Supplementary Analyses for "Investigation of the Convex Time Budget Experiment by Parameter Recovery Simulation"

Yuta Shimodaira*, Kohei Shiozawa†, and Keigo Inukai‡

August, 2022

Contents

*Graduate School of Economics, Osaka University.
†Faculty of Economics, Takasaki City University of Economics.
‡Faculty of Economics, Meiji Gakuin University. E-mail: inukai@eco.meijigakuin.ac.jp

SA.1 Box plots of the $\ln \sigma$ estimates

Fig. S1 shows box plots of the distribution of the estimated curvature parameter $\ln \sigma$, which was omitted from the main text. Each box focuses on a specific ground-truth $\ln \sigma$ and a certain noise size s, and then each plot summarizes 10 replications of all combinations of ground-truth values of δ and β, i.e., 1,000 simulation agents. The two ends of the box represent the first and third quartiles, respectively, and the two ends of the whiskers represent the 5th and 95th percentiles, respectively. For the red line, the error of the estimate is 0.

To ensure that the $\ln \sigma$ estimates converge, we transformed $\ln \sigma$ using the function $\ln \sigma = 4 \tanh(\theta) + 1.5$ and searched the latent variable θ. Through this manipulation, the estimated $\ln \sigma$ ranges between -2.5 and 5.5. The black horizontal dashed lines in the figure indicate the boundaries of $\ln \sigma$.

Contrary to the parameters δ and β, $\ln \sigma$ tends to have lower estimates, especially when the added noise is considerable. In the case of $s = 0.20$, $\ln \sigma$ appears to be saturated at about 2.2.

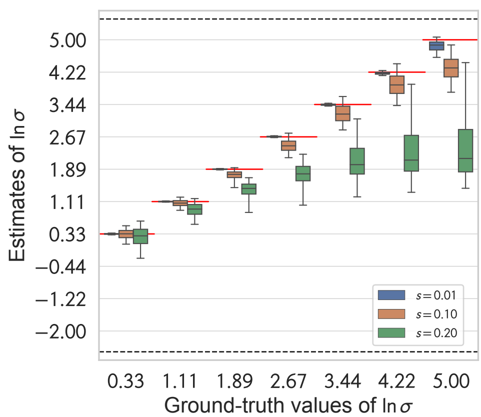

Figure S1: Box plots of the $\ln\sigma$ estimates.

SA.2 Scatter plots of estimated δ–β

Fig. S2 shows scatter plots of the estimated δ and β for each pair of specific ground-truth δ and β. Each δ–β plot includes estimates for all combinations of ground-truth $\ln \sigma$ and s. The darkness of the dots indicates the magnitude of the added noise; the darker the colour, the smaller the noise. The horizontal and vertical axes represent the δ and β estimates, respectively, and the grid shows the ground-truth values. Note that while the ground-truth values are linearly equally spaced, both axes use a logarithmic scale, so the grid lines are not exactly equally spaced. Fig. S2 is a matrix, with columns representing ground-truth δ and rows representing ground-truth β. A red cross indicates each pair of ground-truth values.

We have drawn a red curve (a straight line in log-log graphs) that satisfies $\beta\delta^{70} = \text{const.}$ and passes through the red crosses. The points in each scatter plot are distributed either near the red line or in a more vertical belt than the red line.

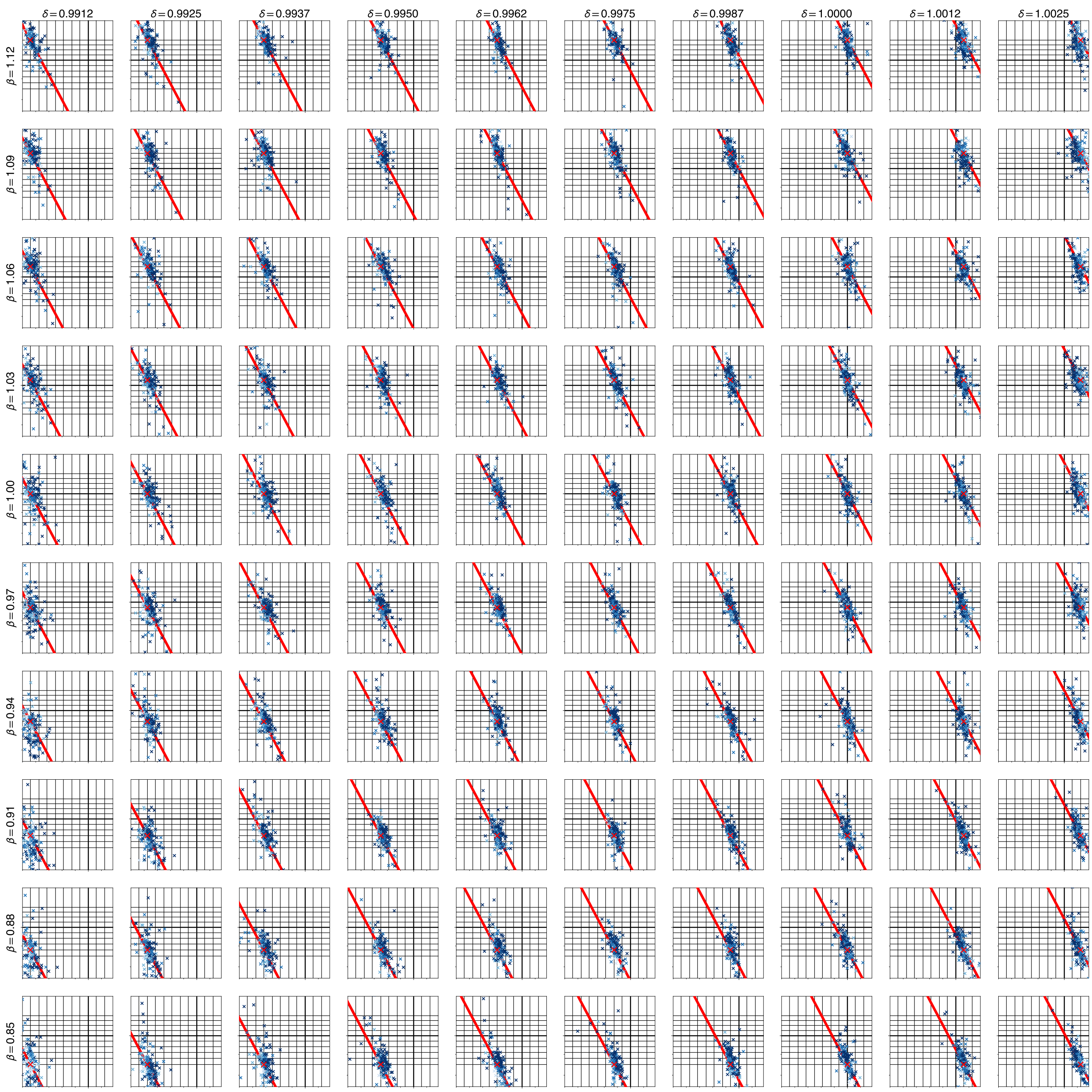

Figure S2: Scatter plots of estimated δ–β.

SA.3 Detect the present bias by using a paired two-sample t-test

We attempted to detect present bias by directly comparing the values of the decision data rather than by parameter β estimation. With two decisions required for the sooner periods, $t = 0$ and $t > 0$ for 21 different prices, we conducted a paired t-test for the two series. Fig. S3 shows the proportion of cases in which the two-tailed test rejected the null hypothesis that there was no difference in decisions at the 5% level. Compared with the success rates using the β estimates in the main analysis, it is more difficult to detect present bias when directly comparing decisions. Individuals whose utility function is almost linear will make drastic decisions, allocating everything to early periods on the one hand and late periods on the other, bounded by a specific price (the switching point). The difficulty of detection occurs because the effect we want to detect is drowned out by noise if we include data in the range where the decision does not change from that in the sooner period t.

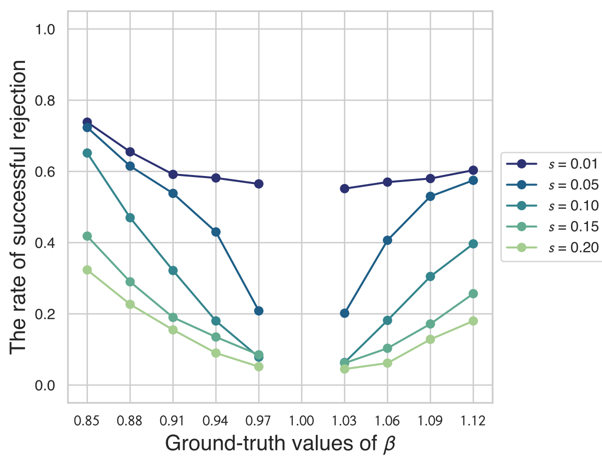

Figure S3: Rate of successful rejection of the null hypothesis that there was no difference in decisions for $t = 0$ and $t > 0$.

SA.4 Re-running the simulation with the expanded ground-truth β spacing

In the main analysis, we concluded that the poor accuracy of the β estimation reflected the required resolution, i.e., the spacing of the ground-truth values, being too narrow to identify in the first place. Here, we show the results of re-running the simulation with the spacing of the ground-truth β expanded by about three times. We used 10 equally spaced values from the range $0.50 \leq \beta \leq 1.40$ for the ground-truth β and the same ground-truth values for δ and $\ln \sigma$ as in the main analysis.

Fig. S4 shows box plots representing the distribution of the estimated δ, β, and $\ln \sigma$. The distributions of the δ and $\ln \sigma$ estimates do not change significantly from those in the main analysis, while the distribution of the β estimates shows that, based on the length of the box, the estimates are accurate enough to identify neighbours even when the noise size is $s = 0.20$.

Fig. S5 shows the rate of successful rejection of the null hypothesis that $\widehat{\delta} = 1$ and $\widehat{\beta} = 1$, respectively. From the graph of the success rates for β, we can see that if the true β is far enough from 1, we can correctly reject the absence of bias.

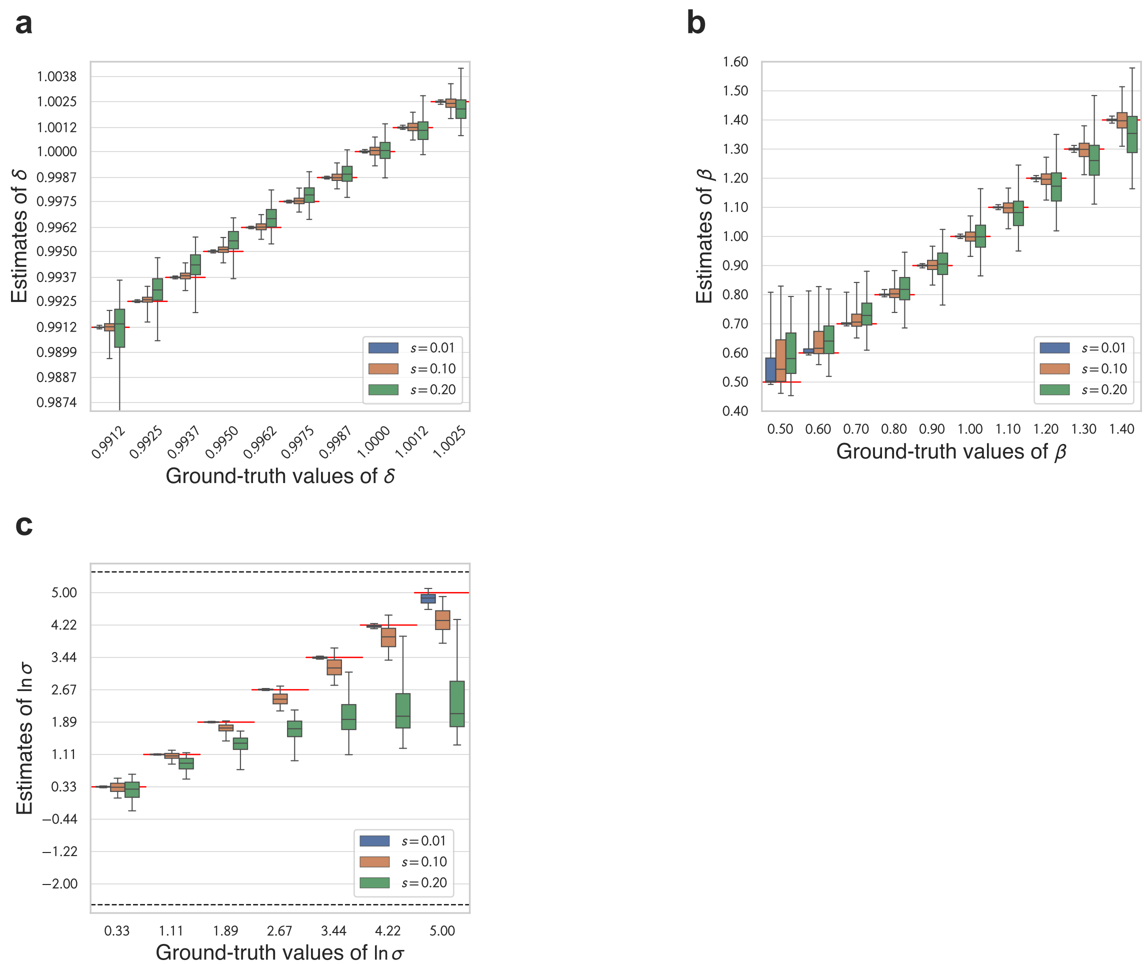

Figure S4: Box plots of **a**) δ, **b**) β, and **c**) $\ln\sigma$ estimates for an expanded ground-truth β range.

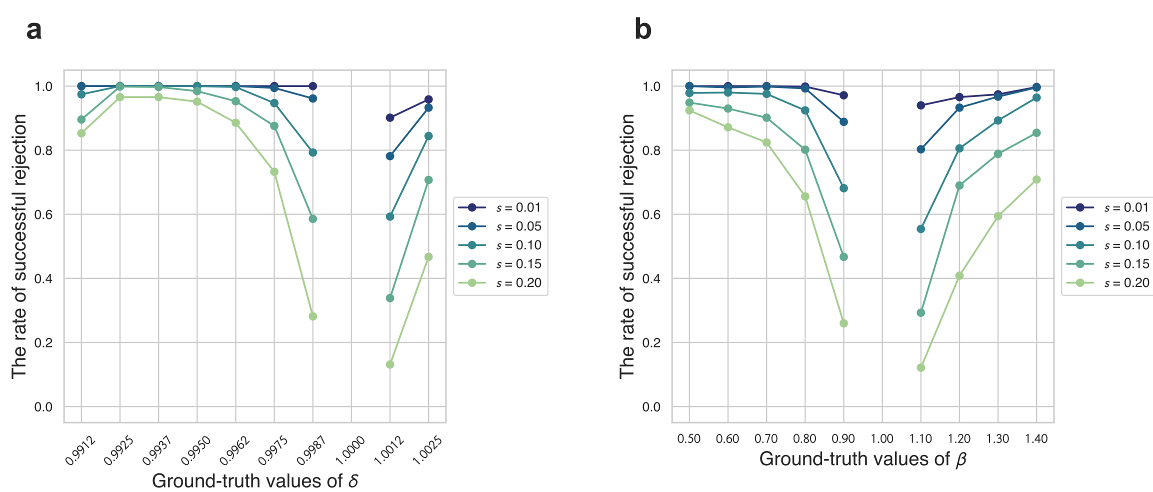

Figure S5: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for an expanded ground-truth β range.

SA.5 Re-running the simulation with $k = 1$

The value of k, which represents the length of the intertemporal period, depends on the scale of δ, and an enormous value of k has no intrinsic effect on the accuracy of parameter estimation. Let us assume that $k = 1$ "day". Figs. S6 and S7 show the results of the recovery simulation, where the data are generated with $k = 1$ for the same ground-truth values as in the main analysis. The resolution is very low for δ estimation. This suggests that, in an experimental setup where there is only 1 day between periods, it is challenging to estimate δ in the ground-truth range we have set. Compared with the figures for β shown in the main analysis, we can see that the estimation accuracy of β is almost the same for $k = 70$ as for $k = 1$.

Next, let us assume that $k = 1$ "period", where one period represents 70 days and δ is not a daily discount factor, but rather, a one "period" discount factor. In other words, set $k = 1$ and generate data using the ground-truth value of δ as the value in the main analysis to the power of 70. Figs. S8 and S9 show the results. Compared with the figures shown in the main analysis, we can see that in addition to the no change in resolution for β, there is little change in δ's resolution over the converted scope range.

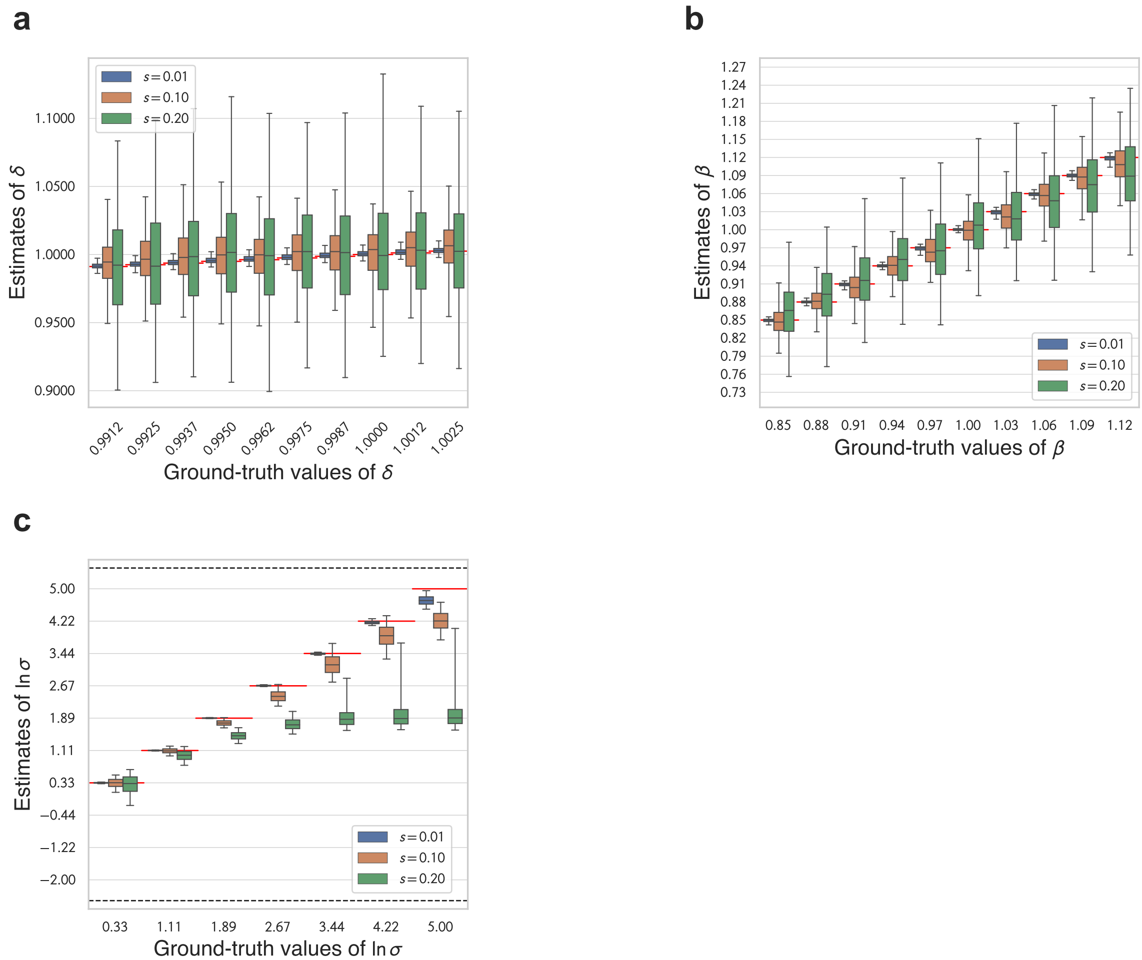

Figure S6: Box plots of **a**) δ, **b**) β, and **c**) $\ln\sigma$ estimates for $k = 1$.

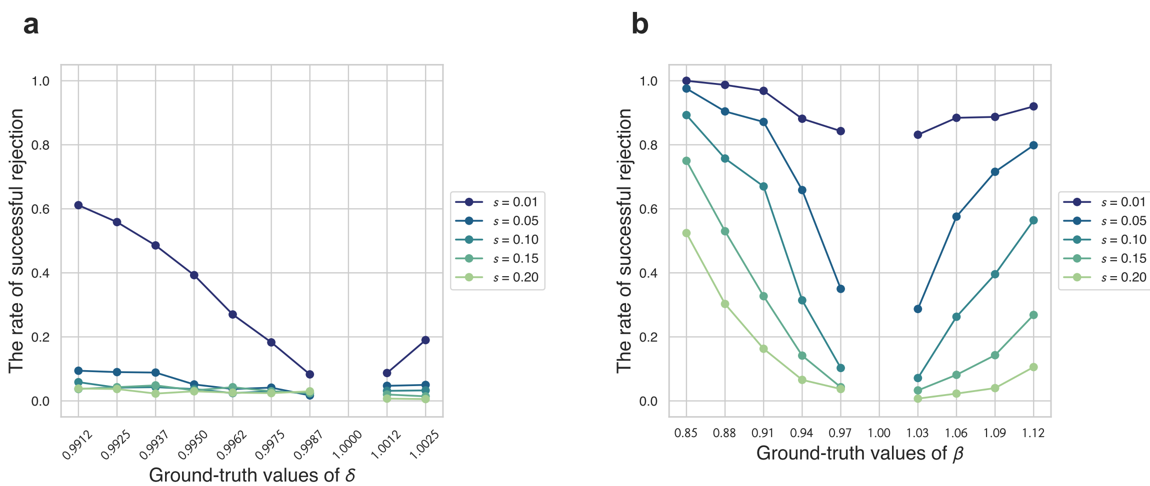

Figure S7: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for $k = 1$.

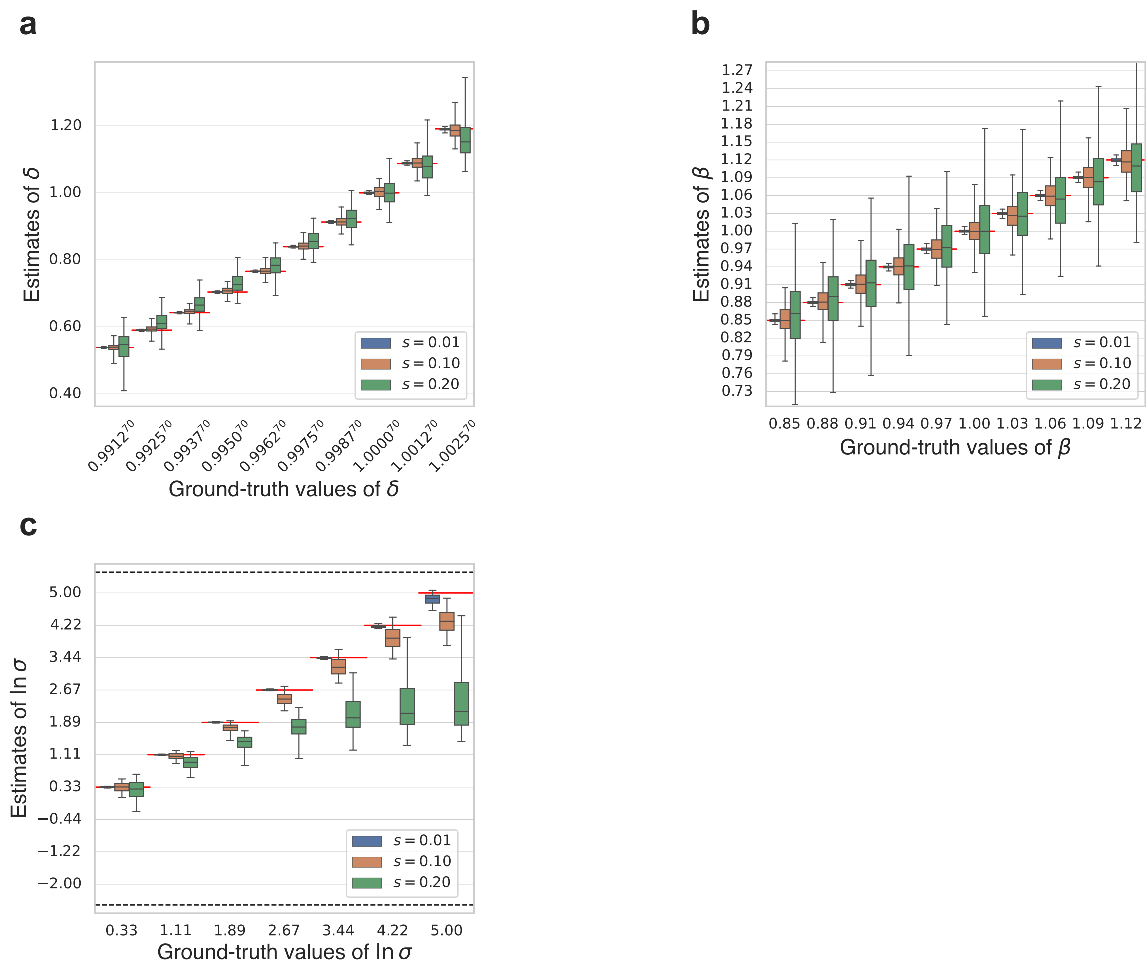

Figure S8: Box plots of **a**) δ, **b**) β, and **c**) $\ln \sigma$ estimates for $k = 1$ with an expanded ground-truth δ range.

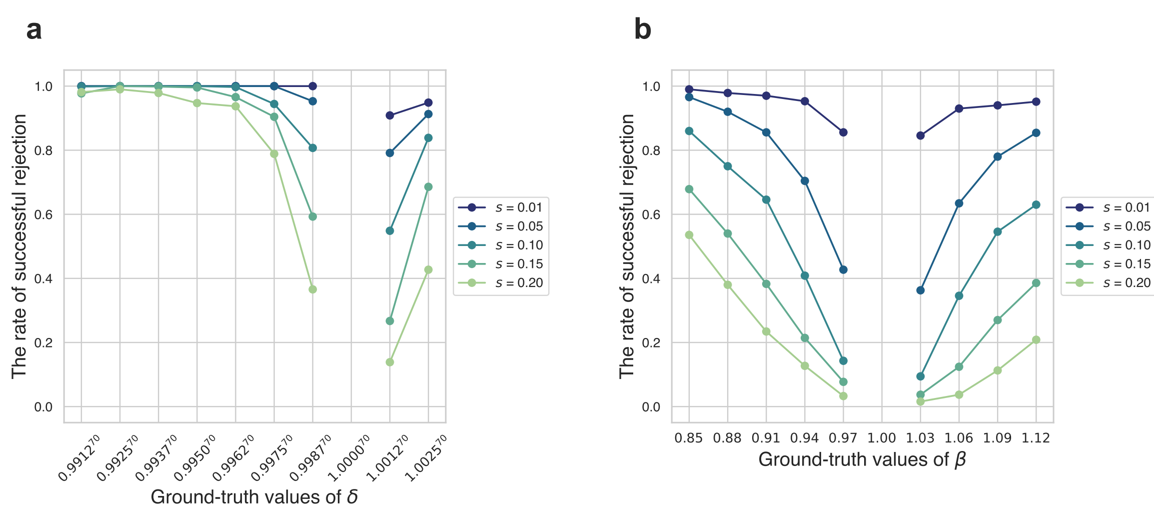

Figure S9: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for $k = 1$ with an expanded ground-truth δ range.

SA.6 Design of problem set

This section analyses how estimation accuracy varies depending on how the problem set is designed. The problem set used in the main analysis is referred to as PS0 in the following. For PS0, m is fixed to 20 and k is fixed to 70, and 21 uniformly spaced prices were drawn from the range $0.6 \leq 1 + r \leq 2$ for $t = 0$ and $t = 1$, respectively, for a total of 42 problems in all.

First, we examine the effect of simply increasing the number of problems. Let PS1 be the problem set that withdraws 42 prices, or twice the number of PS0, from the same range of prices, and let PS2 be the problem set that withdraws 210 prices, or 10 times the number of PS0.

Next, we consider the strategy of changing the number of each variable types without increasing the number of problems. For problem set PS3, instead of increasing the number of k to three—35, 70, and 98—we set the number of prices $1 + r$ to seven from the range $0.6 \leq 1 + r \leq 2$ for each combination of t and k. For problem set PS4, instead of increasing the number of k to seven—14, 28, 42, 56, 70, 84, and 98—we set the number of prices $1 + r$ to three—0.9, 1.2, and 1.5—for each combination of t and k. For problem set PS5, we fixed k to 70 and use 14 prices drawn from the range $0.6 \leq 1 + r \leq 2$, once at $t = 0$ and twice at $t = 1$. This corresponds to setting two different t situations for $t > 0$ (e.g., two decision makings for $k = 70$: one between 7 and 77 days later and the other between 35 and 105 days later). For problem set PS6, we ignored the existence of individuals who place a premium on future payoffs, and did not consider negative interest rates; i.e., we fixed k to 70 and drew 21 prices from the range $1.05 \leq 1 + r \leq 2$ for $t = 0$ and $t = 1$, respectively.

We conducted simulations using the problem sets PS1–PS6 as well as the original problem set used by Andreoni and Sprenger (2012) (henceforth AS). The problem set of AS is summarized in Table S1. For PS1 and PS2, we computed the standard error of the estimates using the inverse of the negative Hessian, whereas for PS3–PS6 and AS, we computed them using the jackknife method, as in the main analysis.

Figs. S10–S23 show the box plots of the estimates and the rate of successful rejection of the null hypothesis for the problem sets PS1–PS6 and AS. We also included the PS0 data as a dashed line in the success rate graph. Figs. S24 and S25 plot the absolute value of the error and the standard error of the estimate by problem set, respectively. We plotted the mean and median with bootstrap 95% confidence intervals. For the plot of the mean value, we used the logarithm scale because some values are enormous.

Comparing PS0 with PS1 and PS2, it is clear that estimation accuracy is improved for PS1 and PS2. Increasing the number of problems improves the accuracy. However, even with PS2, which has 420 tasks in total, it is difficult to reject that the β estimate of an individual, whose actual value is 0.97, is not equal to 1 when the noise size is $s = 0.20$.

For PS3–PS5, it is difficult to conclude that any particular problem set is clearly superior to PS0. As PS6 and AS do not include prices with negative interest rates, the accuracy of δ estimation is inevitably very poor for individuals with $\delta > 1$, but improves in some cases for individuals with $\delta < 1$.

Table S1: Problem set of AS

t	k	$(1+r, m)$				
0	35	(1.05, 20)	(1.11, 20)	(1.25, 20)	(1.25, 25)	(1.43, 20)
0	70	(1.05, 20)	(1.11, 20)	(1.25, 20)	(1.25, 25)	(1.43, 20)
0	98	(1.05, 20)	(1.25, 20)	(1.25, 25)	(1.54, 20)	(2.00, 20)
7	35	(1.05, 20)	(1.11, 20)	(1.25, 20)	(1.25, 25)	(1.43, 20)
7	35	(1.05, 20)	(1.11, 20)	(1.25, 20)	(1.25, 25)	(1.43, 20)
7	70	(1.00, 20)	(1.05, 20)	(1.11, 20)	(1.25, 20)	(1.43, 20)
35	70	(1.05, 20)	(1.11, 20)	(1.25, 20)	(1.25, 25)	(1.43, 20)
35	98	(1.05, 20)	(1.25, 20)	(1.25, 25)	(1.54, 20)	(2.00, 20)
35	98	(1.05, 20)	(1.25, 20)	(1.25, 25)	(1.54, 20)	(2.00, 20)

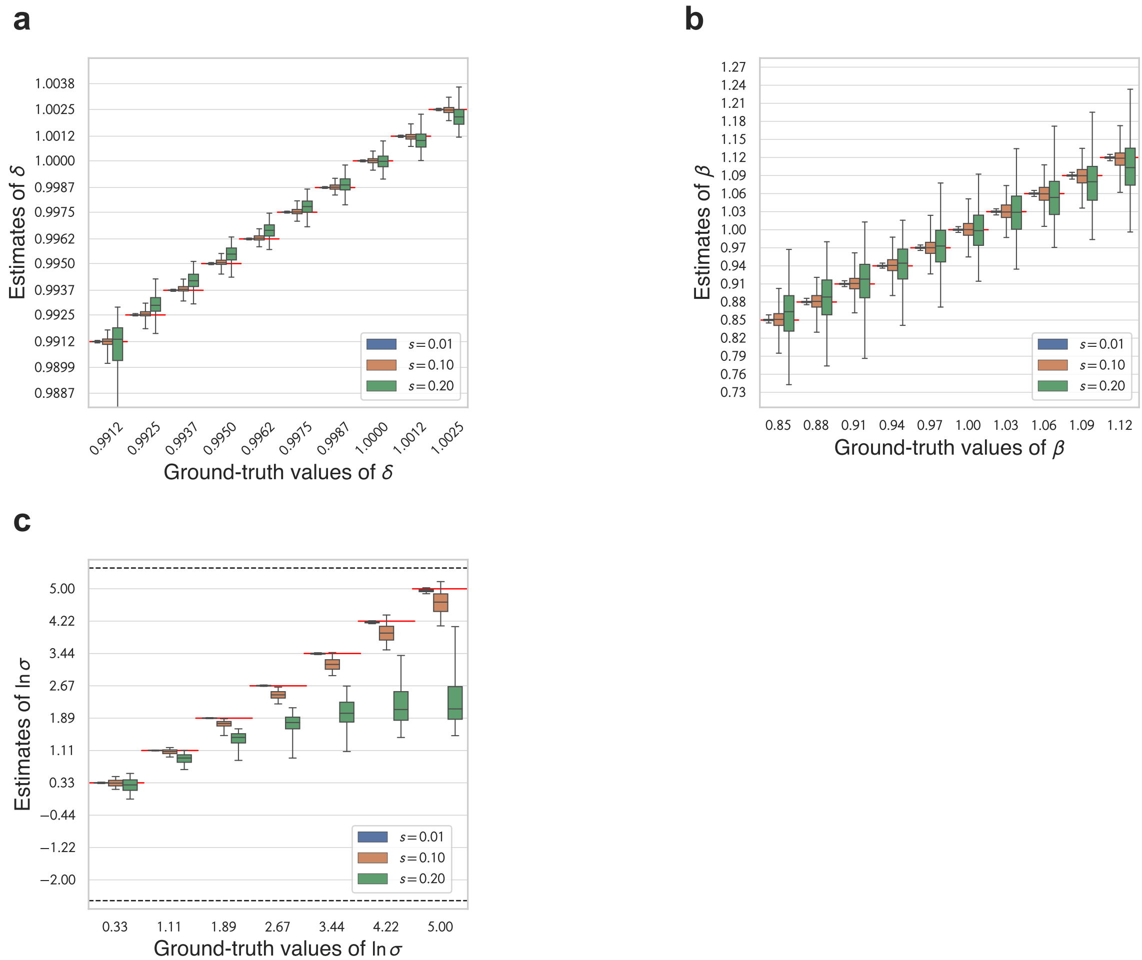

Figure S10: Box plots of **a**) δ, **b**) β, and **c**) $\ln \sigma$ estimates for PS1.

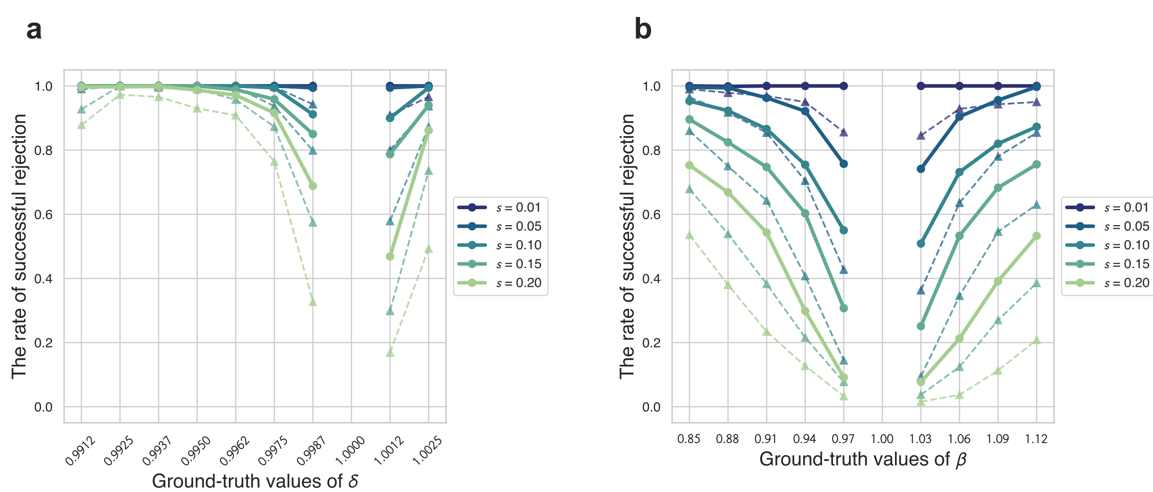

Figure S11: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for PS1.

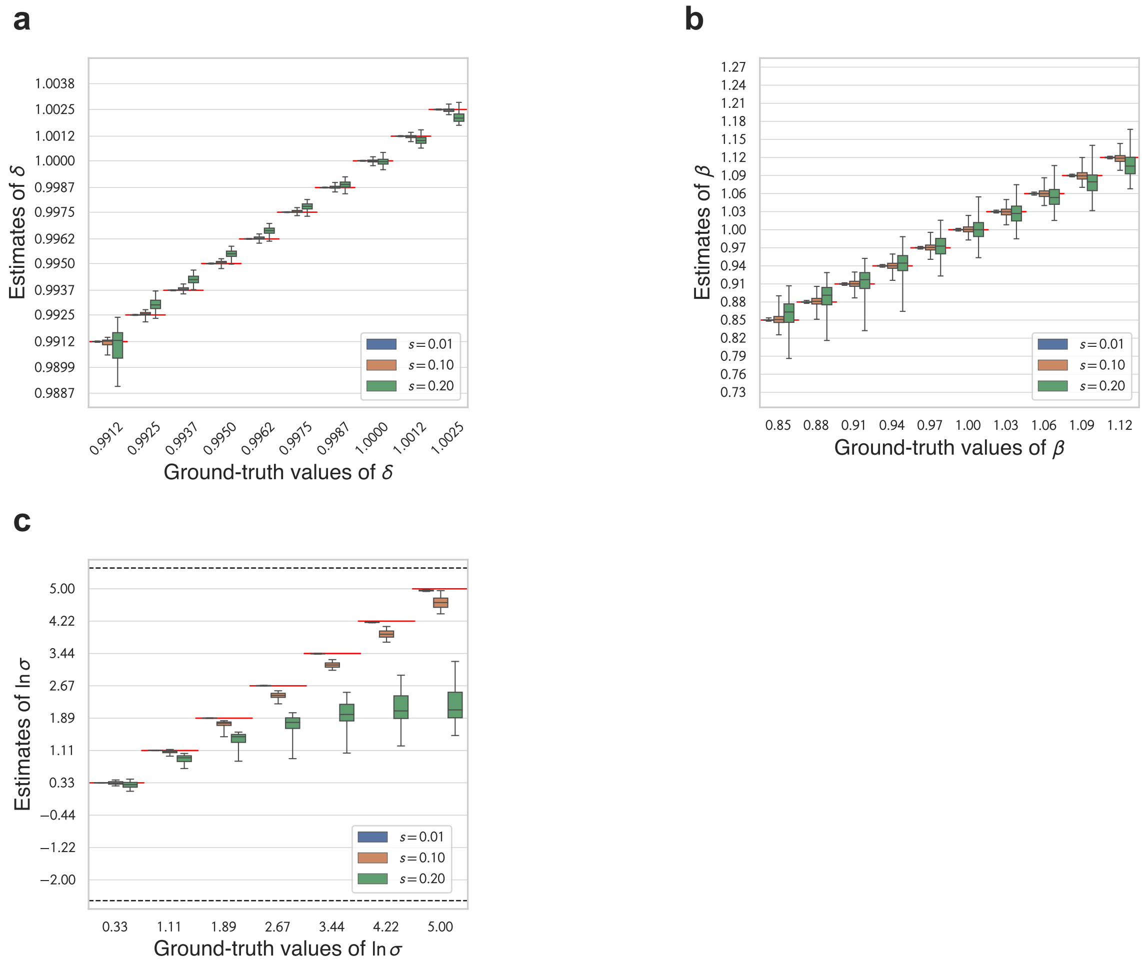

Figure S12: Box plots of **a**) δ, **b**) β, and **c**) $\ln \sigma$ estimates for PS2.

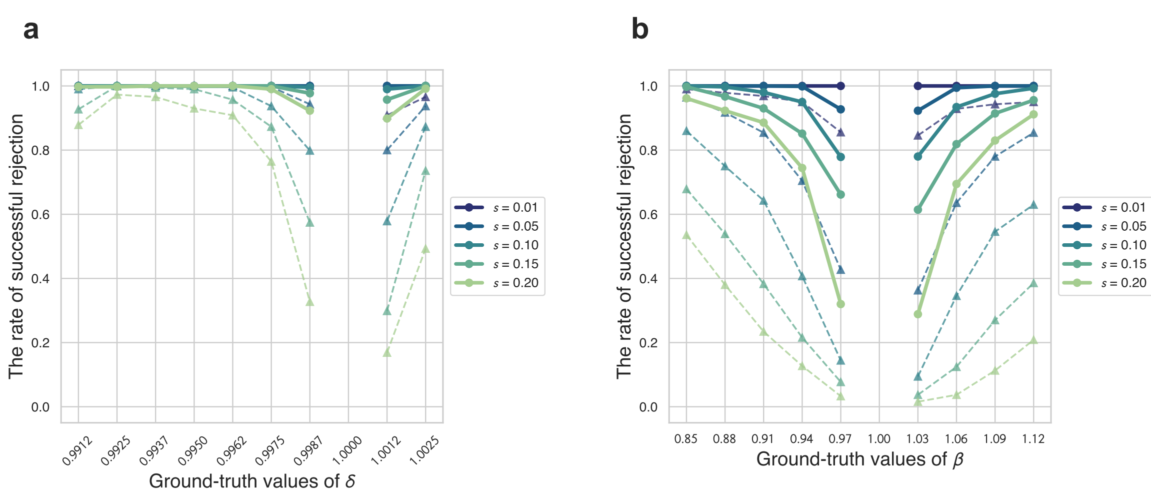

Figure S13: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for PS2.

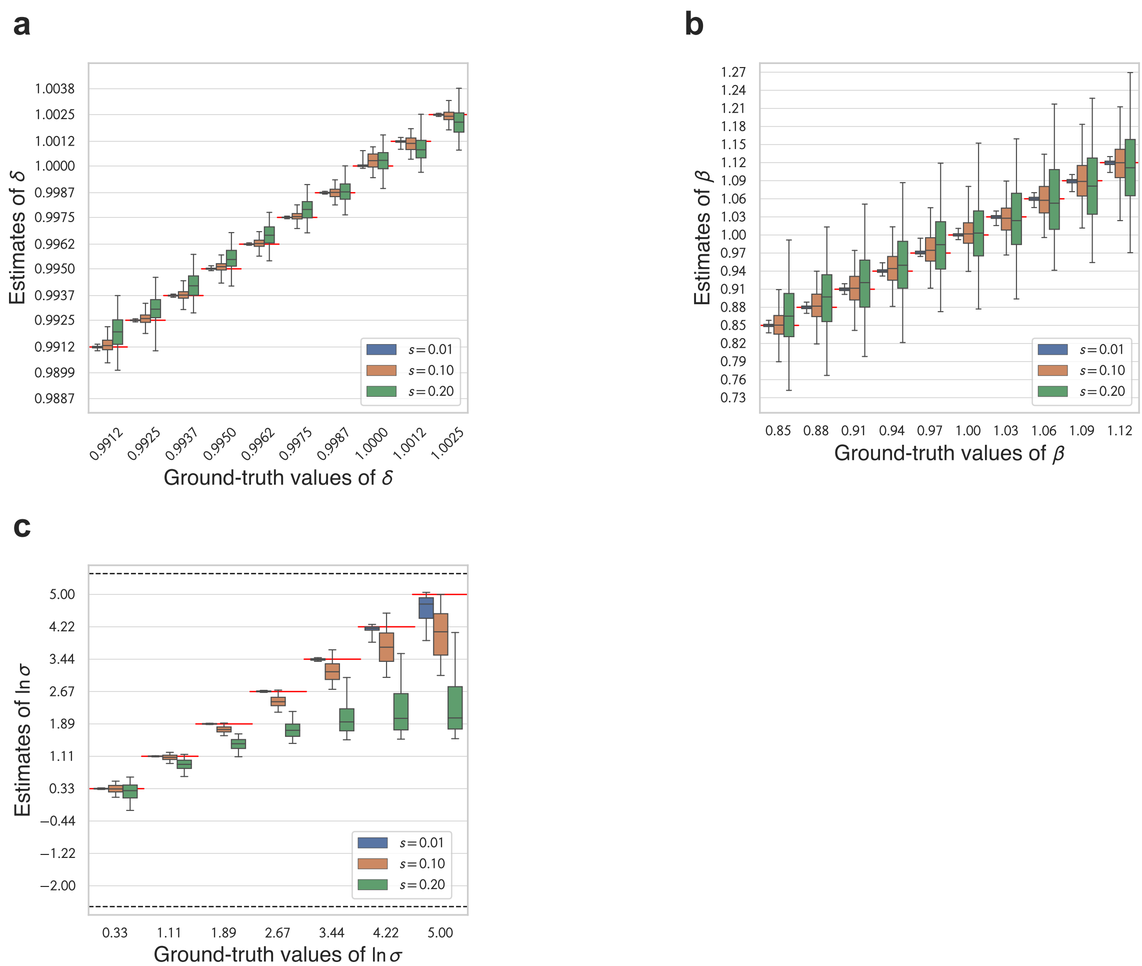

Figure S14: Box plots of **a**) δ, **b**) β, and **c**) $\ln \sigma$ estimates for PS3.

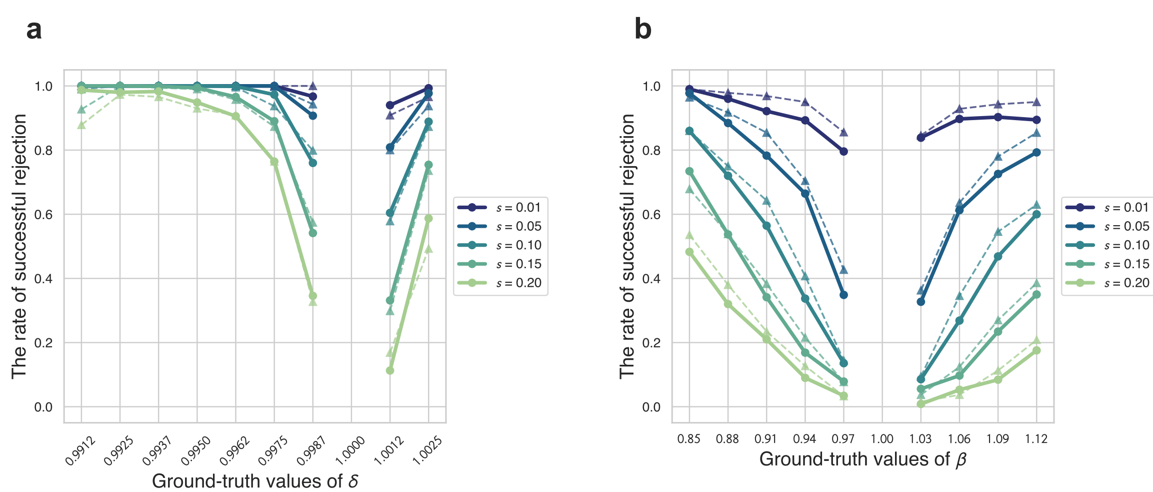

Figure S15: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for PS3.

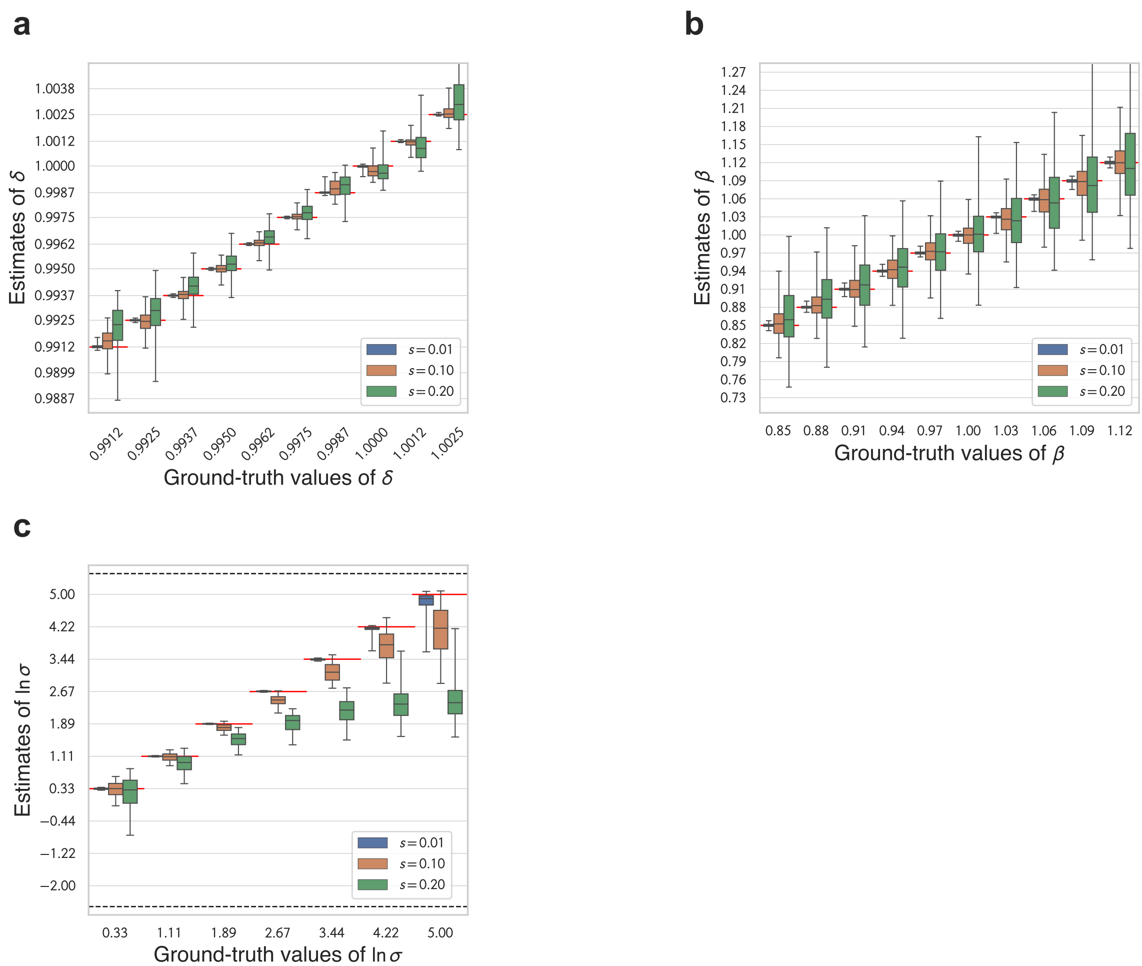

Figure S16: Box plots of **a**) δ, **b**) β, and **c**) $\ln \sigma$ estimates for PS4.

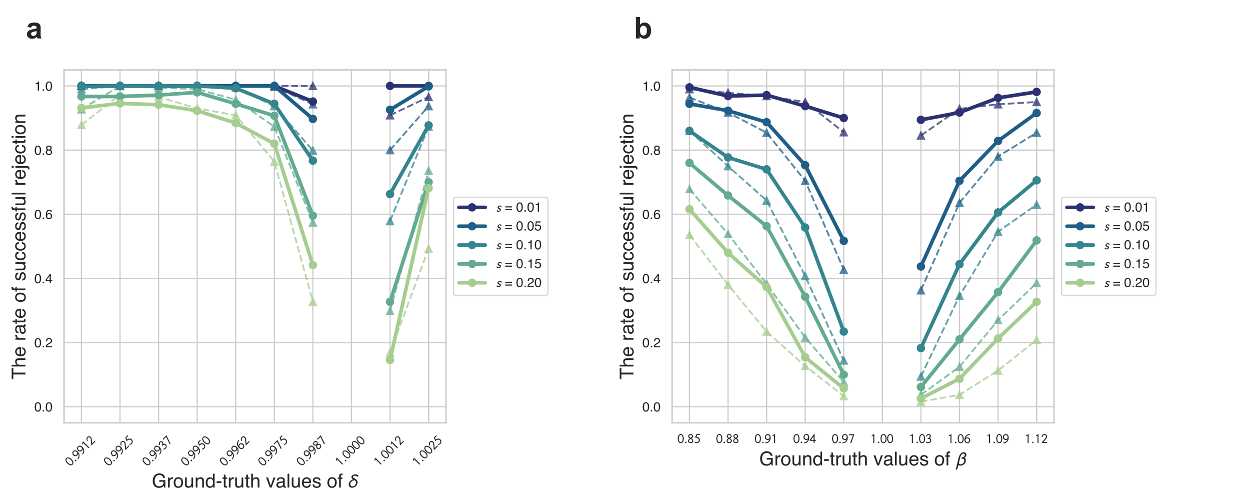

Figure S17: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for PS4.

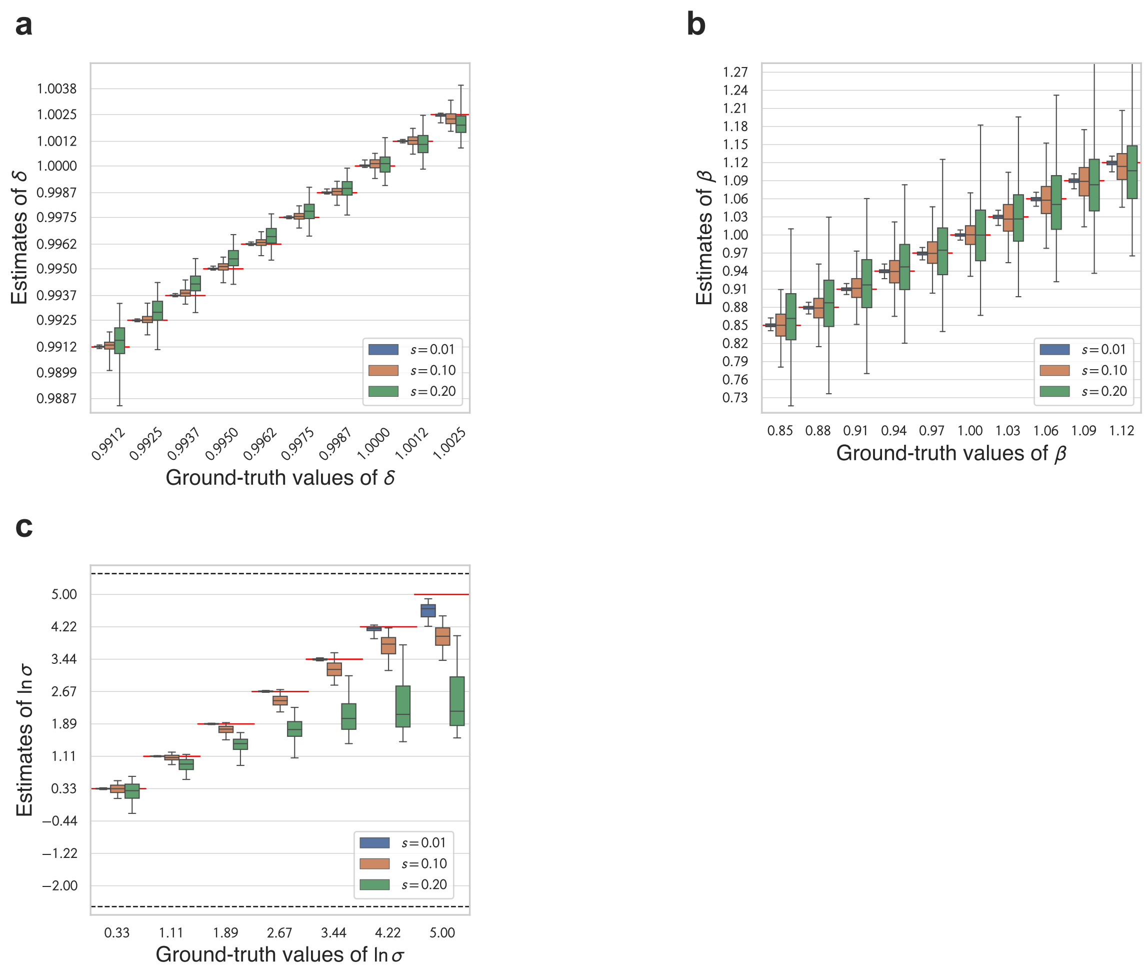

Figure S18: Box plots of **a**) δ, **b**) β, and **c**) $\ln\sigma$ estimates for PS5.

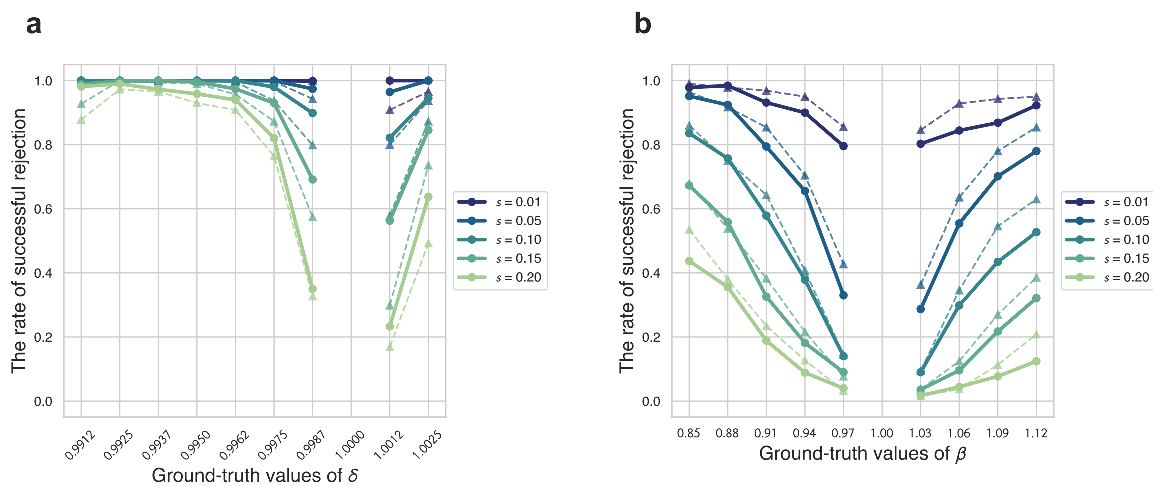

Figure S19: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for PS5.

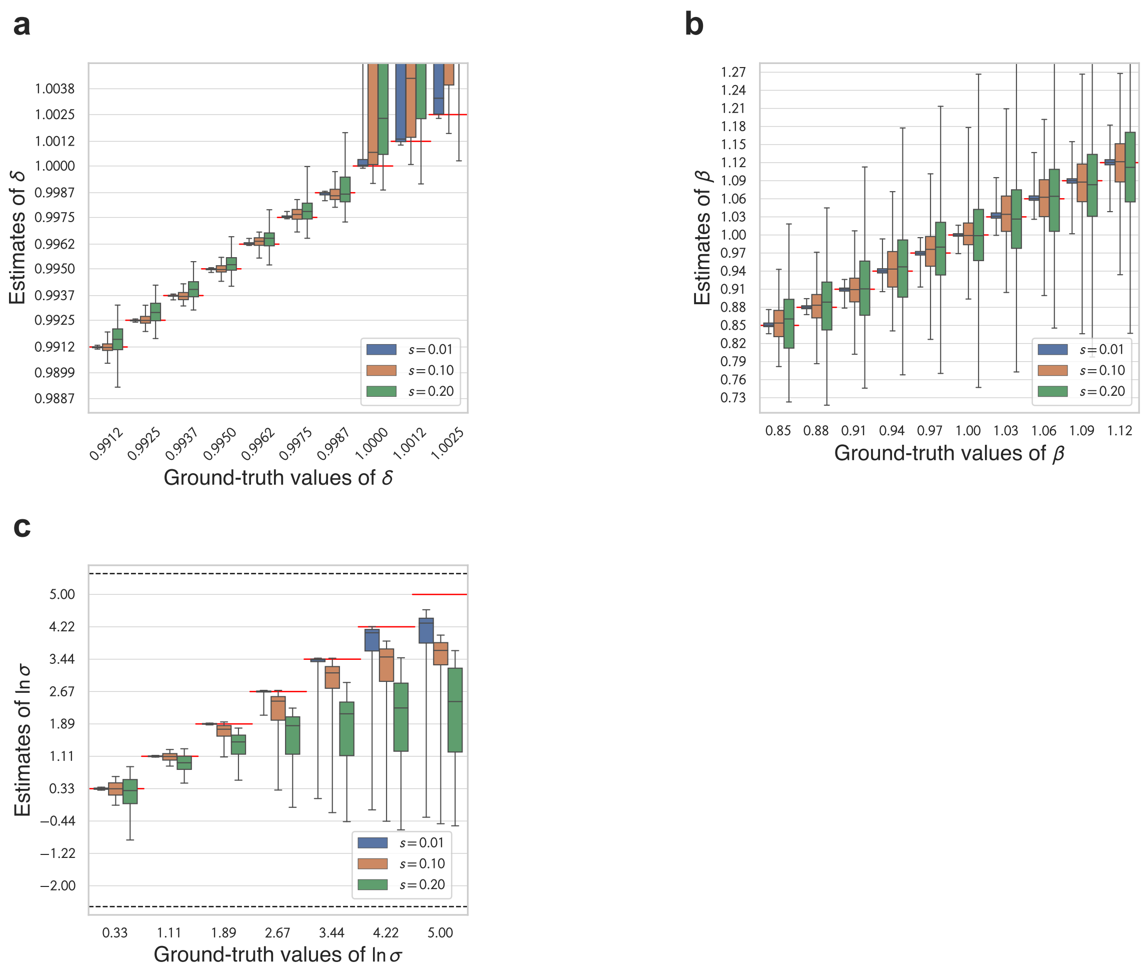

Figure S20: Box plots of **a**) δ, **b**) β, and **c**) $\ln\sigma$ estimates for PS6.

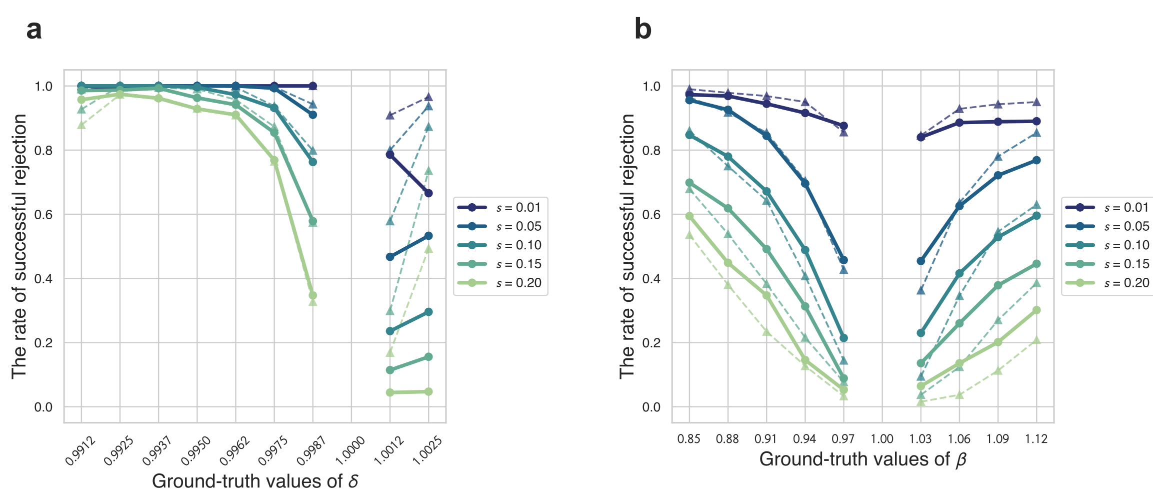

Figure S21: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for PS6.

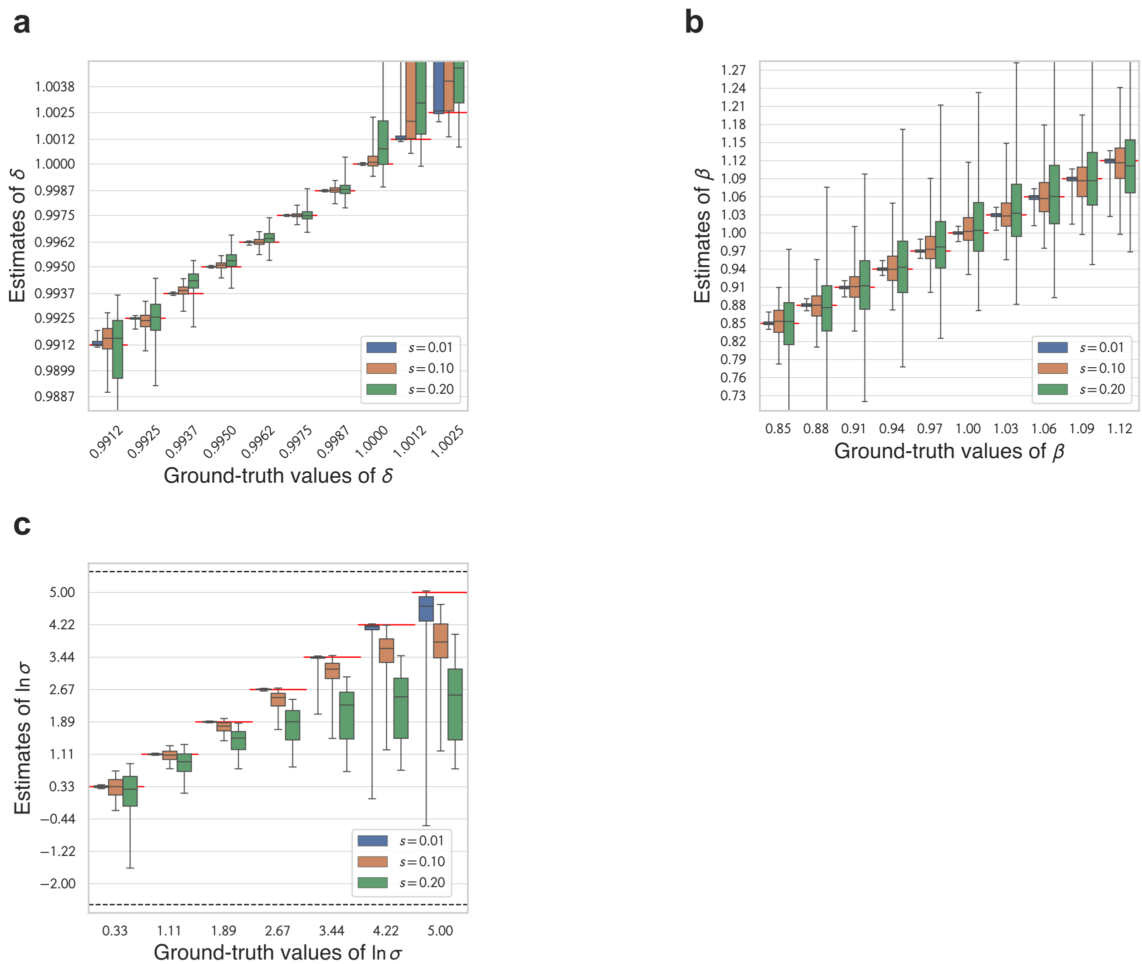

Figure S22: Box plots of **a**) δ, **b**) β, and **c**) $\ln\sigma$ estimates for AS.

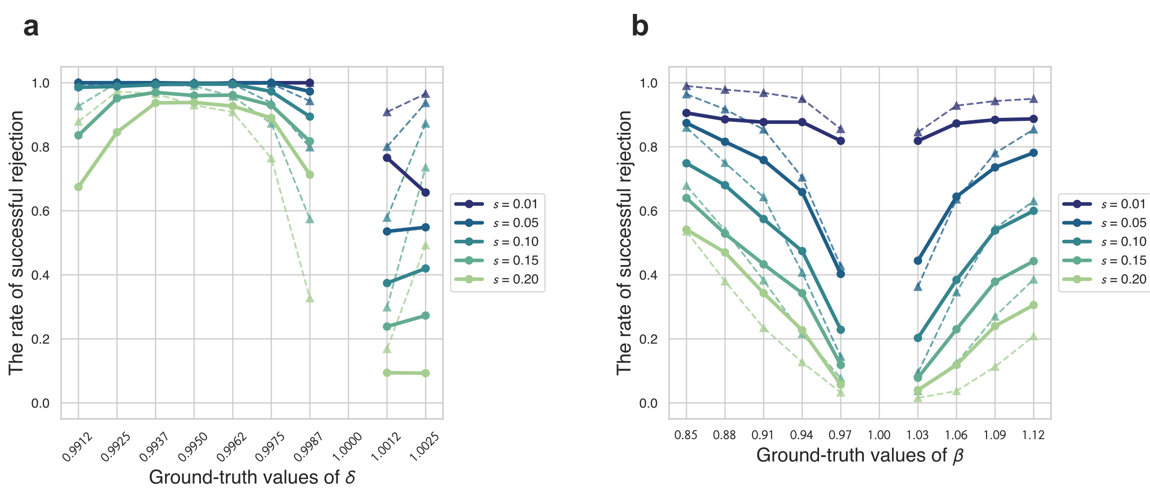

Figure S23: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, for AS.

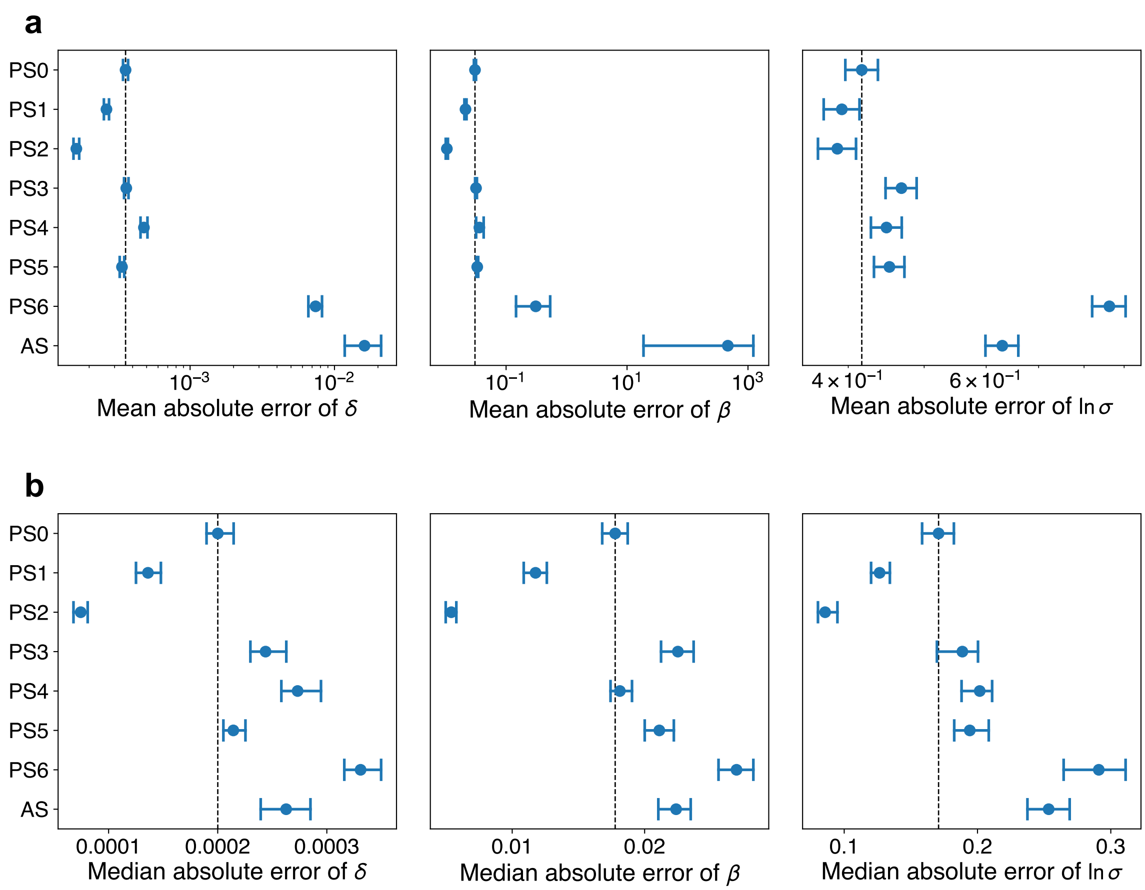

Figure S24: **a**) Mean and **b**) median of the absolute errors of the estimates for each problem set. Error bars represent the bootstrap 95% confidence intervals.

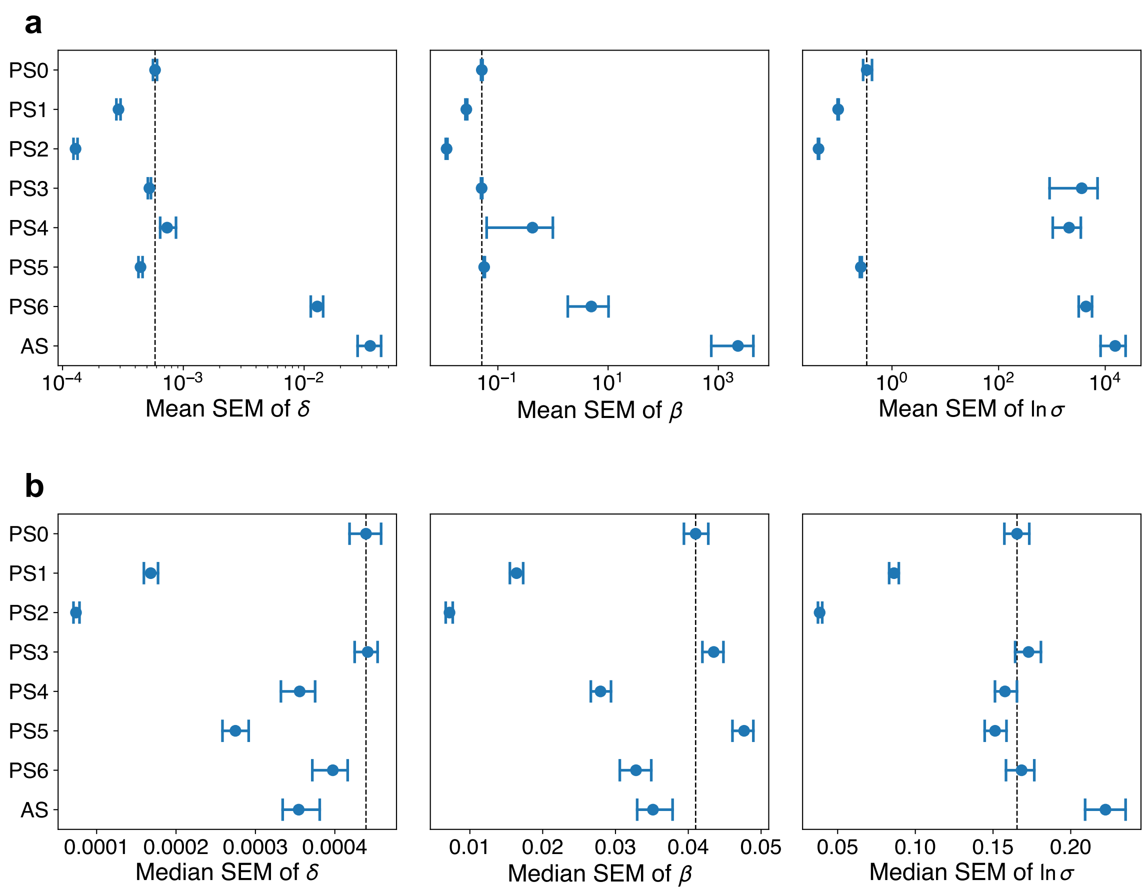

Figure S25: **a**) Mean and **b**) median of the standard errors of the estimates for each problem set. Error bars represent the bootstrap 95% confidence intervals.

SA.7 Case of negative ground-truth $\ln \sigma$

In the analysis in the main text, the ground-truth values of the curvature parameter $\ln \sigma$ were restricted to positive values, but here, we analyse the accuracy of parameter estimation for negative values of $\ln \sigma$ or the case of a lower elasticity. We used ground-truth values for $\ln \sigma$ of -2.00, -1.22, and -0.44, and the same ground-truth values for δ and β as in the main analysis.

Fig. S26 **a**) and **b**) show box plots representing the distribution of the estimated δ and β. Each plot summarizes 10 replications of all combinations of ground-truth values of β (δ) and $\ln \sigma$, i.e., 300 simulation agents. Although the median of the estimates is not significantly different from the case where $\ln \sigma$ is positive, the range of the distribution of the estimates (length of boxes and whiskers) is much larger than that in the case where $\ln \sigma$ is positive.

Fig. S26 **c**) shows box plots representing the distribution of the estimated $\ln \sigma$. Each plot summarizes 10 replications of all combinations of ground-truth values of δ and β, i.e., 1,000 simulation agents. For the case where ground-truth $\ln \sigma$ is positive (Fig. S1), even with a noise size of $s = 0.10$, the estimation was accurate enough not to overlap the whiskers of neighbouring plots, whereas the negative case is poor.

Fig. S27 shows the rate of successful rejection of the null hypothesis that $\widehat{\delta} = 1$ and $\widehat{\beta} = 1$, respectively. Each point contained 10 replications of all combinations of ground-truth values of β (δ) and $\ln \sigma$, i.e., 300 simulation agents. Compared with the case where $\ln \sigma$ is positive, the success rate is generally lower.

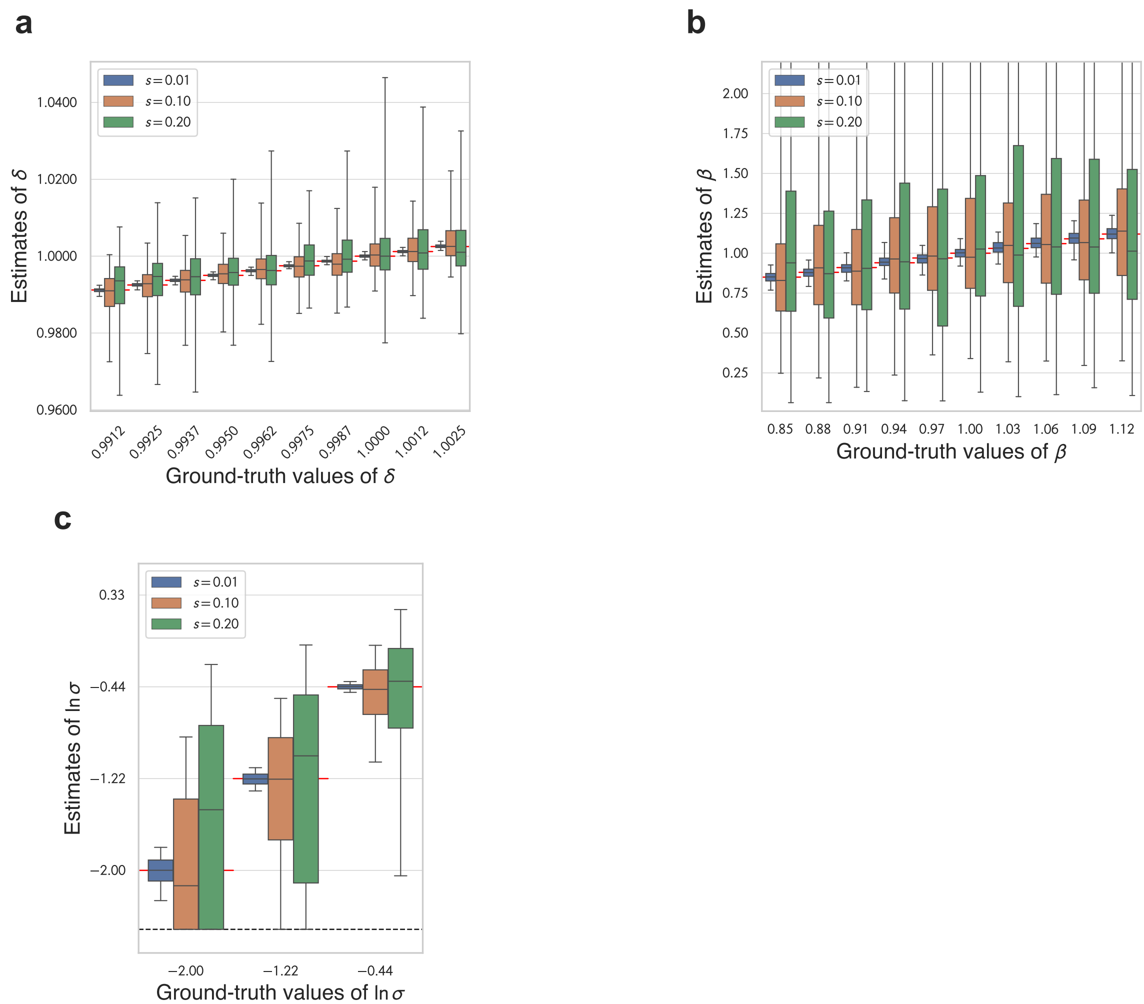

Figure S26: Box plots of **a**) δ, **b**) β, and **c**) $\ln \sigma$ estimates in the case where ground-truth $\ln \sigma$ is negative.

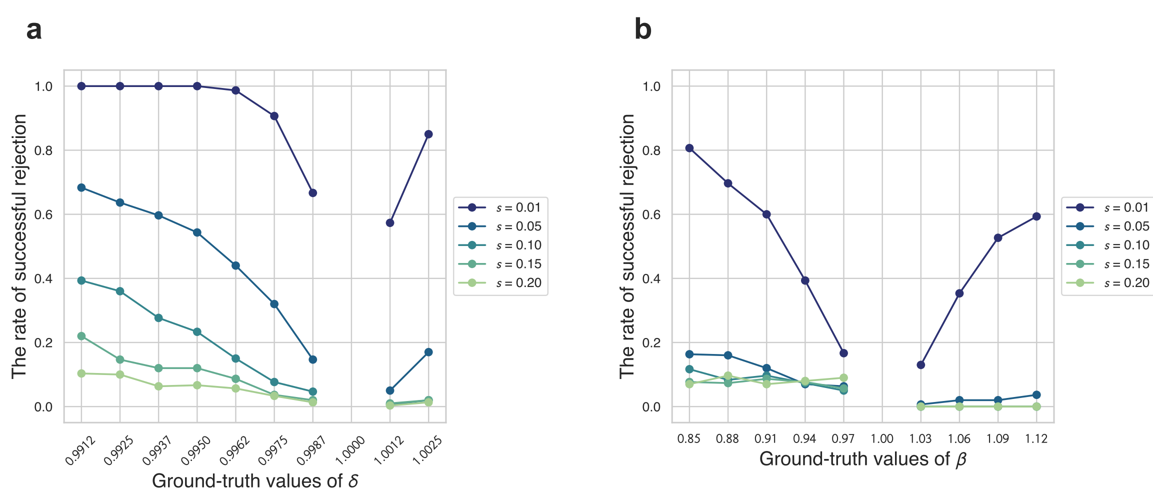

Figure S27: Rate of successful rejection of the null hypothesis that **a**) $\widehat{\delta} = 1$ and **b**) $\widehat{\beta} = 1$, respectively, in the case where $\ln \sigma$ is negative.

SA.8 Demand curve for individuals who has an extreme $\ln\sigma$

Fig. S28 shows demand curves for individuals with very large or small values of the curvature parameter $\ln\sigma$. Both individuals with $\ln\sigma = 6$ and 7 allocate all amount to the sooner period at lower prices than the switching point; otherwise, they allocate all to the later period. There is no difference in behaviour between the two. For $\ln\sigma = 5$, since they allocate an amount greater than 0 to the sooner period at a price $1 + r = 1.44$, which is higher than the switching point, there is a difference in the behaviour compared with $\ln\sigma = 6$. However, there is no difference in behaviour other than at a price $1 + r = 1.44$.

For $\ln\sigma = -2$, -3 and -4, the demand curves do not strictly coincide, so there is a difference in behaviour. However, the differences are slight.

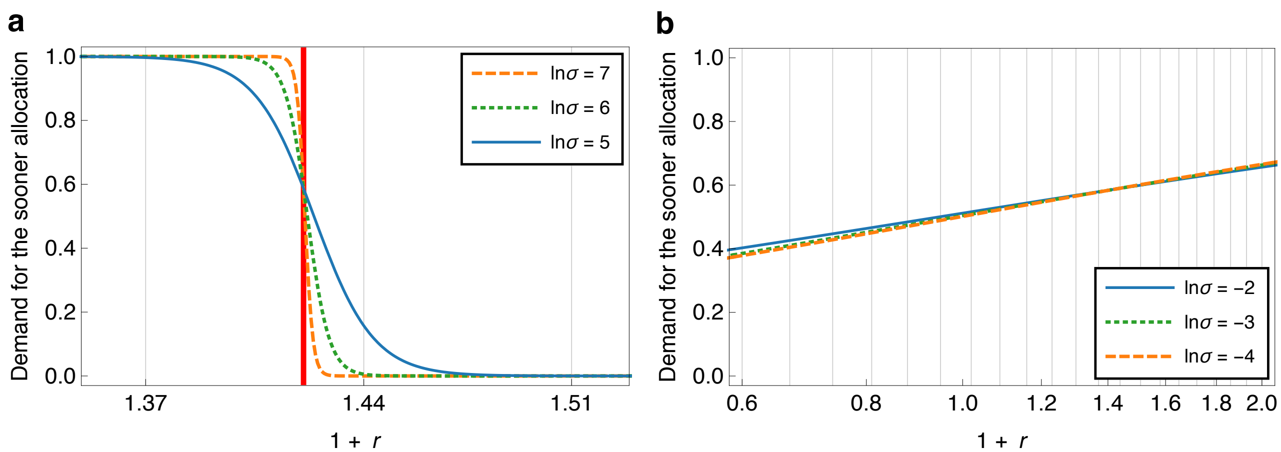

Figure S28: Demand curve for an individual whose curvature parameter is **a**) $\ln\sigma \geq 5$ and **b**) $\ln\sigma \leq -2$, where $\delta = 0.9950$ and $\beta = 1$. The horizontal axis representing price $1+r$ is on a logarithmic scale. Note that the individual faces the decision problem of allocating between now ($t = 0$) and $k = 70$ days later with the prices indicated by the vertical lines in the figure. For Fig. **a**, only the neighbourhood of the switching point (indicated as a red line) is shown.