Crowding-Out Appropriateness Elicitation Experiment: Hypoteses and Tests

Wednesday 17th August, 2022

Contents

1 Hypotheses Tests: Lost wallet situation **4**
1.1 Results . . . 4

2 Hypothesis Tests: Day Care Center situation **8**
2.1 Results . . . 9

3 Hypotheses and Tests: Gender Differences **14**

4 Order Effects **15**
4.1 Social appropriateness . . . 16
4.2 Personal appropriateness . . . 18

5 Figures: distribution of ratings **21**
5.1 Lost wallet situation . . . 21
5.2 Day-care situation . . . 27

List of Tables

1 Inappropriateness of not returning (5 scale) . . . 4
2 Inappropriateness of not returning (7 scale) . . . 4
3 Comparing the inappropiateness of not returning when wallet contains small money or no money (5 scale - t-tests) . . . 5
4 Comparing the inappropiateness of not returning when wallet contains small money or no money (5 scale - rank sum test) . . . 5
5 Inappropiateness of not returning increasing in monetary content (5 scale - t-tests) . . . 5
6 Inappropiateness of not returning increasing in monetary content (5 scale - rank sum test) . . . 6
7 Comparing inappropriateness of not returning the wallet when the wallet contains small money or no money (7 scale - t-tests) . . . 6
8 Comparing inappropriateness of not returning the wallet when the wallet contains small money or no money (7 scale - rank sum test) . . . 6
9 Inappropriateness of not returning increasing in monetary content (7 scale - t-tests) . . . 7
10 Inappropriateness of not returning increasing in monetary content (7 scale - rank sum test) . . . 7
11 Kolmogorov-Smirnoff test results on the pairwise differences (5 scale) . . . 7
12 Kolmogorov-Smirnoff test results on the pairwise differences (7 scale) . . . 8
13 Inppropriateness of arriving late (5scale) . . . 9
14 Inppropriateness of arriving late (7scale) . . . 9
15 Comparing inappropriateness of arriving late under (small or big) fine or no fine (5 scale - t-tests) . . . 9
16 Comparing inappropriateness of arriving late under (small or big) fine or no fine (5 scale - rank sum test) . . . 10
17 Comparing inappropriateness of arriving late under small and big fine - 15 mins of delay (5 scale - t-tests) . . . 10
18 Comparing inappropriateness of arriving late under small and big fine - 15 mins of delay (5 scale - rank sum test) . . . 10
19 Comparing inappropriateness of arriving late under (small or big) fine or no fine - 15 mins of delay (7 scale - t-tests) . . . 11

20 Comparing inappropriateness of arriving late under (small or big) fine or no fine - 15 mins of delay (7 scale - rank sum test) . . . 11
21 Comparing inappropriateness of arriving late under small and big fine (7 scale - t-tests) . . . 12
22 Comparing inappropriateness of arriving late under (small or big) fine or no fine - 15 mins of delay (7 scale - rank sum test) . . . 12
23 Kolmogorov-Smirnoff test results on the pairwise differences - 30 mins of delay (5 scale) . . . 12
24 Kolmogorov-Smirnoff test results on the pairwise differences - 30 mins of delay (7 scale) . . . 13
25 Kolmogorov-Smirnoff test results on the pairwise differences (5 scale) . . . 13
26 Kolmogorov-Smirnoff test results on the pairwise differences (7 scale) . . . 13
27 Gender differences in social appropriateness (5 scale - t-tests) . . . 14
28 Gender differences in social appropriateness (5 scale - rank sum test) . . . 14
29 Gender difference in social appropriateness (7 scale - t-tests) . . . 14
30 Gender difference in social appropriateness (7 scale - rank sum test) . . . 14
31 Gender difference in personal appropriateness (5 scale - t-tests) . . . 15
32 Gender difference in personal appropriateness (5 scale - rank sum test) . . . 15
33 Gender differences in personal appropriateness (7 scale - t-tests) . . . 15
34 Gender differences in personal appropriateness (7 scale - rank sum test) . . . 15
35 Test at level of questions/actions: lost wallet situation (5 scale - t-tests) . . . 16
36 Test at level of questions/actions: lost wallet situation (5 scale - rank sum test) . . . 16
37 Test at level of questions/actions: lost wallet situation (7 scale - t-tests) . . . 16
38 Test at level of questions/actions: lost wallet situation (7 scale - rank sum test) . . . 17
39 Test at level of questions/actions: day care situation (5 scale - t-tests) . . . 17
40 Test at level of questions/actions: day care situation (5 scale - rank sum test) . . . 17
41 Test at level of questions/actions: day care situation (7 scale - t-tests) . . . 18
42 Test at level of questions/actions: day care situation (7 scale - rank sum test) . . . 18
43 Test at level of questions/actions: lost wallet situation (5 scale - t-tests) . . . 18
44 Test at level of questions/actions: lost wallet situation (5 scale - rank sum test) . . . 19
45 Test at level of questions/actions: lost wallet situation (7 scale - t-tests) . . . 19
46 Test at level of questions/actions: lost wallet situation (7 scale - rank sum test) . . . 19
47 Test at level of questions/actions: day care situation (5 scale - t-tests) . . . 20
48 Test at level of questions/actions: day care situation (5 scale - rank sum test) . . . 20
49 Test at level of questions/actions: day care situation (7 scale - t-tests) . . . 20
50 Test at level of questions/actions: day care situation (7 scale - rank sum test) . . . 21

List of Figures

1 Social appropriateness: Not returning the wallet (5 scale) . . . 21
2 Social appropriateness: Not returning the wallet (7 scale) . . . 22
3 Social appropriateness: Cumulative distribution of ratings for returning the wallet (5 scale) . . . 22
4 Social appropriateness: Cumulative distribution of ratings for returning the wallet (7 scale) . . . 23
5 Personal appropriateness: Not returning the wallet (5 scale) . . . 23
6 Personal appropriateness: Not returning the wallet (7 scale) . . . 24
7 Personal appropriateness: Cumulative distribution of ratings for returning the wallet (5 scale) . . . 24
8 Personal appropriateness: Cumulative distribution of ratings for returning the wallet (7 scale) . . . 25
9 Social appropriateness: average ratings (5 scale) . . . 25
10 Social appropriateness: average ratings (7 scale) . . . 26
11 Personal appropriateness: average ratings (5 scale) . . . 26
12 Personal appropriateness: average ratings (7 scale) . . . 27
13 Social appropriateness: Arriving 15 mins late (5 scale) . . . 27
14 Social appropriateness: Arriving 15 mins late (7 scale) . . . 28
15 Personal appropriateness: Arriving 15 mins late (5 scale) . . . 28
16 Personal appropriateness: Arriving 15 mins late (7 scale) . . . 29
17 Social appropriateness: Arriving 30 mins late (5 scale) . . . 29
18 Social appropriateness: Arriving 30 mins late (7 scale) . . . 30

19 Social appropriateness: Arriving 30 mins late (5 scale) 30
20 Social appropriateness: Arriving 30 mins late (7 scale) 31
21 Social appropriateness: cumulative distribution of ratings for arriving 15 mins late (5 scale) 31
22 Social appropriateness: cumulative distribution of ratings for arriving 15 mins late (7 scale) 32
23 Personal appropriateness: cumulative distribution of ratings for arriving 15 mins late (5 scale) 32
24 Personal appropriateness: cumulative distribution of ratings for arriving 15 mins late (7 scale) 33
25 Social appropriateness: cumulative distribution of ratings for arriving 30 mins late (5 scale) 33
26 Social appropriateness: cumulative distribution of ratings for arriving 30 mins late (7 scale) 34
27 Personal appropriateness: cumulative distribution of ratings for arriving 30 mins late (5 scale) 34
28 Personal appropriateness: cumulative distribution of ratings for arriving 15 mins late (7 scale) 35
29 Social appropriateness: average ratings for arriving 15 mins late (5 scale) 35
30 Social appropriateness: average ratings for arriving 15 mins late (7 scale) 36
31 Personal appropriateness: average ratings for arriving 15 mins late (5 scale) 36
32 Personal appropriateness: average ratings for arriving 15 mins late (7 scale) 37
33 Social appropriateness: average ratings for arriving 30 mins late (5 scale) 37
34 Social appropriateness: average ratings for arriving 30 mins late (7 scale) 38
35 Personal appropriateness: average ratings for arriving 30 mins late (5 scale) 38
36 Personal appropriateness: average ratings for arriving 30 mins late (7 scale) 39

1 Hypotheses Tests: Lost wallet situation

Scenario: Individual A works in the reception of a large public office. One day a person comes up to the reception desk, hands over a wallet to Individual A and says: "Hi, I found this on the street around the corner. Somebody must have lost it. I'm in a hurry and must go. Can you please take care of it?" The person who handed in the wallet then leaves. Individual A inspect the wallet and finds that it contains three identical business cards, a grocery list, and a key. Moreover:

- the wallet contains no money, or
- the wallet contains 13.45 $, or
- the wallet contains 94,15 $.

For each of these there variants of the scenario, Individual A must decide whether to:

- notify the owner of the waller; or
- take the wallet, including the money contained in it.

Half of the subjects in the experiment presented with the lost wallet situations has to indicate to what degree the possible behaviors presented above are social appropriate or socially inappropriate on a 5 or 7 point scale. The social appropriate behavior is defined as a behavior that most people agree is the correct/proper or ethical; in other words, if a person were to select a socially inappropriate choice, then someone else might be angry at the person for doing so. The other half are instead asked to indicate the degree to which they each possible behaviour is appropriate or inappropriate based on their personal opinion independently of the opinion of other using the same 5 or 7 point scale.

1.1 Results

Table 1: Inappropriateness of not returning (5 scale)

	Obs	Proportion	Std. Error	P-value (H0: Pr=0.5, H1: Pr>0.5)
Social appropriateness	200	.975	.011	0
Personal appropriateness	194	.9948	.0051	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 2: Inappropriateness of not returning (7 scale)

	Obs	Proportion	Std. Error	P-value (H0: Pr=0.5, H1: Pr>0.5)
Social appropriateness	201	1	0	0
Personal appropriateness	200	.995	.005	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 3: Comparing the inappropiateness of not returning when wallet contains small money or no money (5 scale - t-tests)

	Mean (1)	Mean (2)	Diff.	P-value (H0: diff=0, H1:diff<0)
Social appropriateness				
Small money - No money	-.9167	-.8965	-.0202	.0789
Personal appropriateness				
Small Money - No money	-.9433	-.9046	-.0387	.008

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 4: Comparing the inappropiateness of not returning when wallet contains small money or no money (5 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Social appropriateness					
Small money - No money	198	198	-1.198	.116	.114
Personal appropriateness					
Small Money - No money	194	194	-2.412	.008	.009

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 5: Inappropiateness of not returning increasing in monetary content (5 scale - t-tests)

	Mean (1)	Mean (2)	Diff.	P-value (H0: diff=0, H1:diff<0)
Social appropriateness				
Big money - Small money	-.9369	-.9167	-.0202	.0853
Personal appropriateness				
Big money - Small Money	-.9253	-.9433	.018	.8623

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 6: Inappropiateness of not returning increasing in monetary content (5 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Social appropriateness					
Big money - Small money	198	198	-1.269	.102	.115
Personal appropriateness					
Big money - Small Money	194	194	.559	.712	.783

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 7: Comparing inappropriateness of not returning the wallet when the wallet contains small money or no money (7 scale - t-tests)

	Mean (1)	Mean (2)	Diff.	P-value (H0: diff=0, H1:diff<0)
Social appropriateness				
Small money - No money	-.8773	-.8275	-.0498	.0129
Personal appropriateness				
Small Money - No money	-.8833	-.8633	-.02	.145

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 8: Comparing inappropriateness of not returning the wallet when the wallet contains small money or no money (7 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Social appropriateness					
Small money - No money	201	201	-2.277	.011	.011
Personal appropriateness					
Small Money - No money	200	200	-1.984	.024	.025

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 9: Inappropriateness of not returning increasing in monetary content (7 scale - t-tests)

	Mean (1)	Mean (2)	Diff.	P-value (H0: diff=0, H1:diff<0)
Social appropriateness				
Big money - Small money	-.9171	-.8773	-.0398	.0086
Personal appropriateness				
Big money - Small Money	-.9033	-.8833	-.02	.0956

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 10: Inappropriateness of not returning increasing in monetary content (7 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Social appropriateness					
Big money - Small money	201	201	-2.938	.002	.002
Personal appropriateness					
Big money - Small Money	200	200	-.961	.168	.201

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 11: Kolmogorov-Smirnoff test results on the pairwise differences (5 scale)

	Diff.	P-value
Social appropriateness		
No money vs small money	.0253	1
No money vs big money	.0657	.7869
Small money vs big money	.0404	.997
Personal appropriateness		
No money vs small money	.0515	.9588
No money vs big money	.0515	.9588
Small vs big money	.0155	1

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 12: Kolmogorov-Smirnoff test results on the pairwise differences (7 scale)

	Diff.	P-value
Social appropriateness		
No money vs small money	.0547	.9242
No money vs big money	.1393	.0405
Small money vs big money	.0846	.4685
Personal appropriateness		
No money vs small money	.045	.9874
No money vs big money	.09	.3927
Small vs big money	.045	.9874

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

2 Hypothesis Tests: Day Care Center situation

Scenario: Individual A has a three-year old child who attends a local day-care center. The center is scheduled to operate every day from half past seven in the morning until four in the afternoon, when the parents are supposed to come and collect their children. If a parent is late and arrives after four, a teacher stay at the center with the concerned children. Additionally:

- There is no fine for being late ; or
- Parents have to pay a fine of $1 for a delay of 10 minutes or more.
- Parents have to pay a fine of $10 for a delay of 10 minutes or more.

In each of these there variants of the scenario, Individual A can:

- Arrive before 4:00 pm (not late);
- Arrive at 4:15 pm (15 min late), and pay a fine of $1 or $10
- Arrive at 4:30 pm (30 min late), and pay a fine of $1 or $10

Some subject in the experiment presented with the day-care scenario has to indicate to what degree the possible behaviors presented above are social appropriate or socially inappropriate on a 5 or 7 point scale. The social appropriate behavior is defined as a behavior that most people agree is the correct/proper or ethical; in other words, if a person were to select a socially inappropriate choice, then someone else might be angry at the person for doing so. Other subjects are instead asked to indicate the degree to which they each possible behaviour is appropriate or inappropriate based on their personal opinion independently of the opinion of other using the same 5 or 7 point scale.

2.1 Results

Table 13: Inppropriateness of arriving late (5scale)

	Obs	Proportion	Std. Error	P-value (H0: Pr=0.5, H1: Pr>0.5)
Social appropriateness	200	.93	.018	0
Personal appropriateness	194	.9175	.0198	0

Notes. P-value equals to 0 in the last column means that p < 0.001.

Table 14: Inppropriateness of arriving late (7scale)

	Obs	Proportion	Std. Error	P-value (H0: Pr=0.5, H1: Pr>0.5)
Social appropriateness	201	.9502	.0153	0
Personal appropriateness	200	.96	.0139	0

Notes. P-value equals to 0 in the last column means that p < 0.001.

Table 15: Comparing inappropriateness of arriving late under (small or big) fine or no fine (5 scale - t-tests)

	Mean (1)	Mean (2)	Diff.	P-valule (H0: diff=0, H1: diff<0)
Social appropriateness				
No fine (15) - Small fine (15)	-.2766	-.0736	-.203	0
No fine (30) - Small fine (30)	-.7183	-.5508	-.1675	0
No fine (15) - Big fine (15)	-.2766	-.0635	-.2132	0
No fine (30) - Big fine (30)	-.7183	-.4365	-.2817	0
Personal appropriateness				
No fine (15) - Small fine (15)	-.3686	-.2036	-.1649	0
No fine (30) - Small fine (30)	-.732	-.5928	-.1392	0
No fine (15) - Big fine (15)	-.3686	-.1495	-.2191	0
No fine (30) - Big fine (30)	-.732	-.4845	-.2474	0

Notes. P-value equals to 0 in the last column means that p < 0.001.

Table 16: Comparing inappropriateness of arriving late under (small or big) fine or no fine (5 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Social appropriateness					
No fine (15) - Small fine (15)	197	197	-5.55	0	0
No fine (30) - Small fine (30)	197	197	-4.575	0	0
No fine (15) - Big fine (15)	197	197	-4.874	0	0
No fine (30) - Big fine (30)	197	197	-6.126	0	0
Personal appropriateness					
No fine (15) - Small fine (15)	194	194	-4.569	0	0
No fine (30) - Small fine (30)	194	194	-4.875	0	0
No fine (15) - Big fine (15)	194	194	-4.904	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 17: Comparing inappropriateness of arriving late under small and big fine - 15 mins of delay (5 scale - t-tests)

	Mean (1)	Mean (2)	Diff.	P-valule (H0: diff=0, H1: diff<0)
Social appropriateness				
Small fine (15) - Big fine (15)	-.0736	-.0635	-.0102	.3778
Small fine (30) - Big fine (30)	-.5508	-.4365	-.1142	.0003
Personal appropriateness				
Small fine (15) - Big fine (15)	-.2036	-.1495	-.0541	.0563
Small fine (30) - Big fine (30)	-.5928	-.4845	-.1082	.0006

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 18: Comparing inappropriateness of arriving late under small and big fine - 15 mins of delay (5 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Social appropriateness					
Small fine (15) - Big fine (15)	197	197	-.663	.254	.255
Small fine (30) - Big fine (30)	197	197	-3.363	0	0
Personal appropriateness					
Small fine (15) - Big fine (15)	194	194	-2.137	.016	.016
Small fine (30) - Big fine (30)	194	194	-3.005	.001	.001

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 19: Comparing inappropriateness of arriving late under (small or big) fine or no fine - 15 mins of delay (7 scale - t-tests)

	Mean (1)	Mean (2)	Diff.	P-valule (H0: diff=0, H1: diff<0)
Social appropriateness				
No fine (15) - Small fine (15)	-.34	-.1924	-.1476	0
No fine (30) - Small fine (30)	-.7214	-.6136	-.1078	0
No fine (15) - Big fine (15)	-.34	-.2206	-.1194	.0004
No fine (30) - Big fine (30)	-.7214	-.5622	-.1592	0
Personal appropriateness				
No fine (15) - Small fine (15)	-.38	-.24	-.14	0
No fine (30) - Small fine (30)	-.7	-.57	-.13	0
No fine (15) - Big fine (15)	-.38	-.175	-.205	0
No fine (30) - Big fine (30)	-.7	-.5267	-.1733	0

Notes.

Table 20: Comparing inappropriateness of arriving late under (small or big) fine or no fine - 15 mins of delay (7 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Social appropriateness					
No fine (15) - Small fine (15)	201	201	-5.429	0	0
No fine (30) - Small fine (30)	201	201	-5.203	0	0
No fine (15) - Big fine (15)	201	201	-3.815	0	0
No fine (30) - Big fine (30)	201	201	-5.318	0	0
Personal appropriateness					
No fine (15) - Small fine (15)	200	200	-4.659	0	0
No fine (30) - Small fine (30)	200	200	-4.55	0	0
No fine (15) - Big fine (15)	200	200	-6.067	0	0

Notes. P-value equals to 0 in the last column means that $p < 0.001$.

Table 21: Comparing inappropriateness of arriving late under small and big fine (7 scale - t-tests)

	Mean (1)	Mean (2)	Diff.	P-valule (H0: diff=0, H1: diff<0)
Social appropriateness				
Small fine (15) - Big fine (15)	-.1924	-.2206	.0282	.8303
Small fine (30) - Big fine (30)	-.6136	-.5622	-.0514	.0279
Personal appropriateness				
Small fine (15) - Big fine (15)	-.24	-.175	-.065	.0083
Small fine (30) - Big fine (30)	-.57	-.5267	-.0433	.0343

Table 22: Comparing inappropriateness of arriving late under (small or big) fine or no fine - 15 mins of delay (7 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Social appropriateness					
Small fine (15) - Big fine (15)	201	201	.994	.84	.84
Small fine (30) - Big fine (30)	201	201	-1.962	.025	.025
Personal appropriateness					
Small fine (15) - Big fine (15)	200	200	-2.874	.002	.002
Small fine (30) - Big fine (30)	200	200	-2.808	.002	.002

Table 23: Kolmogorov-Smirnoff test results on the pairwise differences - 30 mins of delay (5 scale)

	Diff.	P-value
Social appropriateness		
No fee vs small fee	.1777	.004
No fee vs big fee	.1929	.0013
Small vs big fees	.0152	1
Personal appropriateness		
No fee vs small fee	.1392	.0467
No fee vs big fee	.1804	.0036
Small vs big fees	.0722	.6933

Table 24: Kolmogorov-Smirnoff test results on the pairwise differences - 30 mins of delay (7 scale)

	Diff.	P-value
Social appropriateness		
No fee vs small fee	.1144	.1438
No fee vs big fee	.1045	.2226
Small vs big fees	.0498	.9647
Personal appropriateness		
No fee vs small fee	.13	.0681
No fee vs big fee	.17	.0062
Small vs big fees	.07	.7112

Table 25: Kolmogorov-Smirnoff test results on the pairwise differences (5 scale)

	Diff.	P-value
Social appropriateness		
No vs small fee	.1574	.0152
No vs big fee	.198	.0009
Samll vs big fee	.1015	.262
Personal appropriateness		
No fee vs small fee	.1289	.0798
No fee vs big fee	.201	.0008
Small vs big fees	.0722	.6933

Table 26: Kolmogorov-Smirnoff test results on the pairwise differences (7 scale)

	Diff.	P-value
Social appropriateness		
No vs small fee	.0945	.3304
No vs big fee	.1592	.0123
Samll vs big fee	.0647	.7945
Personal appropriateness		
No fee vs small fee	.1	.27
No fee vs big fee	.165	.0086
Small vs big fees	.075	.6272

3 Hypotheses and Tests: Gender Differences

The tables below present tests to the hypotheses that women will be more accurate at guessing the social appropriateness rating of others (H0: diff =0, H1: diff<0), and that it is not the case that women's personal ratings are more likely (than men's personal ratings) to be the same as the most frequent appropriateness rating made by others.

Table 27: Gender differences in social appropriateness (5 scale - t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value
Lost wallet situation	200	0.8711	0.8688	0.0023	0.5251
Day care situation	200	0.4560	0.4681	-0.0121	0.4028

* p<0.05, ** p<0.01, *** p<0.001

Table 28: Gender differences in social appropriateness (5 scale - rank sum test)

	Obs (Men)	Obs (Women)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Lost wallet situation	106	94	-.09	.464	.47
Day care situation	106	94	-.24	.405	.405

Notes. P-value equals to 0 in the last column means that p < 0.001.

Table 29: Gender difference in social appropriateness (7 scale - t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value
Lost wallet situation	201	0.7810	0.7205	0.0605	0.8834
Day care situation	201	0.4036	0.3872	0.0164	0.6778

* p<0.05, ** p<0.01, *** p<0.001

Table 30: Gender difference in social appropriateness (7 scale - rank sum test)

	Obs (Men)	Obs (Women)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Lost wallet situation	102	99	1.455	.927	.929
Day care situation	102	99	.435	.668	.668

Notes. P-value equals to 0 in the last column means that p < 0.001.

Table 31: Gender difference in personal appropriateness (5 scale - t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value
Lost wallet situation	194	0.8577	0.8984	-0.0407	0.1482
Day care situation	194	0.4494	0.5381	-0.0887	0.0351

* p<0.05, ** p<0.01, *** p<0.001

Table 32: Gender difference in personal appropriateness (5 scale - rank sum test)

	Obs (Men)	Obs (Women)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Lost wallet situation	89	105	-1.126	.13	.134
Day care situation	89	105	-1.817	.035	.035

Notes. P-value equals to 0 in the last column means that p < 0.001.

Table 33: Gender differences in personal appropriateness (7 scale - t-tests)

	Obs.	Mean (men)	Mean (women)	Diff.	P-value
Lost wallet situation	200	0.7340	0.8053	-0.0713	0.0876
Day care situation	200	0.3434	0.3696	-0.0262	0.2306

* p<0.05, ** p<0.01, *** p<0.001

Table 34: Gender differences in personal appropriateness (7 scale - rank sum test)

	Obs (Men)	Obs (Women)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
Lost wallet situation	99	101	-.956	.17	.169
Day care situation	99	101	-.578	.282	.282

4 Order Effects

For each action in each scenario, compare (i) average rating by subjects who saw that scenario as the first scenario in the set of three scenarios to which it belong, with (ii) average rating by all other subjects who saw that scenario (under the same number of rating alternatives and the same mode of appropriateness). Use a t-test to compare the average ratings of these two groups.
For example, consider the rating of the action of being 15 minutes late in the small money day-care center scenario.

Compare (i) subjects who faced this scenario before they faced any other day-care center scenarios, with (ii) all other subjects who faced these scenarios with the same number of rating alternatives and the same mode of appropriateness.

4.1 Social appropriateness

Table 35: Test at level of questions/actions: lost wallet situation (5 scale - t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-valule (H0: diff=0, H1: diff!=0)
No money	198	-0.8929	-0.9167	0.0238	0.6757
Small money	198	-0.9273	-0.8462	-0.0812	0.1391
Big money	198	-0.9255	-0.9865	0.0610	0.1658

Table 36: Test at level of questions/actions: lost wallet situation (5 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
No money	168	30	.817	.793	.76
Small money	172	26	-1.109	.134	.121
Big money	161	37	1.428	.923	.975

Table 37: Test at level of questions/actions: lost wallet situation (7 scale - t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-valule (H0: diff=0, H1: diff!=0)
No money	201	-0.8140	-0.9080	0.0941	0.1476
Small money	201	-0.8753	-0.8860	0.0107	0.8220
Big money	201	-0.9182	-0.9118	-0.0064	0.8806

Table 38: Test at level of questions/actions: lost wallet situation (7 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
No money	172	29	2.176	.985	.988
Small money	163	38	1.393	.918	.922
Big money	167	34	-.476	.317	.325

Table 39: Test at level of questions/actions: day care situation (5 scale - t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-valule (H0: diff=0, H1: diff!=0)
No fine (15)	197	-0.2807	-0.2500	-0.0307	0.7569
Small fine (15)	197	-0.1071	0.0465	-0.1537	0.1037
Big fine (15)	197	-0.0556	-0.1000	0.0444	0.6714
No fine (30)	197	-0.7251	-0.6731	-0.0521	0.5393
Small fine (30)	197	-0.5422	-0.5814	0.0392	0.6691
Big fine (30)	197	-0.4105	-0.5571	0.1466	0.1995

Table 40: Test at level of questions/actions: day care situation (5 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
No fine (15)	171	26	-.334	.369	.374
r1	154	43	-1.496	.067	.068
Big fine (15)	162	35	.427	.665	.666
No fine (30)	171	26	-.412	.34	.354
Small fine (30)	154	43	.442	.671	.672
Big fine (30)	162	35	1.808	.965	.966

Table 41: Test at level of questions/actions: day care situation (7 scale - t-tests)

No fine (15)	201	-0.3574	-0.2571	-0.1003	0.2114
Small fine (15)	201	-0.2081	-0.1204	-0.0877	0.3220
Big fine (15)	201	-0.2481	-0.0575	-0.1906	0.0472
No fine (30)	201	-0.7349	-0.6571	-0.0778	0.2681
Small fine (30)	201	-0.6242	-0.5648	-0.0594	0.5034
Big fine (30)	201	-0.5678	-0.5287	-0.0391	0.6944

Table 42: Test at level of questions/actions: day care situation (7 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
No fine (15)	166	35	-1.085	.139	.14
r1	165	36	-.721	.236	.237
Big fine (15)	172	29	-2.106	.018	.018
No fine (30)	166	35	-1.631	.051	.053
Small fine (30)	165	36	.212	.584	.583
Big fine (30)	172	29	-.414	.34	.34

4.2 Personal appropriateness

Table 43: Test at level of questions/actions: lost wallet situation (5 scale - t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-valule (H0: diff=0, H1: diff!=0)
No money	194	-0.8975	-0.9394	0.0419	0.3903
Small Money	194	-0.9344	-0.9853	0.0509	0.1226
Big money	194	-0.9146	-0.9833	0.0687	0.1860

Table 44: Test at level of questions/actions: lost wallet situation (5 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
No money	161	33	.635	.737	.798
Small Money	160	34	1.557	.94	.983
Big money	164	30	1.378	.916	.971

Table 45: Test at level of questions/actions: lost wallet situation (7 scale - t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-valule (H0: diff=0, H1: diff!=0)
No money	200	-0.8675	-0.8485	-0.0190	0.6733
Small Money	200	-0.8755	-0.9216	0.0461	0.3952
Big money	200	-0.9051	-0.8952	-0.0098	0.8359

Table 46: Test at level of questions/actions: lost wallet situation (7 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
No money	156	44	.011	.505	.496
Small Money	166	34	.098	.539	.548
Big money	165	35	-.162	.436	.443

Table 47: Test at level of questions/actions: day care situation (5 scale - t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-valule (H0: diff=0, H1: diff!=0)
No fine (15)	194	-0.3829	-0.3056	-0.0774	0.4091
Small fine (15)	194	-0.2037	-0.2031	-0.0006	0.9957
Big fine (15)	194	-0.1333	-0.2414	0.1080	0.3496
No fine (30)	194	-0.7120	-0.8194	0.1074	0.1699
Small fine (30)	194	-0.5727	-0.7069	0.1342	0.2058
Big fine (30)	194	-0.4576	-0.6379	0.1804	0.1272

Table 48: Test at level of questions/actions: day care situation (5 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
No fine (15)	158	36	-.92	.179	.18
Small fine (15)	162	32	-.317	.376	.377
Big fine (15)	165	29	.851	.803	.802
No fine (30)	158	36	.87	.808	.808
Small fine (30)	165	29	1.279	.9	.903
Big fine (30)	165	29	1.368	.914	.916

Table 49: Test at level of questions/actions: day care situation (7 scale - t-tests)

	Obs.	Mean (others)	Mean (seen first)	Diff.	P-valule (H0: diff=0, H1: diff!=0)
No fine (15)	200	-0.4198	-0.2105	-0.2092	0.0112
Small fine (15)	200	-0.2569	-0.1444	-0.1124	0.2377
Big fine (15)	200	-0.1823	-0.1053	-0.0771	0.5076
No fine (30)	200	-0.7263	-0.5877	-0.1386	0.0424
Small fine (30)	200	-0.5599	-0.6667	0.1068	0.3713
Big fine (30)	200	-0.5304	-0.4912	-0.0392	0.7471

Table 50: Test at level of questions/actions: day care situation (7 scale - rank sum test)

	Obs (1)	Obs (2)	Z-stat	P-value	P-value (exact) (H0: diff=0, H1:diff<0)
No fine (15)	162	38	-2.961	.002	.001
Small fine (15)	170	30	-1.139	.127	.129
Big fine (15)	181	19	-.747	.227	.228
No fine (30)	162	38	-2.637	.004	.004
Small fine (30)	181	19	.939	.826	.825
Big fine (30)	181	19	-.817	.207	.209

5 Figures: distribution of ratings

The bar graphs below shows the frequency of response by question, situation and mode of appropriateness.

5.1 Lost wallet situation

Figure 1: Social appropriateness: Not returning the wallet (5 scale)

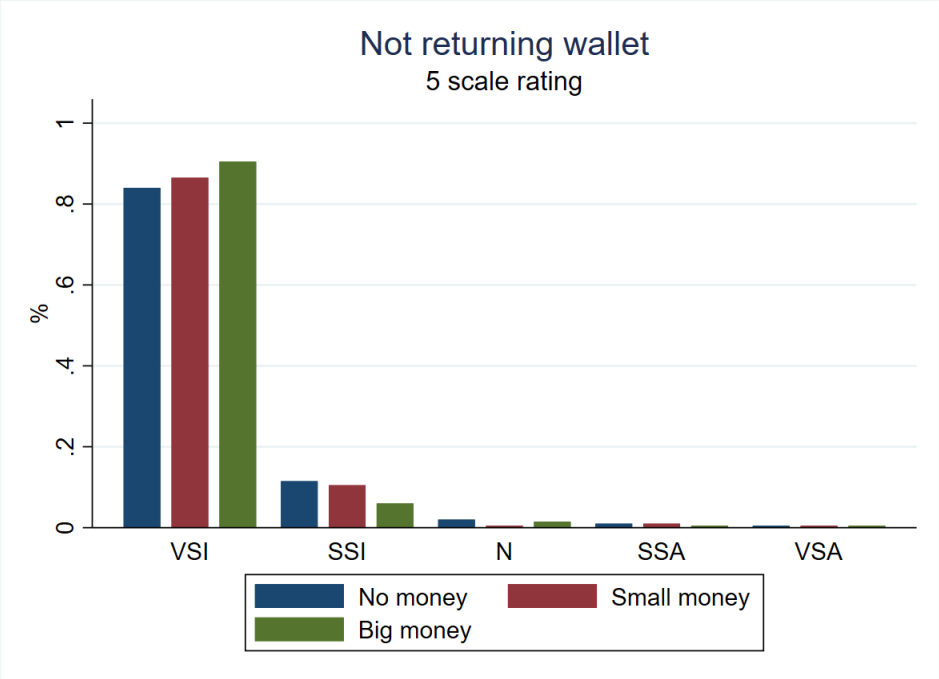


Figure 2: Social appropriateness: Not returning the wallet (7 scale)

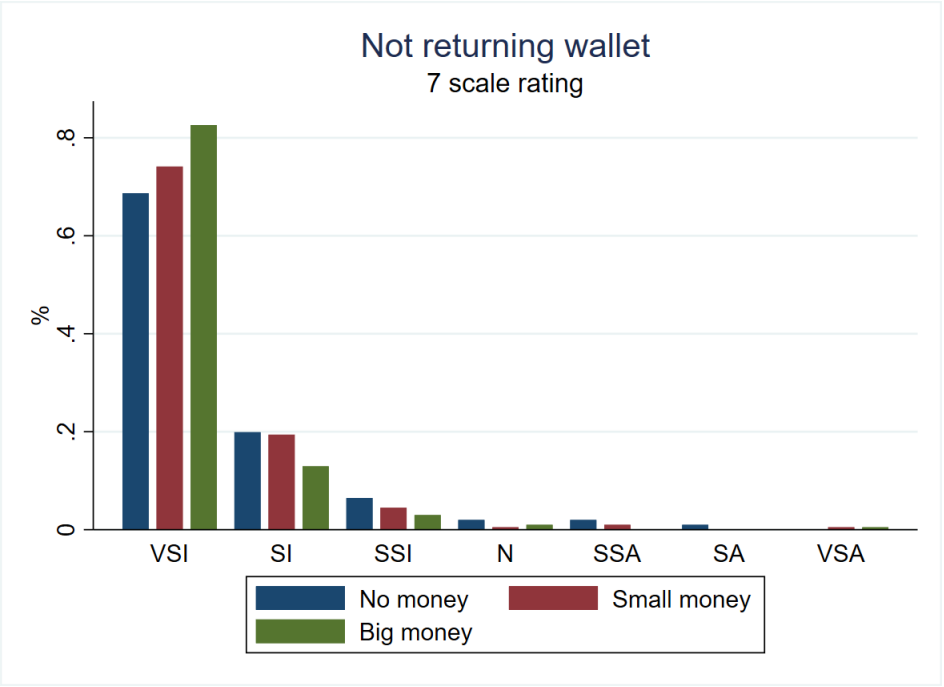


Figure 3: Social appropriateness: Cumulative distribution of ratings for returning the wallet (5 scale)

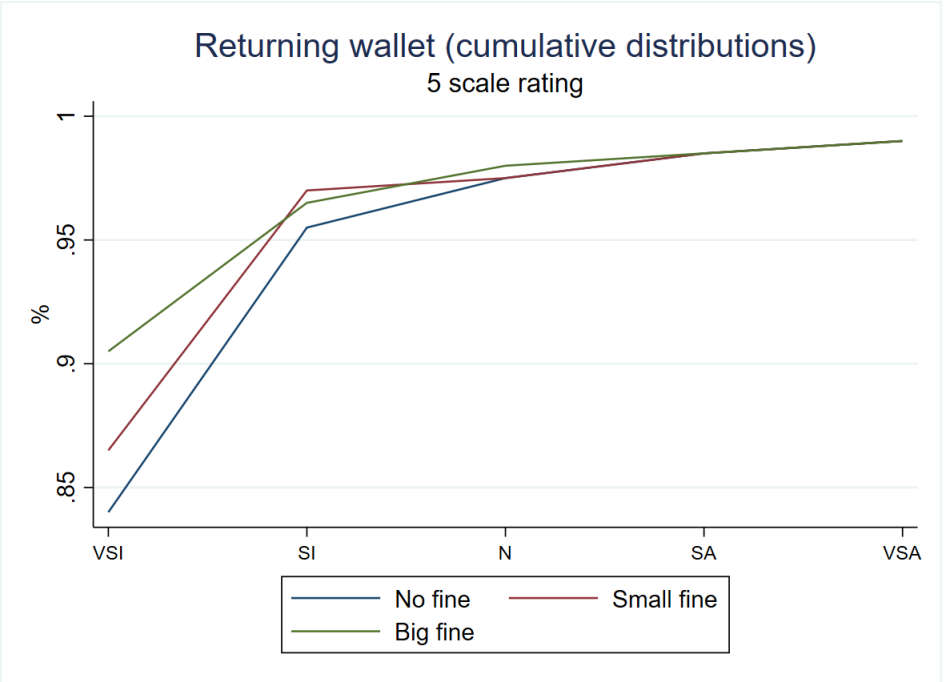


Figure 4: Social appropriateness: Cumulative distribution of ratings for returning the wallet (7 scale)

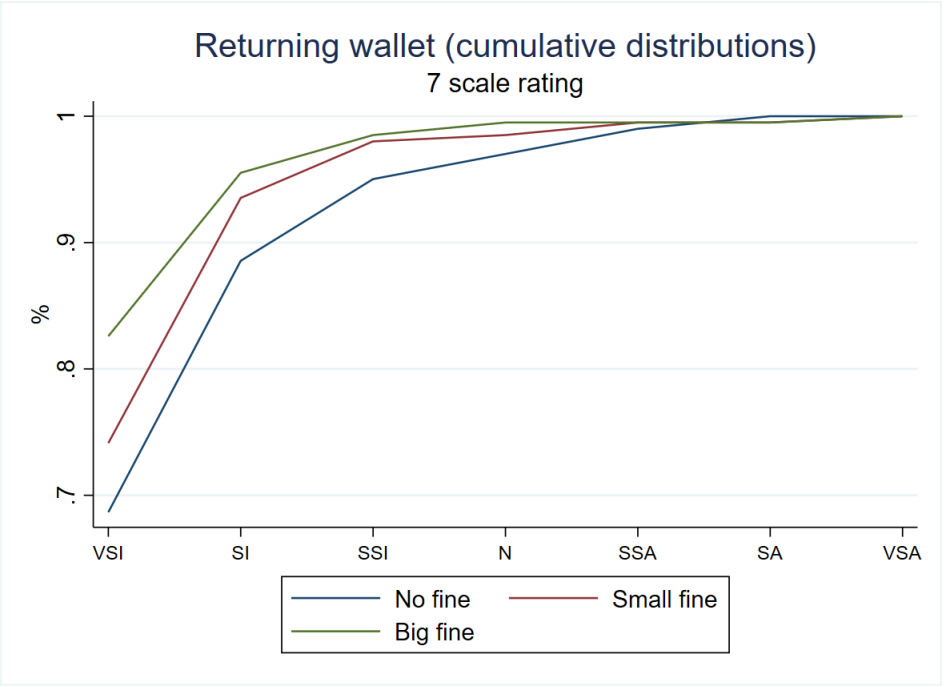


Figure 5: Personal appropriateness: Not returning the wallet (5 scale)

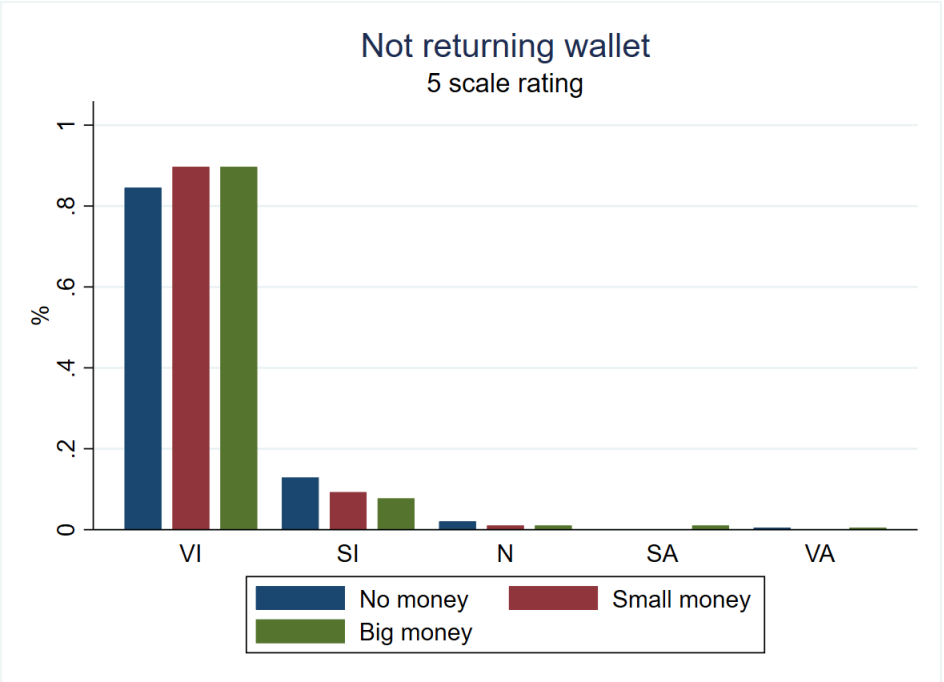


Figure 6: Personal appropriateness: Not returning the wallet (7 scale)

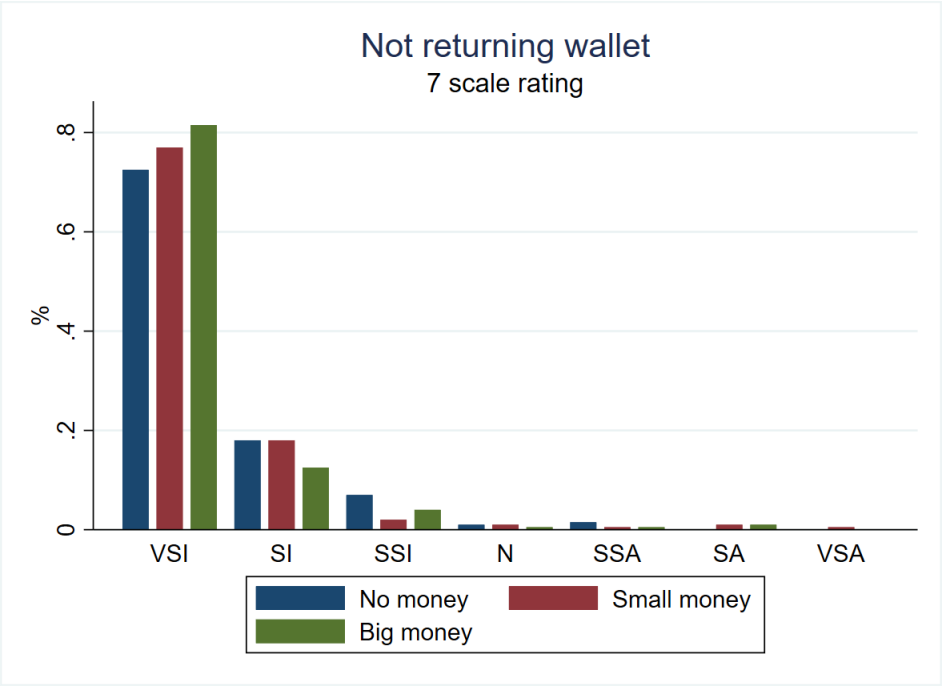


Figure 7: Personal appropriateness: Cumulative distribution of ratings for returning the wallet (5 scale)

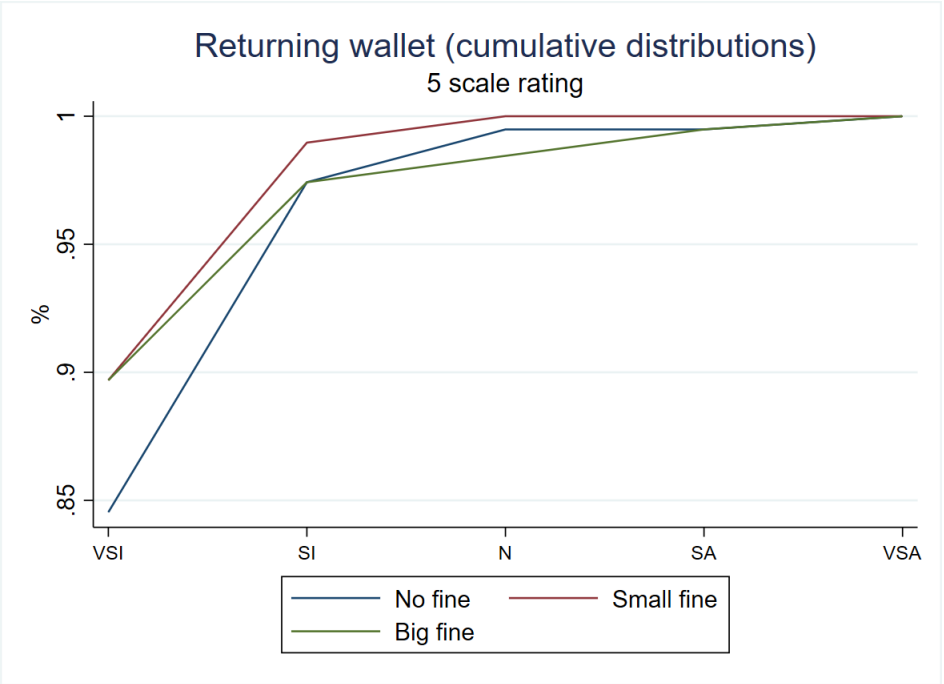


Figure 8: Personal appropriateness: Cumulative distribution of ratings for returning the wallet (7 scale)

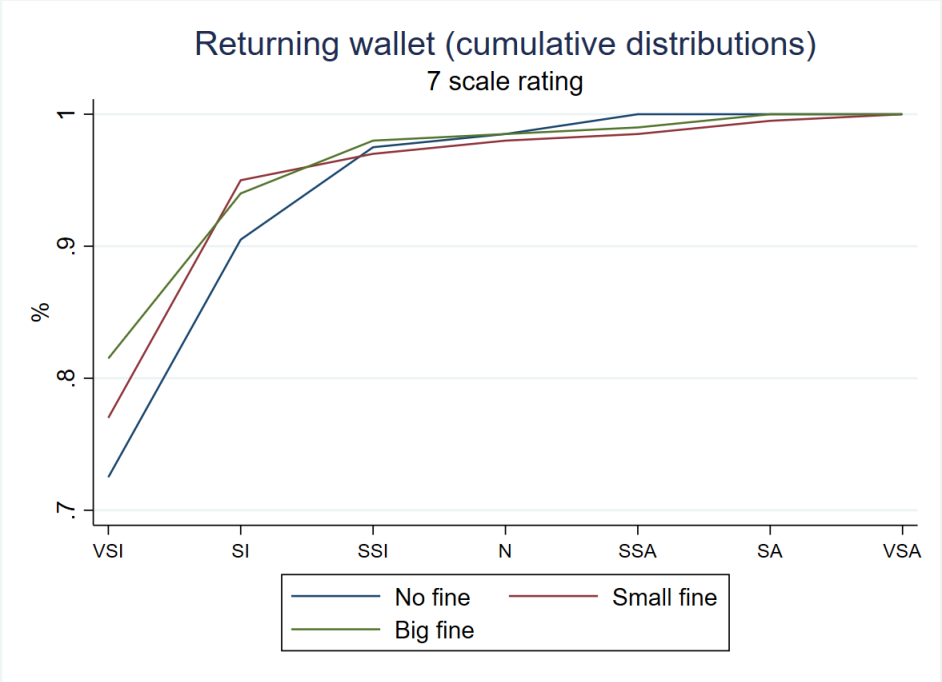


Figure 9: Social appropriateness: average ratings (5 scale)

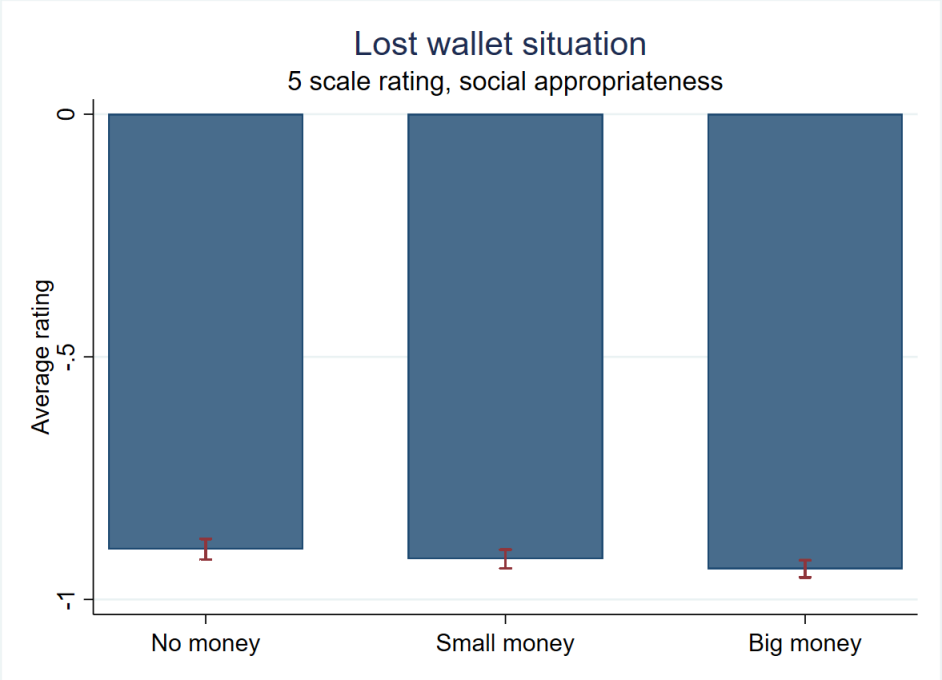


Figure 10: Social appropriateness: average ratings (7 scale)

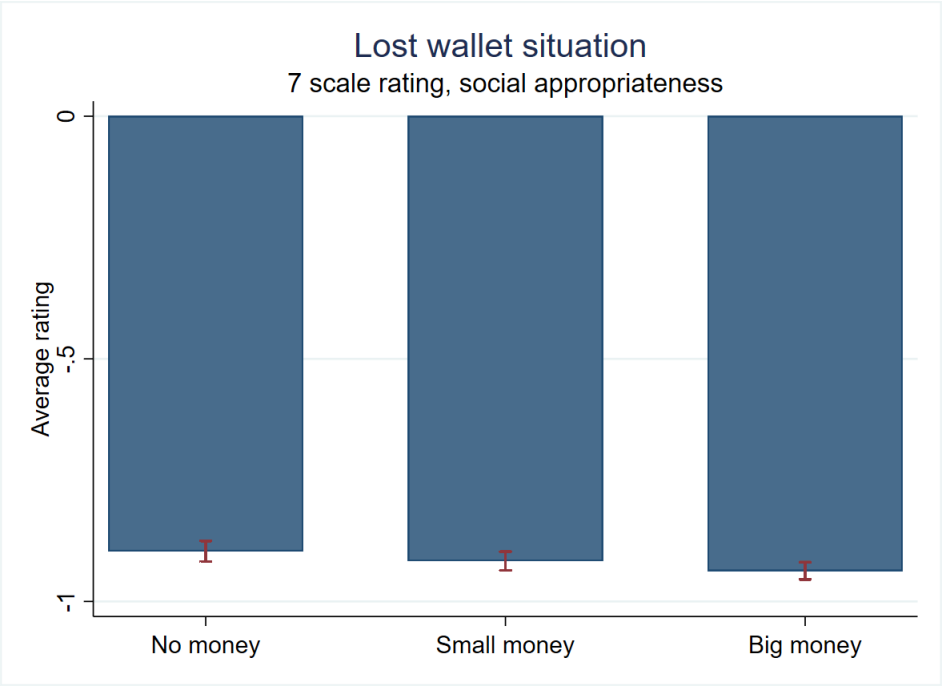


Figure 11: Personal appropriateness: average ratings (5 scale)

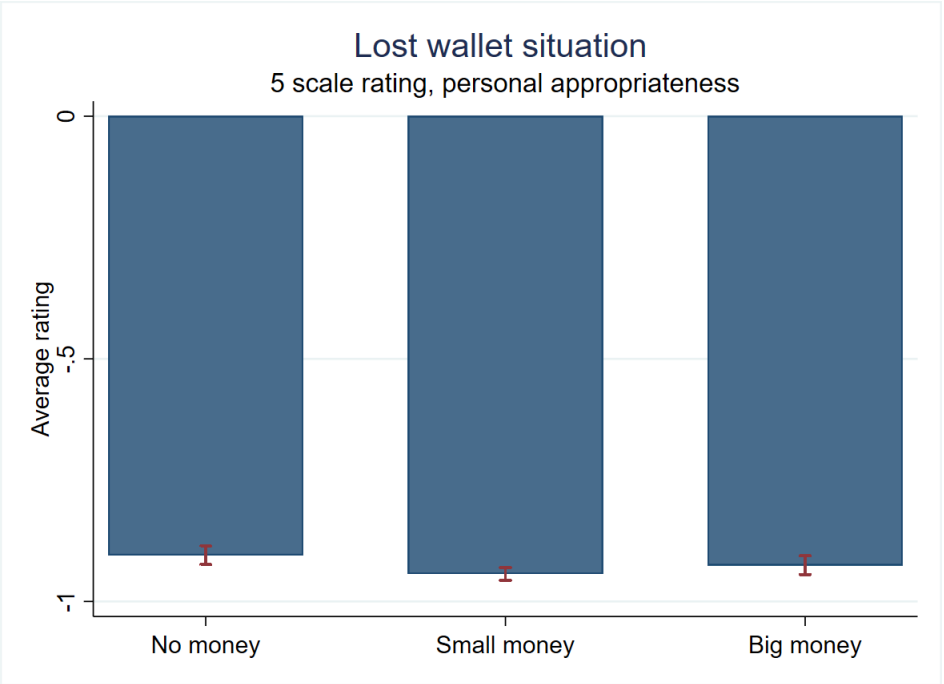
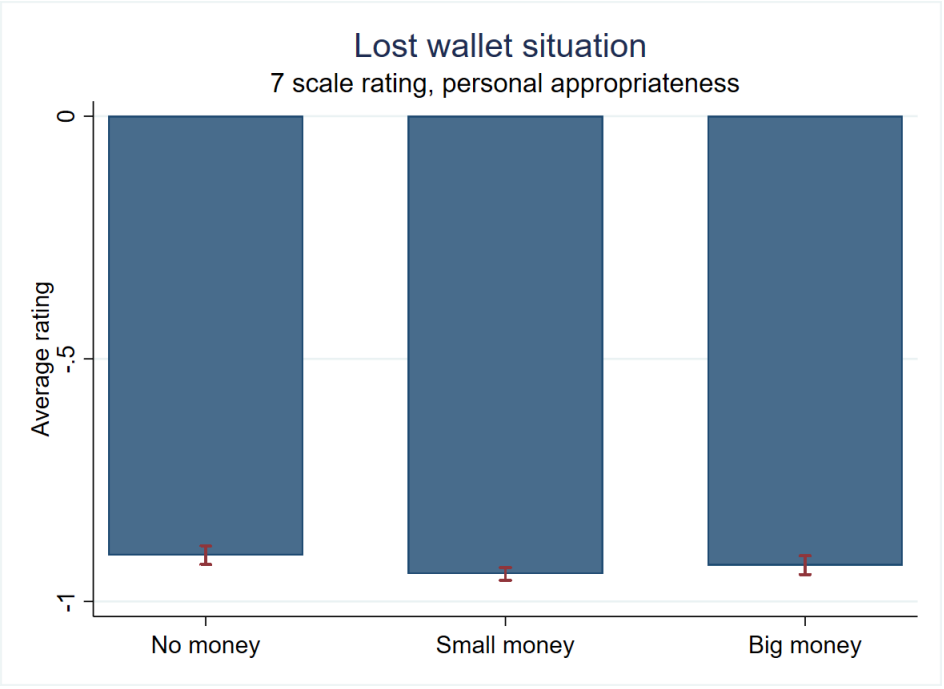


Figure 12: Personal appropriateness: average ratings (7 scale)



5.2 Day-care situation

Figure 13: Social appropriateness: Arriving 15 mins late (5 scale)

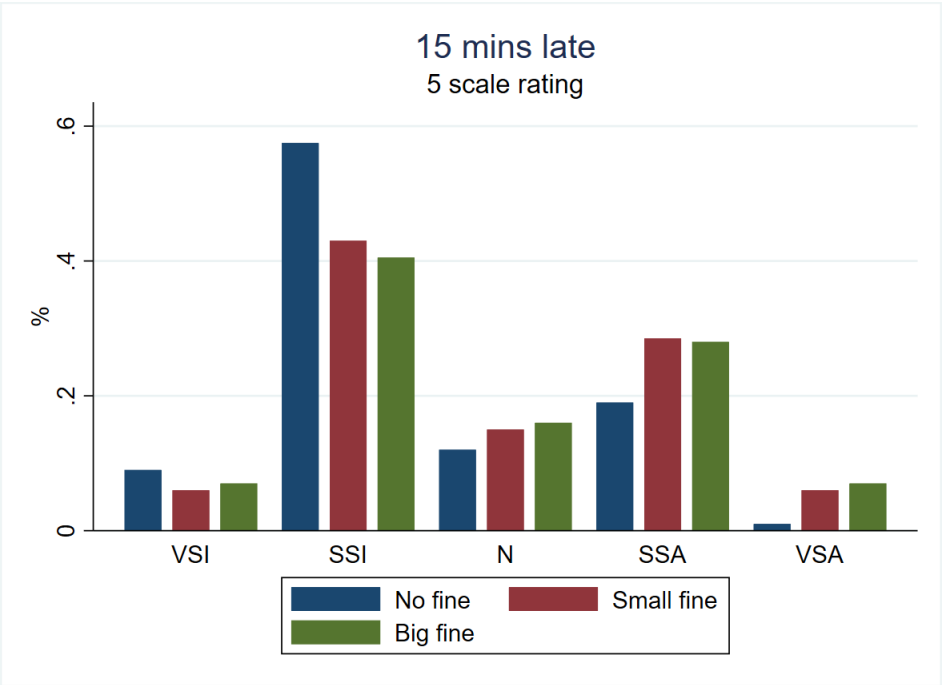


Figure 14: Social appropriateness: Arriving 15 mins late (7 scale)

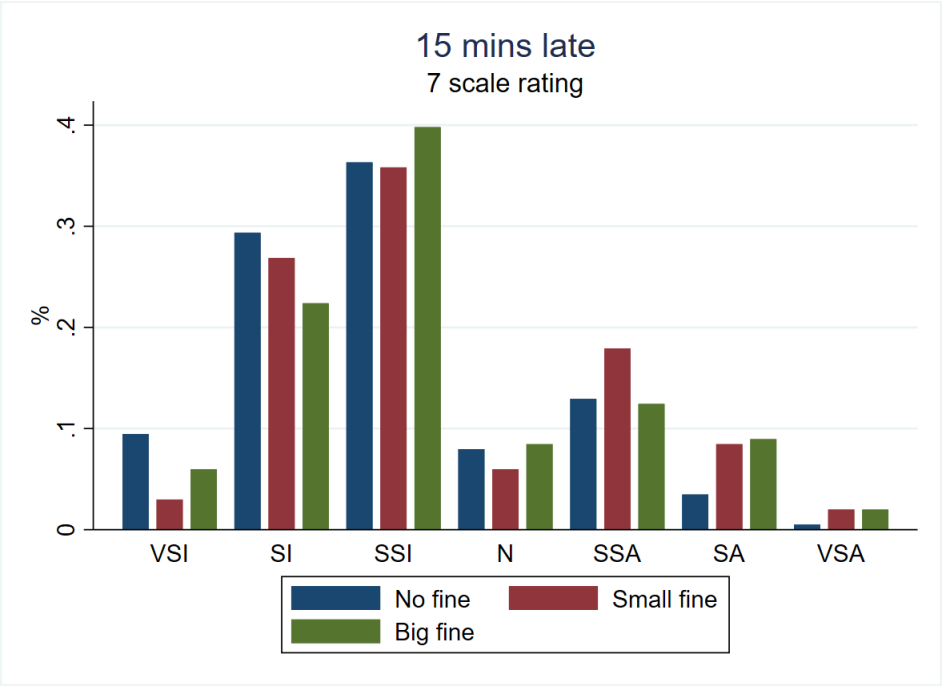


Figure 15: Personal appropriateness: Arriving 15 mins late (5 scale)

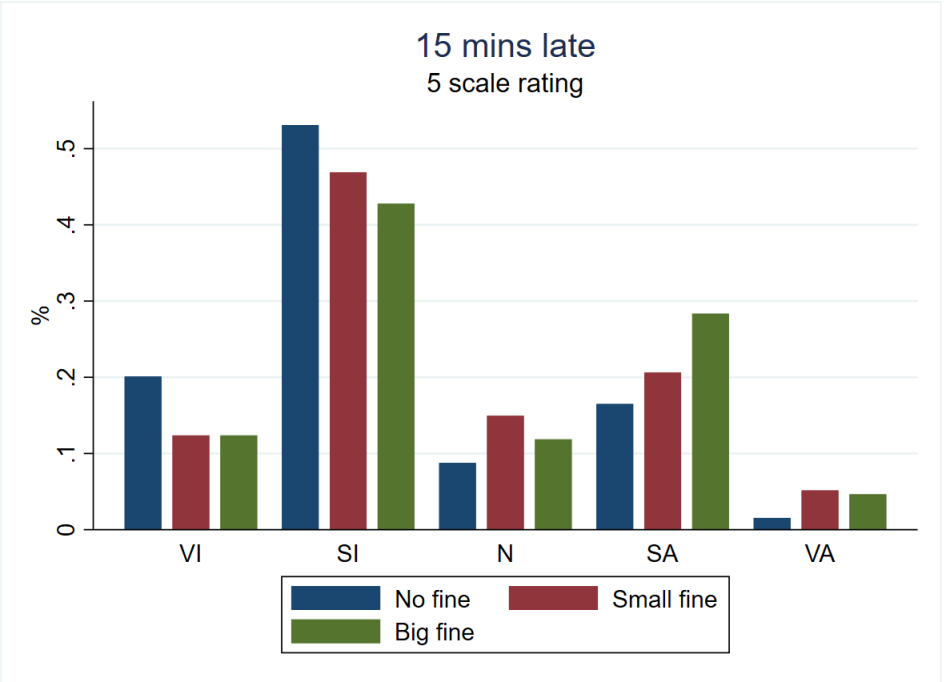


Figure 16: Personal appropriateness: Arriving 15 mins late (7 scale)

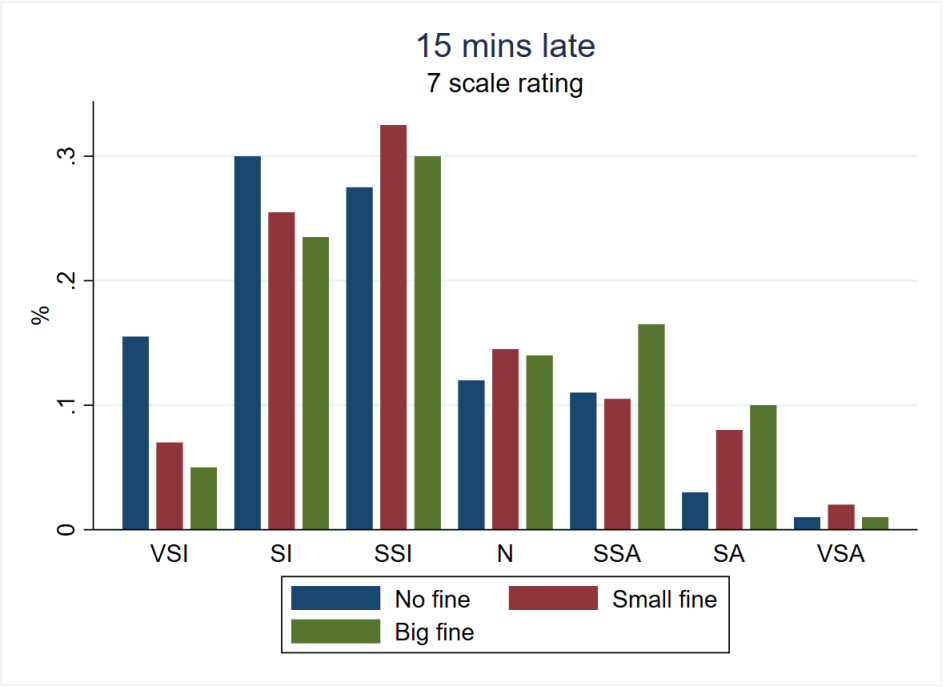


Figure 17: Social appropriateness: Arriving 30 mins late (5 scale)

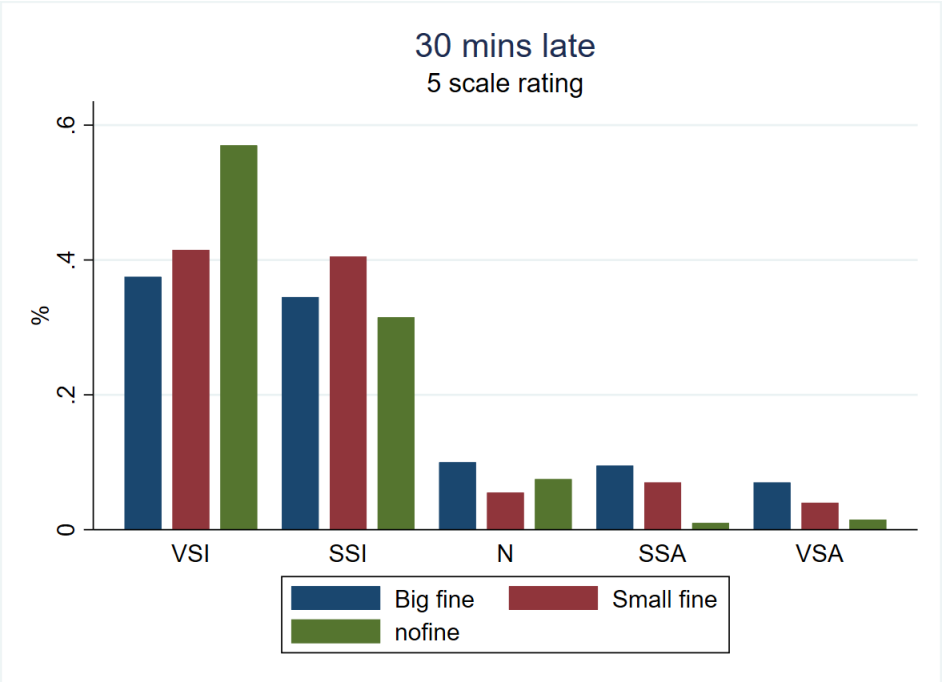


Figure 18: Social appropriateness: Arriving 30 mins late (7 scale)

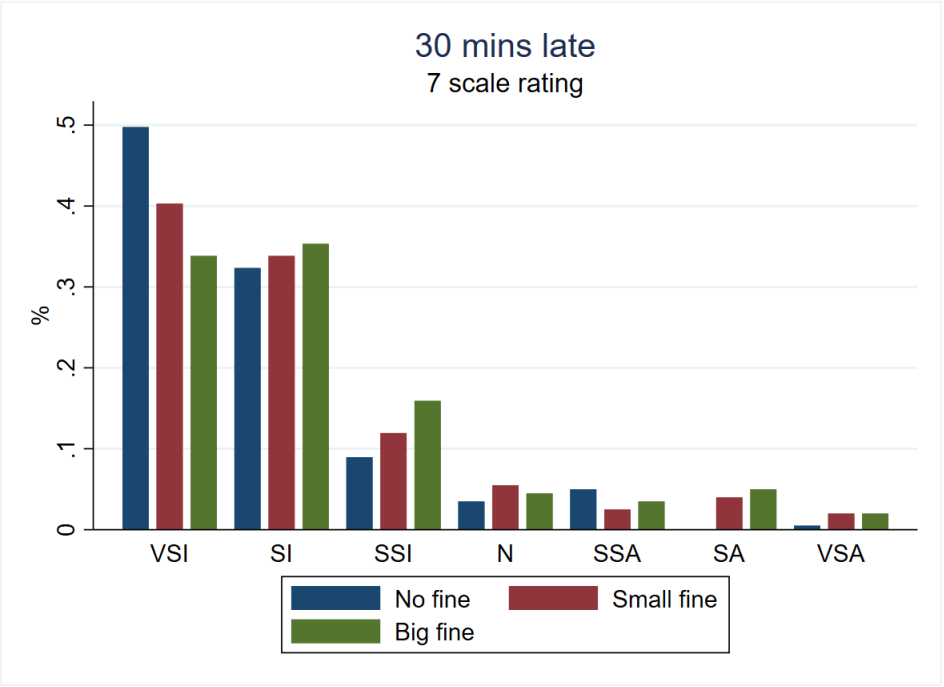


Figure 19: Social appropriateness: Arriving 30 mins late (5 scale)

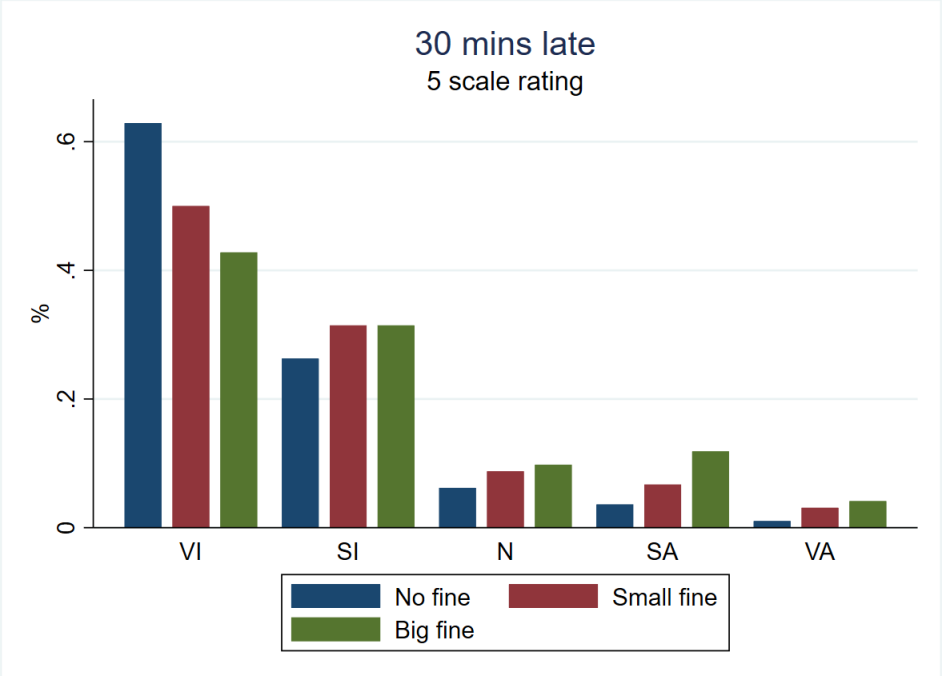


Figure 20: Social appropriateness: Arriving 30 mins late (7 scale)

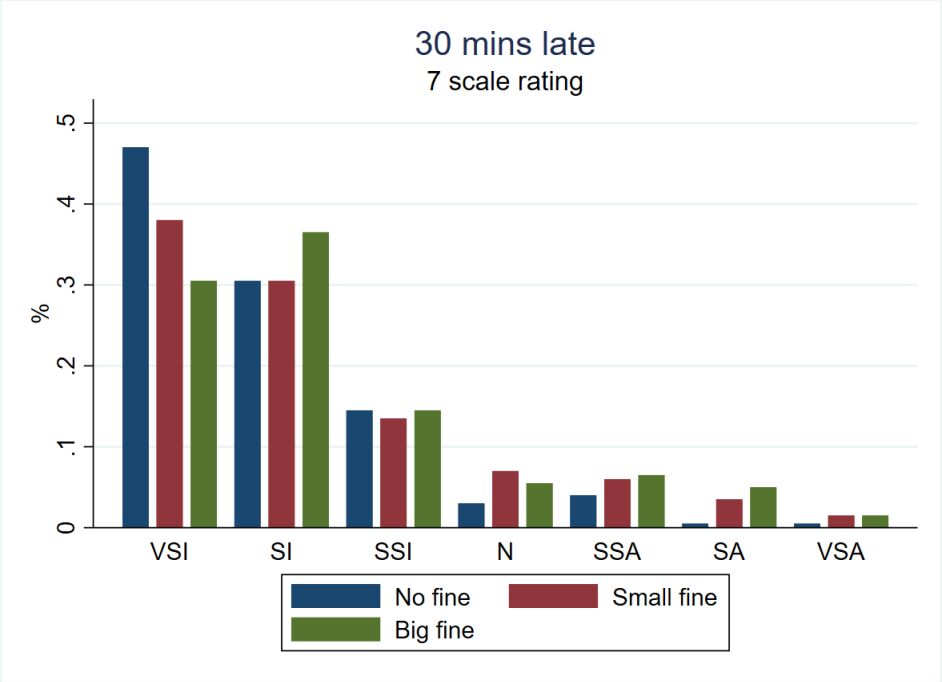


Figure 21: Social appropriateness: cumulative distribution of ratings for arriving 15 mins late (5 scale)

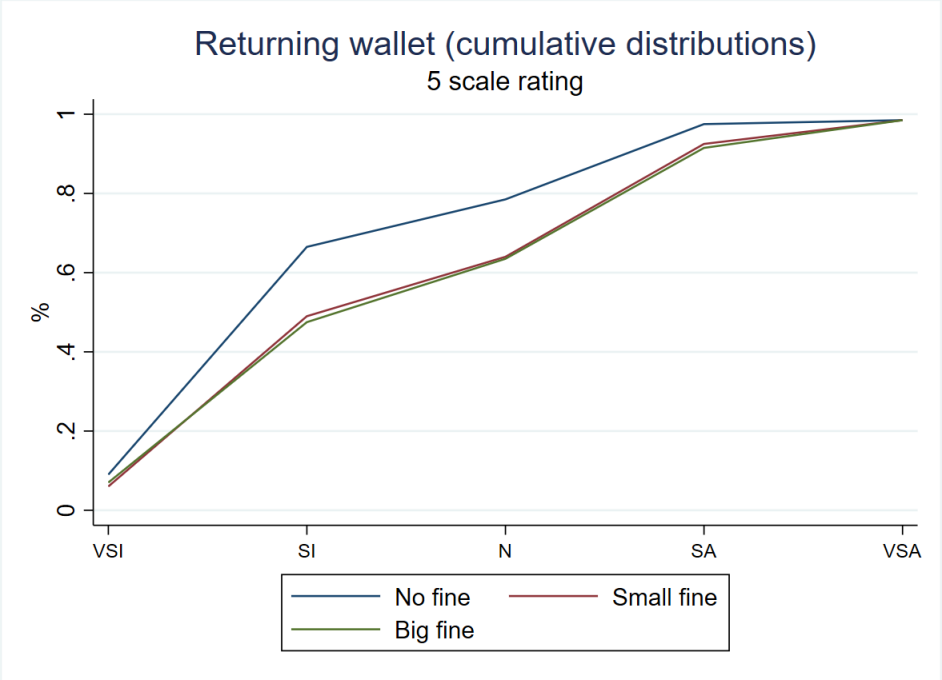


Figure 22: Social appropriateness: cumulative distribution of ratings for arriving 15 mins late (7 scale)

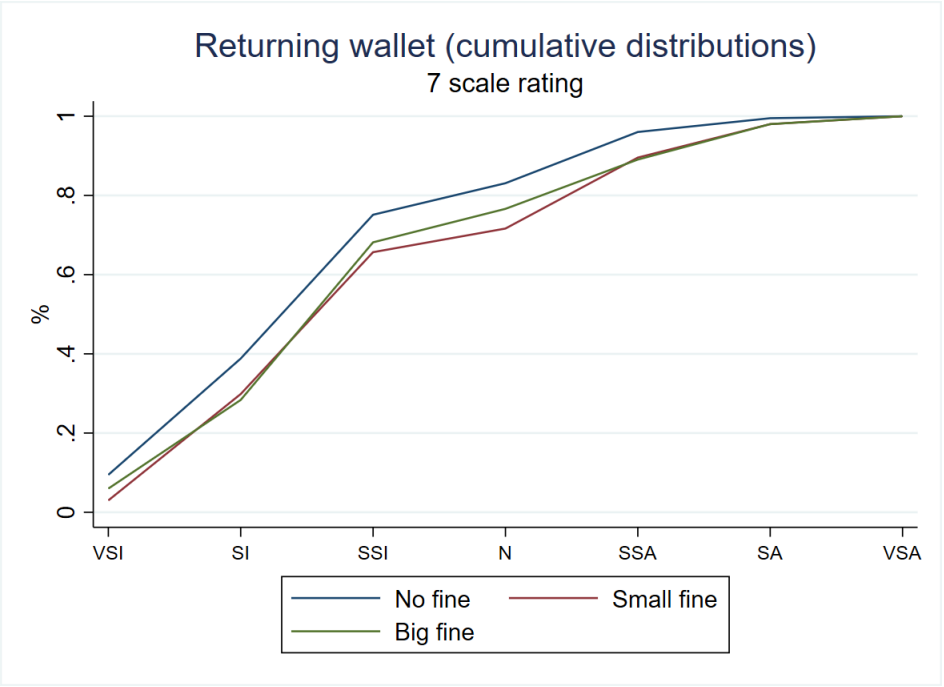


Figure 23: Personal appropriateness: cumulative distribution of ratings for arriving 15 mins late (5 scale)

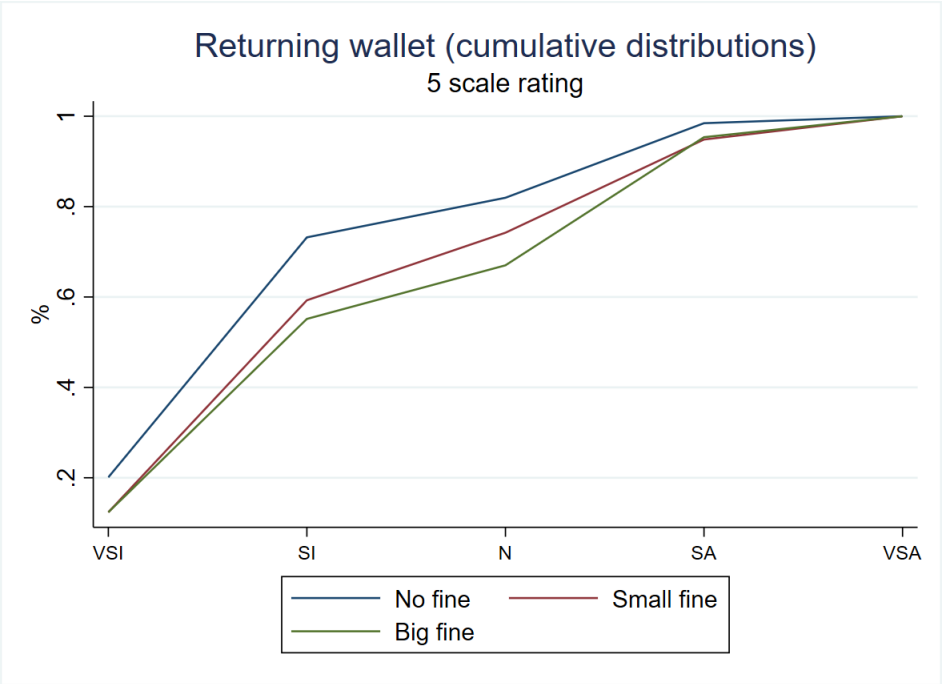


Figure 24: Personal appropriateness: cumulative distribution of ratings for arriving 15 mins late (7 scale)

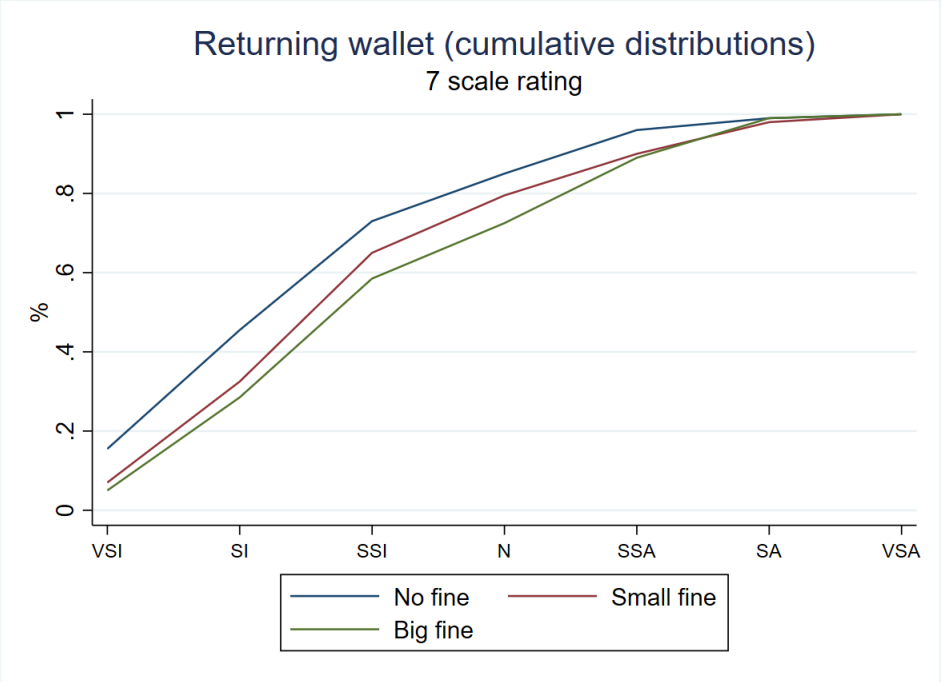


Figure 25: Social appropriateness: cumulative distribution of ratings for arriving 30 mins late (5 scale)

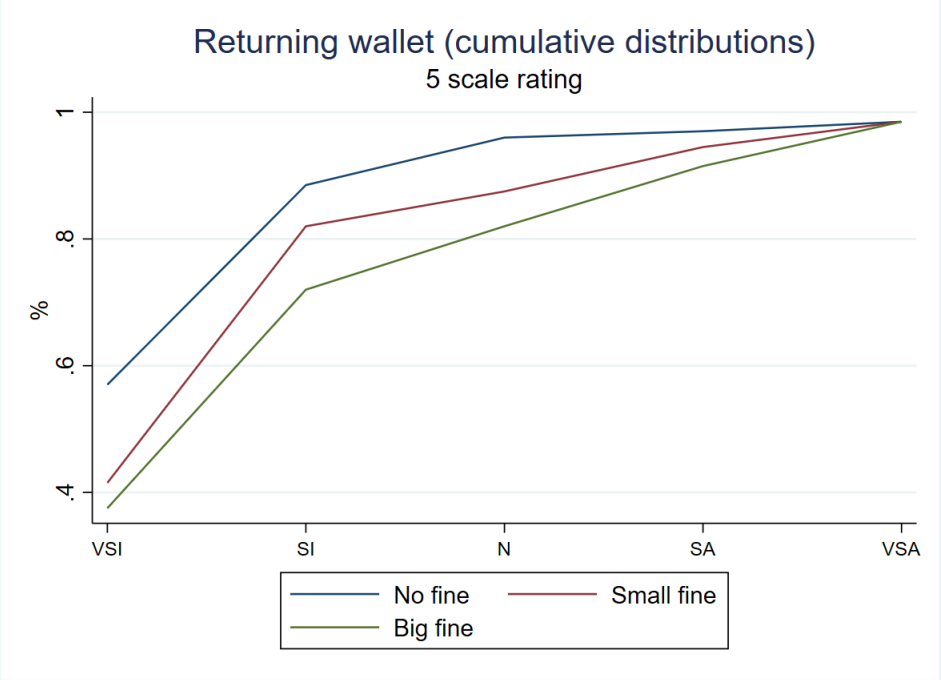


Figure 26: Social appropriateness: cumulative distribution of ratings for arriving 30 mins late (7 scale)

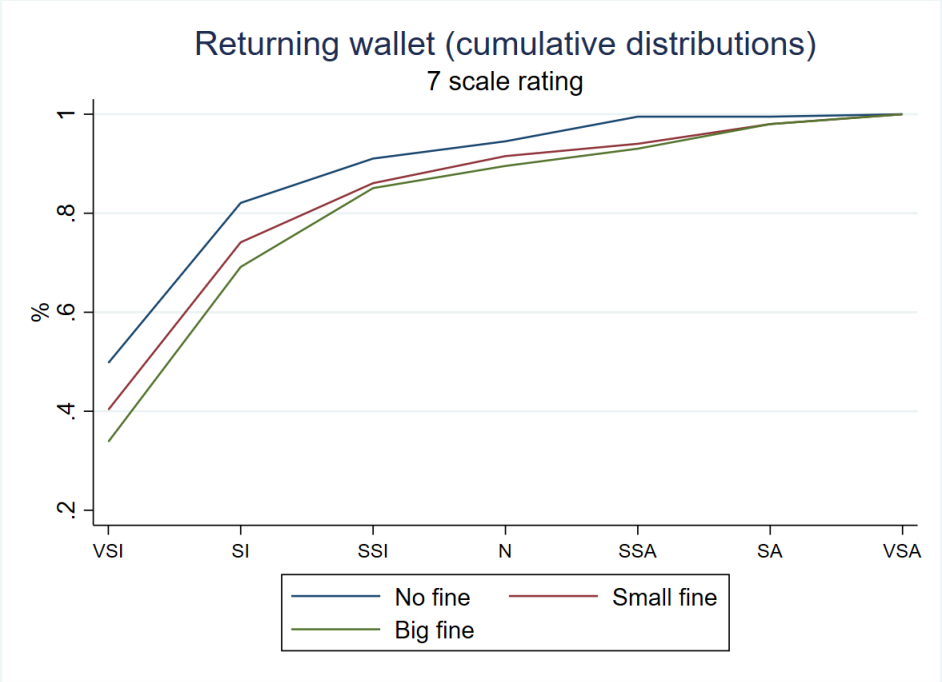


Figure 27: Personal appropriateness: cumulative distribution of ratings for arriving 30 mins late (5 scale)

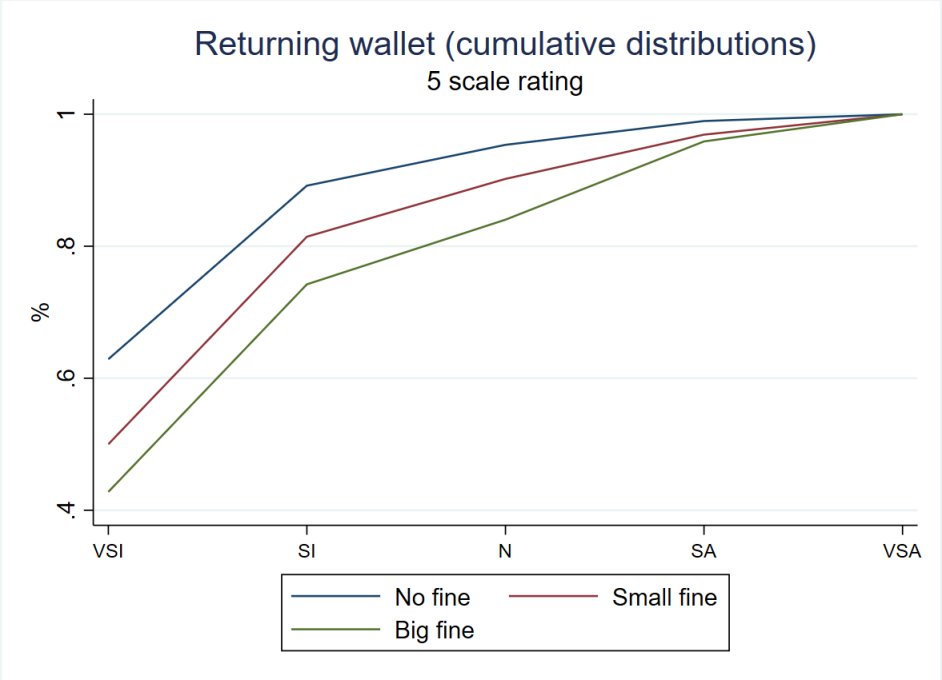


Figure 28: Personal appropriateness: cumulative distribution of ratings for arriving 15 mins late (7 scale)

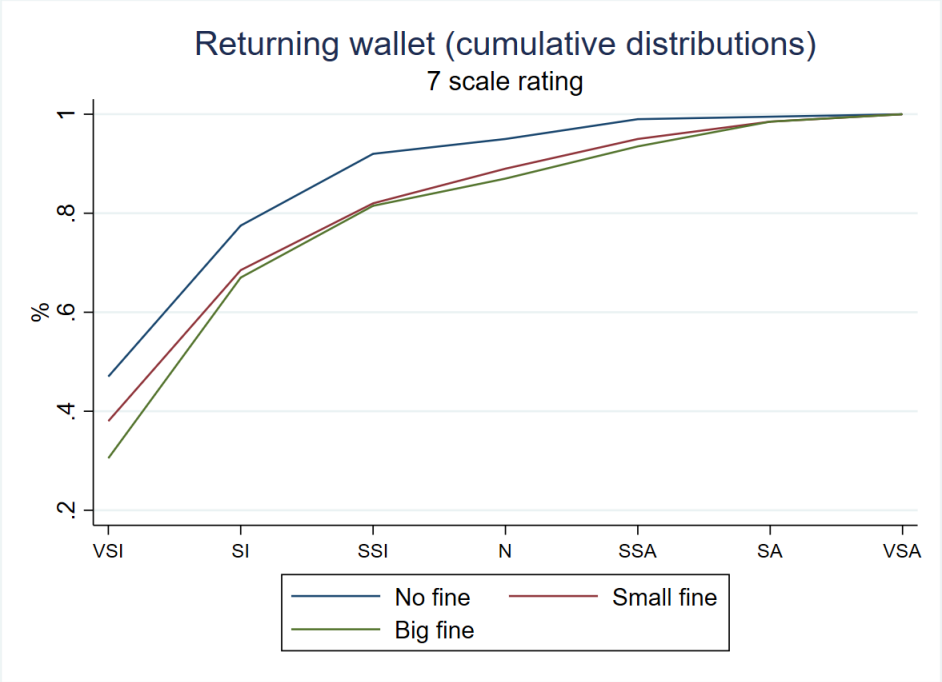


Figure 29: Social appropriateness: average ratings for arriving 15 mins late (5 scale)

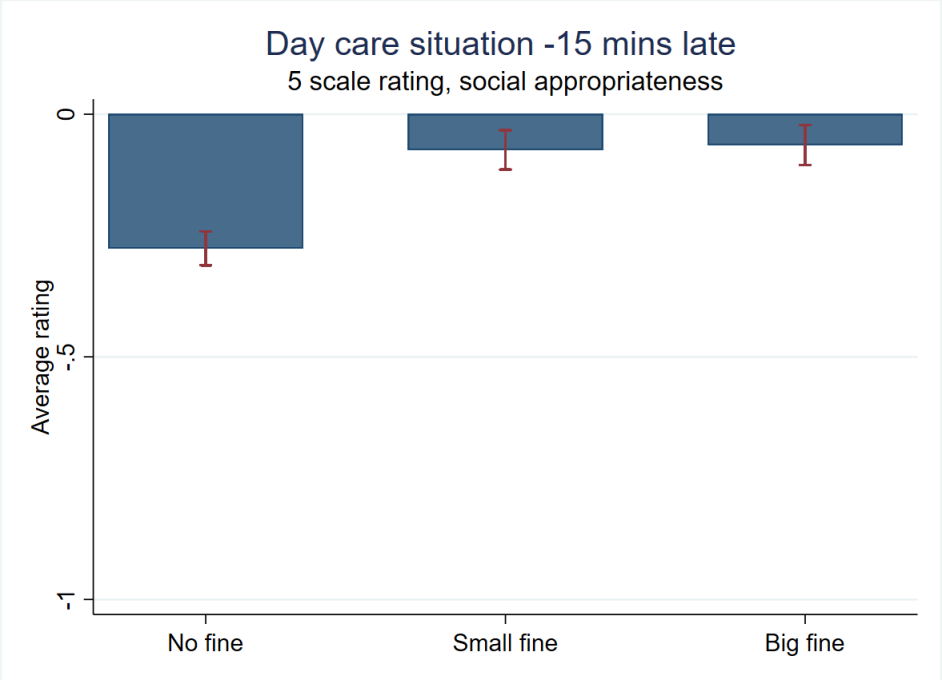


Figure 30: Social appropriateness: average ratings for arriving 15 mins late (7 scale)

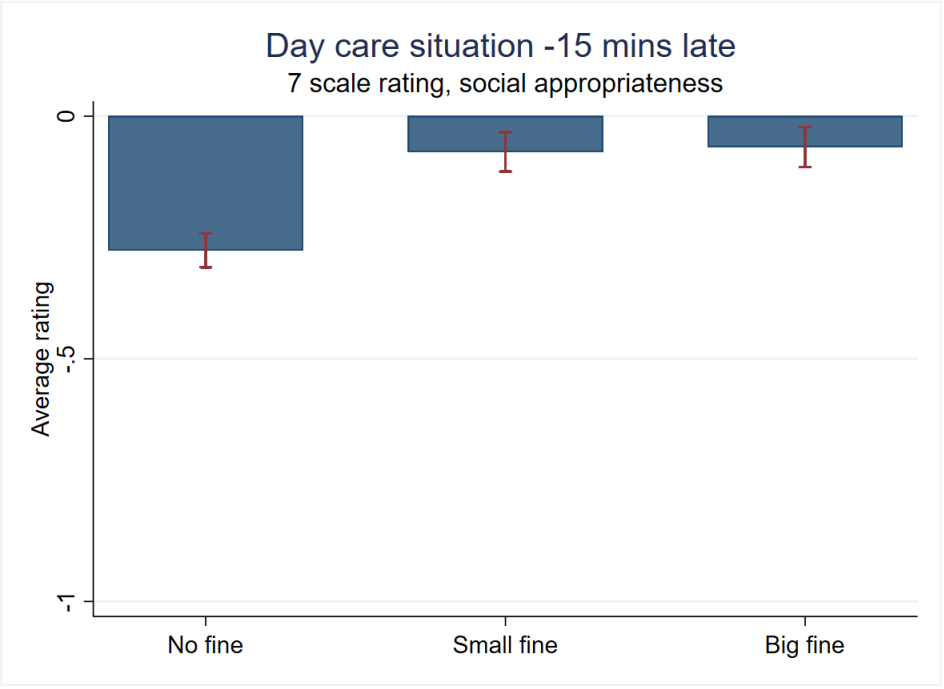


Figure 31: Personal appropriateness: average ratings for arriving 15 mins late (5 scale)

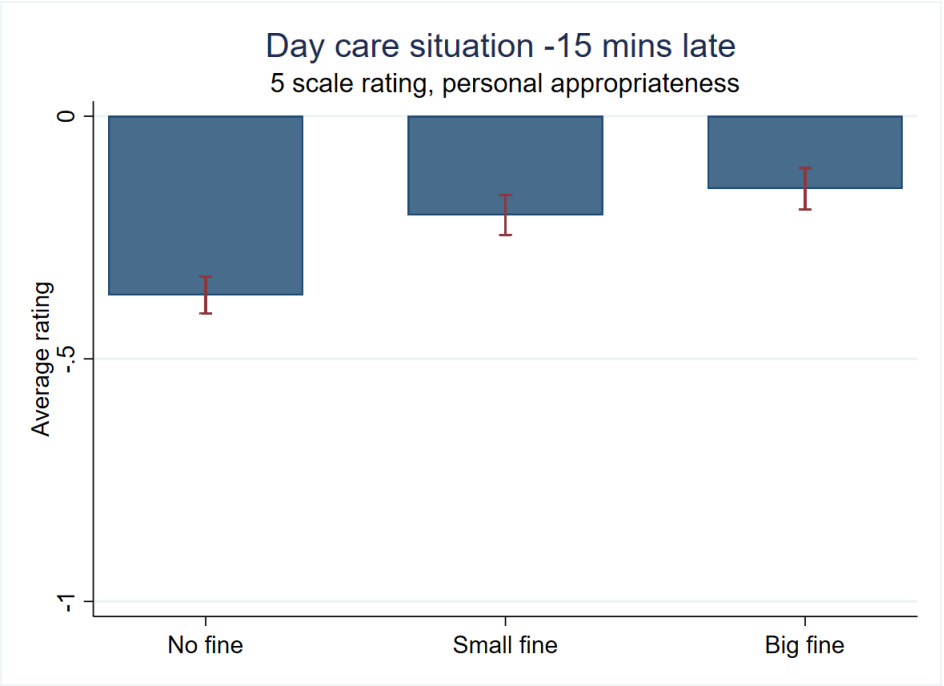


Figure 32: Personal appropriateness: average ratings for arriving 15 mins late (7 scale)

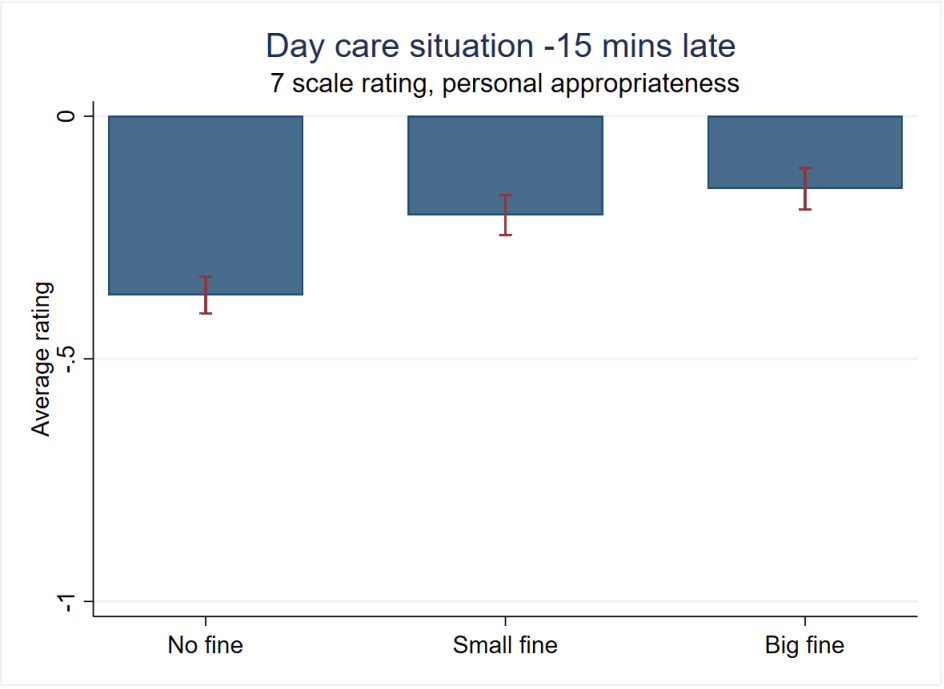


Figure 33: Social appropriateness: average ratings for arriving 30 mins late (5 scale)

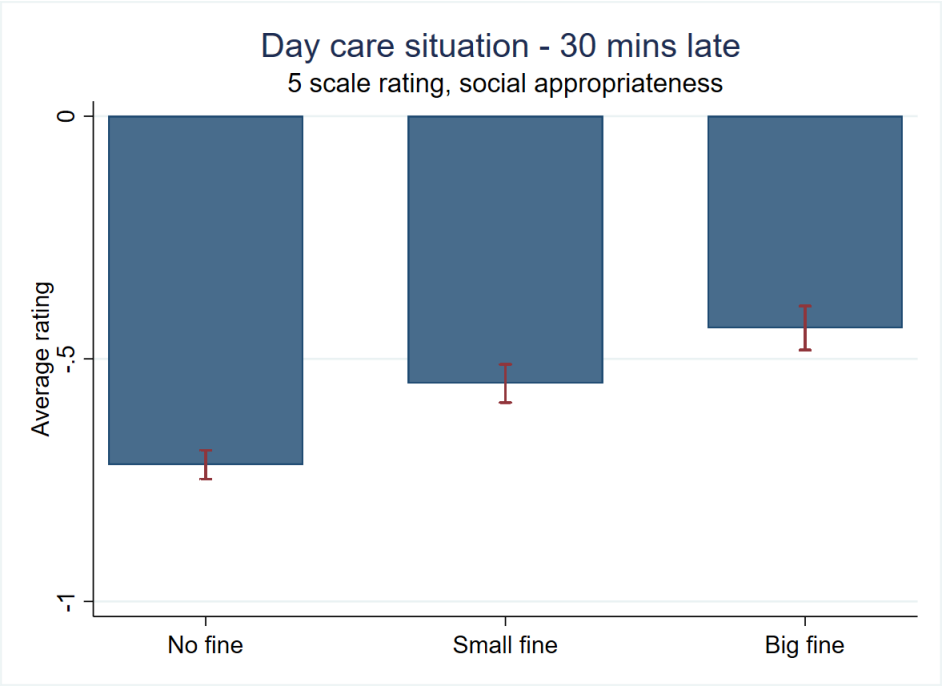


Figure 34: Social appropriateness: average ratings for arriving 30 mins late (7 scale)

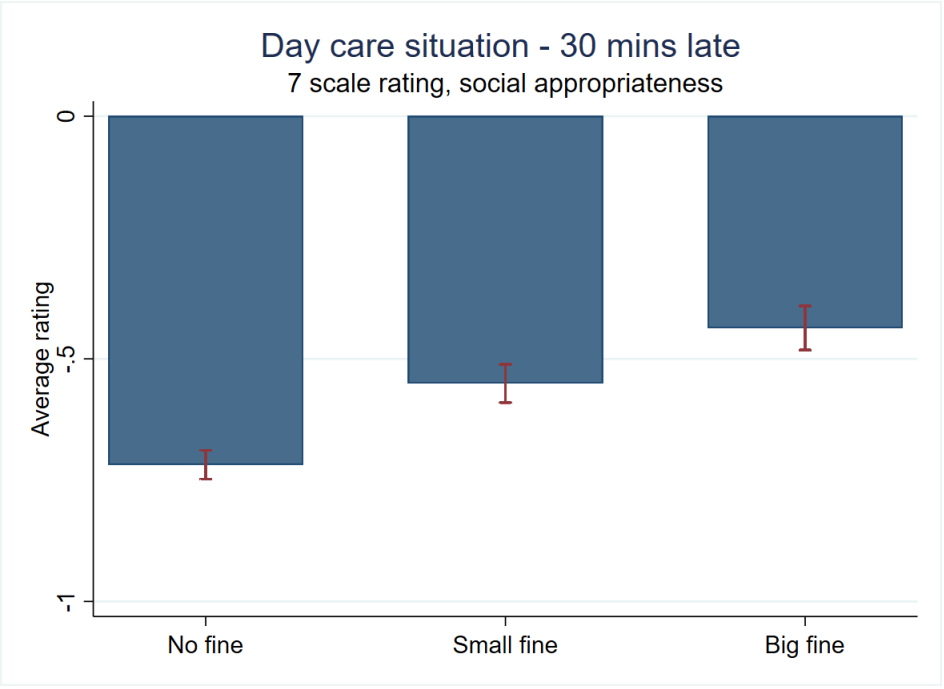


Figure 35: Personal appropriateness: average ratings for arriving 30 mins late (5 scale)

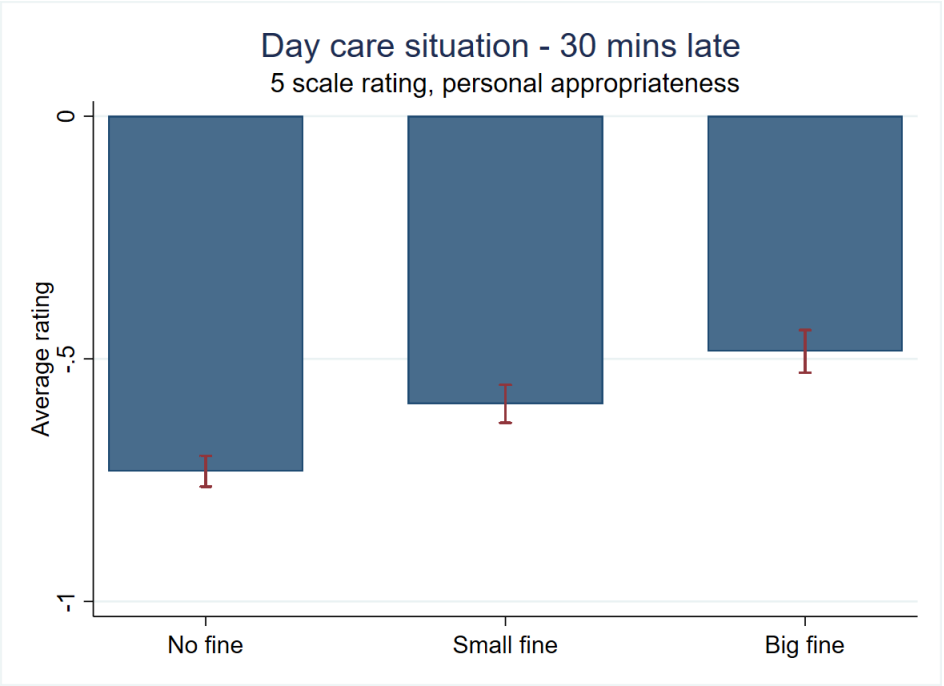
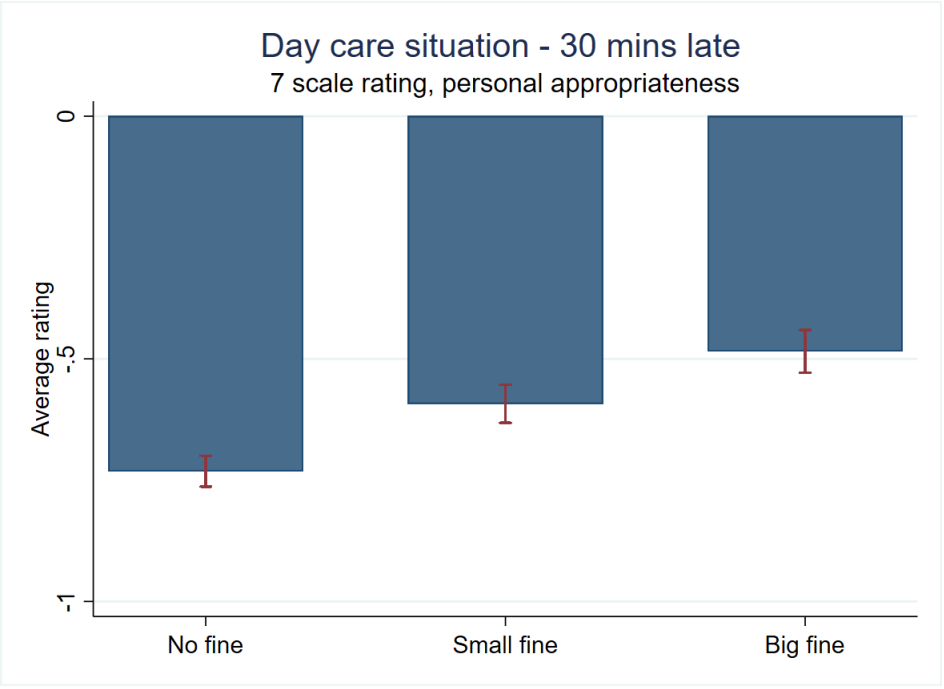



Figure 36: Personal appropriateness: average ratings for arriving 30 mins late (7 scale)