Tests on the aggregation of data[1]

In this appendix we report the test results on session effects, order effects, learning effects and subject pool (economics students versus other subjects) effects. These tests reveal that we can pool the data from sessions that use exactly the same order of treatments and also of those that only differ in the order of the C-levels. However, we do find some significant learning effects. Economics students do not behave differently from other subjects.

Comparing similar sessions For both the unobservable and the observable investment case we have two sessions which use the same (upward) order. With only two group observations per session, no meaningful comparisons can be made at the matching group level. This holds because the smallest significance level that a two-tailed ranksum test can attain equals $\frac{1}{3}$. We therefore only look at tests performed at the level of individual means. Table 1 reports the results. None of the tests reaches significance at the 5%-level. The lowest p-level equals .0725, all others are well above .1. We thus conclude that we can pool the data from sessions that employ the same order of treatments.

Order effects Apart from two sessions with the 'upward' order, we have for each observability case one session with the 'downward' order. For each level of investment costs we compare, across orders, the blocks in which these costs are used for the first time (rounds 1-18), and the blocks in which they are used for the second time (rounds 19-36). Again, with only two group observations for the 'downward' order and four for the 'upward' order, no meaningful comparisons can be made using group level data. We therefore confine ourselves to comparing individual means.

For the unobservable investment case we find only one significant difference, viz. when subjects are confronted with $C = 40$ for the second time (cf. Table 2). Here investment rates are significantly lower in the 'upward' order. Yet for demand behavior no significant differences are found. We conclude that under unobservable investment order effects are almost absent.

When the investment is observable we find five (out of 18) significant differences. Investment rates are significantly lower in the 'upward' order when subjects are confronted with $C = 60$ for the second time. The remaining four significant differences concern demand behavior. In all these cases

[1]This web-appendix belongs to the paper: "Does making specific investments unobservable boost investment incentives?", by Randolph Sloof, Hessel Oosterbeek and Joep Sonnemans.

Table 1: Test results for equality of similar sessions

	first: rounds 1-18			second: rounds 19-36		
	$C = 20$	$C = 40$	$C = 60$	$C = 20$	$C = 40$	$C = 60$
Unobservable						
investment	.3113	.4645	.4815	.2133	.9090	.5897
demand	.4622	.1151	.2922	.7014	.3936	.5891
Observable						
investment	.6039	.7789	.1794	.1961	.0725	.3173
demand: $I = 0$	.8454 (19)	.6388 (33)	.5405	.4793 (35)	.9537	.3173
demand: $I = 1$	.6829	.2005 (35)	.7920 (9)	.4641 (37)	.2419 (14)	n.a.

Remark: In each cell the p-value is given of a ranksum test using individual means. Within parentheses appear the number of observations ($m = n$) when this number deviates from $m = n = 40$.

Table 2: Test results for equality of different orders

	first: rounds 1-18			second: rounds 19-36		
	$C = 20$	$C = 40$	$C = 60$	$C = 20$	$C = 40$	$C = 60$
Unobservable						
investment	1.000	.6261	.8809	.3723	.0318	.2986
demand	.7819	.2933	.0893	.5350	.6549	.0855
Observable						
investment	.3053	.5515	.0894	.1412	.5793	.0073
demand: $I = 0$	.1128 (34)	.0810 (51)	.0023 (59)	.1273 (50)	.5570	.0185
demand: $I = 1$	.9685	.0227 (51)	.8552 (18)	.0220 (57)	.1034 (25)	.1857 (12)

Remark: In each cell the p-value is given of a ranksum test using individual means. Within parentheses appear the number of observations ($m = n = 60$) when this number deviates from $m = n = 60$.

sellers on average demand significantly more under the 'upward' order. Two out of these four cases concern demands after $I = 0$. Although significant, here differences between mean demands are very small in magnitude. Sellers namely almost always demand the complete pie ($P = 50$).

Overall we conclude that some order effects can be detected in the data, especially in the observable investment case. But they are typically only minor. This provides sufficient justification for pooling the results from the two different orders.

Learning effects In both the 'upward' and the 'downward' order every cost level was represented in one block of six rounds in the first 18 rounds and in a second block in the last 18 rounds. To test for learning effects we compare the first block with the second block by means of signrank tests, for all three $C-$levels separately. For the unobservable investment case only three (out of 12) significant differences are found (cf. Table 3). When $C = 20$ both the individual and the group level data indicate that demands significantly increase over time. Mean individual investment levels suggest that for $C = 40$ investment rates decrease over time, yet no significant differences are found when we look at the group level data.

In the observable investment case we observe nine (out of 17) significant differences. There is especially strong evidence that investment rates are decreasing over time. Moreover, mean demands tend to increase over time, although differences are not always significant. Taken together these results imply that, especially when the investment is observable, behavior evolves over time. In the main text we therefore consider the first 18 rounds separately from the last 18 rounds.

Economics students versus other subjects Economics students tend to invest less than other subjects. This is true for all cost levels, both halves of the experiments and for observable and unobservable investments. None of the differences are, however, significant at the 5%-level. Moreover, all the comparative static results with regard to investment levels reported in the main text continue to hold for both types of subjects. Table 4 reports the mean investment levels for the two types of subjects by first/second half of the experiment, cost level and information condition. It also gives the p-values for differences by subject type based on ranksum tests.

Demands are also very similar between subjects studying economics and other subjects. For each information condition, Table 5 reports the mean demands for the two types of subjects by first/second half of the experiment, cost level and information condition. This table also give the p-values for

Table 3: Test results on comparing first half with second half

	$C = 20$	$C = 40$	$C = 60$
Unobservable			
investment	.1118	.0148	.3357
	.2008	.1148	.7489
demand	.0032	.0604	.6261
	.0464	.3454	.9165
Observable			
investment	.0001	.0000	.0947
	.0350	.0350	.0509
demand: $I = 0$	.6315 (29)	.4035 (51)	.0020 (59)
	.8292	.6002	.0350
demand: $I = 1$	.0000 (57)	.1058 (23)	.2914 (9)
	.0277	.0431 (5)	n.a.

Remark: In each cell the upper (lower) p-value is given of a signrank test using individual (group) means. Within parentheses appear the number of matched pairs when this number deviates from 60 for individual data and from 6 for group level data.

Table 4: Mean investment levels by type of subject and tests for equality

	first: rounds 1-18			second: rounds 19-36		
	Econ=0	Econ=1	$p-$values	Econ=0	Econ=1	$p-$values
Obser.:						
$C = 20$	.803	.733	.4935	.606	.552	.7437
$C = 40$	.530	.390	.1963	.227	.171	.5893
$C = 60$	.167	.105	.5593	.106	.067	.5482
Unobs.:						
$C = 20$	.797	.657	.0921	.725	.598	.1154
$C = 40$	.667	.480	.0677	.565	.392	.0779
$C = 60$	.377	.304	.5564	.362	.412	.5306

Remark: $p-$values correspond to a Mann-Whitney ranksum test comparing economics students (Econ=1) with other subjects (Econ=0), based on individual mean investment rates.

differences by subject type based on ranksum tests. The only significant difference occurs in the first half for the observable treatment when $C = 20$ and $I = 0$. The magnitude of this difference is, however, small.

Table 5: Mean demands by type of subject and tests for equality

	first: rounds 1-18			second: rounds 19-36		
	Econ=0	Econ=1	$p-$values	Econ=0	Econ=1	$p-$values
Obser., $I = 0$:						
$C = 20$	46.88	50.00	.0210	48.56	49.91	.1135
$C = 40$	48.82	49.34	.7122	48.33	49.45	.4287
$C = 60$	48.87	48.10	.3791	49.95	49.76	1.000
Obser., $I = 1$:						
$C = 20$	101.29	101.14	.9737	114.48	116.32	.7425
$C = 40$	109.25	104.95	.6680	100.00	108.03	.5873
$C = 60$	77.22	104.86	.2388	93.75	113.13	.3619
Unobs.:						
$C = 20$	90.77	95.32	.5820	102.36	101.47	.9395
$C = 40$	81.84	75.91	.4571	80.43	88.00	.3759
$C = 60$	68.72	66.40	.3112	68.13	67.91	.6769

Remark: $p-$values correspond to a Mann-Whitney ranksum test comparing economics students (Econ=1) with other subjects (Econ=0), based on individual mean demands.