Supplemental materials

A: Experiment1

Sequential investment task

In this task, participants make predictions about the return of a price sequence each period and decide whether to invest in a single stock or a safe asset. The participant's goal is to maximize the expected return. The fundamental return of the stock is randomly set at the beginning of each price sequence. The safe asset return is always set to zero. This fundamental return is fixed and unchanging throughout the price sequence, but the actual return of the stock observed each period is composed of the fundamental return plus white noise caused by exogenous factors such as the market environment. Therefore, participants cannot directly observe the fundamental returns when making investment decisions. However, they can calculate a sufficiently large number of past returns from observable price sequences and average them to get the fundamental return. If the inferred fundamental return is higher than the safe asset return, the rational decision in this setting is to invest fully in the stock, otherwise not.

In this experiment, we have two treatments for the variance of exogenous noise, one large and one small. Each price column was prepared with 10 paths with different fundamental returns. Five of these paths used white noise with large variance, and the other five paths used white noise with small variance (All figures of the 10 paths are available in Supplemental materials).

Let f be the return determined from the fundamentals of the stock and be constant over time. The return observed in each period t is defined by $r_t = f + \varepsilon_t$, where $\varepsilon_t \sim N(0, \sigma_N^2)$ $i.i.d.$. The return on the safe asset is always assumed to be $r^S = 0$. The price sequence presented to the participant's GUI is calculated as $P_{t+1} = (1 + r_t)P_t$. Therefore, when the fundamental return is zero, $f = 0$, the price will randomly walk.

The participant can predict the fundamental return f from the observed returns. If there is no cognitive constraint on information processing, fundamental returns f can be identified from the law of large numbers by using a large enough number of past returns and taking the average of the time series. However, if there is a cognitive constraint, participants may not reasonably formulate a prediction because it takes a certain amount of calculation to obtain each period's returns from prices and to take their average.

To fit the setting of the model to be analyzed, the decision-making of the participants was assumed to be static. That is, participants are given 100 units of points at the beginning of each period, which is cleared at the end of the period. Thus, the participant decides each period whether to invest in a stock or safe asset to maximize the return of the period. However, participants can use the historical price sequence to identify the fundamental return in this experiment. Participants are first presented with a 100 periods price sequence. The reason for presenting 100 periods price in advance is to give a sufficient sample to infer the fundamental return. Participants then invest, the next price sequence is revealed, and liquidation for that period takes place. Price updates are made every second. Participants made this independent choice 180 times per price

sequence.

Experimental setup

The left panel of Fig.1 shows the experimental GUI that participants faced during the decision-making process. To avoid biometric artifacts caused by the button selection behavior, the experiment used a cylinder-type input to allow the investment rate to be varied continuously. However, in the analysis, the investment rate of each participant was converted to binary based on the time-averaged investment rate following the model setup.

Fig. 1. The left figure is the experimental setup the right figure is Brain regions of interest

The biometric information used in this analysis was the change in blood hemoglobin concentration in the prefrontal area. Functional NIRS (BriteMKII supplied by Artinis Medical Systems) was used to measure the blood hemoglobin concentration in the prefrontal area. The right panel of Fig.1 shows the brain regions that we focused on in this study. We will focus on the dorsolateral, ventral, and rostral regions, which are considered to be closely related to costly cognition, working memory, and reasoning. The prefrontal cortex may be roughly divided into the orbitofrontal cortex ((Brodmann Area [BA]) 11, 12, and 13), medial prefrontal cortex ((Brodmann Area [BA]) 24, 25, 32, and mesial portions of 10), and dorsolateral cortex ((Brodmann Area [BA]) 8, 9, and 46). Each region has a distinct cytoarchitecture and function as well as distinct connections. Briefly, the orbitofrontal cortex is involved in decision making, processing awards and punishment; and the medial prefrontal cortex, particularly the anterior cingulate cortex, mediates emotional monitoring and self-regulation. The dorsolateral prefrontal cortex (including Brodmann areas 46, 9) is involved in working memory. Working memory is the ability to hold a limited amount of information in mind for a short period. For example, working memory is necessary for holding a phone number 'in mind,' or keeping track of geographical locations as someone gives you multistep directions to a location across town. This type of memory is critical to bridging temporal gaps so that the information can be 'worked' with or mentally manipulated for a short period. This ability to hold representations in mind is critical to other complex cognitive functions, such as decision-making, planning, and problem-solving. Area 8A can be considered as a key area for the top-down control of attentional selection, it is also a very important region for this experiment, but this time we used NIRS as experimental equipment, so it was difficult to measure

Area 8A ([1], [2], [3], [4], [5], [6], [7]).

Functional NIRS measures the difference in optical absorption characteristics between oxygenated and non-oxygenated hemoglobin by irradiating near-infrared region light from an external source. Therefore, compared to fMRI, it is more difficult to measure deep brain regions, but on the other hand, it has the advantage of being able to measure in a non-constrained and realistic environment. The sampling rate is relatively fast (50 Hz), but the spatial resolution is not so high (20 mm-30 mm). This is non-invasive equipment. These characteristics make it suitable for decision analysis in more realistic environments as in the present financial task.

Participants

This experiment was conducted under the approval of the Ethics Committee of Tokyo University of Science. The participants of this experiment were 10 undergraduate and 30 graduate students (40 samples) of the School of Business, Tokyo University of Science. None of them had any investment experience. We explained the details to all participants in advance of the experiment and obtained their written consent. The number of valid samples was 39. Participants made 180 investment decisions for one price sequence and were rewarded up to 1000 yen (including 500 yen for participation) according to their points. Participants were randomly assigned to a task for each treatment. 19 samples were assigned to the large white noise condition and 20 to the small condition. Participants were briefed on the GUI operation and the investment task setting and then engaged in the task.

Fitting to the observed data

The model parameters were estimated using the maximum likelihood method. The model parameters to be estimated are the weights λ of the selection function. Since $\lambda > 0$ it is estimated by substituting $\lambda = e^{\tau}$ in the calculation. Let I_t denote the index function that is 1 if the stock is invested in the target stock price in periods t and 0 otherwise, the binomial distribution is used for the likelihood function, and the log-likelihood is defined as

$$\ln\left[\prod_t p_t(f, r_{safe}, \lambda)^{I_t}(1 - p_t(f, r_{safe}, \lambda))^{1-I_t}\right] \text{(Formula ID: LLH)}$$

The model parameters are thus determined to maximize the log-likelihood defined from the observed data. The SCG method was used to compute the optimization.

B: Experiment2

Experimental setup

To clarify the causal relationship between the amount of information processed and its cognitive cost,

we conducted an experiment in which we controlled for the amount of information available. Figure 2 shows the PC screens presented to the experimental participants. In this experiment, the participants are presented with a series of prices of stocks to be invested in (center of the screen, target stock price) and a series of prices of stocks that have some correlation with that stock price (two columns on the left of the screen, signal stock prices). In the experiment, the signal stock prices are presented as leading indicators one period earlier, and participants in the experiment can predict the target stock price to invest in from these signal stock price columns. People choose whether to invest in the target stock price or hold it in safe assets based on this prediction.

Three Treatments were conducted in the experiment: Treatment 1 was the control condition, in which only one signal stock price was presented; the number of signal stock prices presented in Treatment 2 and Treatment 3 were increased to 4 and 8, respectively. Table 1 summarizes the tasks assigned to the participants. The more stock prices that serve as signals, the more information is available, and more accurate predictions of target stock prices can be made. But on the other hand, there is more information to process and cognitive costs are expected to be incurred. In the RI model, information costs increase dramatically when processing an amount of information above a certain level due to human cognitive limitations.

In each Treatment, a 30-period target and signal stock price sequence is initially presented, followed by investment selection for 170 periods.

Fig. 2. GUI image

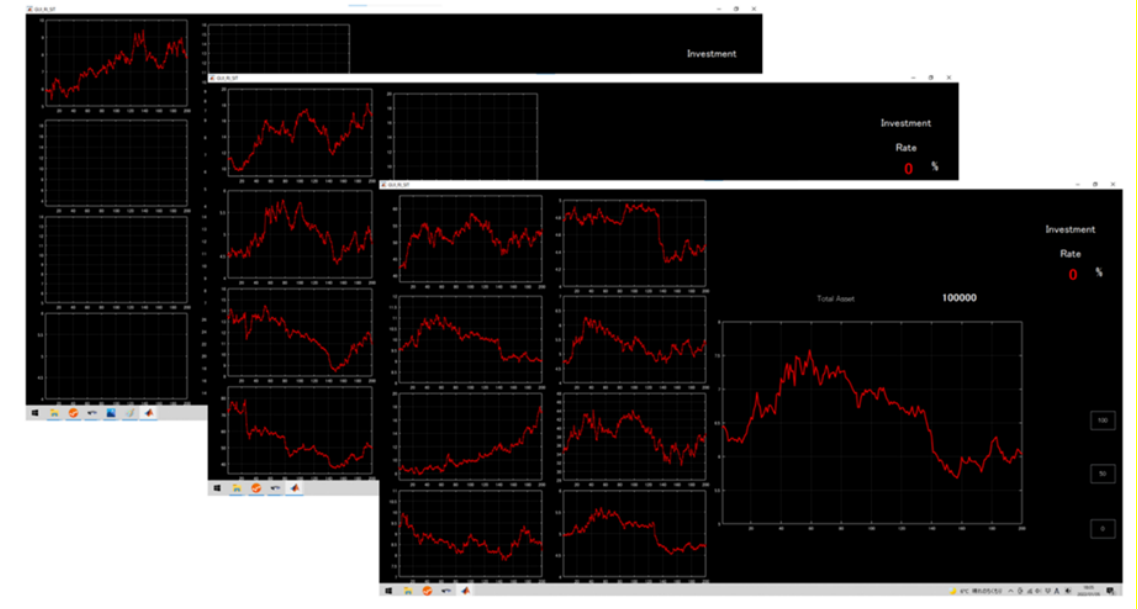


Table 1. Task assign

participant id	price path id	treatment id	number of signals
1	1	1	1
2	2	2	4
3	3	3	8
4	1	2	4
5	2	1	1
6	3	3	8
7	1	3	8
8	2	2	4
9	3	1	1
10	1	2	4
11	2	1	1
12	3	3	8
13	1	2	4
14	2	1	1
15	3	3	8
16	1	3	8
17	2	2	4
18	3	1	1
19	1	3	8
20	2	2	4
21	3	1	1
22	1	2	4
23	2	1	1
24	3	3	8
25	1	3	8
26	2	2	4
27	3	1	1
28	1	1	1
29	2	2	4
30	3	3	8
31	1	1	1
32	2	2	4
33	3	3	8
34	1	3	8
35	2	2	4
36	3	1	1
37	1	1	1
38	2	2	4
39	3	3	8
40	1	2	4
41	2	1	1
42	3	3	8

Model Details

Fundamental returns were not specified a priori in this experiment because real stock prices were used. In the following analysis, the fundamental returns were the expected value of the forecasts made by Kalman filter in each period according to Sims ([10]). This is because if participants are rational, they are expected to make the statistically best forecasts from the available signals. Specifically, we modeled the following.

Let S be the number of signal stock prices. Let the relationship between the target stock price return $r_{m,t}$

and the signal stock price return $r_{s,t}, s \in [1,2,\cdots,S]$ in any t period be as follows.

$$r_{s,t} = \alpha_s + \beta_s r_{m,t} + \varepsilon_s, \quad i.i.d.\, \varepsilon_s \sim N(0, \eta_s^2) \quad \text{(CAPM)}$$

This is an analogy to the CAPM ([8], [9]) when the target stock price is the market portfolio. The update formula for the target stock return was set to

$$r_{m,t} = \rho r_{m,t-1} + \omega, \quad i.i.d.\, \omega \sim N(0, \gamma^2) \quad \text{(Random Walk)}$$

Real stock prices were used in the experiment. The stocks used are all from the Shanghai Stock Exchange and the target stocks used are Henan Taloph Pharmaceutical Stock Co., Ltd.(600222), Air China(601111), and China Construction Bank(601939) for the period December 26, 2020, to October 15, 2021. Before the experiment, the parameters appearing in the above return equations for the stock price series used in the experiment were estimated. The parameters of the equations used in the experiment are summarized in Table 2. The return on safe assets was set to $r_{safe} = 0$ in the experiment.

Table 2. Experimental Parameters

price path id 1	signals							
alfa	6.1E-16	2.35E-16	-5.1E-17	-4.8E-15	6.23E-16	7.93E-16	-1.7E-15	1.53E-16
beta	0.696284	0.886104	0.564184	0.714747	0.640495	0.715025	0.723156	-0.87085
eta^2	0.517791	0.215904	0.68514	0.491607	0.592744	0.491207	0.479454	0.242846
rho	0.987297							
gamma^2	0.023938							
price path id 2	signals							
alfa	2.54E-15	-2E-16	3.74E-16	-9.7E-17	-5.6E-16	-2.5E-17	1.22E-15	-1.7E-15
beta	0.508538	0.853711	0.442223	0.047526	0.671316	0.61826	0.727377	0.767742
eta^2	0.745134	0.272548	0.808501	1.00278	0.552109	0.620875	0.473302	0.412646
rho	0.97533							
gamma^2	0.050565							
price path id 3	signals							
alfa	1.49E-16	-9.6E-17	-4.1E-16	9.83E-16	-2.8E-15	1.12E-15	1.54E-15	5.63E-16
beta	0.368513	0.914274	-0.63677	-0.03112	0.753969	0.57394	0.722625	0.92507
eta^2	0.868563	0.164933	0.597523	1.004077	0.433709	0.67398	0.480226	0.144975
rho	0.987966							
gamma^2	0.029421							

Formation of people's decision rule

A stochastic choice model was employed in the empirical analysis. Since the fundamental returns are not

given a priori in Experiment 2, we replace them with reasonable expectations. That is, we used the expected return μ_t which is updated sequentially using the Kalman filter. Given the rationally expected return, the stochastic choice type RI model derives people's investment rule as the following SoftMax type function.

$$p_t(\mu_t, r_{safe}, \lambda) = \frac{\exp\left\{\frac{\mu_t}{\lambda}\right\}}{\exp\left\{\frac{r_{safe}}{\lambda}\right\} + \exp\left\{\frac{\mu_t}{\lambda}\right\}}$$

Fitting to the observed data

The model parameters were estimated using the maximum likelihood method. The model parameters to be estimated are the weights λ of the selection function. Since $\lambda > 0$ it is estimated by substituting $\lambda = e^{\tau}$ in the calculation. Let I_t denote the index function that is 1 if the stock is invested in the target stock price in periods t and 0 otherwise, the binomial distribution is used for the likelihood function, and the log-likelihood is defined as

$$\ln\left[\prod_t p_t(\mu_t, r_{safe}, \lambda)^{I_t}(1 - p_t(\mu_t, r_{safe}, \lambda))^{1-I_t}\right]$$

The model parameters are thus determined to maximize the log-likelihood defined from the observed data. The SCG method was used to compute the optimization.

Participants

This experiment was conducted under the approval of the Ethics Committee of the Tokyo University of Science. The participants of this experiment were 9 undergraduate and 12 graduate students of the School of Business, Tokyo University of Science. None of them had any investment experience. We explained the details to all participants in advance of the experiment and obtained their written consent. The number of valid samples was 42. The participants were rewarded up to 1000 yen (including 500 yen for participation) according to their points.

C: The Kalman filter type RI model and model selection

In the RI model by Sims ([10]), people attempt to remove the noise in the observed signal (past returns in the present task) by using the Kalman filter, but the accuracy of the filter is constrained by their information processing capacity. If the capacity is small, the information in the signal cannot be utilized well (Kalman gain becomes small), which means that there is momentum in investment.

More specifically, it is modeled as follows. We write the subjective variance of the noise in the observed signal as σ_N^2. The conditional mutual information of the observed return r_t in each period t is $I_{t-1}(f) = \frac{1}{2}\log(1+\sigma_N^{-2}\sigma_{t-1}^2)$, and the constraint $\kappa = \frac{1}{2}\log(1+\sigma_N^{-2}\sigma_{t-1}^2)$ holds when the cognition capacity of people is κ. From the constraint, the observed noise becomes $\sigma_N^2 = \frac{\sigma_{t-1}^2}{\exp(2\kappa)-1}$. In this model, the noise appearing in the observation equation is therefore subjective; if κ is ∞, then the variance of the observation noise is 0, and if κ is 0, then the variance of the observation noise is ∞. Given the subjective observation noise, the update of people's expectations follows;

$$\mu_t = \mu_{t-1} + \frac{1}{1+\sigma_N^2(\kappa)}(r_{t-1} - \mu_{t-1}), \quad (1)$$

$$\sigma_t^2 = \sigma_{t-1}^2 - \left(\frac{\sigma_{t-1}^2}{\sigma_{t-1}^2 + \sigma_N^2(\kappa)}\right)\sigma_{t-1}^2.$$

Selection Rules

People invest one unit in stocks or safe assets based on this expectation formation. In this paper, we examine the logit or softmax choice models, which are widely used in the decision-making field. Specifically, the following six variations of the choice model are considered, including loss aversion.

Model1: $p_t = \exp(\mu_t) / \{1+\exp(\mu_t)\}$.

Model2: $p_t = \exp\left(\frac{\mu_t}{\sigma_t^2}\right) \Big/ \left\{1+\exp\left(\frac{\mu_t}{\sigma_t^2}\right)\right\}$.

Model3: $p_t = \exp\left(\frac{\mu_t}{\delta}\right) \Big/ \left\{1+\exp\left(\frac{\mu_t}{\delta}\right)\right\}$, where δ is an additional parameter.

Model4: $p_t = \exp\left(\frac{\mu_t}{\gamma\sigma_t^2}\right) \Big/ \left\{1+\exp\left(\frac{\mu_t}{\gamma\sigma_t^2}\right)\right\}$, where γ is the absolute risk aversion. This is a choice function obtained from Grossman-Stiglitz type exponential utility maximization.

Model5: The choice function is the same as in Model 1, but the model takes into account the loss aversion in equation (1), which takes different values of κ when the observed return is positive and when it is negative.

Model6: The choice function is the same as in Model 2, but the model takes into account the loss aversion in equation (1), which takes different values of κ when the observed return is positive and when it is negative. Table 3 clearly shows that the stochastic choice type is superior for this experimental data.

Analysis

Fig. 3 shows the average hemoglobin concentration in the blood of each group in dichotomous analysis, and Fig. 4 shows the average hemoglobin concentration in the blood of each group in the trichotomous analysis. The figure on the left shows the hemoglobin concentration in the blood of each sample, divided into two or three groups according to the size of the estimated model parameters (the left bar shows a small value group and the right bar is a large value group). From left to right, they correspond to rostral, dorsolateral, and ventral lateral regions. The right figure is a scatter plot of the parameter for each sample and the blood hemoglobin concentration in the extraperitoneal region.

Fig. 3. Correlation with κ, dichotomous analysis

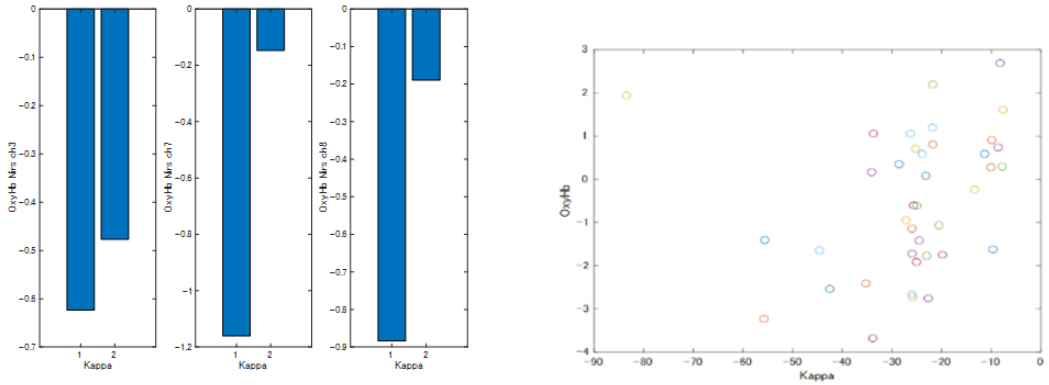
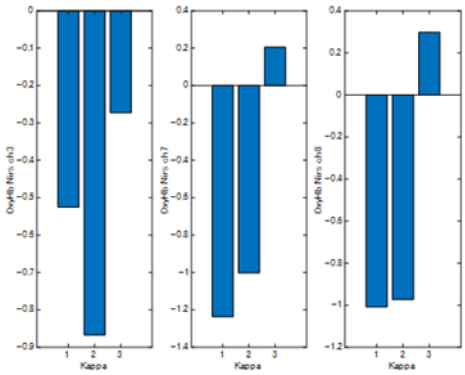


Fig. 4. Trichotomous analysis of κ



From these figures, we can see that brain activity is consistent with the assumptions of the RI models. In other words, the larger the capacity κ, the more activated the brain regions involved in costly cognition. However, the differences between these groups are not significant for κ in all brain regions. As can be seen from the scatter plots in each figure, the correlations are not sufficiently clear due to the large individual differences in biological data. This may be partly because the Kalman filter type RI model does not fit well in the present sample, which will be revealed in the next section.

Model selection

Finally, we select a more appropriate model based on the behavioral data, using the maximum likelihood method for model fitting. Table 3 summarizes the log-likelihood (llh), Akaike's information criterion (AIC), and R^2 (R2) for the Kalman filter type and stochastic choice type RI models and their variations. R2 is the improvement in the log-likelihood from a random prediction, defined as follows, $R^2 = \frac{\log M - \log M_{rand}}{-\log M_{rand}}$, where M is the likelihood of the model concerned, and M_{rand} is that of the random prediction model.

The following six variants of the Kalman filter type RI models and the stochastic choice type RI model are examined. Since the fundamental returns were fixed in this experiment, the prediction converged at a relatively early stage, and therefore the characteristics of the Kalman filter type model may not have been fully utilized. As for the change in blood hemoglobin concentration, the correlation between the parameters

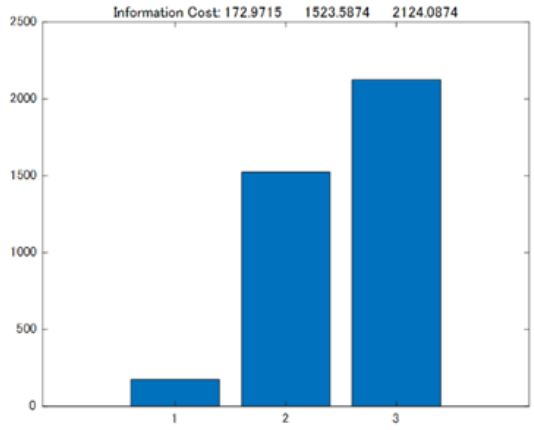
of the stochastic choice type model was clearer than that of the Kalman filter type model, and the same tendency was observed for the behavioral data.

Table 3. Model selection

	Stochastic Choice Type	Kalman Filter Type					
		model1	model2	model3	model4	model5	model6
llh(mean)	-106.8063496	-123.925	-114.441	-118.459	-122.949	-123.221	-112.993
llh(std)	28.19307684	0.619818	20.88867	13.03432	1.804096	1.17822	23.27223
AIC	215.6126993	249.8496	230.8825	240.9178	249.8971	250.4413	229.9866
R2	0.139167648	0.001197	0.077632	0.009066	0.009066	0.006873	0.089302

D: Relationship between changes in investment rate and changes in oxidized blood hemoglobin concentration

Fig. 5. Information cost of each treatment(The horizontal axis is Treatment1,2,3 from left to right, and the vertical axis is the average of the information costs measured by the RI model over the entire sample.)



Scatter plot of the magnitude of the variation in the investment rate (variance) and the change in the oxidized blood hemoglobin concentration (period mean) of all subject data.

Fig. 6. The horizontal axis is the variance of the investment rate by session, and the vertical axis is the change in the oxidized blood hemoglobin concentration of the rostral region.

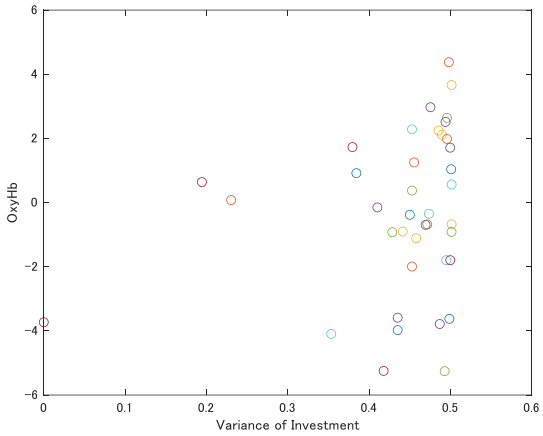


Fig. 7. The horizontal axis is the variance of the investment rate by session, and the vertical axis is the change in the oxidized blood hemoglobin concentration of the dorsolateral region.

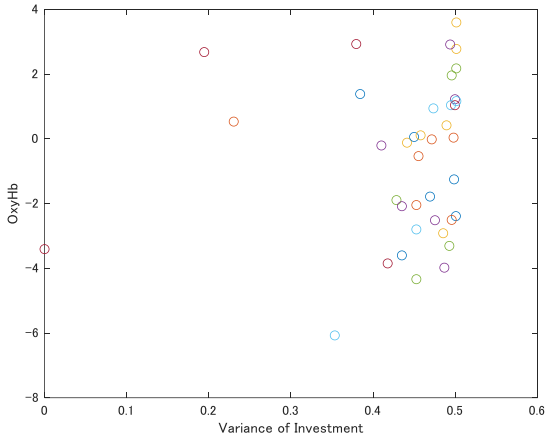
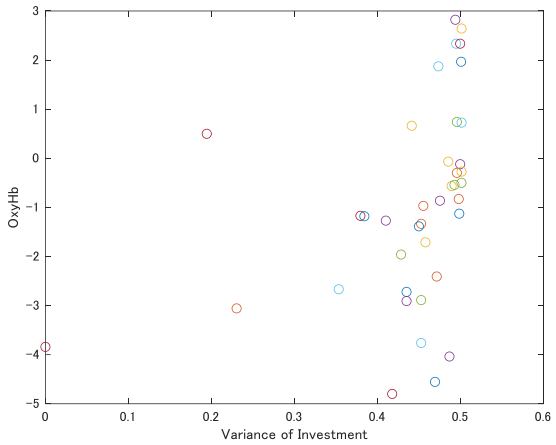


Fig. 8. The horizontal axis is the variance of the investment rate by session, and the vertical axis is the change in the oxidized blood hemoglobin concentration of the ventral lateral region.



References

1. Filip Matejka: Rigid pricing and rationally inattentive consumer. Journal of Economic Theory 158,656-678(2015).doi: 10.1016/j.jet.2015.01.021
2. D'Esposito, M., Detre, J. A., Alsop, D. C., et al.： The neural basis of the central executive system of working memory. Nature 378(6554),279-281(1995).doi: 10.1038/378279a0
3. Smith, Edward E., and John Jonides: Neuroimaging analyses of human working memory. Proceedings of the National Academy of Sciences 95(20),12061-12068(1998).doi: 10.1073/pnas.95.20.12061
4. Smith, Edward E and Jonides, John: Working memory: A view from neuroimaging. Cognitive psychology 33(1),5-42(1997).doi: 10.1006/cogp.1997.0658
5. Gupta, Rashmi, and Tranel, Daniel: Memory, neural substrates. Encyclopedia of human behavior, 593—600(2012) DOI: 10.1016/B978-0-12-375000-6.00230-5
6. Germann, J. and Petrides, M: Area 8A within the posterior middle frontal gyrus underlies cognitive selection between competing visual targets. Neuro 7(5),(2020) doi: 10.1523/ENEURO.0102-20.2020
7. Hamada Hamid: Networks in Mood and Anxiety Disorders. Neuronal Networks in Brain Function, CNS Disorders, and Therapeutics 327-334(2014) doi: 10.1016/B978-0-12-415804-7.00024-1
8. Sharpe, William F. (1964), "Capital asset prices: A theory of market equilibrium under conditions of risk", The Journal of Finance 19 (3): 425-442,
9. Sharpe, William F. "Capital asset prices with and without negative holdings." The Journal of Finance 46.2 (1991): 489-509.
10. Christopher A. Sims: Implications of rational inattention. Journal of Monetary Economics 50 (3), 665–690(2003).doi: 10.1016/S0304-3932(03)00029-1