

State Dependent Price Setting Rules Under Implicit Thresholds: An Experiment - Supplementary Material

Justin Deloy LeBlanc
University of Arkansas

Andrea Civelli
University of Arkansas

Cary Deck
University of Arkansas
Economic Science Institute, Chapman University
University of Alaska Anchorage

Klajdi Bregu
Univeristy of Arkansas

Abstract

This document provides additional material to the submitted version of the paper *State Dependent Price Setting Rules Under Implicit Thresholds: An Experiment* for on-line publication.

A Learning Analysis

In order to explore whether or not subjects' learning over the course of the experiment was affecting our results, extensive analysis was performed to ascertain any evidence of this occurring. This analysis included two specific approaches to the data. The first was to investigate if subjects that saw a given treatment in the first period performed statistically different from other individuals that saw that same treatment in the last period. The second was to investigate if over the course of the experiment, subjects' behavior changed from the first half of the a treatment to the second half. In the two subsections that follow, we present the findings of our analysis, which revealed that there is no indication that learning—or lack thereof—has an impact on our results.

A.1 Ordering Effects Analysis

Figure S1 shows the average total profits for a given treatment as it was experienced in a given period for each of the subjects. For instance, for Treatment 1, if subjects saw that treatment in the first period, their data are represented at 1 on the x-axis. If those subjects saw that treatment in the second period, those data are indicated at $x = 2$, and so on. As can be seen by the standard error bars, only one of the data points appears to be statistically different from any other data point for a given treatment. When Treatment 1 is experienced in the second period, it appears that it could possibly be statistically different because the error bars do not overlap. To test this, a z -test was performed, yielding a z -statistic of 2.4. As a rule of thumb, a z score in the range of 2.0 – 2.5 is considered to indicate two samples' means are only marginally statistically different. Thus, we conclude that with regard to average total profit, there is no indication that subjects are able to achieve statistically significant higher average profits across all treatments regardless of in which period the subjects encountered the treatment.

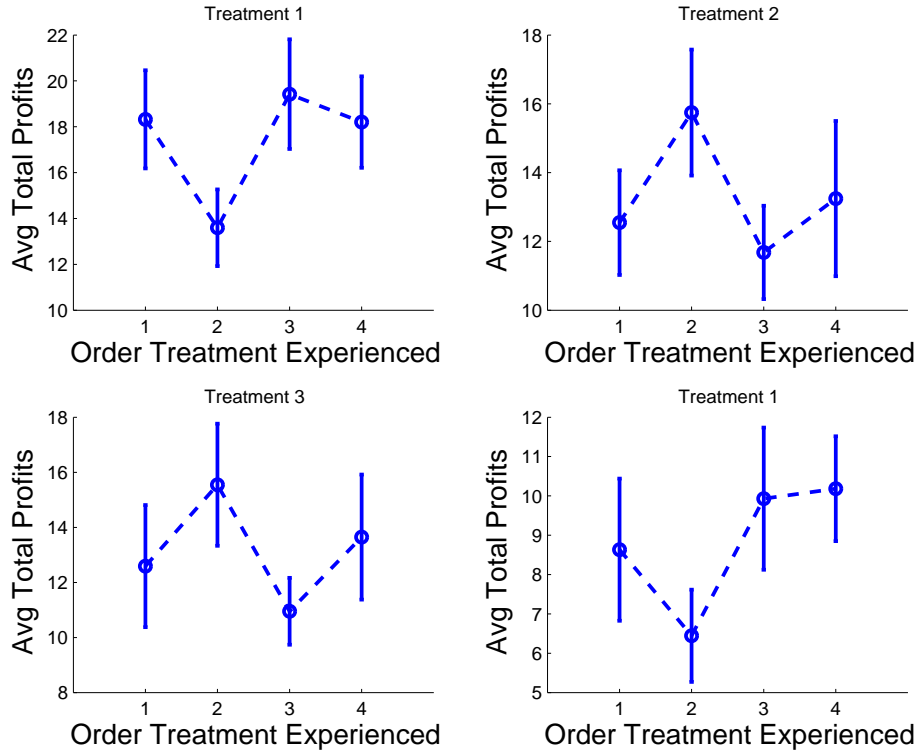


Figure S1: Plots indicating the average total profits earned by subjects that saw a particular treatment in order from first to last. The way to read this is: For the panel for Treatment 1, the data point at 1 on the x-axis means that that subject pool saw Treatment 1 in the first period, at $x = 2$, the subjects saw that treatment in the second period, etc.

Figure S2 shows the average number of pizzas baked by subjects across treatments as a function of period in which they saw that treatment. As can be seen, there are a couple of instances where there are statistical differences between how many pizzas are baked when a given treatment is seen in a particular period. In particular this seems to occur when given treatments are seen in the second and third periods. However, there are a few things to note regarding these data. First, in no treatment was there a statistical difference between the number of pizzas baked when any of the treatments were seen in the first period compared to the fourth period. Further, notice that there is no systematic pattern as to which period subjects would bake more in the second period rather than the third period. Finally, the fact that for any given treatment only one of the periods has an average that statistically differs from one of the other periods is telling that, on the whole, we do not see any substantive evidence that there is a change in behavior over the course of the four periods.

Finally, Figure S3 shows the most important data from this analysis: The evolution of the number of price updates performed by treatment as a function of the period in which the

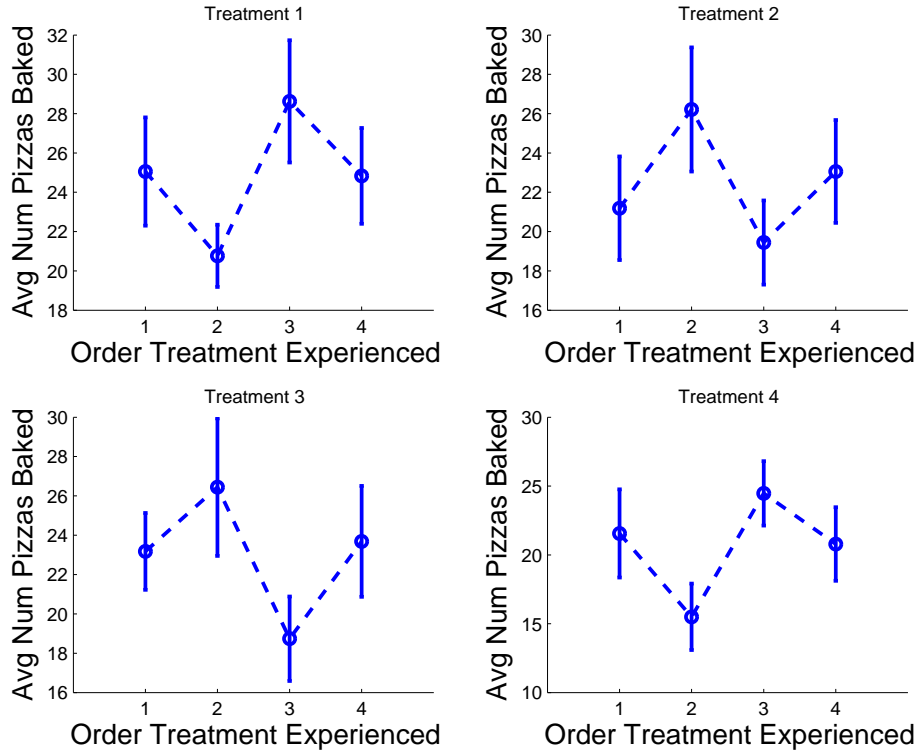


Figure S2: Plots indicating the average number of pizzas baked by subjects that saw a particular treatment in order from first to last. The way to read this is: For the panel for Treatment 1, the data point at 1 on the x-axis means that that subject pool saw Treatment 1 in the first period, at $x = 2$, the subjects saw that treatment in the second period, etc.

subject saw the treatment. As can be seen, when Treatment 1 is experienced in the second period is the only case that might cause us to think there is any statistical significance across treatments and orders. However, after performing a z -test, we find a z -statistic of 1.3, which indicates that we cannot conclude that the average number of price updates when Treatment 1 was experienced second is statistically different from when Treatment 1 is seen in the first period.

We turn now to the evidence of a change in the subjects' ability of inferring the profit thresholds in function of the ordering at which they face a specific treatment. We use the strategy described in the main paper, and in Section C of this online Appendix, to estimate the kernel densities of the baking and updating decisions of each treatment separately for the two sets of subjects who see the treatment either in the first or last period of the experiment. Figure S4 shows that it is possible to observe a visually clear difference between the two groups only for the updating decisions of Treatment I. On the contrary, the densities of the two groups of subjects for the other treatments and decisions are very close to each other.

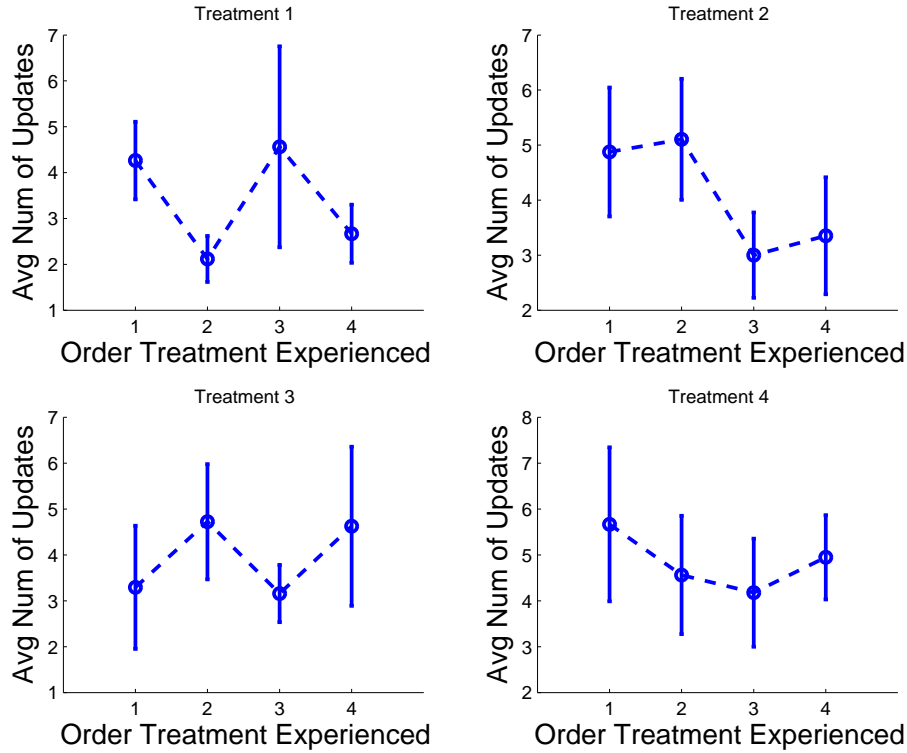


Figure S3: Plots indicating the average number of price updates by subjects that saw a particular treatment in order from first to last. The way to read this is: For the panel for Treatment 1, the data point at 1 on the x-axis means that that subject pool saw Treatment 1 in the first period, at $x = 2$, the subjects saw that treatment in the second period, etc.

The density of the price updating decisions when Treatment I is seen as last by subjects exhibits a much smaller variance and it is better centered around the theoretical threshold; moreover, the density of the updating tasks when the treatment comes first in the ordering attributes a lot of probability mass to very high values of profits above 0.80, whereas the other density is basically zero for that region of profits. We assess the statistical significance of these differences by using the two-sample Kolmogorov-Smirnov test that tests the null hypothesis that two samples come from a common empirical distribution. The test is based on the maximum distance between the empirical cumulative densities of two samples, and it relies on a statistic that asymptotically follows the Kolmogorov-Smirnov distribution. The test statistic is .07, and it does not reject the null hypothesis at very large level of significance (the p-values is .95). Therefore, there is no particular evidence to conclude that different orderings cause a change in the decision rules of the subjects that play a treatment as last relative those who play it first in the experiment.

Although the densities of each decision type are not significantly different from each other,

there might still be the possibility that the thresholds estimates implied by specific pairs of baking and updating densities are different in a systematic way that suggests learning effects in subjects. We provide evidence against this possibility looking at the threshold estimates for the two groups of subjects of this analysis in Figure S5. We find that the thresholds decreases for Treatment I and III if the treatment is seen last; despite the change going in the right direction, these values are way above the optimal values from the simulations. For Treatment II there is no effect, while for Treatment IV we actually find the opposite effect than what we would expect from learning. Overall, we find enough evidence to exclude learning within the experiment time horizon.

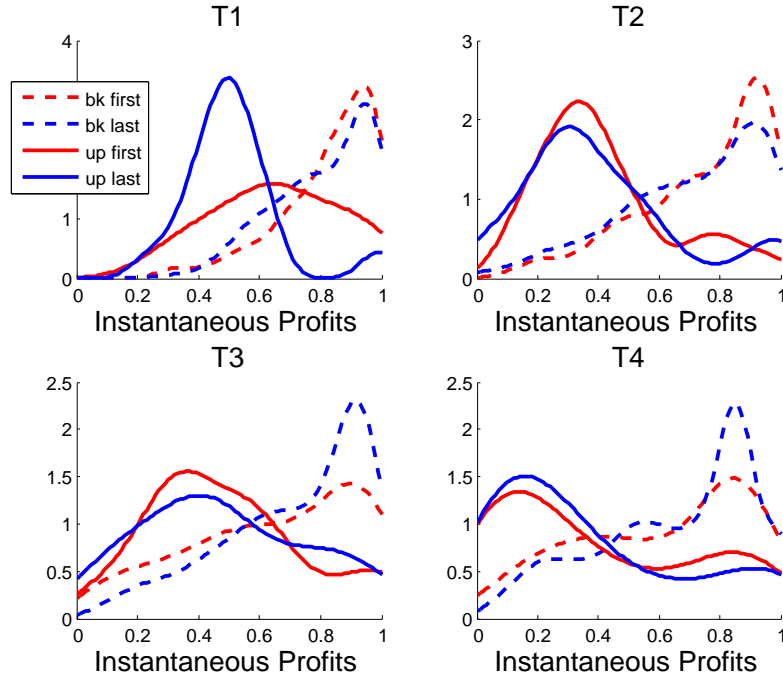


Figure S4: Empirical kernel distribution of the decisions of price updating and baking tasks. Gaussian kernel function for the aggregate set of decisions across subjects that saw a particular treatment either in the first or the last period of the experiment.

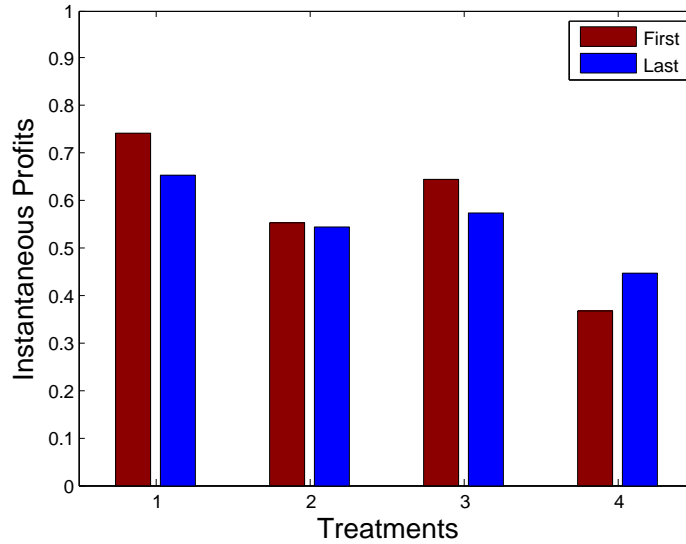


Figure S5: Estimated thresholds in the four treatments; aggregate set of decisions across subjects that saw a particular treatment either in the first or the last period of the experiment.

A.2 Time-binned Analysis

Figure S6 shows plots akin to the plots above, but these data show data for average total profits that have been split by actions that happened in the first half of the period and those that happened in the second half. As can be seen, there is no dramatic difference between the two, and all of the actions follow the same pattern across treatment ordering.

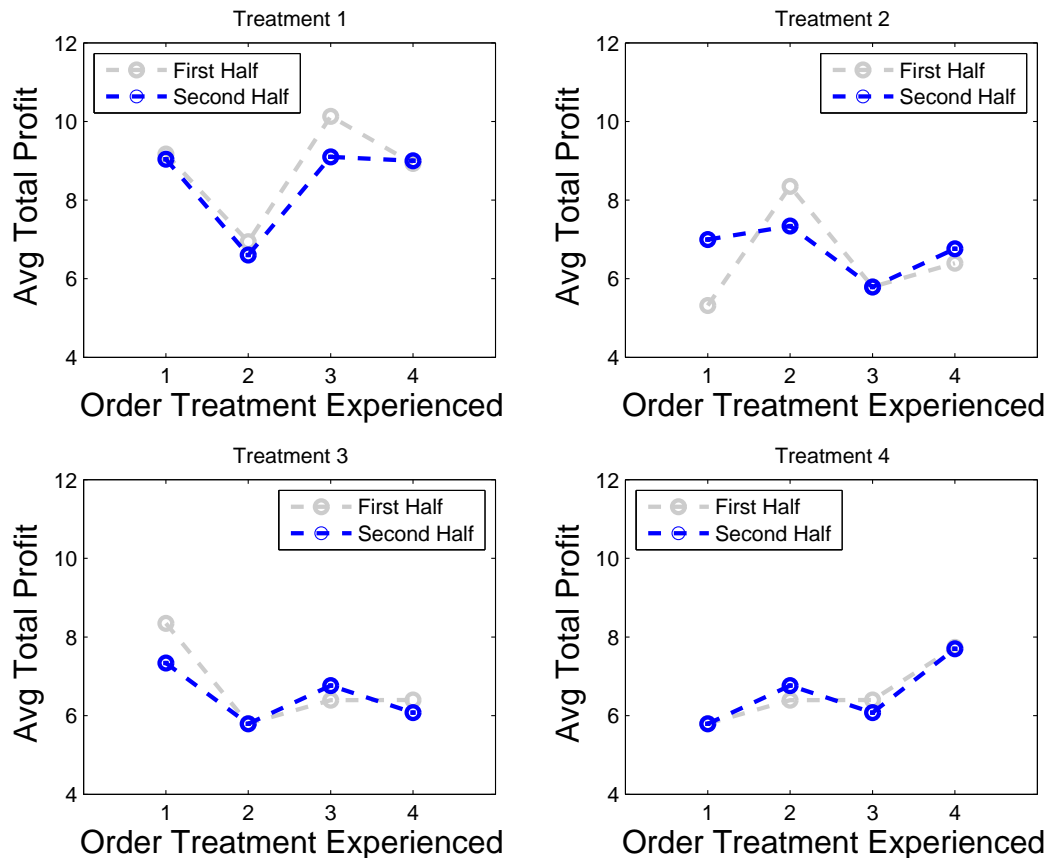


Figure S6: Plots indicating the average total profits earned by subjects that saw a particular treatment in order from first to last. The way to read this is: For the panel for Treatment 1, the data point at 1 on the x-axis means that that subject pool saw Treatment 1 in the first period, at $x = 2$, the subjects saw that treatment in the second period, etc. The grey line indicates average total profits earned in the first half of the experiment, and the blue line indicates the average total profits earned in the second half of a given period.

Figure S7 shows plots of the average number of pizzas baked in each treatment again ordered by the period in which subjects experienced the treatment. The grey line indicates the average number of pizzas baked in the first half of the treatment, and the blue line indicates the average number of pizzas baked in the second half of the treatment. As can be seen, the subjects behave quite similarly in the first half and the second half. Moreover, both the averages are directionally consistent while moving from one period to the next.

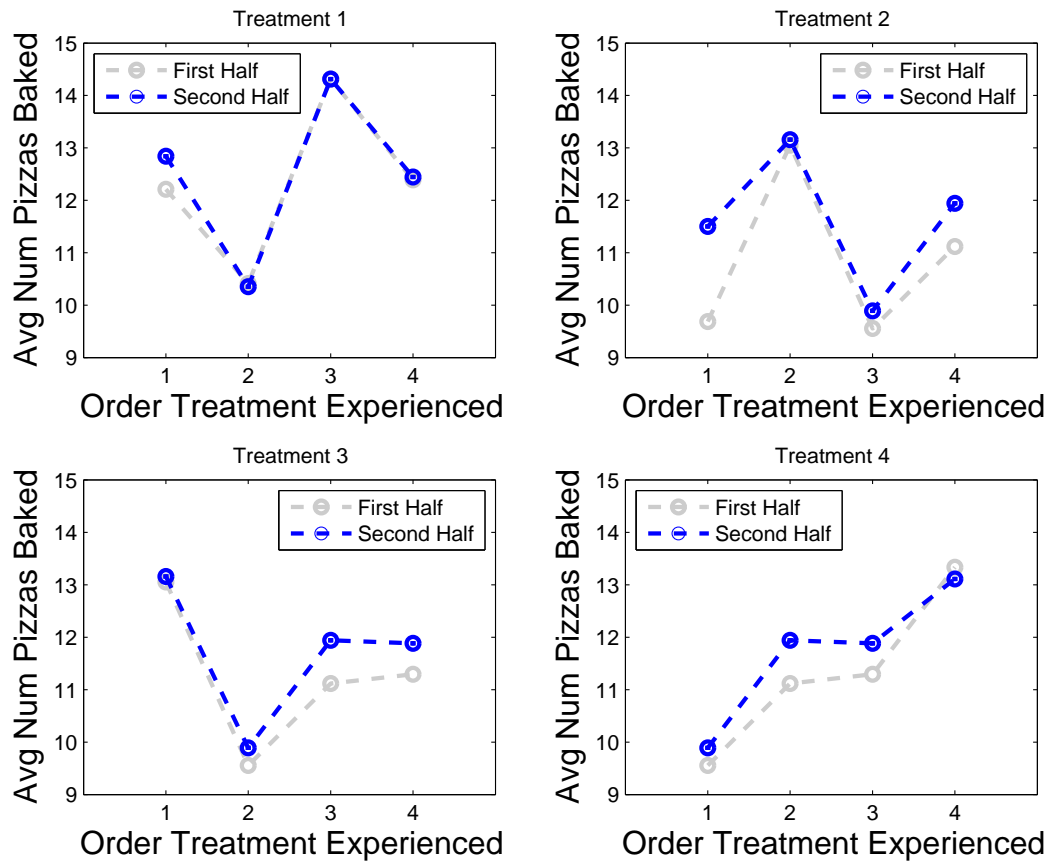


Figure S7: Plots indicating the average number of pizzas baked by subjects that saw a particular treatment in order from first to last. The way to read this is: For the panel for Treatment 1, the data point at 1 on the x-axis means that that subject pool saw Treatment 1 in the first period, at $x = 2$, the subjects saw that treatment in the second period, etc. The grey line indicates average number of pizzas baked in the first half of the experiment, and the blue line indicates the average number of pizzas baked in the second half of a given period.

Figure S8 shows plots of the average number of price updates in each treatment again ordered by the period in which subjects experienced the treatment. The grey line indicates the average number of price updates in the first half of the treatment, and the blue line indicates the average number of price updates in the second half of the treatment. As can be seen, subjects' behavior in a given period is very similar in the first half when compared to the second half, and the pattern of subjects' first and second half averages follows a similar pattern from seeing a given treatment in a given period.

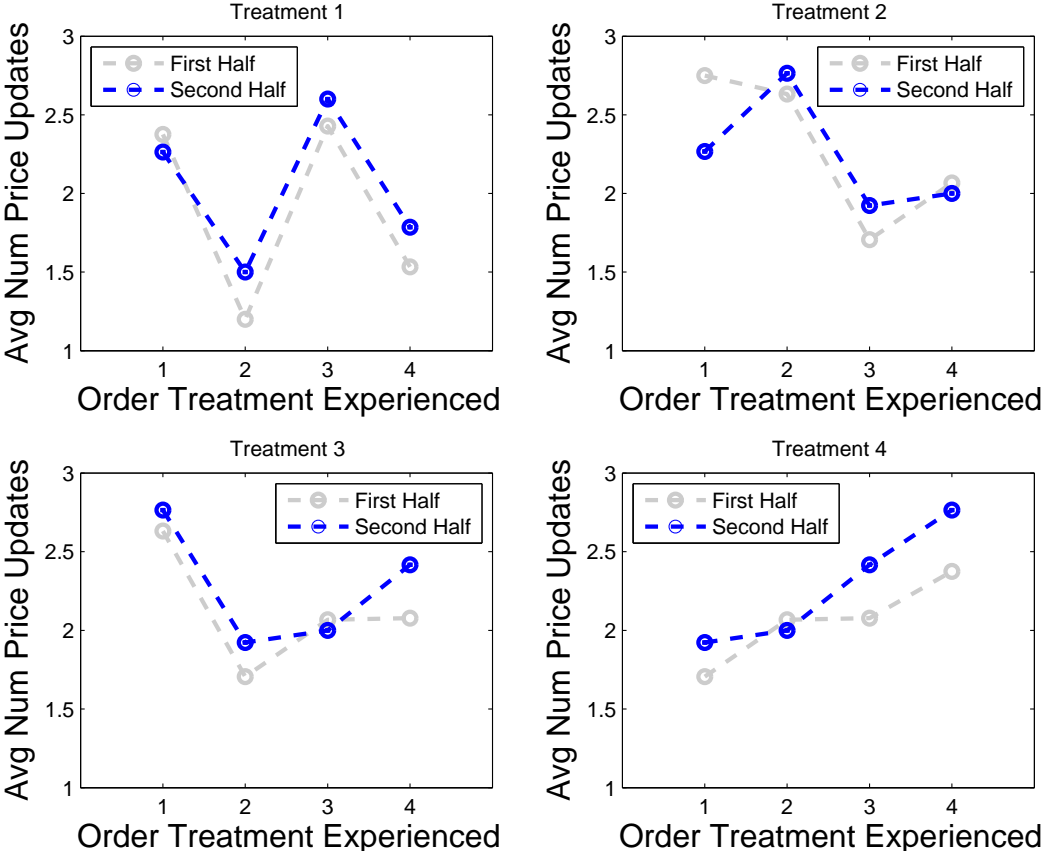


Figure S8: Plots indicating the average number of price updates by subjects that saw a particular treatment in order from first to last. The way to read this is: For the panel for Treatment 1, the data point at 1 on the x-axis means that that subject pool saw Treatment 1 in the first period, at $x = 2$, the subjects saw that treatment in the second period, etc. The grey line indicates average number of price updates in the first half of the experiment, and the blue line indicates the average number of price updates in the second half of a given period.

We end this subsection again with the output from the thresholds estimation strategy. As done for the treatments ordering, we estimate the kernel densities of the baking and updating decisions of each treatment for the decisions taken in the first and second half of the

treatment period. Figure S9 illustrates that large differences between the earliest and latest sets of decisions are not found for the baking nor the updating tasks. The densities exhibit similar dispersions and locations across treatments, and there is no particular evidence to conclude that decisions taken in the second half of the treatment follow different rules than those taken at the beginning of the treatment period.

Figure S10 checks for learning effects in the thresholds used by the subjects in the two periods. As above, we provide evidence against this possibility. The threshold estimated for Treatment I is higher in the second half of the treatment period. The thresholds are lower for Treatment III and IV, while it is the same for Treatment II. All the values are, however, way above the optimal values from the simulations and there is no systematic gain from being engaged in the same treatment environment for a longer period. Overall, we do not find evidence of any particular improvements in the thresholds identified by the subjects in the second halves of the treatment periods and we can exclude learning within the experiment time horizon also in this case.

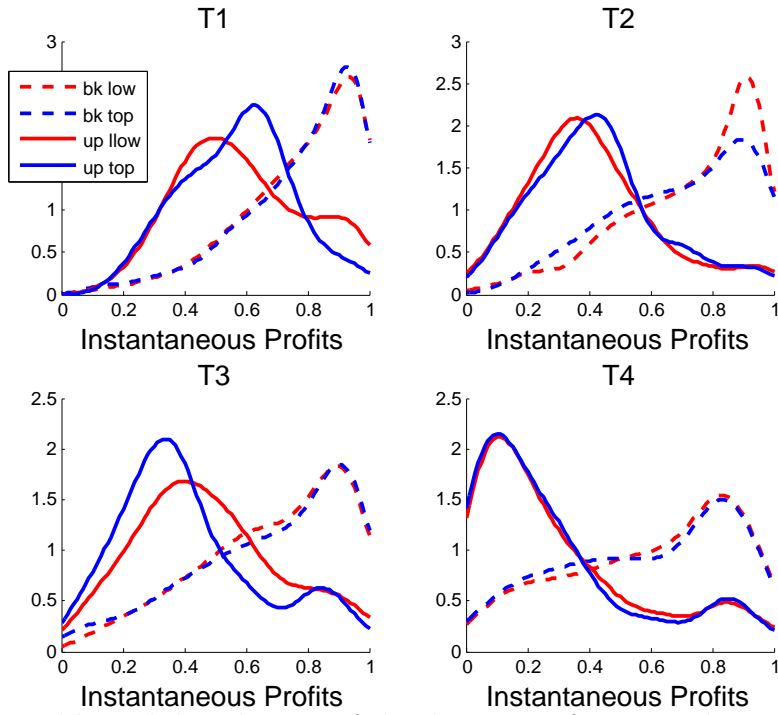


Figure S9: Empirical kernel distribution of the decisions of price updating and baking tasks. Gaussian kernel function for the aggregate set of decisions taken in the first and second half of each treatment period respectively.

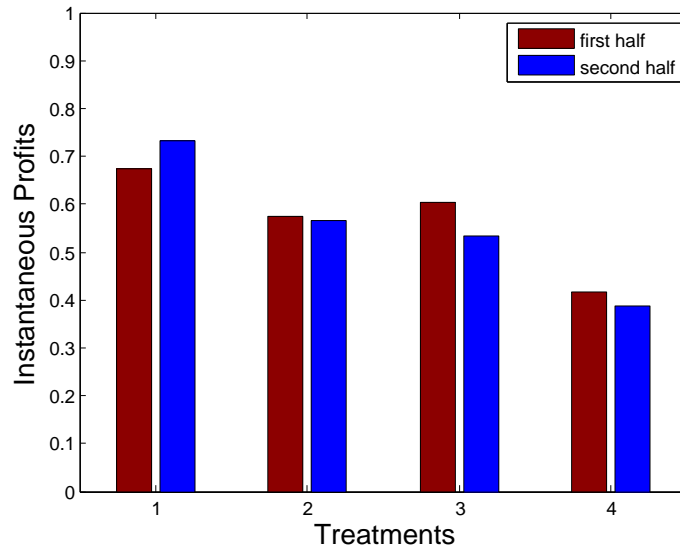


Figure S10: Estimated thresholds in the four treatments; aggregate set of decisions taken in the first and second half of each treatment period respectively.

B Supplemental Data Analysis

In the subsections that follow, we present additional data analysis as well as graphical analysis that have been performed to ensure the robustness of our results and to clarify some points in the paper. Specifically, Section B.1 looks at possible effects of outliers on the experimental results; we find the means reported in the paper are quite robust. Section B.2 provides some more technical details about the implementation of the simulation. Section B.3 illustrates the diffusion process of profits after a price update in the four treatments. Finally, Section B.4 reports the histories of profit realizations, baking and price updating for all the subjects in the experiment.

B.1 Median/Mean Comparison

In the main body of the paper we present results in terms of means. One of the possible problems with using means—as opposed to medians, for instance—is that means are possibly distorted by heavy outliers. To be thorough, we also examined the main results in terms of medians, and in Table S1 we present the medians of the results side by side with the means. As can be seen, we do not see any substantive difference between the two representations, which indicates to us that outliers are not driving the results presented in the primary text.

	Treatment I	Treatment II	Treatment III	Treatment IV
Experiment (mean/median)				
Number of Price Updates	3.4/3	4.0/4	3.9/4	4.8/4
Number of Pizzas Baked	24.7/24	22.5/22.5	22.9/23	20.7/20.5
Total Profits [ELD]	17.4/16.9	13.3/13.2	13.1/12.8	8.8/8.5
Simulation (mean/median)				
Number of Price Updates	2.9/3	4.3/4	4.3/4	5.3/5
Number of Pizzas Baked	29/29	27.7/28	27.7/28	26.7/27
Total Profits [ELD]	21.4/21.4	16.7/16.7	16.7/16.8	11.6/11.6

Table S1: Summary of pertinent data collected from the experiment and the simulation. Both means and medians are presented where applicable.

B.2 Task Completion Times

As mentioned in Section 5.1 in the primary text, we found an extremely prominent mode in the data on task completion times. This can be seen in the histograms provided below in Figures S11 and S12. As enumerated in the aforementioned section in footnote 11, we explored the use of varying times for the use of simulations, but we found negligible differences. Table S2 shows data reflective of comparing the simulation outcome from using 10 seconds as the duration for both tasks, as well as data from the simulation with noise added to the task times. Noise was added by drawing randomly from a uniform distribution. Thus, for every simulation, the baking task and updating task duration were $t_{\text{task}} = 10 + \varepsilon$, where $\varepsilon \in [-3, 3]$. As can be seen, there is no substantive difference between the two.

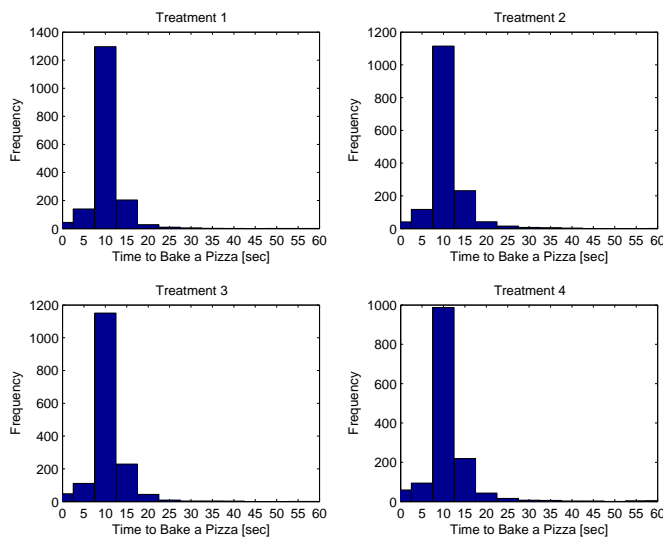


Figure S11: Histograms showing the distribution of the time it took subjects to complete the task of baking a pizza

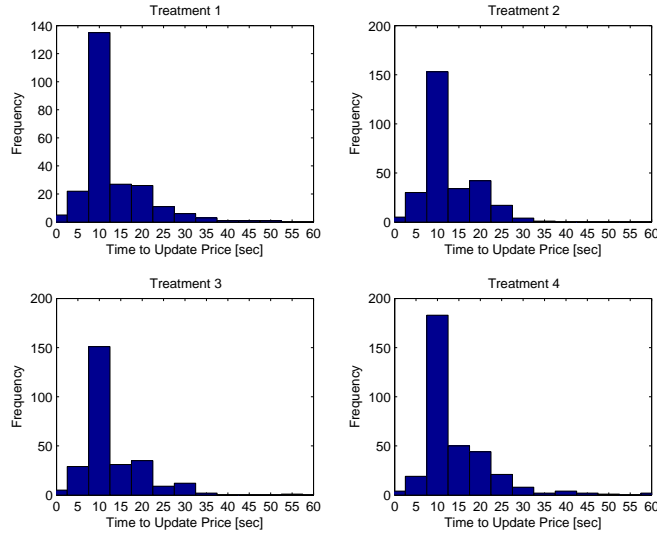


Figure S12: Histograms showing the distribution of the time it took subjects to complete the task of updating their price

	Treatment I	Treatment II	Treatment III	Treatment IV
Simulation (standard/with noise)				
Number of Price Updates	2.9/2.8	4.3/4.1	4.3/4.1	5.3/5.0
Number of Pizzas Baked	29/29.7	27.7/28.2	27.7/28.2	26.7/27.1
Total Profits [ELD]	21.4/21.9	16.7/17.1	16.7/17.1	11.6/11.8

Table S2: Summary of pertinent data collected from the the simulation. The standard simulation (task time equal to 10 seconds) is compared to the results obtained from adding noise from a uniform distribution from $[-3, 3]$.

B.3 Profit Diffusion

Figure S13 supplements the basic information reported in Figures 6 and 7 of the main paper with a more complete illustration of the profits distribution in function of the time elapsed from a price update, formally the profit “diffusion”. The average time necessary to reach a given level of instantaneous profit and thresholds lines correspond to the dotted red lines (from Figure 7 of the main paper).

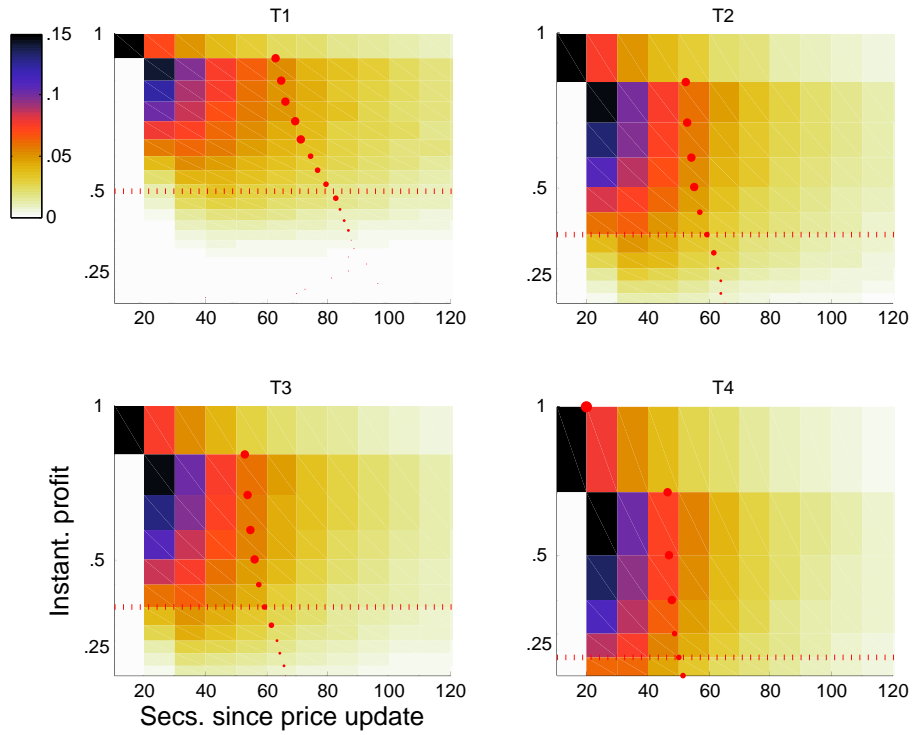


Figure S13: Profit diffusion over time since last price update occurred by treatment. Average time to reach a given level of instantaneous profit and thresholds lines are reported from Figure 7 of the main paper.

B.4 Subject History Plots

We presented only a few examples of the history of subjects' behavior over the course of the experiment in the main body of the paper in Figure 8. Below in Figure S14, we plot the history for each of the 70 subjects.

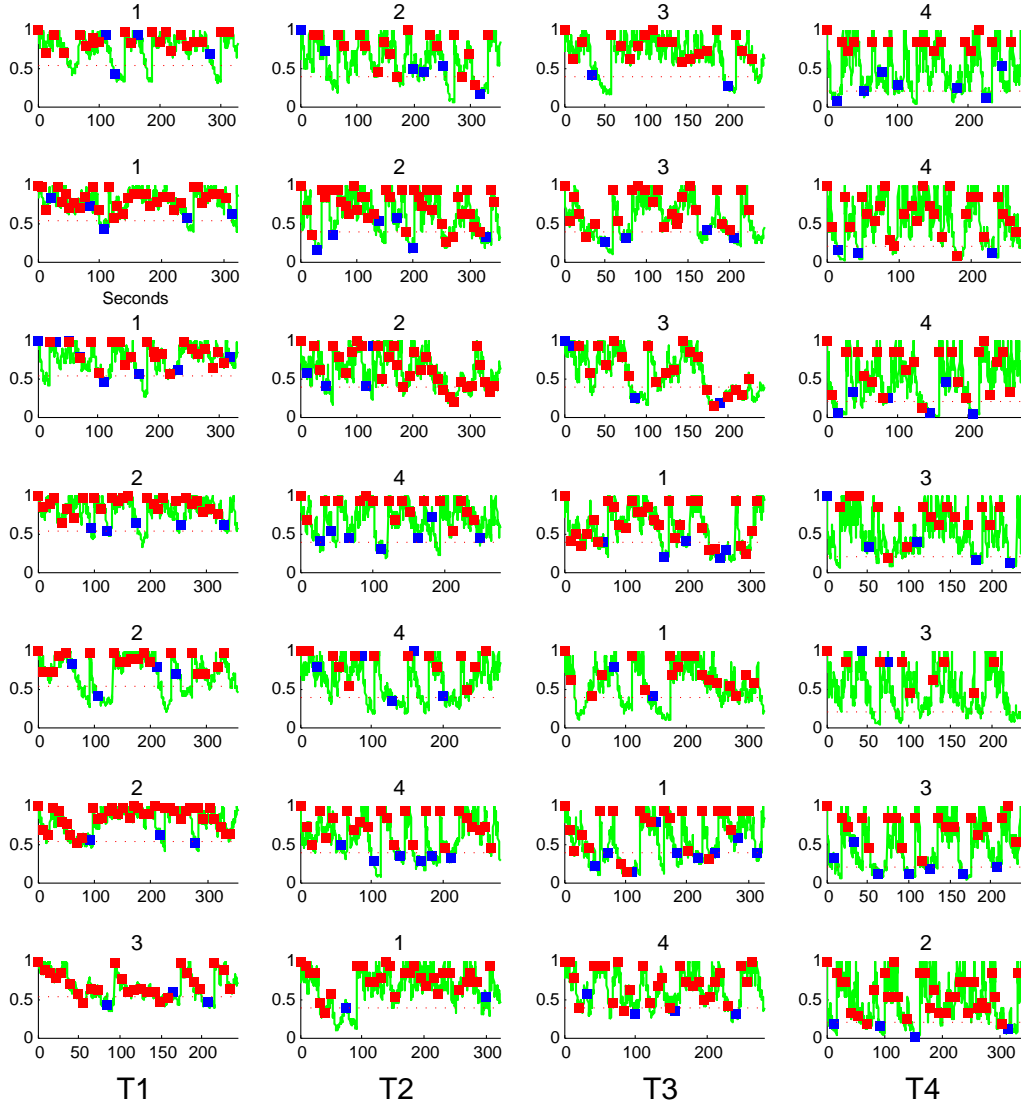


Figure S14: Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

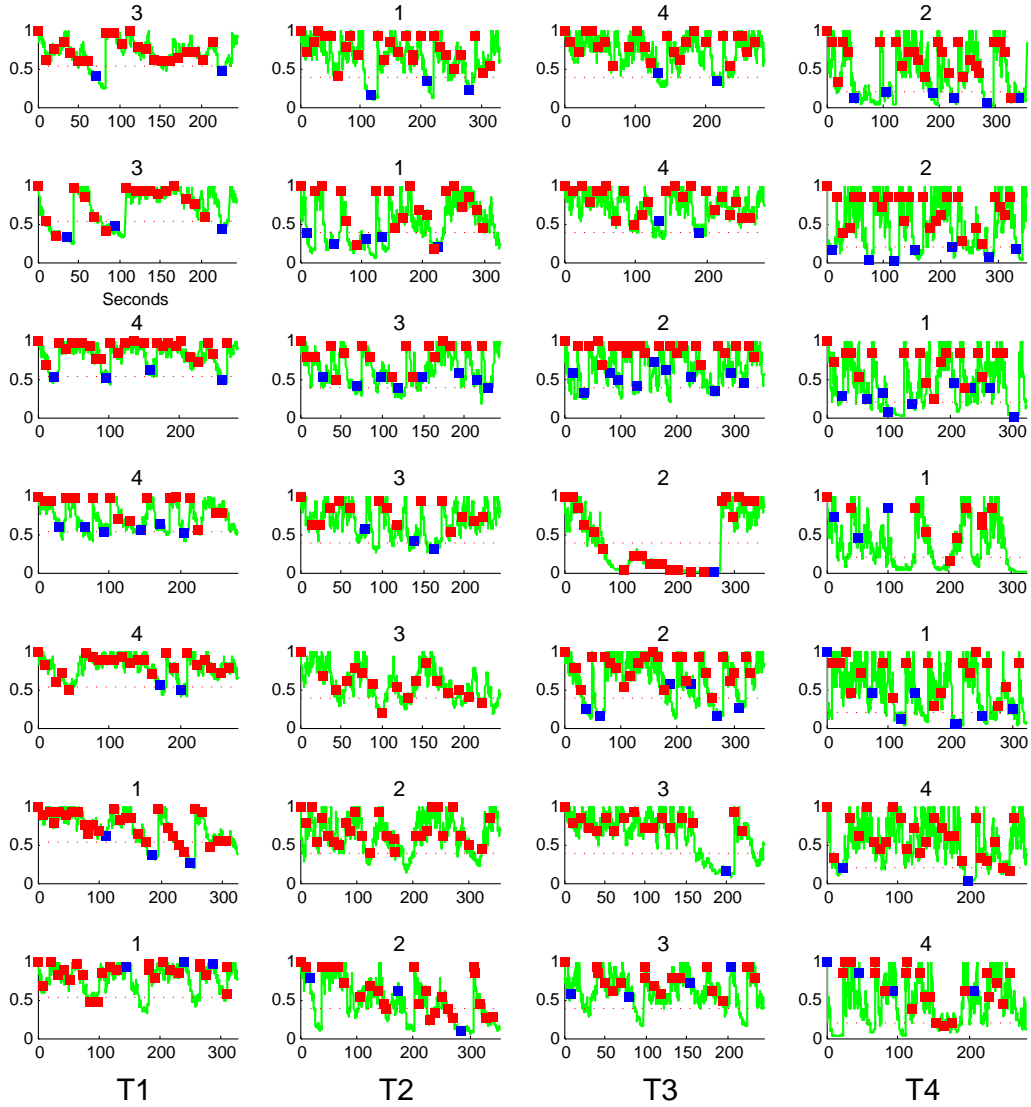


Figure S14 (cont.): Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

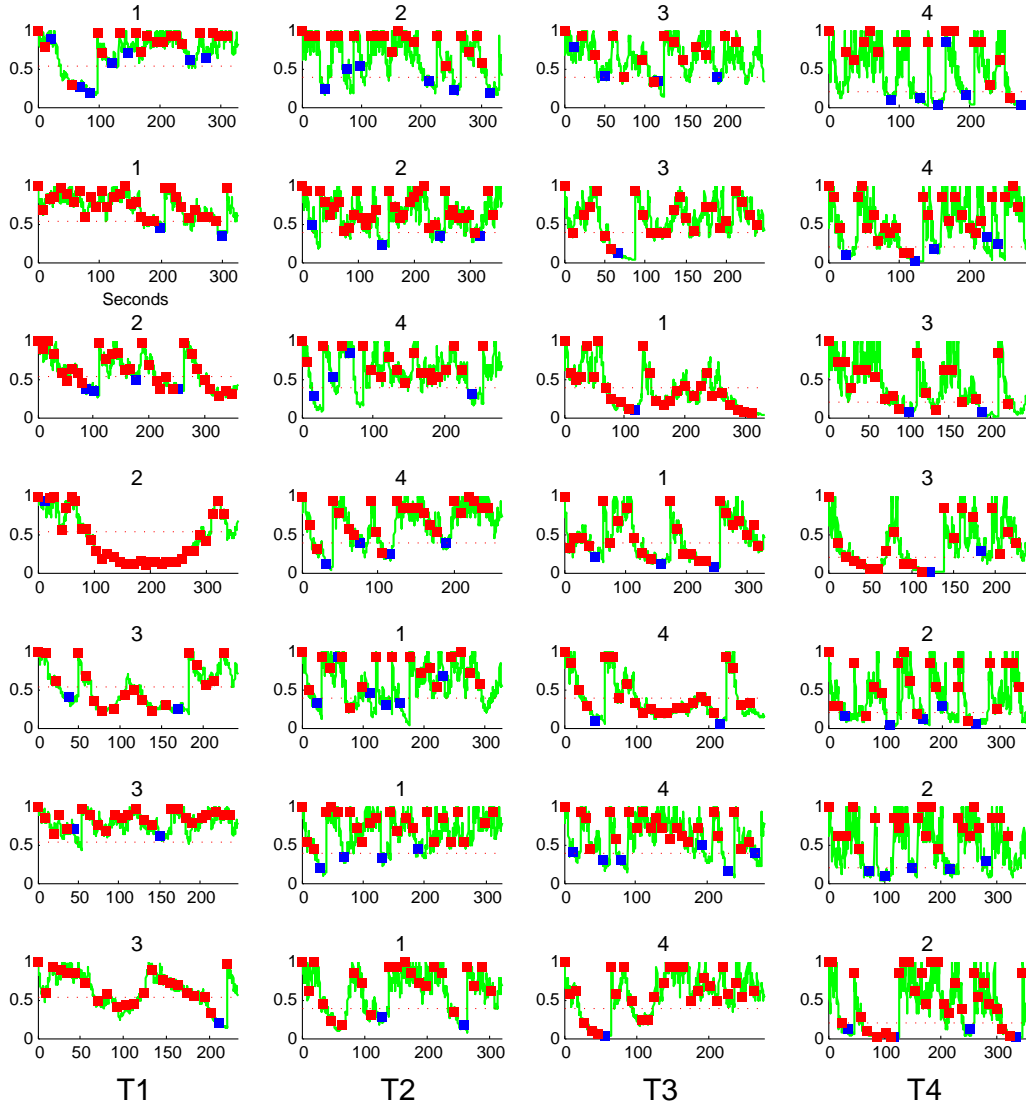


Figure S14 (cont.): Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

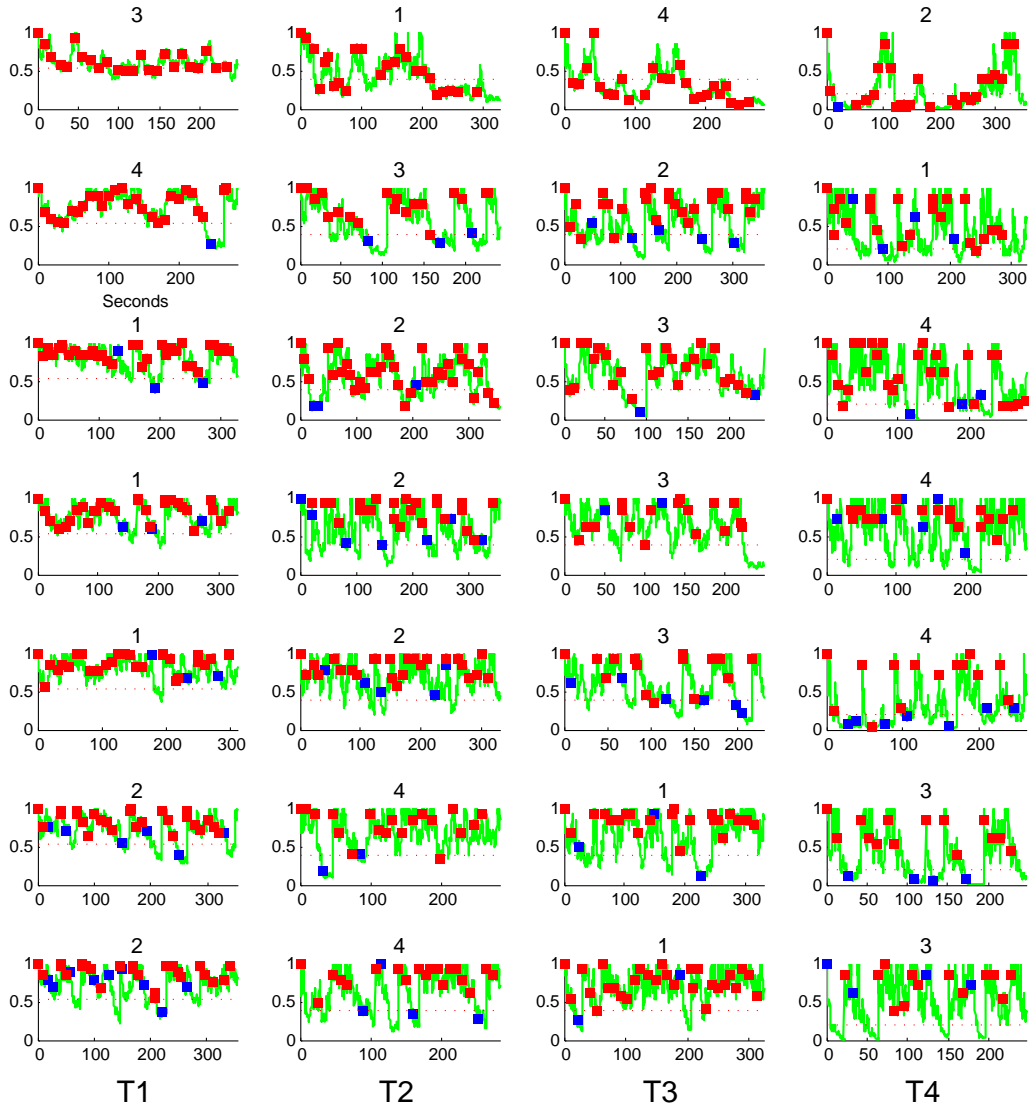


Figure S14 (cont.): Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

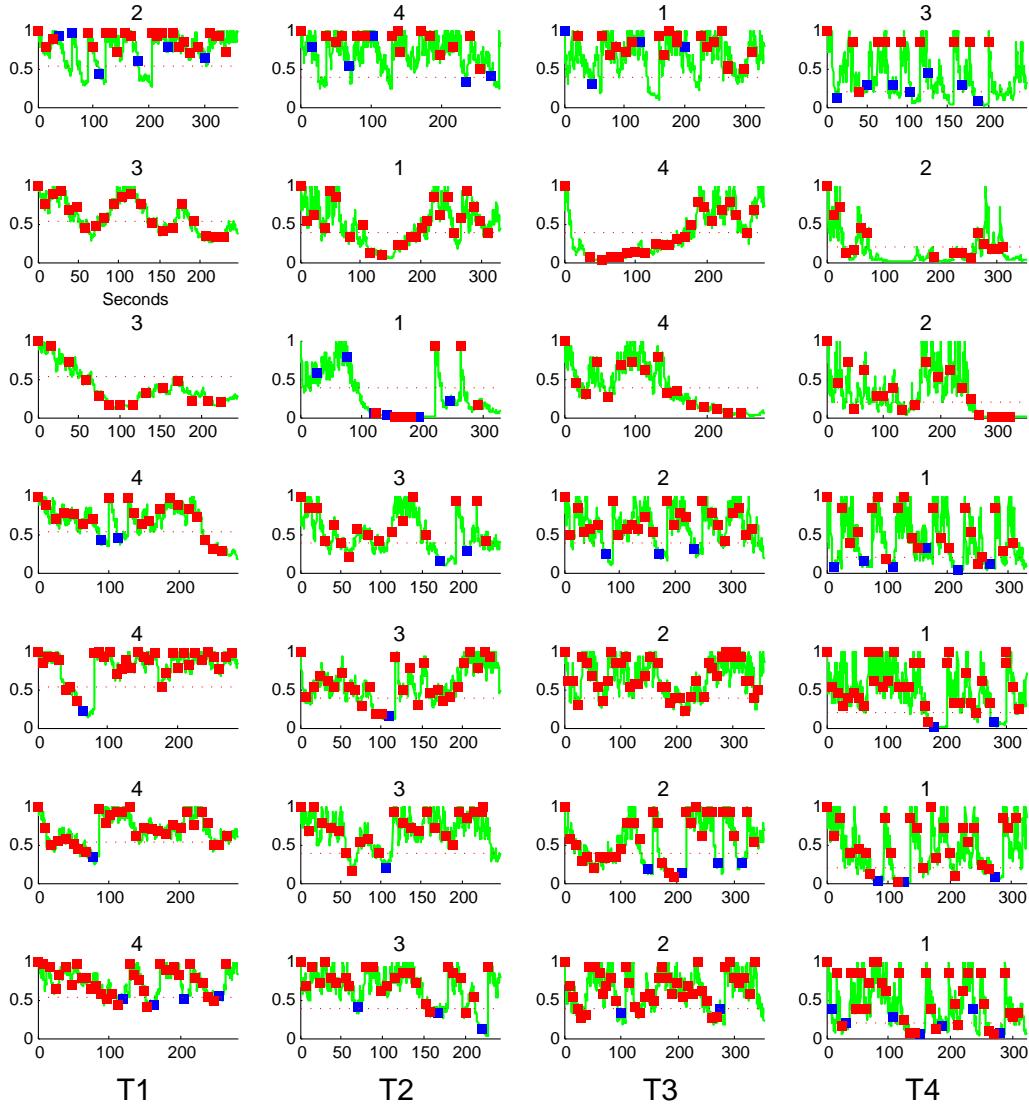


Figure S14 (cont.): Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

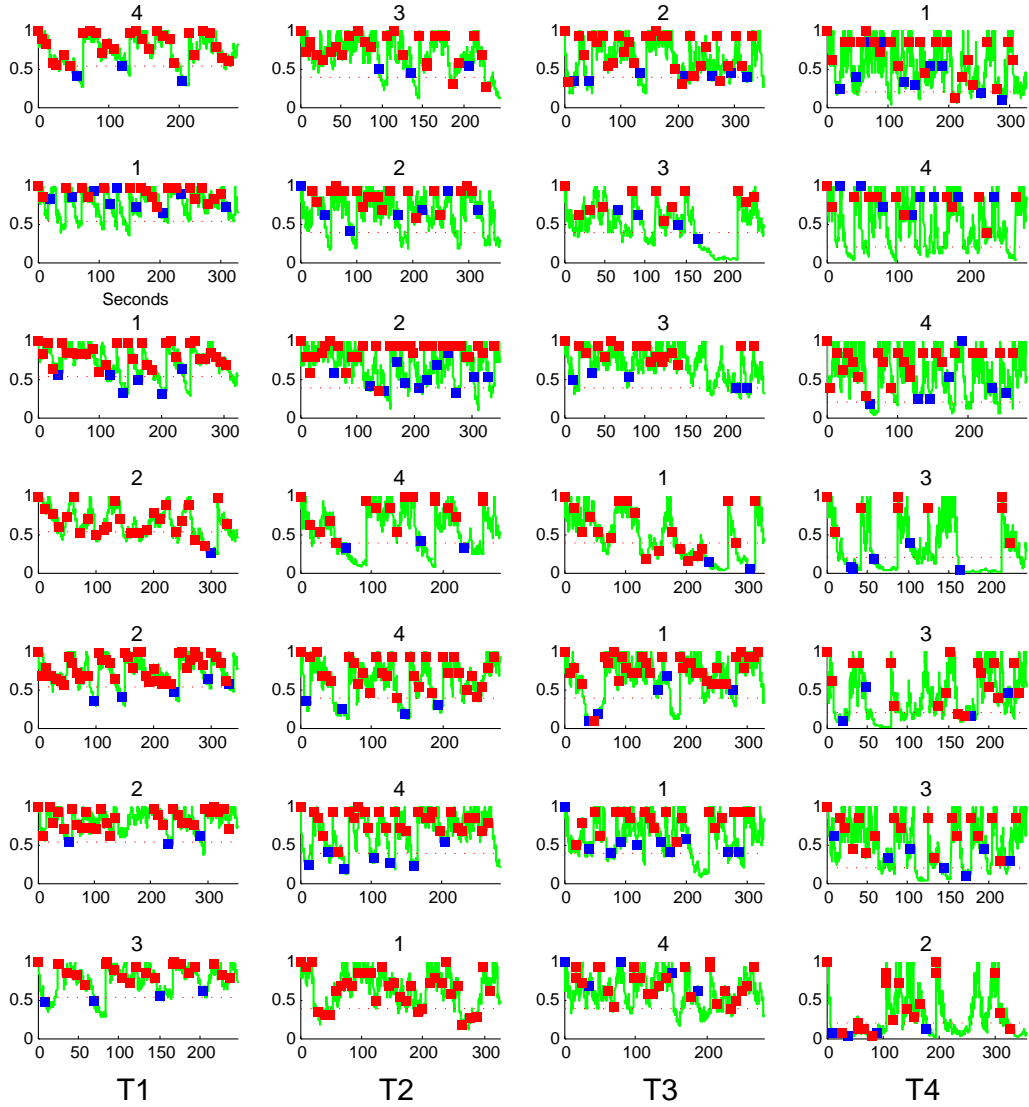


Figure S14 (cont.): Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

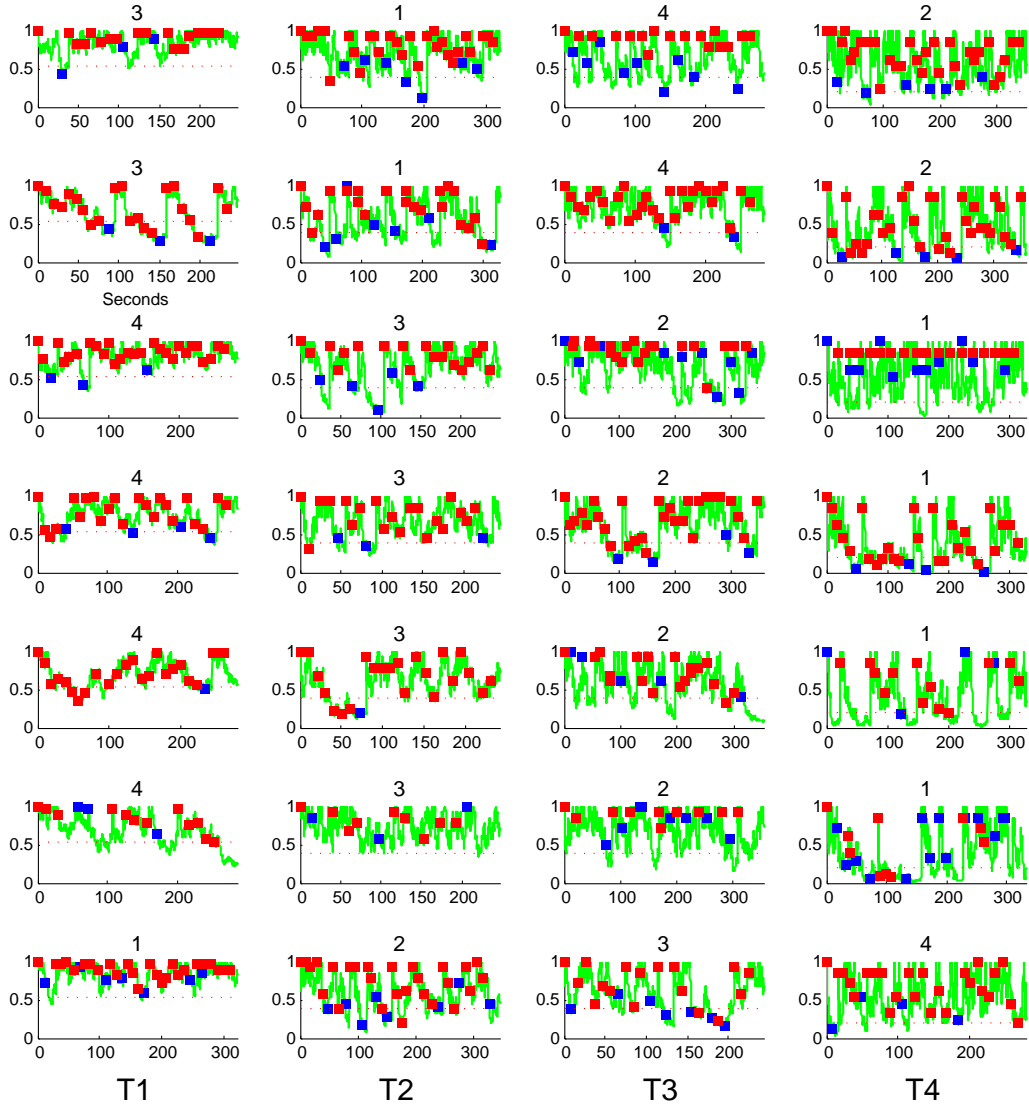


Figure S14 (cont.): Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

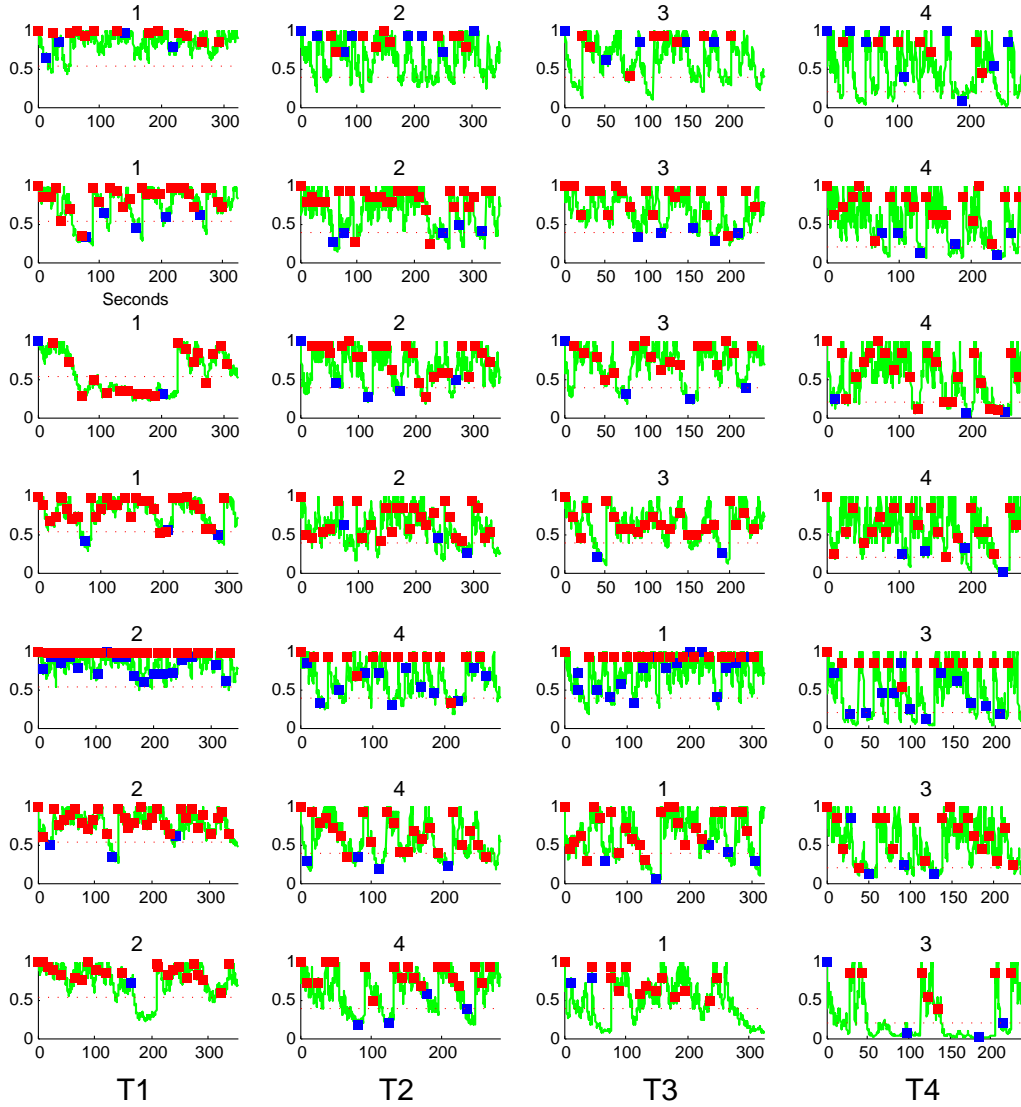


Figure S14 (cont.): Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

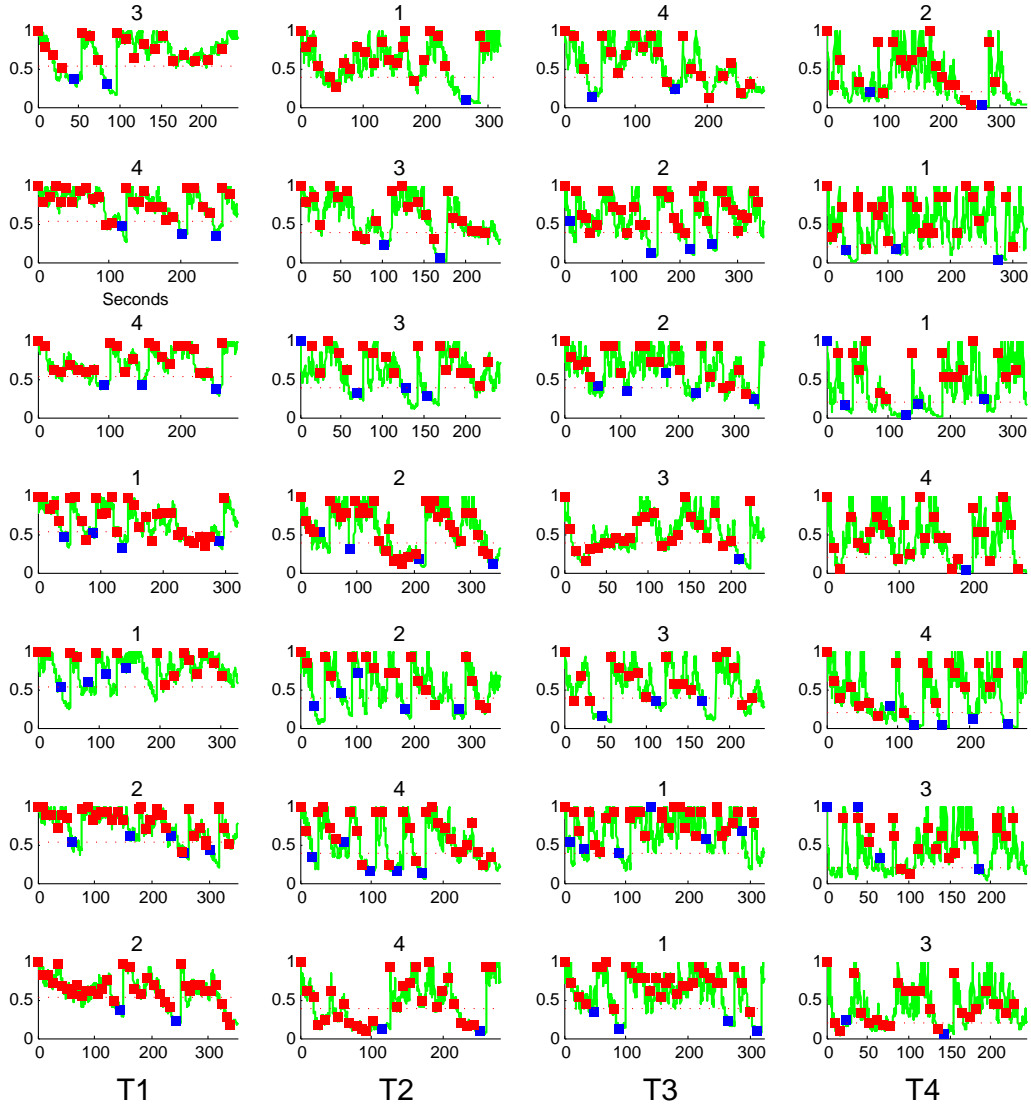


Figure S14 (cont.): Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

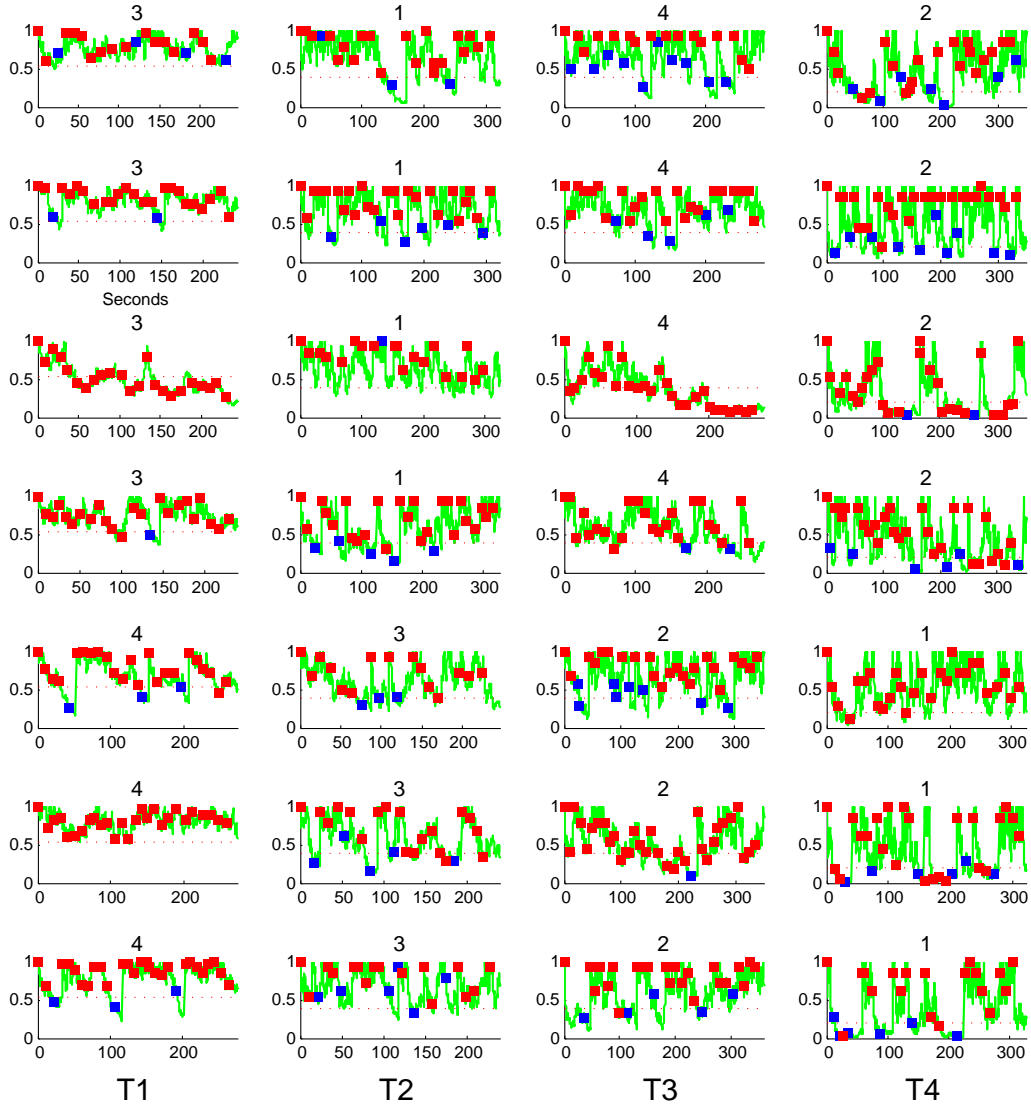


Figure S14 (cont.): Plots showing the updating and baking history for each of the 70 subjects for all four treatments.

C Assessment of Thresholds Estimation

In this section we show that the empirical strategy adopted in the paper to infer the thresholds used by the subjects in the experiment is able to estimate the thresholds with high accuracy. First, we report in Figures S15–S17 the simulation-based plots equivalent to Figures 9–11 in the paper for an aggregate set of 10,000 simulations (virtual simulated subjects).

The estimated thresholds are [.52 .37 .37 .18], while the optimal thresholds were [.52 .38 .38 .20]. The estimation strategy correctly recovers the threshold for Treatment I, and it is fairly close to the correct values for Treatment II and III; however, the precision slightly decreases for Treatment IV. The principle on which the strategy builds is illustrated by the two Figures S15 and S16. The empirical distribution of baking decisions in Figures S15 is similar to a uniform distribution between the threshold profit and the upper bound of the support (profit equal to 1): baking can occur at any of those profits. The distribution of the updating decisions by optimal subjects are made as soon as profits below the threshold are observed; the density of this distribution spikes at the threshold and then declines relatively quickly as less and less updating episodes occur for lower values of profits. On the contrary, the baking density goes to zero just after the threshold. The threshold in theory corresponds to the point where the peak of the updating distribution on the one hand and the zero lower bound of the baking distribution on the other hand are achieved. In terms of the CDF's in Figure S16, the switch in the decision is identified by the largest distance between the two cumulative distributions, where the survivor function of baking leaves the upper bound of 1 and the survivor function of updating switches from positive to zero. In practical terms, the non-parametric kernel estimation (which embeds some degree of smoothness in it) can still attribute some small positive probability to values of profits below the threshold when some observations of baking occur just before the threshold. This is more likely to be the case in the higher treatments, where the profit diffusion is steeper and deeper (see Figure S13 in this Appendix), which can push the estimate of the threshold somewhat to the left of the theoretical value. This effects would be harder to precisely quantify with the experimental data, but it seems small and it would not significantly affect the main interpretation of the paper results; so, we decided to not apply any further correction to the our estimation strategy.

Figures S18–S21 are additional simulation-based plots that replicate the output of Figure 12, and Figures A3–A4 in the paper. The threshold estimation analysis was repeated at sub-

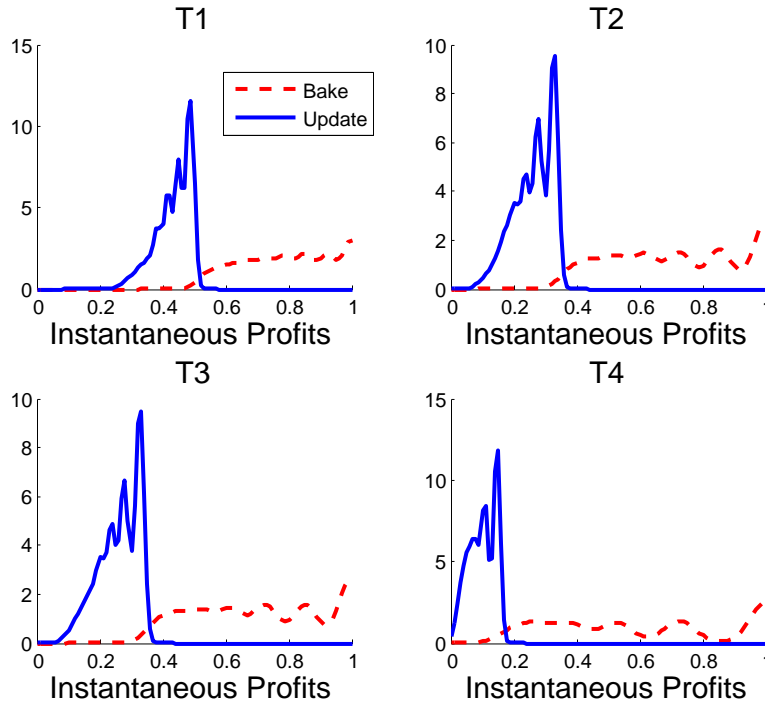


Figure S15: Kernel distribution of the decisions of price updating and baking tasks. Gaussian kernel function for the aggregate set of simulations.

ject level, for each individual virtual subject from the optimal simulation. More specifically, Figures S18–S20 refer to the seven subjects identified by Figure A2 in Appendix D of the paper. These three Figures show the applicability of the estimation procedure at individual level as well. Clearly, the smaller amount of observations available for each subject might affect the precision of the estimation in some specific cases, but overall the estimation strategy reveals to be very robust and effective. Figure S21 reiterate on this point by showing that the distributions of individual thresholds are suitable to identify the theoretical thresholds and are quite narrowly centered around the target values (represented by the vertical dashed lines).

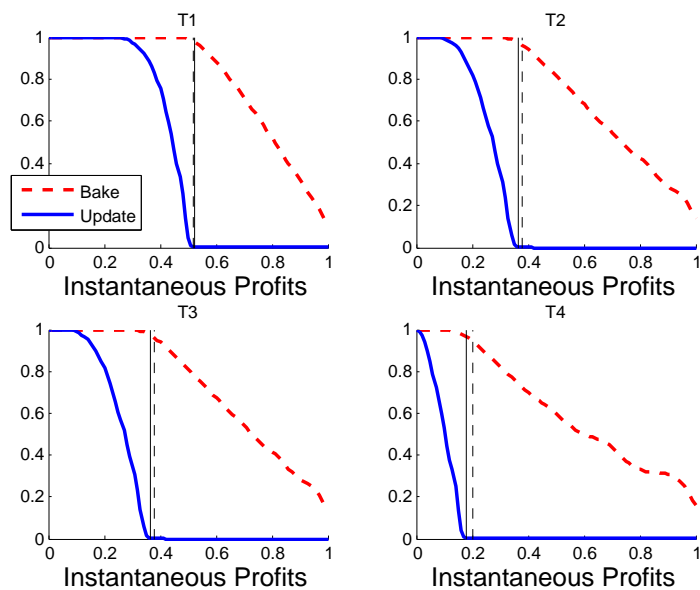


Figure S16: Survivor functions (1-CDF) of the two types of decisions. The dashed vertical lines correspond to the optimal thresholds; the solid vertical lines are the estimated thresholds implied by our empirical strategy. Aggregate set of simulations.

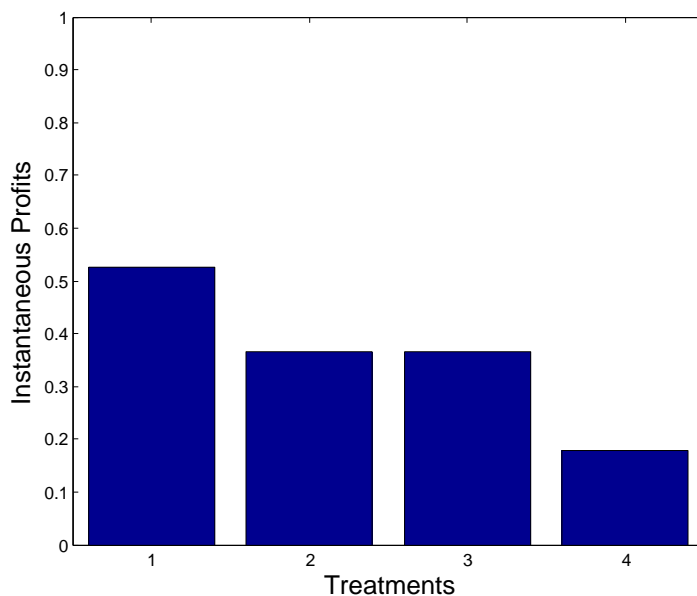


Figure S17: Estimated thresholds in the four treatments; aggregate set of simulations.

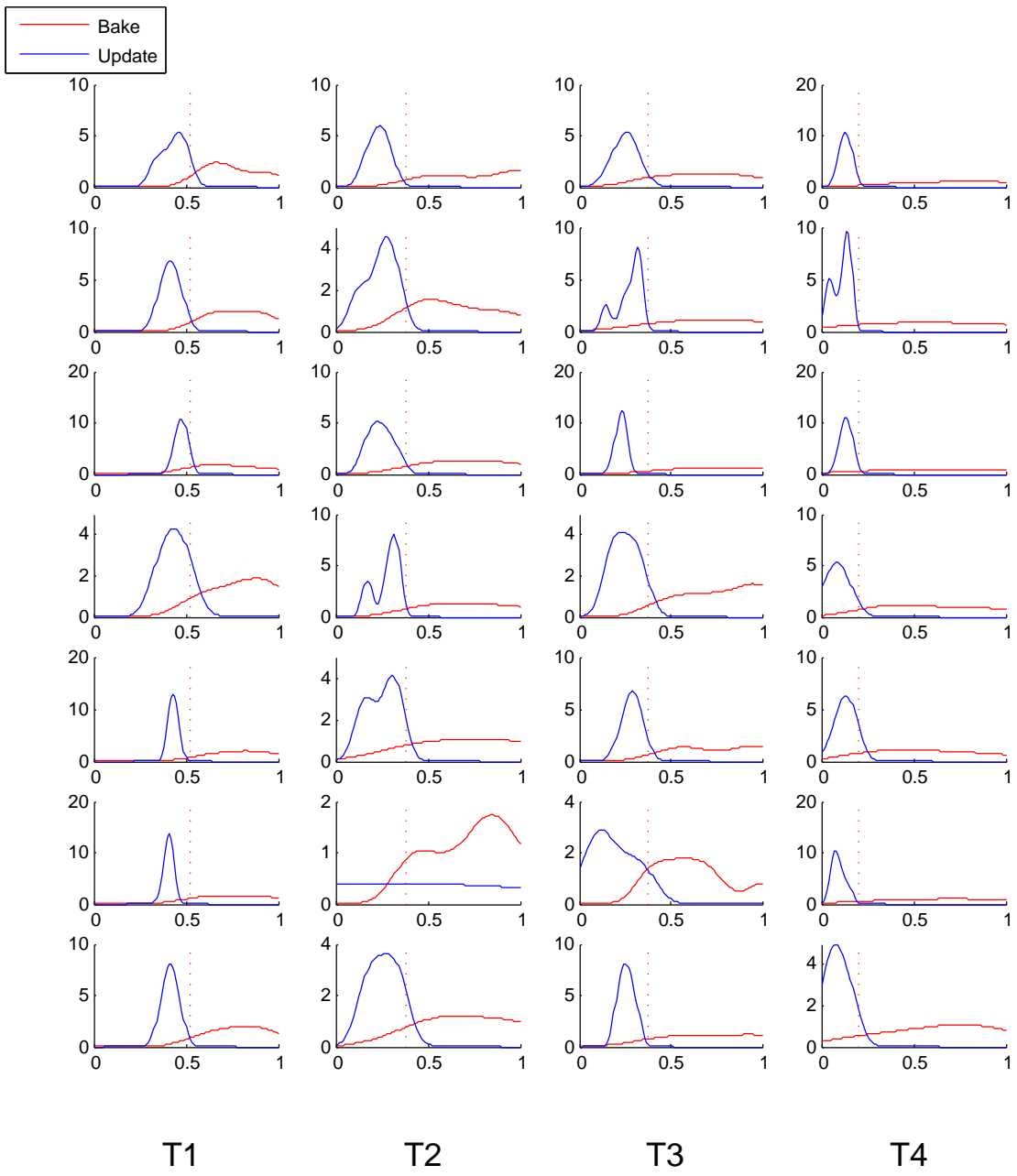


Figure S18: Kernel distribution of the decisions of price updating and baking tasks for a sample of virtual subjects from optimal simulation. Gaussian kernel function for the aggregate set of simulations.

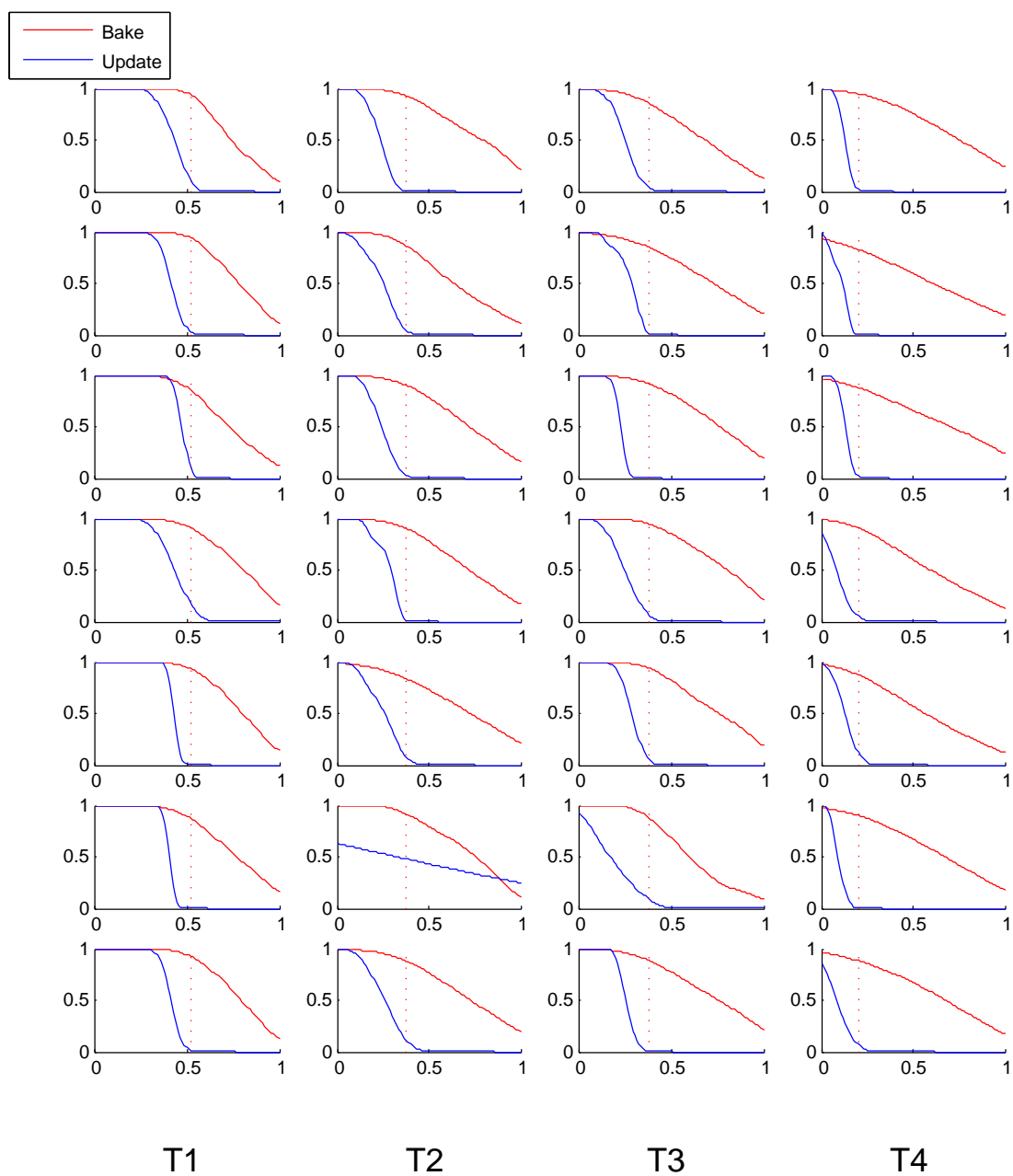


Figure S19: Survivor functions (1-CDF) of the two types of decisions for a sample of virtual subjects from optimal simulation. Instantaneous profits on the vertical axis. Optimal thresholds represented by the dashed vertical lines.

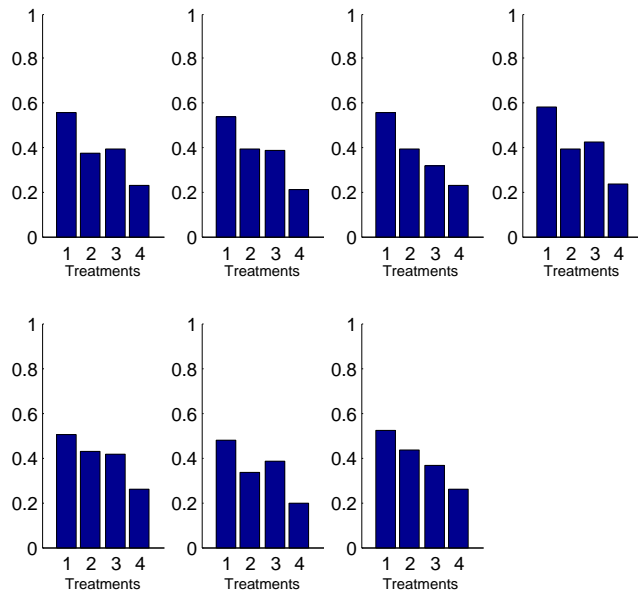


Figure S20: Estimated thresholds in the four treatments for a sample of virtual subjects from optimal simulation. Each panel represents a subject; bars correspond to the four treatments. Instantaneous profits are on the vertical axis and treatments on the horizontal one.

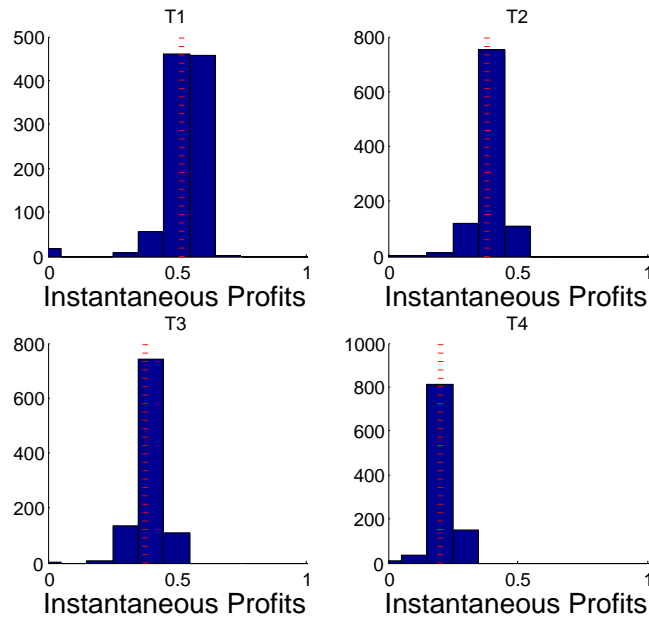


Figure S21: Dispersion of the individual estimated thresholds by treatment – simulation based.