Joint Consumption Smoothing in the Lab*

Logan Miller **Ryan Rholes**

University of Arkansas Texas A&M University

October 2020

Abstract

This paper revisits the learning-to-optimize literature to ask whether or not the pervasive practice of using individuals as representative decision-makers is the best approach to testing the potential limits of rationality in the context of dynamic optimization. To do this, we study the relative ability of individuals and pairs who must spend, borrow, and save to solve a finite lifecycle problem featuring a stochastic income process. We find that joint decisions are significantly better aligned with the rational, representative-agent benchmark than individual decisions. This performance difference leads subjects participating in the Pairs treatment to earn about 40% more than subjects in the Individuals treatment. Chat data reveals that pairs frame their discussions in terms of spending, rather than saving or borrowing, and develop simple consumption heuristics that are largely invariant to past errors and to wealth balances. We then classify subjects into heuristic types and find that pairs and individuals both use constant MPC heuristics most of the time. However, pairs have a much higher MPC than individuals.

JEL classifications:

Keywords:

*We thank Luba Petersen, Catherine Eckel, Don Koh, Sherry Li, Pete McGee, and seminar participants at the 7th Annual TExAS conference for valuable feedback. We also thank Tenley Getschman and Caleb Hill for outstanding research assistance.

† Rholes: Department of Economics, Texas A&M, Allen Building, 3035 Research Parkway, College Station, TX 77845, United States, rar145@tamu.edu.

∓ Miller: Department of Economics, University of Arkansas, Walton College of Business, Fayetteville, AR, Lmiller@walton.uark.edu.

1 Introduction

Modern macroeconomic theory typically models decision-makers as rational individuals capable of solving complex, dynamic choice problems without influence from, and independent of, other economic agents. Indeed, this representative agent assumption critically influences how monetary and fiscal policy operate in these models. Macroeconomists, who entered the lab in earnest following Lucas (1986), have taken this representative agent assumption seriously when designing experiments to test the micro-foundations of macroeconomic theory (Duffy, 2006).

However, using individuals as representative decision-makers in the lab may not fully capture the behavior modeled in macroeconomic theory. Among the macroeconomic models that do assume a representative household or firm, most do not explicitly rule out the notion of joint decision making within that household or firm. In fact, theory sometimes explicitly models this. For example, many theories include the assumption of multiple decision roles in the household, like the worker-shopper pair introduced by Lucas and Stokey (1987). Suggestive evidence using Panel Study of Income Dynamics (PSID) data shows that there are many households making these sorts of joint decisions and joint-decision households respond quite differently to negative income shocks than individual-decision households.[1] To understand this heterogeneity, it is critical to understand both joint and individual decision-making.

This paper revisits dynamic optimization in the laboratory. Our objective is to better understand differences in how individuals and pairs, who form joint decisions, solve dynamic optimization problems. To do this, we explore a dynamic consumption problem in a simple, finite-period environment that features a stochastic income process and allows for both borrowing and saving. Pairs in our experiment engage in unrestricted communication via a chat window to form joint decisions.

We find that allowing pairs to communicate to form joint decisions significantly improves decisions, compared to individuals, relative to the rational, representative benchmark. This is true when we measure performance along both the unconditionally- and conditionally-optimal consumption paths. On average, pairs earn about 40% more than individuals after accounting for fixed show-up fees paid to all subjects.

Chat data collected from our pairs treatment provides valuable insight into how subjects in those treatments think through dynamic choice problems. Textual analysis reveals that subjects almost exclusively frame discussions in terms of spending even though saving and borrowing are important components of earnings maximization. Further, subjects develop simple, invariant heuristics that can lead to persistent and compounding errors. This textual analysis corroborates our classification of individuals and pairs into different consumption heuristics.

We find that both pairs and individuals adhere primarily to a constant marginal propensity to consume (MPC) heuristic, but find that pairs have higher MPCs on average. This leads to fewer instances of under-spending, mitigates the compounding nature of conditional errors, and also leads to less severe and less frequent consumption binges relative to individuals.

[1] We provide details about this suggestive evidence in Section 7.3 of the Appendix.

Finally, we see that pairs often include a participant who nudges the initial joint decision toward the rational benchmark, most often by advocating for higher spending.

2 Literature Review

There is an extensive literature, thoroughly discussed in Duffy (2016) and summarized here, that studies the ability of individuals to solve dynamic stochastic optimization problems.

Hey and Dardanoni (1988) study dynamic optimization in a pure exchange economy in which individual subjects, acting as representative agents, earn a constant return on savings, face a no-borrowing constraint, and receive a per-period stochastic income. They find that consumption decisions deviate significantly from the rational benchmark and that consumption behavior seems to be largely time-dependent. Carbone and Hey (2004) and Carbone (2006) simplify this design by eliminating discounting, using a finite horizon, 25-period model, and by simplifying the stochastic income process by using a two-state Markov process to determine whether a subject is employed and receives and income or is unemployed and does not. They find that subjects overreact to an increased probability of remaining employed, under-react to the probability of becoming employed, are myopic, and that current consumption is too dependent upon current income. Carbone and Infante (2014) study behavior in a dynamic optimization game under certainty, risk, and ambiguity. The authors find that subjects significantly under consume when faced with ambiguity relative to risk and certainty.

Others have studied various types of learning in dynamic optimization by allowing individual decision-makers to interact in various capacities. Ballinger, Palumbo, and Wilcox (2003) study inter-generational learning by exposing subjects to the prior decisions of another subject and, in some treatments, by allowing communication with earlier generations. The authors find that later generations, in treatments featuring both high- and low-income uncertainty, do better at solving this lifecycle problem. Brown et al. (2009) study learning in the context of dynamic optimization by exposing current subjects to the decisions of subjects from previous 'private learning' treatments. Subjects in both treatment types solved the same buffer-stock model many times, while subjects in social learning treatments were able to view the decisions of the best, the worst, and one randomly selected subject from a private-learning treatment before making their own decision. This process of social learning drastically improved the speed of own-learning compared to rates of own-learning from subjects in private-learning treatments. Carbone and Duffy (2014) study the impact of contemporaneous social learning by revealing to individuals the average level of past consumption. This revelation causes subjects to deviate further from both the conditionally- and unconditionally optimal consumption path. Bao, Duffy, and Hommes (2013) test the ability of subjects to jointly optimize and forecast. They find that pairing subjects together and having each subject either forecast or optimize leads to quicker convergence to the rational expectations equilibrium than does having a single subject perform both tasks.

Ubiquitous across these previous studies is the use of individual decision-makers. However, there are also studies comparing the behavior of groups in macroeconomic settings. For example, Blinder and Morgan (2005) show that groups outperform individuals setting monetary policy to maintain to stabilize an experimental economy around inflation and employment targets. This finding was corroborated by Lombardelli et al. (2005) who also show that groups outperform individuals as policymakers because groups can strip out the effect of bad play in a given period, and because group members are able to share information and learn from each other's interest rate decisions. Similarly, Rholes and Petersen (2020) show in a learning-to-forecast experiment that aggregating over group expectations produces more stable inflation dynamics than do individual expectations.

Most closely related to our work are Carbone and Infante (2015), and Carbone, Georgalos, and Infante (2019), which both study the differences between pairs and individuals in a dynamic optimization setting. Carbone and Infante (2015) conclude that stable pairs perform no differently than individuals in solving the life cycle problem. We find the exact opposite – pairs in our experiments consistently outperform individuals as planners. Carbone, Georgalos, and Infante (2019) compare group and individual performance in an optimization task while facing either risk or ambiguity. Here, the authors see differences between individuals and pairs in some settings but not in others. Both Carbone and Infane (2015), and Carbone, Georgalos, & Infante (2019) consider an environment that includes a positive interest rate, does not allow borrowing within a period, and features a bimodal income distribution. The confluence of these design choices yields an environment in which it is optimal for subjects to accumulate wealth and increase spending toward the end of the lifecycle. Our work is complementary in that it provides some insight into when, and possibly why, individual and joint-decision makers behave differently when making dynamic choices.

Finally, we also contribute to the extensive literature that studies differences between groups and individuals. Examples are Cooper and Kagel (2005) who find that teams play more strategically than individuals and Kugler, Bornstein, Kocher, and Sutter (2007) who have both pairs and individuals play a trust game and find that groups are less trusting than individuals but are equally trustworthy. Charness and Sutter (2012), note that group choices better align with standard game-theoretic predictions, while individual choices often exhibit biases, cognitive limitations, and social consideration. Kagel and McGee (2016) also note that in finite, repeated prisoner dilemma games, two-person teams start with significantly less cooperation than individuals, however this quickly switches to teams cooperating more than individuals. Team dialogue shows increased payoffs from cooperation, the anticipation of opponents' recognition of the same strategies, and provides the basis for cooperation despite anticipating defection near the end of the game.

3 Theory

Subjects in both our Individual and Pairs treatments maximize their discounted lifetime utility, subject to an intertemporal budget constraint:

$$max\ \mathbb{E}_0 \sum_{t=1}^{t=T} \beta^t U(c_t) \tag{1}$$

$$s.t. \ \sum_{t=1}^{t=T} c_t \leq \sum_{t=1}^{t=T} w_t + a_0 \tag{2}$$

where c_t is consumption, a_0 is initial wealth, and w_t is an i.i.d. per-period stochastic income with $w \sim U(\underline{w}, \overline{w})$. Subjects in our experiment save freely and borrow up to $\underline{w}$ in all but the final decision period. We denote saving and borrowing throughout as s_t.

We induce the quadratic utility function

$$U(c_t) = \phi c_t - \frac{1}{2}c_t^2. \tag{3}$$

This functional form is useful for several reasons. First, it allows subjects to consume zero in any period without incurring negative utility. Second, it is concave across the action space, which induces a consumption smoothing motive.[2] Finally, combining this functional form with equations above yields Hall's (1978) stochastic equation:

$$c_t = (1-\kappa)\phi + \kappa \mathbb{E}_t c_{t+1} \tag{4}$$

where $\kappa \equiv \beta(1+r)$. Setting $\beta = 1, \ r = 0$ allows us to reduce Equation (4) to:

$$c_t = \mathbb{E}_t c_{t+1}. \tag{5}$$

Solving by backward induction yields our unconditionally-optimal consumption path [3]

$$\begin{cases} c_{T-j} = y_{T-j} + s_{T-j-1}, & j = 0 \\ c_{T-j} = \frac{j}{j+1}\mu + \frac{j-1}{j+1}(y_{t-j} + s_{t-j-1}) & j \in (1, 2, ..., T-1) \end{cases}$$

This solution indicates that optimal consumption is a linear function of the mean of the income distribution, μ, and period wealth. Intuitively, subjects should focus less on the income distribution and more on wealth as the game nears completion.

We also consider subjects' decisions relative to a conditionally-optimal level of consumption, $\hat{c_t^*}$, which accounts for past consumption errors by recalculating optimal consumption for each remaining period conditional on past mistakes.

$$\hat{c_t^*} = c_t^* + \frac{(y_t - c_t^*) + s_{t-1}}{T-(t-1)}, \ \forall \ t \ \in \ \{2, ..., T-1\} \tag{6}$$

[2]Restrictions on ϕ are such that, across the feasible action space, the first derivative of $u(c_t)$ is strictly positive and the second derivative is strictly negative.

[3]Notice that if $r > 0$ then per-period consumption is lower and per-period savings are higher in most periods. This might lead to behavior similar to that found in Carbone and Infante (2015).

We plot the unconditionally-optimal consumption path alongside the income processes used in all experimental sessions in Figure 1. The unconditionally-optimal path is the same for all subjects because we hold the stochastic income processes used in blocks 1 and 2 constant across all experimental sessions. Conversely, the conditionally-optimal path depends on individual deviations from the unconditionally-optimal consumption path. Therefore, we do not plot those here.

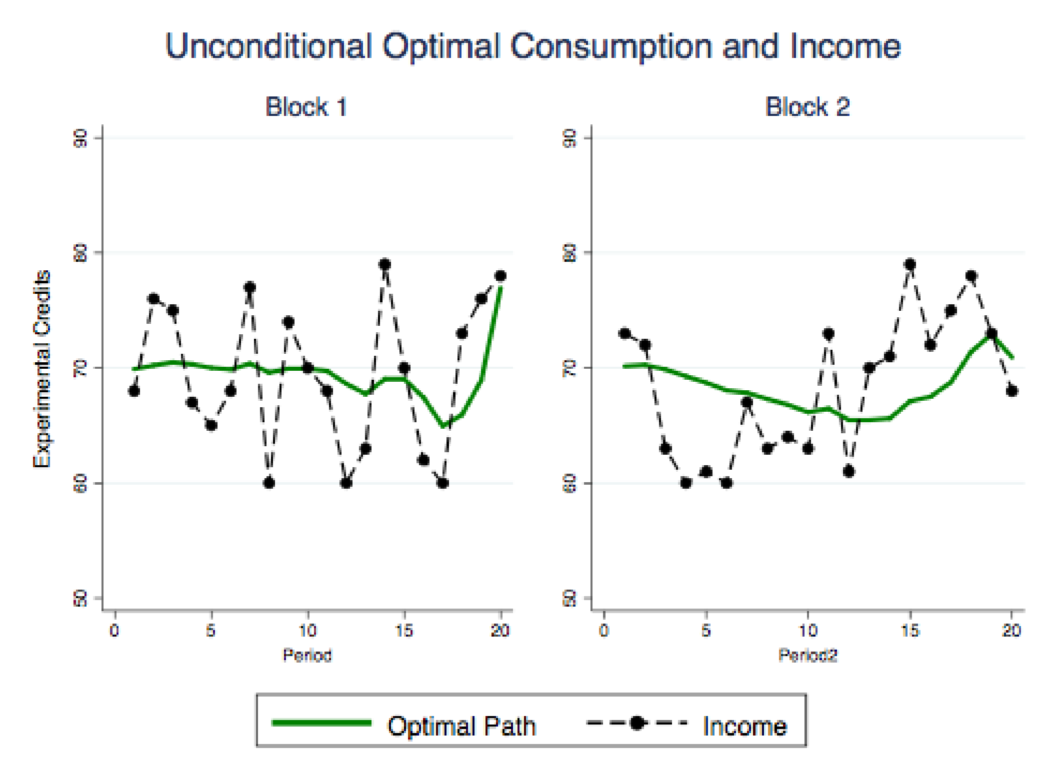

Figure 1: This figure shows the unconditionally-optimal consumption path for decision blocks 1 and 2 of all experimental sessions. The graph also includes the pre-drawn stochastic income processes used for blocks 1 and 2 in all experimental sessions.

4 Experimental Design

We use a simple 2×1, between-subjects design built around a standard learning-to-optimize (LTO) framework where we compare the ability of individuals and pairs to solve the finite-period smoothing problem outlined in Section 3. We set $\phi = 1,600,\ \overline{w} = 80,\ \underline{w} = 60,\ \beta = 1,\ r = 0$ for all sessions. The consumption smoothing motive in our setting comes from the concavity of the induced quadratic utility function. Subjects in both our Individual and Pairs treatment solved two, twenty-period lifecycle problems. We used two different pre-drawn stochastic income processes for these two lifecycle problems so that we could hold the income process constant across all periods.

Sessions began with a 6-question, individual-level Cognitive Reflection Test (CRT) (Frederick, 2005) as a proxy for cognitive ability. Subjects had 90 seconds to answer each CRT question and earned $.25 for each correct answer. We followed this with an individual-level

Eckel-Grossman test of risk preferences (Eckel and Grossman, 2002). Following these two tasks, subjects in individual sessions worked alone to solve both lifecycle problems. We randomly formed individuals into stable pairs for the Pairs treatment. Subjects in each pair then worked together to solve both lifecycle problems, using a chat box to engage in unrestricted communication. Subjects were not time-constrained when solving the lifecycle problem in either treatment. We ended each session with a demographic survey that also included a survey of attitudes toward debt and spending.

Instructions provided detailed information about the utility function, income process, lifecycle duration, and borrowing and saving so that they had sufficient information to fully solve the lifecycle problem. Further, we provided subjects with information about their per-period income, and their current bank account balance to help them keep track of their borrowing/savings. We also provided subjects with a consumption smoothing tool to reduce the cognitive complexity of the problem. To use the tool, subjects could propose a hypothetical level of consumption and learn the corresponding levels of utility (we called this consumption points in the game), savings or debt, and the marginal utility of consumption (we called this the 'marginal increase' in the game). Subjects could use this tool as many or as few times as desired. We provide an example of the decision screen for an individual in Figure 8 and for pairs in Figure 9 in the Appendix.

We converted individual consumption points to U.S. dollars at the rate of 50 points per \$1 and pair consumption points at the rate of 25 points per \$1. This conversion scheme holds individual-level incentives constant across treatments. Subjects also received a \$10 show-up fee. We conducted all sessions at the University of Arkansas's Behavioral Business Research Laboratory. All subjects were undergraduates and were inexperienced. We have 26 individual observations and 27 pair observations. We implemented our experiment using zTree (Fischbacher, 2007).

5 Results

We begin by considering the average consumption decisions of both individuals and pairs, which we graph in Figure 2.[4]

Worth noting in Figure 2 is that average consumption increases over time for both individuals and pairs in both decision blocks. Also, note in the top right panel of Figure 2, which shows average consumption decisions during the second decision block, that both pairs and individuals, on average, consume too little for approximately the first 10 periods relative to the unconditionally-optimal consumption path. This period of significant under consumption corresponds to a time where subjects must borrow to behave optimally. To see this, refer to the stochastic income process depicted in Figure 1, and note that per-period income was consistently below the unconditionally-optimal level of consumption.

Additionally, we consider the per-period level of within-group heterogeneity in consumption decision, which we graph in Figure 2, and note that participants in our Individual treatment exhibit considerably more heterogeneity in consumption than do pairs.

[4]See the appendix for graphs depicting individual consumption choices and absolute consumption errors for subjects in both Individual and Pair treatments.

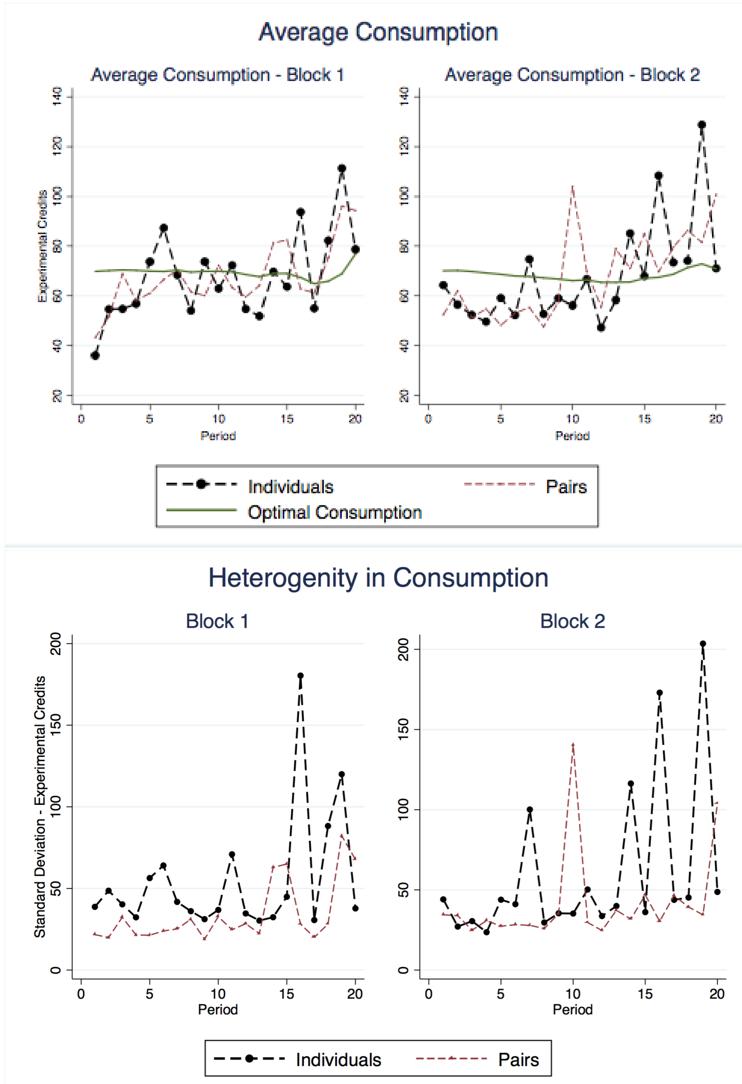

Figure 2: The top panel of this figure shows the average level of consumption expressed as experimental credits in each period by treatment. The bottom panel shows within-treatment heterogeneity in consumption by period.

Next, we average consumption errors within-period by treatment and plot this in Figure 3. Both a Mann-Whitney U test and a two-sample t-test confirm statistically significant differences in errors between pairs and individuals. The Mann-Whitney U test is highly significant ($P < .001$). This is true when considering the full sample of decisions and also when considering both blocks independently.[5] The two-sample T-test, where our null hypothesis is that the mean of the consumption error of individuals and pairs is zero, rejects the null hypothesis and confirms the mean consumption error of pairs (29.36) is statistically significantly lower ($p < .001$) than individuals (36.67).

We use a mixed-effects model to estimate performance differences. Though random effects models are common in this literature (examples are Carbone and Duffy (2014), Ballinger et al. (2011)), a Hausman test indicates the need to control for fixed effects, which perhaps

[5]Considering the decisions in block 2 reduces the level of significance to $p = .002$.

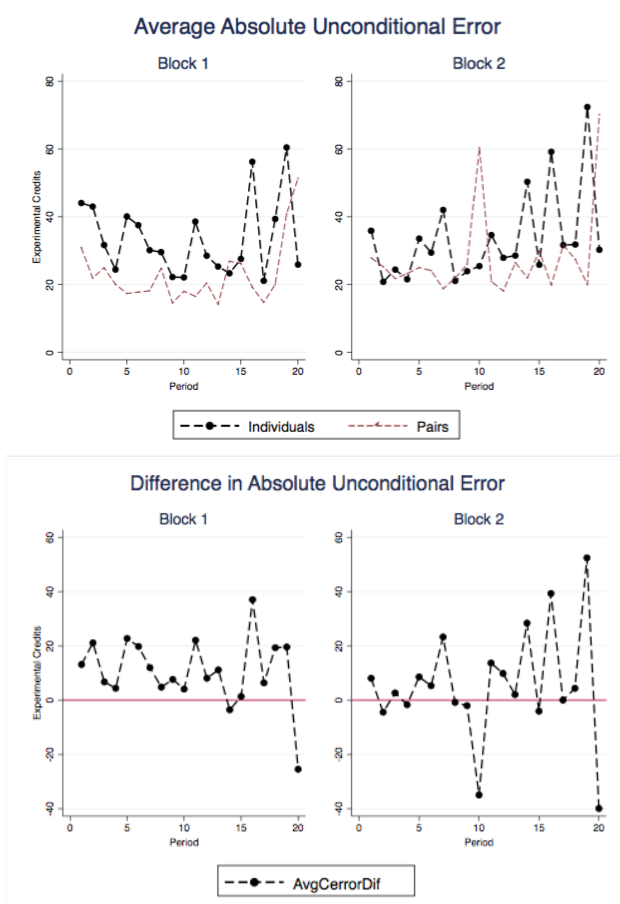

Figure 3: Panel (a) of this figure depicts the average unconditional consumption error by period and treatment for decision blocks one and two. Panel (b) plots the difference between the absolute value of the average unconditional individual consumption error and the absolute value of the average unconditional pair consumption error. Thus, values above (below) zero indicate that pairs (individuals) outperformed individuals. For example, a value of 20 here would imply that the average unconditional consumption error of individuals was 20 experimental credits larger than pairs in that period.

result from static session effects (Fréchette, 2012).[6] We project absolute unconditional errors onto a set of controls and a dummy that indicates whether or not a pair made a given consumption decision. The results of this are reported in Table 1 using robust standard errors.

Each column in Table 1 includes additional controls relative to previous columns. Column 4 includes controls for risk preferences measured using the Eckel-Grossman risk elicitation task both before and after completing the dynamic optimization game. Regardless of specification, we see in Table 1 that pairs significantly outperform individuals. Our preferred specification, shown in column 4, indicates that subjects making joint decisions in our Pairs treatment were 9.56 EC's closer, on average, to the unconditionally-optimal level of consumption.

Next, we analyze differences in conditional consumption errors. This approach, as discussed in Section 3, accounts for initial errors by distributing accrued savings/debt over remaining decision periods. We show deviations from the conditionally-optimal consumption path in

[6]We also include results from a random-effects model in our appendix.

Table 1: Regression Table: Absolute Unconditional Errors

	(1)	(2)	(3)	(4)
Pairs	-8.05***	-8.06***	-8.07***	-9.56***
	(0.001)	(0.001)	(0.001)	(0.000)
Wealth		0.04***	0.05***	0.03***
		(0.000)	(0.000)	(0.000)
Income		0.37**	0.40**	0.39**
		(0.024)	(0.017)	(0.019)
CRT Score		0.95	1.51	1.67
		(0.406)	(0.204)	(0.156)
Block			3.95	3.77
			(0.094)	(0.107)
Risk				✓
Constant	33.54***	5.29	1.06	5.08
	(0.000)	(0.632)	(0.928)	(0.666)
N	2040	2040	2040	2040

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 4. Both a Mann-Whitney U test ($p < .01$) and a two-sample t-test ($p < .01$) confirm that the average conditional consumption error for pairs (39.28) is statistically significantly lower than for individuals (57.94). However, it is worth noting that there is no statistically significant difference between the average conditional error of pairs and individuals whenever we only consider the first ten periods of block 2 ($p = 0.5516$).

Also worth noting in Figure 4 is the gradual buildup of errors toward the end of each decision block. This is likely due to the adoption of simple and invariant consumption heuristics. This is because the conditionally-optimal path assumes subjects will account for previous mistakes in remaining decisions, these invariant heuristics are increasingly penalized when moving along the conditionally optimal consumption path. We discuss this further in our textual analysis section.

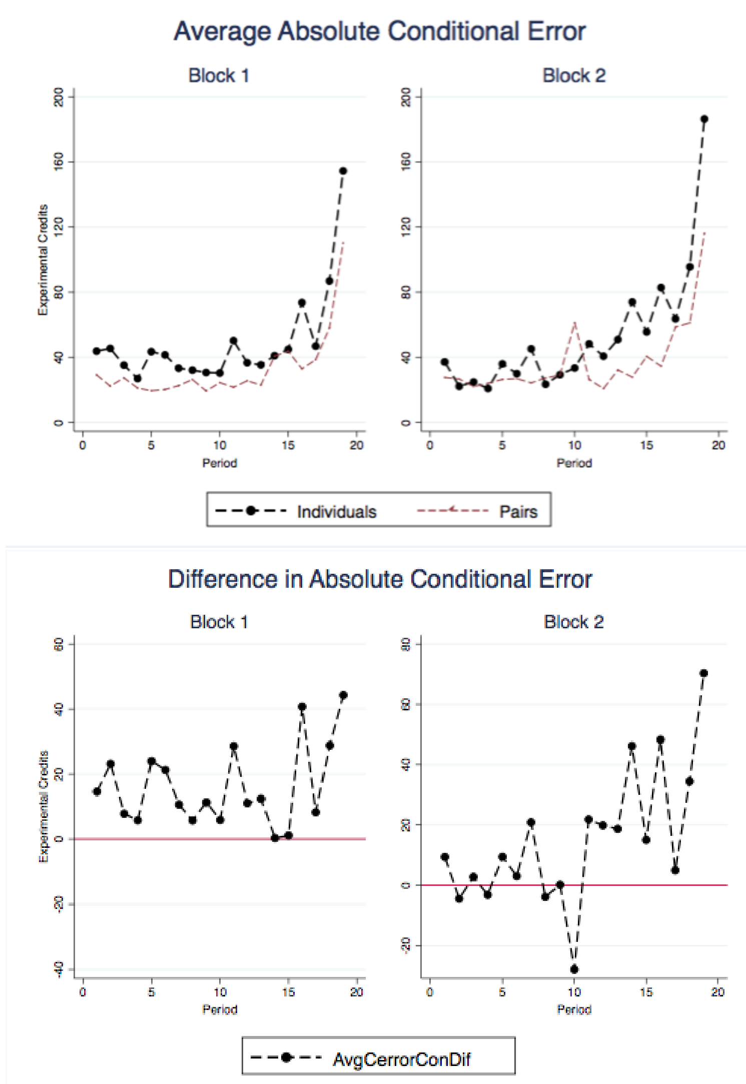

Figure 4: Panel (a) of this figure depicts the average unconditional consumption error by period and treatment for decision blocks one and two. Panel (b) plots the difference between the absolute value of the average conditional individual consumption error and the absolute value of the average conditional pair consumption error. Thus, values above (below) zero indicate that pairs (individuals) outperformed individuals. for example, a value of 20 here would imply that the average conditional consumption error of individuals was 20 experimental credits larger than pairs in that period.

Table 2 reports the results of using a mixed-effects model to estimate treatment-level differences in absolute conditional consumption errors. Similar to Table 1, we project the absolute value of conditional consumption errors onto a pairs dummy and a set of controls using a mixed-effects model. Each column includes additional controls relative to the previous column. Column 4 controls for risk preferences. Results from this exercise indicate that subjects making joint decisions in our Pairs treatment were more than 15 ECs closer, on average, to the conditionally-optimal level of consumption. This suggests that pairs do a better job than individuals of adjusting consumption behavior to account for past deviations from the rational benchmark.

Since subjects in our experiment are concerned with earnings maximization, it perhaps makes

Table 2: Regression Table: Absolute Conditional Error

	(1)	(2)	(3)	(4)
Pairs	-15.58***	-15.56***	-15.58***	-15.67***
	(0.000)	(0.000)	(0.000)	(0.000)
Wealth		0.14***	0.15***	0.13***
		(0.000)	(0.000)	(0.000)
Income		1.73***	1.77***	1.76***
		(0.000)	(0.000)	(0.000)
CRT Score		-1.361	-0.34	0.01
		(0.268)	(0.793)	(0.992)
Block			7.30***	7.17***
			(0.009)	(0.009)
Risk				✓
Constant	50.83***	-77.06***	-83.30***	-77.72***
	(0.000)	(0.000)	(0.000)	(0.000)
N	2040	2040	2040	2040

p-values in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the most sense to consider average earnings differences between subjects in our Pairs and Individuals treatments. Subjects in the Individual treatment earned \$20.20 on average, while subjects in the Pairs treatment earned an average of \$24.34. Because we are concerned with earnings differences that result from differences in decisions, we subtract from these averages the fixed show-up fee of 10. we see that pairs earned approximately $\frac{\$14.34-\$10.20}{\$10.20} = 40.59\%$ more, on average, than individuals. Without making this adjustment, earnings differences are still quite large: pairs earn approximately 20% more than individuals.

Finally, we quantify differences between pairs and individuals by comparing pairs to synthetic pairs formed using subjects in our Individuals treatment. To do this, we randomly match individuals into stable synthetic pairs and assume these pairs consume the average of what the two individuals consumed. Next, we assess the degree to which we can positively handicap synthetic pairs before the errors of synthetic and real pairs are no longer significantly different at the 10% level.

We find that we can reduce the conditional consumption error of synthetic pairs by approximately 37%, on average, before the performance of real and synthetic pairs becomes indistinguishable. In level terms, this reduces the average conditional consumption error of synthetic pairs from 57.94 to approximately 36.5 experimental credits. We use the same method to quantify how pairs and individuals improve with experience. We can positively handicap synthetic pairs by approximately 36.5% over the first ten periods of play and by approximately 47% in the last ten periods of play. This suggests that pairs and individuals learn differently with experience in the dynamic optimization task, or that pairs are better than individuals at choosing use heuristics that mitigate compounding errors.

5.1 Textual Analysis

Because groups in our experiment engaged in unrestricted chat to form a joint strategy, we are able to use textual analysis to gain deeper insight into how people frame the dynamic optimization problem and how they develop heuristics.

Following Cooper and Kagel (2005), we establish a set of categories we use to classify the language used by subjects in our group treatments when discussing consumption and savings strategies. We describe these categories in Table 3. These categories are neither exhaustive nor meant to be mutually exclusive. Rather, the categories are meant to be complementary and allow for some nuance in classification despite the binary coding system. We trained two research assistants independently and then had them work alone to classify team chats into our pre-selected categories using a binary coding system. For example, if a team discussed how to allocate resources in terms of spending but never in terms of savings, the research assistants would likely code 'Discuss Savings' as a zero and 'Discuss Spending' as a one.

Table 3: Categories for Coding Team Dialogues

Category	**Description**
Discuss Saving	Team frames discussion in terms of saving
Discuss Spending	Team frames discussion in terms of spending
Save More	Someone proposes saving more relative to previous suggestion/period
Save Less	Someone suggests saving less relative to previous suggestion/period
Spend More	Someone proposes spending more relative to previous suggestion/period
Spend Less	Someone proposes spending less relative to previous suggestion/period
Nominal Target	Team discusses a nominal target (i.e. consumption points)
Real Target	Team discusses a real target (i.e. total dollar earnings)
Marginal Target	Team targets a 'marginal increase' target
Savings Target	Team tries to maintain a certain amount of savings
Period Earnings Target	Team discusses a per-period earnings target
Total Earnings Target	Team discusses a lifetime earnings target
Proportion of Income to Spend	Team discusses spending a proportion of per-period income
Proportion of Income to Save	Team discusses saving a proportion of per-period income
Proportion of Wealth to Spend	Team discusses spending a proportion of total wealth
Proportion of Wealth to Save	Team discusses saving a proportion of total wealth
Borrow	Team discusses borrowing against future income
Constant Spending	Team discusses spending a constant amount
Save & Binge	Team discusses saving heavily to spend a large lump sum later

In Table 4 we report a measure that captures, on average, how often team chat aligns with a given category. We construct this measure by first summing over all periods, sessions, and teams for both research assistants and then dividing this sum by two times the total number of periods times the total number of teams. Thus, we report a number bounded between zero and one where a value of one means all teams used language compatible with that category in all periods. Anything less than one means that there are at least some teams who do not use that language in at least some periods. To calculate agreement percentages, we divide the number of times they disagree by the number of opportunities to code a discussion category, subtract this from one, and then convert to percentage terms.

Notice in Table 4 the relatively high frequency of the "Discuss Spending" category, which indicates that groups mostly frame discussions around spending experimental credits rather than saving or borrowing them. Though it is true that subjects must spend credits to earn money, the stochastic income process used here, coupled with the consumption smoothing motive, makes saving and borrowing important components of earnings maximization. Despite this, our subjects rarely discuss saving or borrowing. We also see that subjects, explicitly or implicitly, discuss spending strategies that fix consumption either in levels or as a proportion of wealth or income. This makes sense, given that we find in Section 5.2 that about half of our pairs are best described as using a constant marginal propensity to consume heuristic.

Table 4: Coding Frequencies of Team Dialogue Categories

	Mean	Agreement(%)
Discuss Saving	.070	97.69
Discuss Spending	.873	97.87
Save More	.025	99.91
Save Less	.004	100.00
Spend More	.054	99.35
Spend Less	.046	97.41
Nominal Target	.091	93.70
Real Target	.017	99.72
Marginal Target	.048	100.00
Savings Target	.006	99.35
Period Earnings Target	.014	97.41
Total Earnings Target	.011	98.43
Proportional Spender	.053	99.24
Borrow	.045	99.63
Constant Spending	.038	97.13
Save & Binge	.044	99.91
N	2160	1080

These sorts of simple heuristics greatly reduce the cognitive load of the optimization task but might fail subjects whenever saving or borrowing is necessary for optimization. For example, teams spending a fixed proportion of the per-period endowment would not borrow whenever necessary to spend at the unconditionally- or conditionally-optimal level. This aligns with Carbone and Hey (2001) and Hey and Knoll (2011) who study the process of dynamic decision making and find that subjects are more likely to develop simple decision criteria and adopt strategies aimed at reducing the cognitive complexity of the choice task.

This also relates to Meissner (2016), who demonstrates that individuals perform worse when solving dynamic optimization problems that require borrowing relative to tasks that require saving. Our evidence may suggest that people under borrow because they employ heuristics that are blind to borrowing – maybe because they are less cognitively taxing – and not necessarily because people are innately averse to debt.

We also note that the adoption of these simple time- and wealth-invariant consumption heuristics might explain the upward trend in conditionally-optimal consumption errors that we do not see in the unconditionally-optimal consumption errors. This is because a heuristic that leads to an error in one period will, on average, lead to a similar error in a later period and the invariant nature of the heuristic prevents subjects from avoiding errors in the current period and correcting for errors in past periods.

Additionally, see that subjects are much more likely to frame discussions in nominal rather than real terms. This is not surprising, given that our pairs tend toward simple heuristics that collapse the complexity of this optimization problem. Since nominal and real earnings are isomorphic, it might be the case that subjects prefer nominal framing because it avoids the added complexity of real framing. This aligns with Petersen and Winn (2014), who find that nominal inertia arising in a choice task results from cognitive complexity and that money illusion exerts only second-order effects in the same task.

Finally, we see that our pairs discuss saving and binging as a strategy with surprising frequency. It is easy to assume that such behavior, first documented by Noussair and Matheny (2000), is reactionary because it clearly demonstrates that subjects do not understand the consumption smoothing motive. However, we see here that this behavior can be thoughtful, planned, and forward-looking. This is interesting since the literature so often associates these sorts of consumption smoothing failures with boundedly-rational agents whose failures to optimize stem from myopia. However, we observe here that this behavior, albeit a poor approach to consumption smoothing, may be the result of some innate preference for binge-type behavior.

Next, we consider textual classifications by period, depicted in Figure 5, in order to better understand how communication evolves. To do this, we focus on six thematic categories and eliminate categories that are most infrequent in our data. We eliminate categories that are not discussed by two or more teams at least once in a single period.[7] The proportional strategy indicates that a pair discussed spending or saving a proportion of their per-period endowment or accumulated wealth. The nominal and marginal strategies indicate a pair focusing on a nominal target or a marginal target to guide its consumption decision. The constant spending strategy emerges when pairs discuss spending a fixed level of EC's in each period. Finally, the save and binge strategy involves pairs saving EC's with the intention of 'binging' a large amount of EC's later in a single period. Similar to what we did in Table 4, we first sum the occurrences of a given category for all teams and both research assistants and divide this number by two times the number of teams. We then plot this number for each of our six categories by period.

We note a few things about the evolution of communication. First, the cyclical pattern of these frequencies seems to suggest that not all pairs discuss a strategy in all periods. Instead, it seems like pairs discuss a strategy, follow it for some time, and then reaffirm or discuss the strategy again after a few additional periods of play. Second, we note that the frequency of discussion for all strategies, excluding the nominal target strategy, falls in block 2 relative to block 1, which could indicate that pairs settle into a heuristic with more experience. In particular, we see that teams gradually think less about the marginal benefit of consumption when making decisions. This is interesting insofar as it indicates that pairs are relying less on the information we provide about the marginal benefit of consumption since this information is an important component of optimizing. Finally, we note that the discussion of proportional strategies spikes at the end of block 1, but not block 2. This is

[7] We include in the appendix a similar graph for all categories not shown in Figure 5.

likely because pairs realized at the end of block 1 that they needed to spend all remaining savings. The absence of this same spike at the end of block 2 matches the marked decrease in binging behavior at the end of block 2 relative to block 1. Interestingly, we see that several teams discussed saving and binging in the early periods of block 2, which could explain the large spike in consumption errors that we see around period 30 (period 10 of block 2).

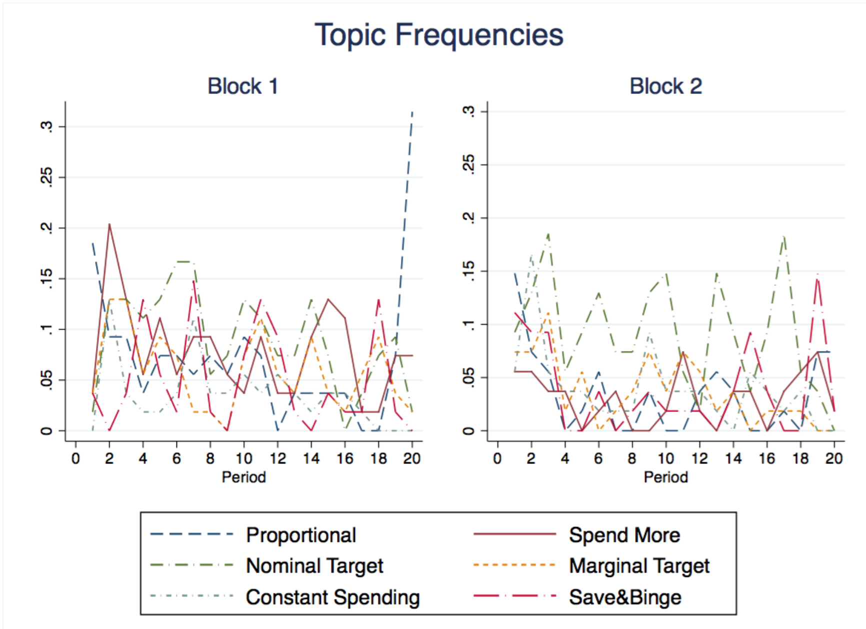

Figure 5: This figure shows the frequency of discussion topics across periods.

5.2 Consumption Heuristics

We now consider the heuristics used by individuals and pairs to make consumption decisions. To do this, we construct a set of 5 possible heuristics that may describe a subject's (or pair's) decisions (see Tasneem and Engle-Warnick (2018), and Fenig and Petersen (2020) for other examples of consumption heuristics.). We describe these five heuristics in Table 5.

H1 assumes that a subject consumes all of her income in each period. This is equivalent to having a fixed marginal propensity to consume (MPC) of 1 in each period. A real-world equivalent is an individual or family that lives paycheck-to-paycheck. H2 assumes that subjects optimize perfectly along the unconditionally-optimal path. This heuristic model captures the behavior of a fully rational agent in the context of our finite lifecycle problem. H3 assumes that subjects optimize along the conditionally-optimal path. H4 assumes that a subject i (or pair i) spends a fixed proportion α_i of income in each period. We consider a range of values for $\alpha_i = \frac{n}{30}$ for $n = \{1, 2, ..., 29\}$. Note that H1 is equivalent to H4 whenever $n = 30$ since this yields $MPC = 1$.

We classify a subject (or pair) to one of H1-H5 by comparing consumption in each period to what a subject would consume according to each possible heuristic and selecting the heuristic that minimizes the sum of absolute differences.

Initial results from this classification exercise, shown in Figure 6, indicate that a smaller

Table 5: Consumption Heuristics

Model Class	Heuristic Name	Model
H1	Hand-to-Mouth	$c_{i,t} = y_t$
H2	Unconditional Optimizer	$c_{i,t} = c_t^*$
H3	Conditional Optimizer	$c_{i,t} = \hat{c}_{i,t}^*$
H4	Constant MPC	$c_{i,t} = \alpha_i y_t$
H5	Constant Spending	$c_{i,t} = c_{i,t-1}$

proportion of teams than individuals are using a hand-to-mouth strategy and instead use a constant spending strategy. This aligns with our textual analysis, where we see many teams setting nominal and marginal spending targets.

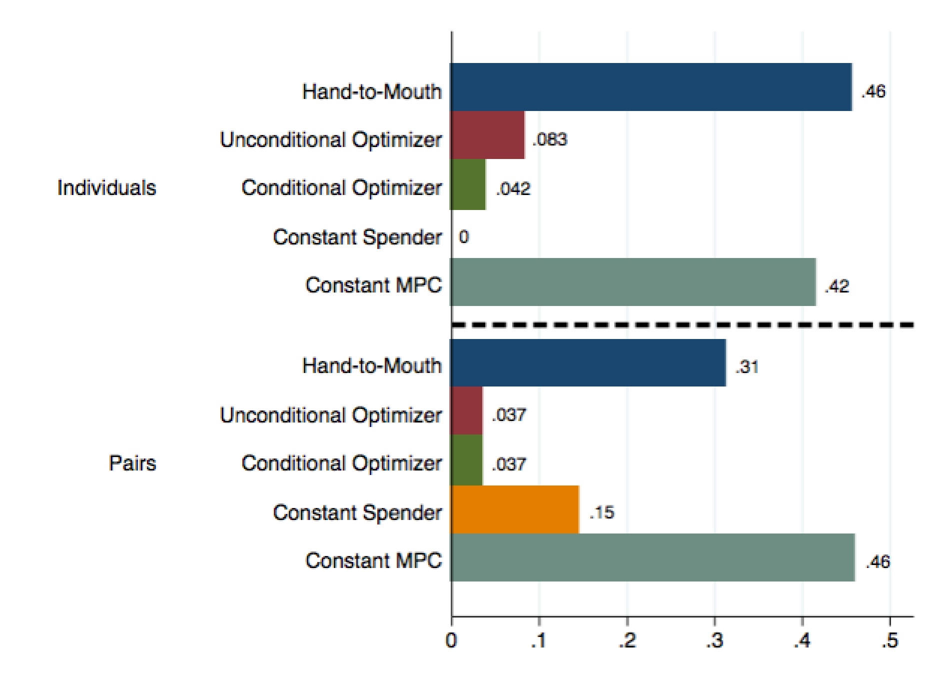

Figure 6: This figure shows the proportion of individuals and pairs using each of the five consumption heuristics.

However, examining the distribution of α reveals systematic differences between pair and individual heuristics. We plot the distribution of MPC values that best describe decisions for pairs and individuals classified to H4 in Figure 7. This shows that, on average, pairs adhere to a higher MPC than do individuals, which is consistent with the idea that pairs engage in less precautionary saving than do individuals. This consistently higher level of spending leads to less buildup of excess savings among groups, which is why pairs engage in fewer and in less severe consumption binges.

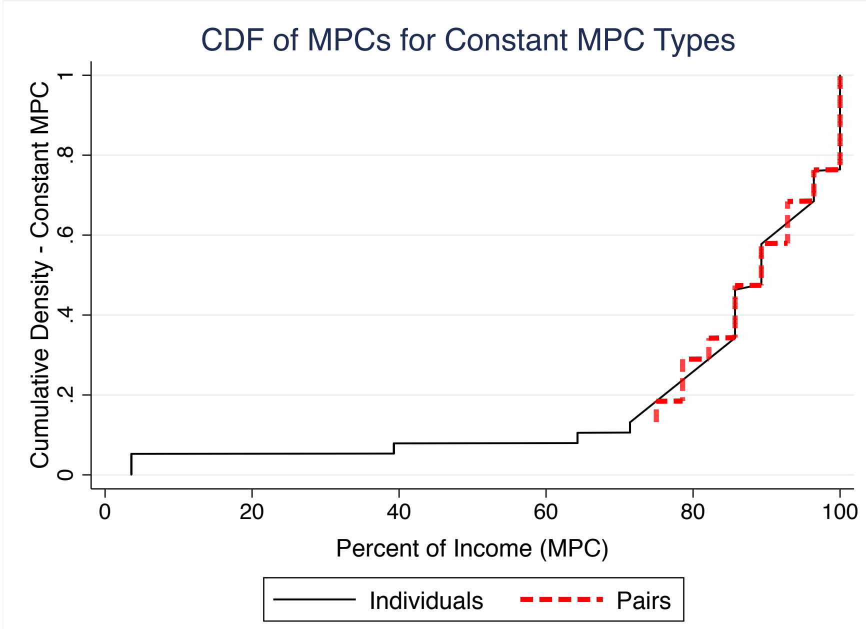

Figure 7: This figure plots the cumulative distribution of α_i for both pairs (red dashed line) and individuals (black solid line). A two-sample Kolmogorov − Smirnov test confirms that the CDF of MPCs for pairs is shifted significantly to the right relative to individuals (P<.001).

6 Conclusion

This paper examines the use of individuals as representative decision-makers in laboratory experiments designed to test dynamic optimization. We do this by comparing the ability of pairs and individuals to solve a finite-period, dynamic optimization problem. We find that allowing for communication and joint decision making leads pairs to significantly outperform individuals along both the unconditionally- and conditionally-optimal consumption paths. This performance gap, on average, leads to subjects in our Pairs treatment earning about 40% more than subjects in our Individuals treatment. We try to rationalize this performance difference by analyzing chat data from our pairs and by examining the consumption heuristics used by subjects in both treatment types.

Textual analysis reveals that pairs almost exclusively frame discussions in terms of spending even though the stochastic per-period income process, coupled with the consumption smoothing motive, makes saving and borrowing important components of earnings maximization. Further, we see that pairs develop simple, invariant heuristics that can lead to persistent errors that compound over time, which explains why our subjects perform worse when measured along their conditionally-optimal consumption paths. Finally, we see that saving and binging is the outcome of thoughtful, forward-looking behavior and not the result of extreme myopia or lack of a strategy entirely.

Our heuristic classification reveals pairs and individuals predominately adhere to a constant MPC heuristic, but that pairs exhibit less heterogeneity in the measure of MPC used and have higher MPCs on average. These higher MPCs help mitigate the compounding nature of conditional errors because they lead to few and less-severe instances of under spending.

References

Ballinger, T. P., Hudson, E., Karkoviata, L., Wilcox, N. T. (2011). Saving behavior and cognitive abilities. Experimental Economics, 14(3), 349-374.

Ballinger, T. P., Palumbo, M. G., Wilcox, N. T. (2003). Precautionary saving and social learning across generations: an experiment. The Economic Journal, 113(490), 920-947.

Bao, T., Duffy, J., Hommes, C. (2013). Learning, forecasting and optimizing: An experimental study. European Economic Review, 61, 186-204.

Blinder, A. S., Morgan, J. (2005). Are two heads better than one? Monetary policy by committee. Journal of Money, Credit and Banking, 789-811.

Brown, Alexander L., Zhikang Eric Chua, and Colin F. Camerer. "Learning and Visceral Temptation in Dynamic Savings Experiments." Quarterly Journal of Economics, 2009, 124(1): 197-231.

Carbone, E. (2006). Understanding intertemporal choices. Applied Economics, 38(8), 889-898.

Carbone, E., Duffy, J. (2014). Lifecycle consumption plans, social learning and external habits: Experimental evidence. Journal of Economic Behavior Organization, 106, 413-427.

Carbone, E., Georgalos, K., Infante, G. (2019). Individual vs. group decision-making: an experiment on dynamic choice under risk and ambiguity. Theory and Decision, 87(1), 87-122.

Carbone, E., Hey, J. D. (2004). The effect of unemployment on consumption: an experimental analysis. The Economic Journal, 114(497), 660-683.

Carbone, E., Infante, G. (2015). Are groups better planners than individuals? An experimental analysis. Journal of Behavioral and Experimental Economics, 57, 112-119.

Charness, G., Sutter, M. (2012). Groups make better self-interested decisions. Journal of Economic Perspectives, 26(3), 157-76.

Cooper, D. J., Kagel, J. H. (2005). Are two heads better than one? Team versus individual play in signaling games. American Economic Review, 95(3), 477-509.

Duffy, J. (2016). Macroeconomics: a survey of laboratory research (pp. 1-90). eScholarship, University of California.

Farhi, Emmanuel, and Iván Werning. 2019. "Monetary Policy, Bounded Rationality, and Incomplete Markets." American Economic Review, 109 (11): 3887-3928.

Fehr, E., Zych, P. K. (2008). Intertemporal choice under habit formation. Handbook of experimental economics results, 1, 923-928.

Fischbacher, U. z-Tree: Zurich toolbox for ready-made economic experiments. Exp Econ 10, 171–178 (2007). https://doi.org/10.1007/s10683-006-9159-4.

Frederick, S. (2005). Cognitive reflection and decision making. Journal of Economic perspectives, 19(4), 25-42.

García-Schmidt, M., Woodford, M. (2019). Are low interest rates deflationary? A paradox of perfect-foresight analysis. American Economic Review, 109(1), 86-120.

Gabaix, X. 2020. "A Behavioral New Keynesian Model," American Economic Review, American Economic Association, vol. 110(8), pages 2271-2327, August.

Hey, J. D. (1988). Experimental investigations into economic behaviour under uncertainty. In Risk, Decision and Rationality (pp. 147-162). Springer, Dordrecht.

Hey, J. D., Dardanoni, V. (1988). Optimal consumption under uncertainty: An experimental investigation. The Economic Journal, 98(390), 105-116.

Hommes, C. H. (Forthcoming). Behavioral experimental macroeconomics and policy analysis: a complex systems approach.

Kaplan, G., Moll, B., Violante, G. L. (2018). Monetary policy according to HANK. American Economic Review, 108(3), 697-743.

Lombardelli, Clare, James Proudman and James Talbot (2005), 'Committees versus Individuals: An Experimental Analysis of Monetary Policy Decision Making', International Journal of Central Banking, 1(1), 181–205.

Lucas Jr, R. E., Stokey, N. L. (1983). Optimal fiscal and monetary policy in an economy without capital. Journal of monetary Economics, 12(1), 55-93.

Mankiw, N. G., Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. The Quarterly Journal of Economics, 117(4), 1295-1328.

Meissner, T. (2016). Intertemporal consumption and debt aversion: an experimental study. Experimental Economics, 19(2), 281-298.

Noussair, C., Matheny, K. (2000). An experimental study of decisions in dynamic optimization problems. Economic Theory, 15(2), 389-419.

Petersen, L., Winn, A. (2014). Does money illusion matter? Comment. American Economic Review, 104(3), 1047-62.

Rholes, R., Petersen, L. (2020). Should central banks communicate uncertainty in their projections?. Available at SSRN 3529832.

Sims, C. A. (2003). Implications of rational inattention. Journal of monetary Economics, 50(3), 665-690.

Stiglitz, J. E. (2018). Where modern macroeconomics went wrong. Oxford Review of Economic Policy, 34(1-2), 70-106.

Vines, D., Wills, S. (2018). The rebuilding macroeconomic theory project: an analytical assessment. Oxford Review of Economic Policy, 34(1-2), 1-42.

7 Appendix

7.1 Tables and Figures

Here we show the decision screens used by subjects in their respective treatments.

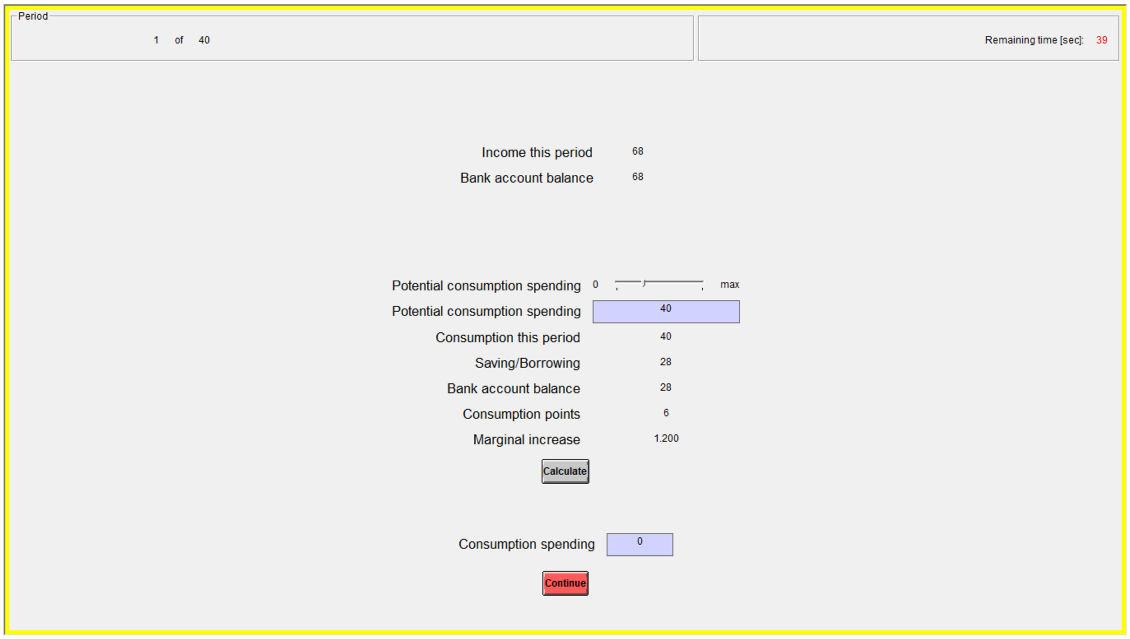

Figure 8: Decision screen for Individuals treatment.

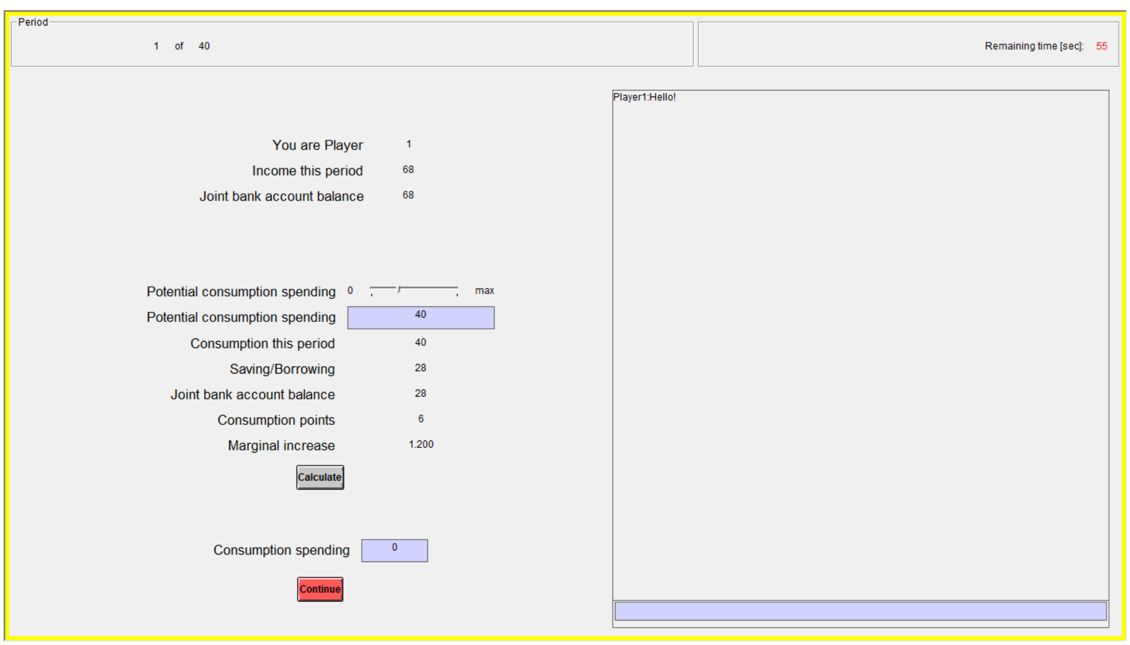

Figure 9: Decision screen for Pairs treatment.

We provide an example, corresponding to Figure 8, that explains how an individual might use the consumption tool and the available information to play this game.

Notice under 'Income this period' that our hypothetical subject has received an endowment of 68 experimental credits (ECs) in period 1. This is reflected in the "Bank account balance," which updates each period to account for per-period and previous saving/borrowing. The subject may then explore the outcome of all possible consumption decisions using the 'Potential consumption spending' slider or by entering hypothetical levels of consumption in the gray box labeled 'Potential consumption spending'.

For this example, our subject could spend between 0 and 128 ECs, since subjects could borrow up to 60ECs in all but the final period of a lifecycle. Moving the slider or entering a value in the box and clicking calculate will update all other variables. In Figure 8, our hypothetical subject has selected a potential consumption value of 40. Notice that all available information has been updated to reflect this. "Consumption this period" is updated to reflect the chosen value of 40.

The "Saving/Borrowing" field updates to 28 to reflect the 28 ECs that would remain in the subject's bank account after spending 40 of the available 68 ECs[8] This balance is also shown in the "Bank account balance" field within the consumption calculator.[9] Further, the "Consumption points" field updates to show the consumption points earned under a choice of spending 40 ECs on consumption, which is 6.

The subject is also shown the marginal utility from using one more EC on consumption in the "Marginal increase" field, which is 1.200. The subject then enters their chosen value for consumption in the 'Consumption spending' box and presses the red button labeled 'Continue' to proceed.

7.2 Empirical Data

We analyze data from the Panel Study of Income Dynamics (PSID) at the family-level, or household level, to evaluate consumption smoothing. The PSID is a longitudinal household survey that has been conducted by the Institute for Social Research at the University of Michigan since 1968.[10] .

Administrators of the PSID survey ask participants to report on their consumption expenditure totals across a large range of items over the course of the previous year. Examples of these items include food, utilities, transportation, education, childcare, and health care. Because respondents self-report data, people may over- or under-report their incomes and consumption expenditures for personal reasons, or due to memory lapse. Additionally, the PSID over-samples low-income families. However, we control for this in our analysis.

We restrict our sample to the years 1999 - 2017. We do this to account for two major changes in PSID data collection that came in 1999. First, surveyors began collecting data biannually instead of annually. Second, the PSID became a richer data source as surveyors

[8] This number would be negative if the subject decided to spend more than 68 ECs.

[9] These numbers match because this is period 1. They would not necessarily match in later periods.

[10] Surveyors collect data on a range of topics including education, employment, income, wealth, and expenditures, which makes it well-suited for the study of consumption smoothing.

began collecting additional information about household consumption and income.

For the purposes of this exercise, we made certain sample selection decisions when cleaning the PSID data. We restrict the sample to household heads aged 20 to 65. We used the OECD-modified adult equivalence scale to adjust for the increase that is proportionate per adult necessary to maintain some standard of living given a change in demographic circumstances, like the birth of a new child. We then adjusted all consumption and income measures by the personal consumption expenditure (PCE) index, to account for changes in prices, and by the OECD-modified equivalence scale. We drop all observations from the original Survey of Economic Opportunity (SEO) sample and the branches of this original sample to avoid the bias that would be introduced from the over sample of poor households, restricting our sample to just the Survey Research Center (SRC) sample. We drop observations where the household head reported working more than 5,200 hours or the household head reported working more than 520 hours at half of the minimum wage. We also drop observations where consumption expenditures are reported to be zero or negative. Thus, we restrict the sample to observations that only report positive consumption expenditures. Finally, we restrict the sample to the lowest income quartile.

7.3 Empirical Motivation

This section provides some suggestive empirical evidence that further motivates our laboratory experiment. Table 6 reports the results of regressing four types of consumption growth (food, non-durable, durable, and total) on a number of household characteristics, conditional on the household head receiving an income shock of spending some months unemployed.

Table 6: Empirical Evidence

Variables	Food	Non-durable	Durable	Total
Single	-0.1781***	-0.0934**	-0.0697	-0.1112**
	(0.0719)	(0.0536)	(0.0613)	(0.0440)
Year FE	✓	✓	✓	✓
Y Group FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	848	857	830	857
R^2	0.0806	0.0179	0.0322	0.0538

We do this using feasible GLS and include both year and income group fixed effects. Additionally, we control for the reported sex of the head of household, the number of children living in the household, the number of adults living in the household, educational level of the head of household, and the reported race of the head of household. Columns two through four in Table 6 show the effect of a negative income shock on consumption across three categories: food consumption, consumption of non-durable goods, and durable goods. Column five reports the effect of a negative income shock on total consumption.

The coefficients reported in Table 6 show the percentage point response of consumption growth, by category, to an unemployment shock. Thus, relative to a married household, a single household experiences a 17.81 percentage point decrease in food consumption, a 9.34 percentage point decrease in non-durable consumption, and an 11.12 percentage point decrease in total consumption. If markets were complete, meaning consumers are able to insure against all possible states of the world, these coefficients should take a value of zero, which would indicate no change in consumption growth in response to an unemployment shock.

7.4 Ledger Treatment

A possible concern is that two things could contribute to differences between individual and pair outcomes. First, subjects in a pair are able to discuss strategies and exchange ideas. Second, subjects in a pair articulate the logic of their decisions while communicating. Thus, one could question if pairs do better because they are are making a joint decision or instead because they are articulating the logic of the consumption/savings decision.

We test this by implementing a third treatment, which we call the Ledger treatment. The Ledger treatment is identical to the individual treatment, except that subjects in the Ledger treatment have access to the same chat window as do subjects in group treatments. Subjects can use this window as a sort of journal to articulate the logic underlying their individual decisions, which could lead to more careful thought and therefore better decisions. In order to be consistent with the Pairs treatment, we neither require subjects in the Ledger treatment to use the ledger nor do we allow them access to ledger entries from previous periods.

Of the 10 subjects who participated in our Ledger treatment, one subject used the ledger one time. A two-sided Student's T-test shows that the mean absolute consumption error (both unconditional and conditional) for subjects in the Ledger and Individual treatments is not significantly different. However, the same test shows that decisions from subjects in the Pairs and Ledger treatments are highly significantly different.