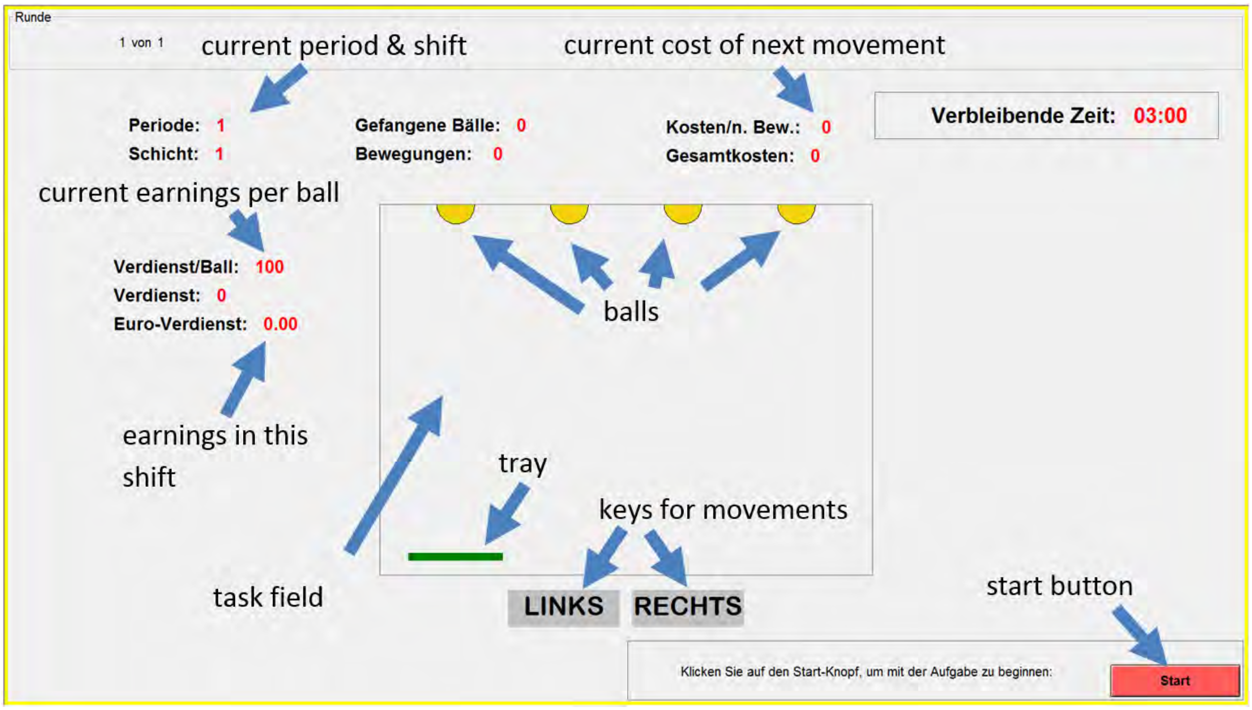
By working on the assignment you can earn **points**. Points are the currency of this experiment. The points you earn during one shift will be converted into euros.

**Round 1**

In the first round, you are working on the assignment consisting of two shifts. In this round, each shift consists of a period, which always lasts 180 seconds. *A* ***period*** *is the time during which a particular earning is paid.*

**Your assignment**

While working on the assignment you will see a **task field** in the middle of the screen, similar to the following figure.




Left of the task field you can see in which **period** and which **shift** you are currently in. As soon as you click the **start button**, the countdown starts and balls start falling randomly from the upper part of the task field. The remaining time is shown in the upper right corner of the screen. The catching tray can be moved by clicking **"LEFT"** or **"RIGHT"** at the bottom part of the task field, in order to catch the balls. To catch a ball, the **catching tray** has to be positioned right underneath the ball, at the moment the ball touches the tray. As soon as the ball touches the tray, the number of balls caught increases by one. The **number of the balls caught so far** and the **number of the current moves** are shown above the task field.

Each move of the catching tray generates costs. Each ball caught generates earnings. The **cost of the next move** is shown above the task field. Underneath you can find the **current overall costs**. The **current earnings per ball** are shown left of the task field. Underneath you can see your earnings in this shift in points and in euros.

Earnings in points are calculated as followed:

Earnings = Number of balls caught * Earnings per ball caught – Sum of the costs of the moves

**Earnings per ball caught**

In each period you will be informed about the **earnings per ball caught**. Your earnings per ball caught are always 100 points throughout the first period. In the **second period**, your earnings per ball caught are determined **randomly**. The earnings may either be 180 points or 20 points. Both values occur with equal probability of 50 percent. In the second period, the point and euro earnings for both 20 and 180 points can be found on the left and the right side of the task field. Only at the end of the experiment you will learn which earning will be paid in the second period.

It is important to understand that your earnings per ball caught are randomly generated in the second period. Which value your earnings have in one period, **neither** depends on the value your earnings had in previous periods **nor** on the way you behaved in the previous periods. Only at the very end of the experiment you will be informed about the actual value of your earnings in the second period. That implies that for the duration of the task, you do not know which earnings are relevant for payoff, 20 or 180 points.

**Costs for moves**

At the start of each shift the **cost for a move** is always zero points. The cost per move increases in the number of moves:

Cost per move = 0.1 * (number of moves so far)$^2$

The cost per move is rounded to the closest integer. A table with chosen function values is included in the instructions.

Example: Supposing the number of your current moves is 30. The costs per move are calculated as 30*30*0.1 = 90. The next click on “LEFT” or “RIGHT” consequently costs 90 points. After the next click, the number of your current moves increases by one. The costs per move are calculated as 31*31*0.1 = 96.1. The result is rounded to the next integer, 96.

**Shift result**

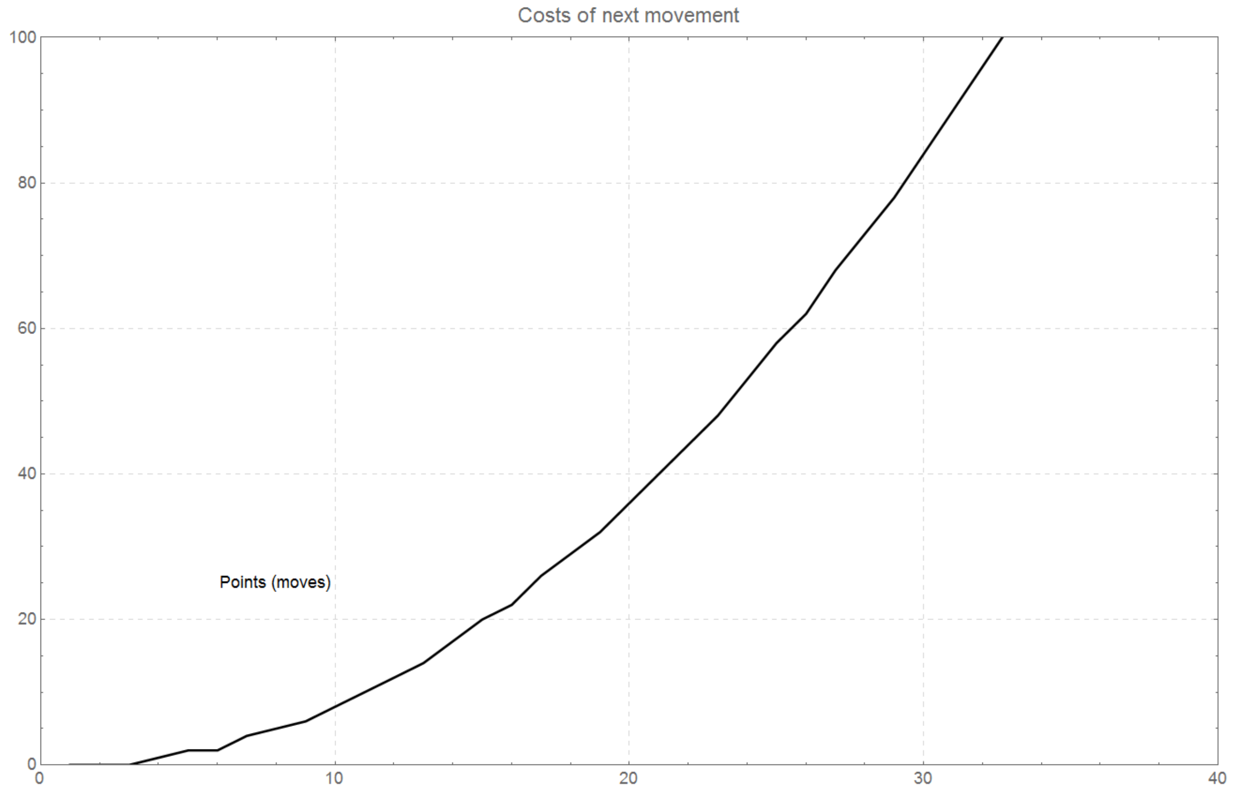
The sum of all the points you earned in one shift is your **shift result**. The higher your shift result, meaning the sum of all points earned in one shift, the higher is the payoff in this particular shift. The shift result is converted to euros as followed:

Shift result in euros = 4*[ln(shift result in points) – 7].

The following illustration shows the shift result in euros, depending on the points earned. A table with chosen function values is included in the instructions.

Example: Suppose the number of points you earned in the first shift of a round is 6400. Your result for this shift equals 4*[ln(6400) – 7) = 7.21 euros. In case you earn 100,000 points in the second shift, your result for this shift is 4*[ln(100,000) – 7] = 18.32 euros.

**Rounds 2 to 4**




The following sections inform you how rounds 2 to 4 differ from round 1.

**Round 2**

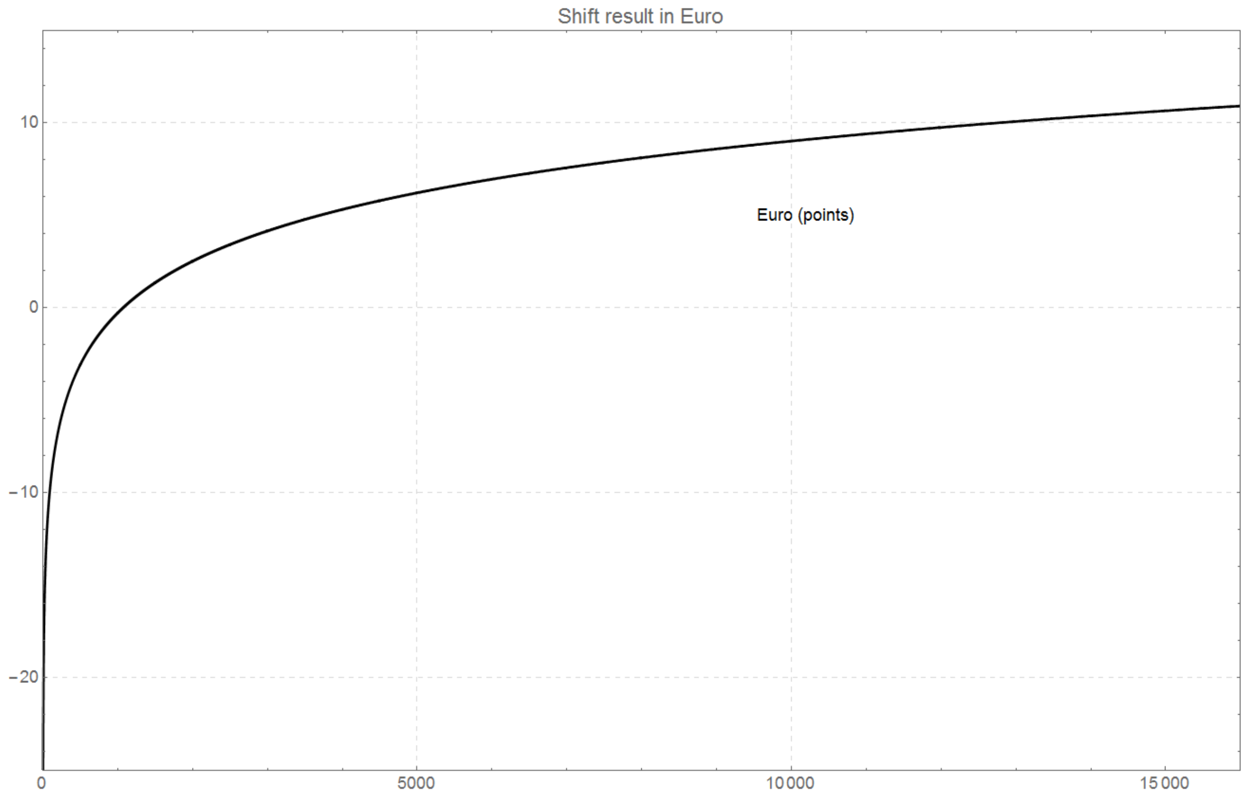
In round 2 you work on the task the same way you did in round 1, for two shifts (which correspond to the periods that last 180 seconds each). Now you have the opportunity to **save** points after the first shift. You can transfer points from the first shift to the second shift. Points that you save are subtracted from the first shift result (accordingly, you earn a lower euro amount in the first shift). The saved points are added to your second shift result (thus, resulting in a higher shift result in euros).

You can save at most so many points that your euro earnings in the first shift are zero. You cannot save a negative amount of points.

**Round 3**

In round 3 you can decide **how much time** you want to spend in each shift. Overall, you have **360 seconds** at your disposal. The earnings per ball caught in the first period (the first 180 seconds) remain 100 points and the earnings in the second period (the following 180 seconds) remain either 20 or 180 points.

With a button under the task field, you can decide when to end the first shift. After that, the second shift begins.




Example 1: Suppose you end the first shift after 120 seconds. Your shift result for the first shift will be calculated based on the earnings and costs for these 120 seconds. (During these 120 seconds, your earnings per ball caught equal 100 points since you are in the first period.) In the following shift 2, you work on the task for 240 seconds (360 minus 120 seconds). In the first 60 seconds of the second shift, you are still in period 1, meaning you earn 100 points per ball caught. In the following 180 seconds, you are in period 2 and earn either 20 or 180 points per ball caught. Your shift result in points in shift 2 is the sum of the earnings of both periods minus the cost for moves.

Example 2: Suppose you end the first shift after 240 seconds. During the first shift, you are in period 1 during the first 180 seconds and earn 100 points per ball caught. In the following 60 seconds, you are in period 2 and earn either 20 or 180 points (of which the costs are then subtracted). Throughout the second shift (which only lasts 120 seconds) you are in period 2 and earn either 20 or 180 points per ball caught.

**Round 4**

In round 4 you can save points after the first shift (just as in round 2) as well as decide on the time you want to spend in each shift (just as in round 3).

**Part 3**

The third part is with regards to content completely unrelated to the first two parts. The instructions for the third part will be shown only on your screen.

**Overall pay-out in euros**

The result for a round equals the sum of both shift results.

Round result = shift 1 result in euros + shift 2 result in euros.

The overall payoff is calculated as followed:

Overall payoff = result of a random period of part 1 + result of a random round of part 2 + amount earned in part 3

The payoff of the random round is rounded to cents. This amount can drop under zero euros, meaning your payoff might be **negative**. In this case, the loss will be settled with the earnings of the other parts. You will not leave this experiment with a loss: Should the overall payoff be negative, you do not get a pay-out.

**Questions**

Now please answer the quiz questions about the contents of these instructions. Please raise your arm once you are done. In case you have any questions, please also raise your arm. A person in charge will come to you and answer the question.

# C Tables with Selected Values of the Consumption and Cost Function (Part of the Printed Instructions)

Table C.1: Cost Function

| Costs | |
|---|---|
| Number of movements so far | Cost of next movement in points |
| 0 | 0 |
| 2 | 0 |
| 4 | 2 |
| 6 | 4 |
| 8 | 6 |
| 10 | 10 |
| 12 | 14 |
| 14 | 20 |
| 16 | 26 |
| 18 | 32 |
| 20 | 40 |
| 22 | 48 |
| 24 | 58 |
| 26 | 68 |
| 28 | 78 |
| 30 | 90 |
| 32 | 102 |
| 34 | 116 |
| 36 | 130 |
| 38 | 144 |
| 40 | 160 |
| 42 | 176 |
| 44 | 194 |
| 46 | 212 |
| 48 | 230 |
| 50 | 250 |
| 52 | 270 |
| 54 | 292 |
| 56 | 314 |
| 58 | 336 |
| 60 | 360 |

Table C.2: Consumption Function

| Shift earnings | |
|---|---|
| Earned points | Value in euros |
| 0 | -25.00 |
| 1000 | -0.37 |
| 2000 | 2.40 |
| 3000 | 4.03 |
| 4000 | 5.18 |
| 5000 | 6.07 |
| 6000 | 6.80 |
| 7000 | 7.41 |
| 8000 | 7.95 |
| 9000 | 8.42 |
| 10000 | 8.84 |
| 11000 | 9.22 |
| 12000 | 9.57 |
| 13000 | 9.89 |
| 14000 | 10.19 |
| 15000 | 10.46 |
| 16000 | 10.72 |

## D Translation of the Quiz Questions (With Correct Answers)

QUIZ QUESTIONS

Please answer the following questions before the experiment starts. With these questions we merely intent to make sure that you understand the instructions properly.

1. True or false? Your earnings in period 1 are always 100 points.
   ☒True ☐False

2. What is the probability that your earnings per ball caught in period 2 are 180 points?
   50%

3. True or false? In rounds 3 and 4 you can influence the total duration for which you earn 100 points per ball caught.
   ☐True ☒False

4. True or false? Each time a new shift begins the costs per movement are reset to zero.
   ☒True ☐False

5. In each shift increase the costs per movement in the number of movements so far. But this increase becomes flatter in the number of movements so far.
   ☐True ☒False

6. Suppose you earned 10,000 points in the first shift and 1,000 points in the second shift. What are your euro earnings in each shift and in the round?
   10,000 points = 8.84 euros; 1,000 points = -0.37 euros; together 8.47 euros

7. Suppose that you (based on the earnings given under 6.) saved 2,000 points. What are your euro earnings in each shift and in the round?
   8,000 points = 7.95 euros; 3,000 points = 4.03 euros; together 11.98 euros

8. Suppose that you spent 100 seconds in the first shift.

   a) How many seconds will you spend in shift 2?
      260 seconds

   b) For how many seconds will you earn 100 points per ball caught in shift 2?
      80 seconds

c) For how many seconds will you earn either 20 or 180 points per ball caught in shift 2?
   180 seconds

9. True or false? You will not learn your payoff during the entire experiment. Only at the very end you will learn this.
   ☒True ☐False

# E Example Screen-Shots of the Computer Interface

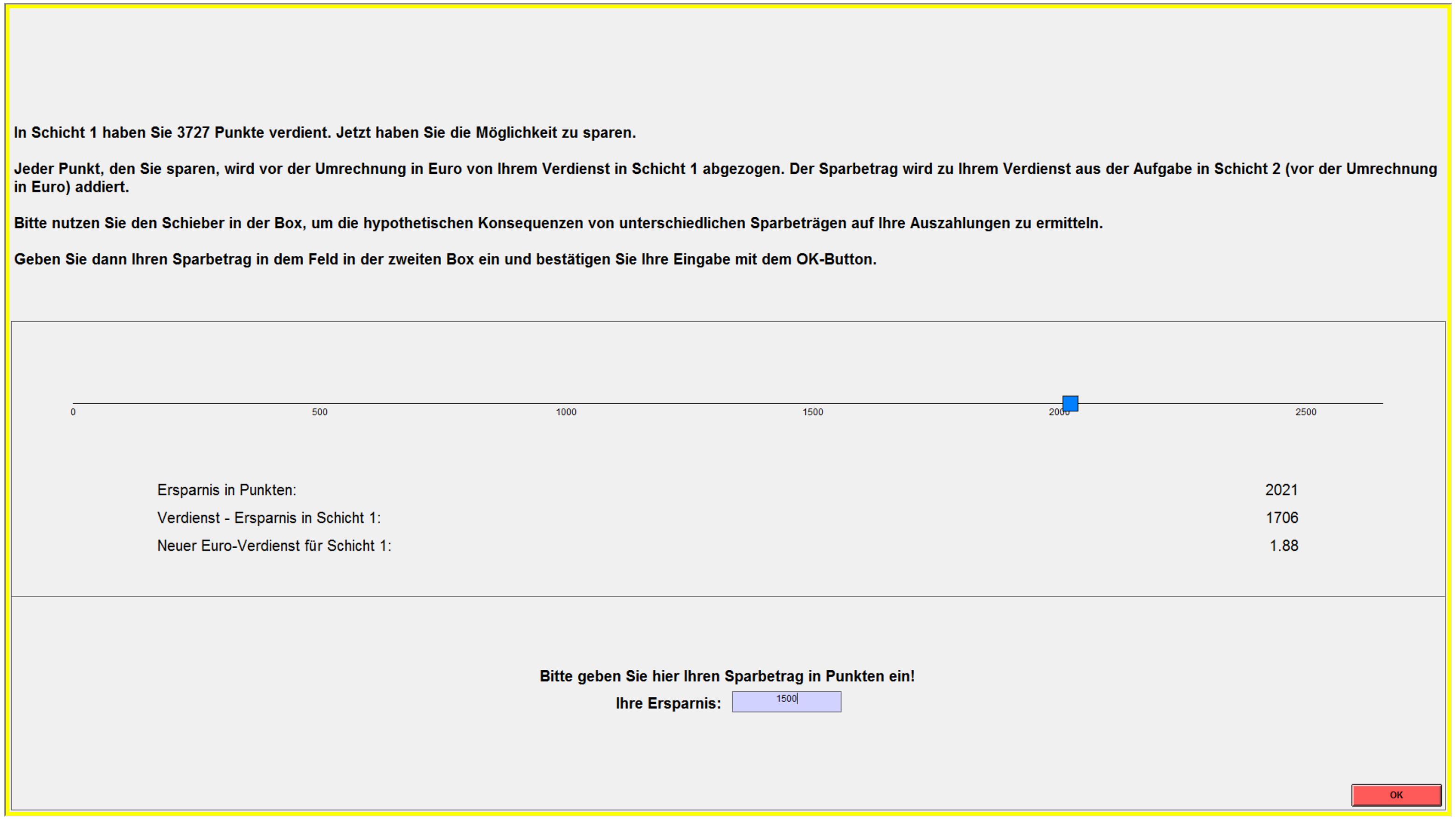



Figure E.1: Screenshot of Saving Screen.

*Source:* Own interface based on z-Tree.

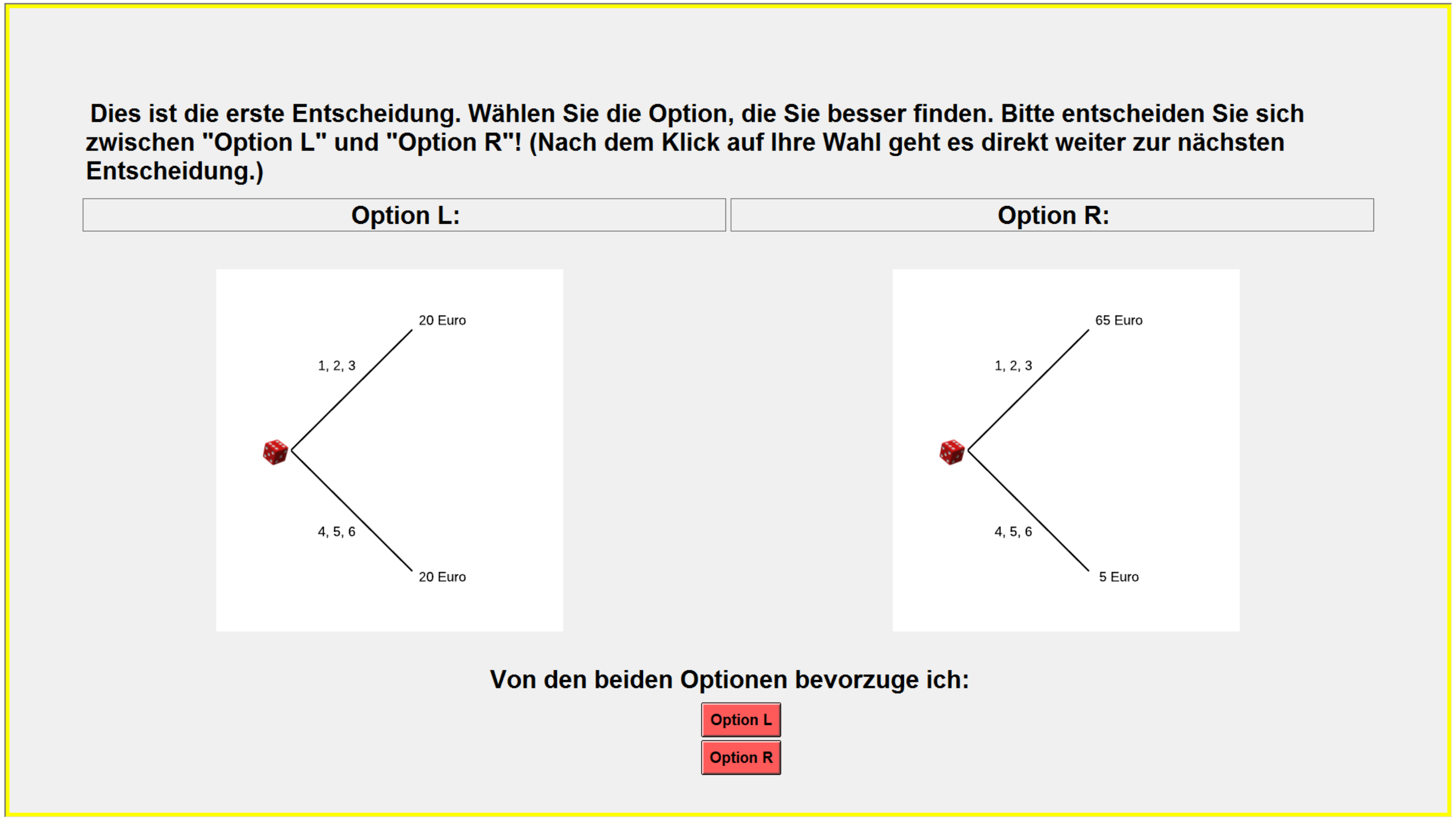



Figure E.2: Screenshot of the Experimental Interface with the Elicitation of the Risk Aversion and Prudence.

*Source:* Own interface based on z-Tree.

# F Sample Characteristics

Table F.1 presents summary statistics of our sample. The median person switched to the certain amount of 25 euros instead of the lottery between 65 euros or 5 euros with expected value of 35 in the questions for risk aversion. For both measures of prudence the median person preferred the most prudent option. Subjects classified as other indicated multiple switching points. This means that for all three measures, 85.4, 82.2, and 84.4 percent chose consistently to switch only once. However, they seem not all to follow expected utility theory: while almost 90 percent of subjects behaved according to a coefficient of relative prudence of greater than 2, only about 41 percent chose consistently with a coefficient of relative risk aversion of greater than 1.

Table F.1: Summary of Subjects' Observable Characteristics

| | % | SD | | % |
|---|---|---|---|---|
| Age | 23.0 | (3.90) | *Field* | |
| Female | 60.9 | (48.92) | Psychology | 1.56 |
| Semester | 5.0 | (3.84) | Other | 8.85 |
| Extremely risk averse | 42.2 | | Economics | 10.42 |
| Very, very risk averse | 10.9 | | Humanities | 10.42 |
| Very risk averse | 15.6 | | Sciences | 12.5 |
| Risk averse | 9.4 | | Other social science | 17.19 |
| Not risk averse | 4.7 | | Law | 18.75 |
| Risk loving | 2.6 | | Business | 20.31 |
| Other | 14.6 | | *Subjective Effort* | |
| *Variance* | | | Not demanding at all | 6.25 |
| Extremely prudent | 65.1 | | Not demanding | 28.65 |
| Very prudent | 7.3 | | Not demanding, not effortless | 35.42 |
| Prudent | 4.7 | | Somewhat demanding | 21.35 |
| Not prudent | 4.2 | | Quite demanding | 6.77 |
| Other | 18.8 | | Very demanding | 1.56 |
| *Stakes* | | | *Attention to Risk* | |
| Extremely prudent | 68.2 | | Inattentive | 7.29 |
| Very prudent | 7.8 | | Risk pessimist | 59.38 |
| Prudent | 3.6 | | Risk realist | 24.48 |
| Not prudent | 4.7 | | Risk optimist | 8.85 |
| Other | 15.6 | | | |
| RRA greater 1 | 46.9 | | | |
| RP greater 2 | 89.6 | | | |
| RRA greater 1 and RP greater 2 | 41.1 | | | |

*Source:* Authors' calculations.

# G Expected Payoff by Treatment

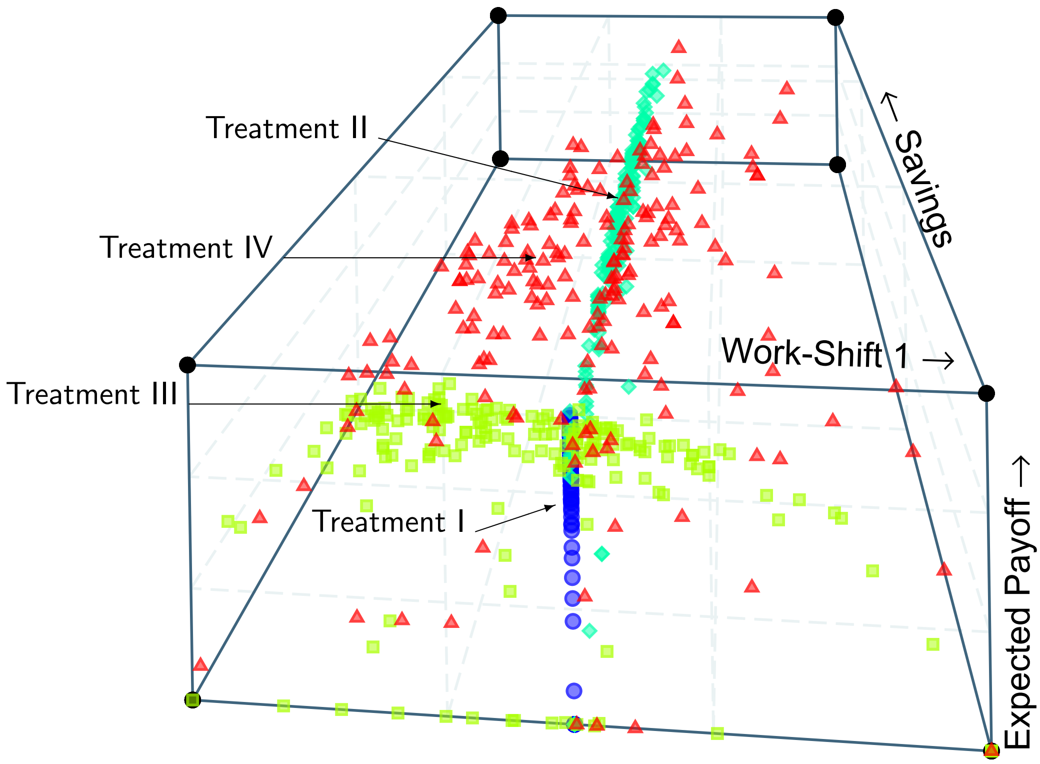



Figure G.1: Work-Shift-Savings-Payoff Space

*Legend:* Each color and marker represents observations from a given treatment. Expected payoffs range from 0 euros to 14.89 euros. The fraction allocated to work-shift 1 ranges from 0 to 100 percent, savings from 0 points to 6,473 points.

*Source:* Own presentation using GRAPH3D.

## H Estimation of Production Functions

Figure H.1 provides an overview of the means, histograms, and kernel density distributions of the number of caught balls divided by the number of movements (balls per movement$_i$ = caught balls$_i$/ movements$_i$) in the two shifts for all four treatments, pooled for all subjects. The distributions appear to be very similar and display a fair amount of dispersion. We conduct pairwise Kolmogorov-Smirnov tests between the four treatments for the variable balls per movement in the two shifts: all pairwise comparisons cannot reject the equality of the distributions at the 5 percent percent level. We take this as evidence that we do not observe much learning in our treatments after the previous three trial periods.

Table H.1 displays pairwise correlations between balls per movement in the different treatments and shifts. All correlation coefficients lie between 0.420 and 0.729 and are significantly different from zero at the 1 percent level. We take this as evidence for behavioral consistency as the dispersion observed in Figure H.1 is driven by the subjects' heterogeneity of ability to perform the real-effort task: subjects who perform well in the real-effort task (and catch many balls with one movement) in one treatment and shift, do so in all the other treatments and shifts as well.

Figure H.2 displays the two estimated production functions and the observations in all four treatments.

Table H.1: Pairwise correlations of balls per movement in the two work-shifts in all treatments

| | TI, shift 1 | TI, shift 2 | TII, shift 1 | TII, shift 2 | TIII, shift 1 | TIII, shift 2 | TIV, shift 1 |
|---|---|---|---|---|---|---|---|
| TI, shift 1 | 1 | | | | | | |
| TI, shift 2 | 0.548*** | 1 | | | | | |
| TII, shift 1 | 0.598*** | 0.542*** | 1 | | | | |
| TII, shift 2 | 0.464*** | 0.451*** | 0.525*** | 1 | | | |
| TIII, shift 1 | 0.503*** | 0.420*** | 0.605*** | 0.521*** | 1 | | |
| TIII, shift 2 | 0.547*** | 0.474*** | 0.586*** | 0.421*** | 0.564*** | 1 | |
| TIV, shift 1 | 0.550*** | 0.462*** | 0.615*** | 0.477*** | 0.729*** | 0.512*** | 1 |
| TIV, shift 2 | 0.553*** | 0.570*** | 0.597*** | 0.429*** | 0.533*** | 0.626*** | 0.620*** |

Significantly different from zero at $p < 0.01$: ***.
*Source:* Own calculations.

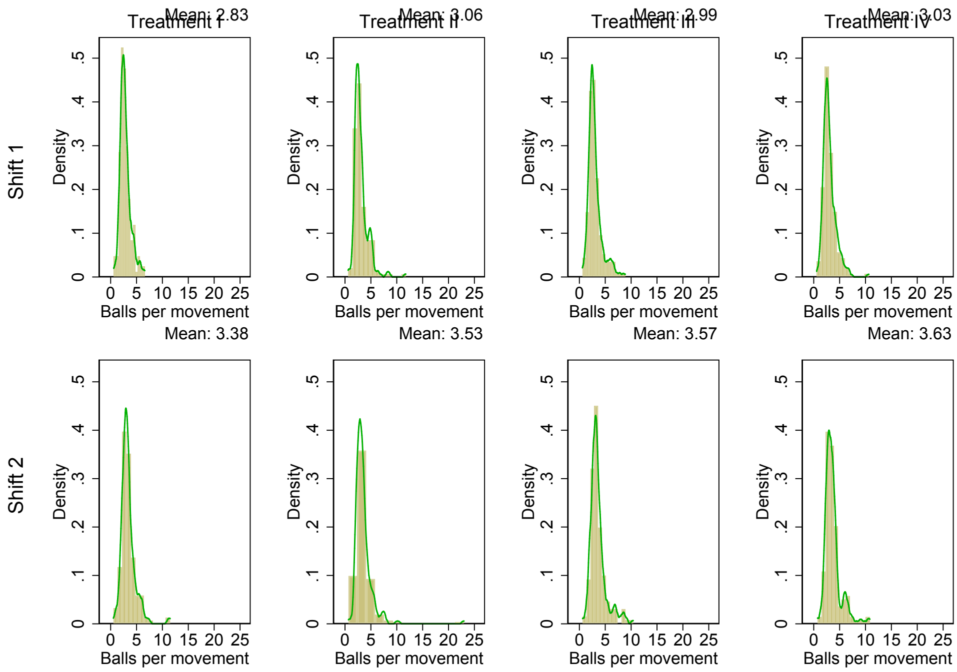
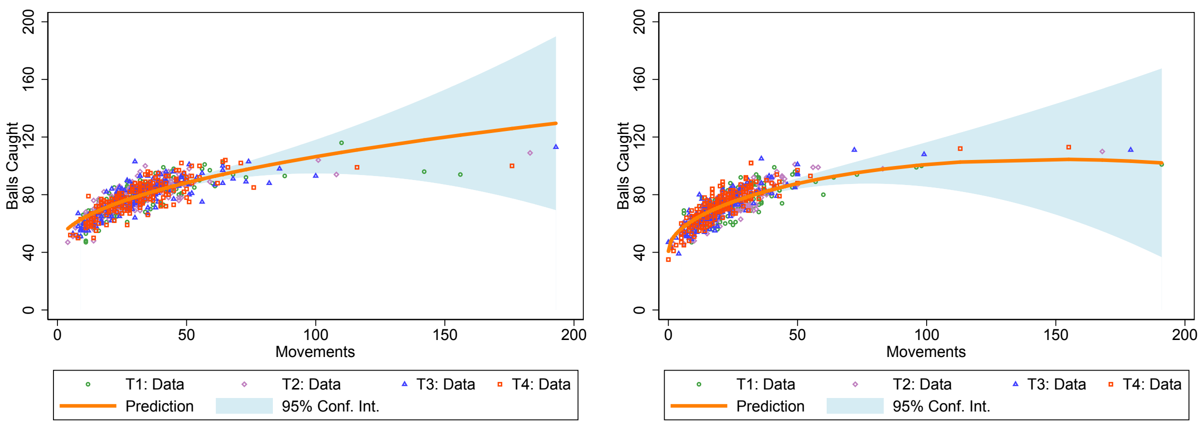



Figure H.1: Means, distributions and kernel density distributions of balls per movement in the two shifts of all four treatments

*Source:* Authors' presentation.



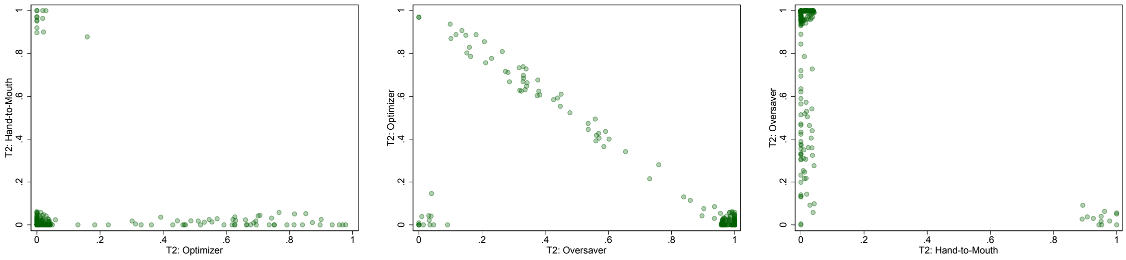

(a) Period 1

(b) Period 2

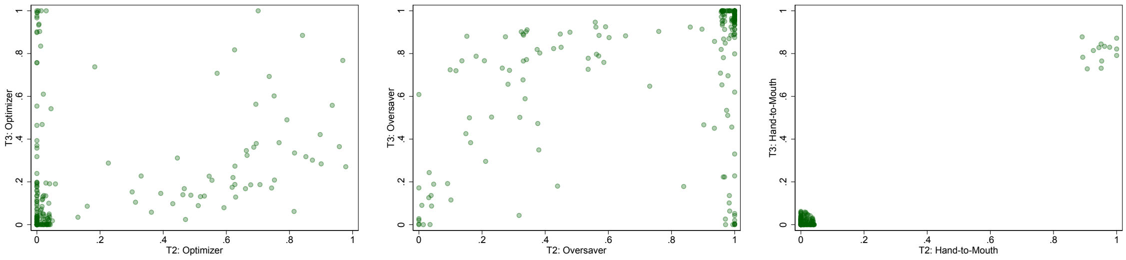
Figure H.2: Separate estimates of the production functions and observations in all four treatments

*Source:* Authors' presentation.

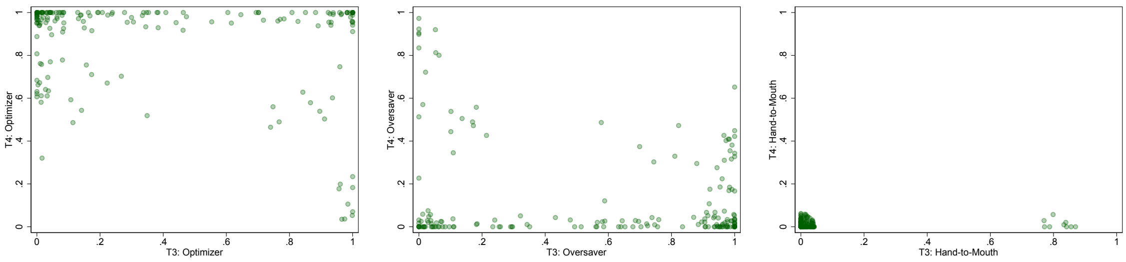
# I Posterior Probabilites



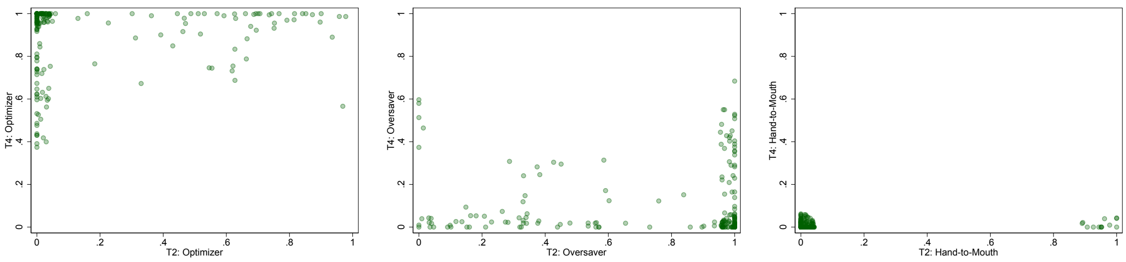

(a) Assignment Probabilities in Treatment II




(b) Group Probabilities in Treatment II and Treatment III




(c) Group Probabilities in Treatment III and Treatment IV




(d) Group Probabilities in Treatment II and Treatment IV

Figure I.1: Posterior Probabilities within Treatment II and Over Treatments

*Notes:* First row: Mass in upper left and lower right corners indicates successful classification. Rows two to four: Mass in upper right corner indicates that a subject followed the same strategy in two treatments, lower right or upper left indicates change in strategies.

*Source:* Authors' presentation.