Appendix

A Additional tables and figures

TABLE A.1: Participant Attributes across treatments

	No Advice	AI Advice	Human-AI Advice
Fixed Pay			
Avg. Age	24.80	25.34	24.47
% Female	62.27	58.86	64.52
% Univ.Degree	58.58	58.68	54.84
% Austrian	54.44	52.10	47.85
N	169	167	186
Performance Pay			
Avg. Age	25.31	25.11	24.40
% Female	58.14	56.47	62.68
% Univ.Degree	54.65	47.65	48.59
% Austrian	59.30	58.82	54.93
N	172	170	142
Tournament Pay			
Avg. Age	24.83	24.69	24.20
% Female	57.86	64.88	58.97
% Univ.Degree	56.60	60.71	44.23
% Austrian	45.91	53.57	54.49
N	159	168	156

TABLE A.2: OLS Regressions predicting the true price based on participant's initial estimate and provided algorithmic advice

	No Advice		AI Advice		HumanAI advice	
Initial Estimate	0.026***	0.021***	0.053***	0.056***	0.050***	0.046***
	(0.005)	(0.007)	(0.009)	(0.010)	(0.009)	(0.010)
IniE × PerfPay		0.005		0.020		0.031
		(0.011)		(0.018)		(0.021)
IniE × Tournament		0.013		-0.018		-0.004
		(0.010)		(0.018)		(0.019)
Alg. Advice	0.871***	0.877***	0.849***	0.847***	0.853***	0.856***
	(0.006)	(0.010)	(0.010)	(0.010)	(0.009)	(0.011)
AlgA × PerfPay		-0.006		-0.022		-0.034
		(0.014)		(0.019)		(0.023)
AlgA × Tournament		-0.015		0.019		0.005
		(0.013)		(0.019)		(0.020)
Observations	5,000	5,000	5,050	5,050	4,840	4,840
R-squared	0.937	0.937	0.938	0.938	0.938	0.938
N Participants	500	500	505	505	484	484

Notes: The dependent in all regressions is the true price of the apartment, and the constant is omitted. Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

TABLE A.3: OLS regressions of weight of advice across different periods

	Task 1		Tasks 2-5		Tasks 6-10	
	(1)	(2)	(3)	(4)	(5)	(6)
PerformancePay	0.143***	0.136***	0.093***	0.090***	0.090***	0.090***
	(0.035)	(0.035)	(0.024)	(0.024)	(0.024)	(0.024)
Tournament	0.045	0.048	0.057**	0.059**	0.070***	0.071***
	(0.035)	(0.035)	(0.024)	(0.024)	(0.023)	(0.023)
HumanAIAdvice	0.089***	0.089***	0.039	0.039	0.052**	0.050**
	(0.034)	(0.034)	(0.024)	(0.024)	(0.024)	(0.024)
HumanAIAdvice × PerfPay	-0.111**	-0.109**	-0.045	-0.044	-0.042	-0.041
	(0.049)	(0.049)	(0.035)	(0.035)	(0.034)	(0.034)
HumanAIAdvice × Tourn	-0.049	-0.058	-0.026	-0.032	-0.051	-0.051
	(0.048)	(0.048)	(0.033)	(0.033)	(0.033)	(0.033)
IsFemale		-0.017		-0.015		-0.005
		(0.021)		(0.014)		(0.014)
Age		0.001		0.001		0.000
		(0.002)		(0.002)		(0.002)
HasUnivDegree		-0.056**		-0.037**		-0.016
		(0.022)		(0.016)		(0.015)
IsAustrian		0.009		-0.006		-0.034**
		(0.021)		(0.015)		(0.015)
Constant	0.420***	0.422***	0.358***	0.359***	0.303***	0.333***
	(0.043)	(0.069)	(0.030)	(0.049)	(0.027)	(0.047)
Observations	972	972	3,886	3,886	4,853	4,853
R-squared	0.033	0.041	0.012	0.015	0.038	0.040
N Participants	972	972	989	989	989	989

Notes: Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for apartment, order, and subject pool fixed effects. Models 1 and 2 have fewer than 989 participants since for some of them WOA is not defined (initial estimate = algorithmic advice).

TABLE A.4: OLS REGRESSIONS OF TIME SPENT ON TASK ACROSS DIFFERENT PERIODS

	Task 1		Tasks 2-5		Tasks 6-10	
	All	Advice only	All	Advice only	All	Advice Only
PerformancePay	15.428***	24.012***	5.195***	8.136**	3.049**	3.552
	(3.706)	(6.920)	(1.791)	(3.272)	(1.369)	(2.357)
Tournament	13.451***	20.327***	4.163***	4.925*	3.750**	4.020*
	(3.823)	(7.699)	(1.598)	(2.745)	(1.493)	(2.324)
HumanAIAdvice		4.440		-1.579		0.015
		(5.079)		(2.262)		(2.150)
HumanAIAdvice × PerfPay		-21.442**		-1.915		1.156
		(9.398)		(4.551)		(3.426)
HumanAIAdvice × Tourn		-9.780		-0.239		0.128
		(10.200)		(3.984)		(3.799)
Constant	77.588***	72.329***	45.909***	47.650***	35.644***	35.369***
	(5.595)	(7.455)	(2.061)	(2.785)	(1.777)	(2.392)
Observations	1,489	989	5,956	3,956	7,445	4,945
R-squared	0.032	0.049	0.014	0.018	0.015	0.020
N Participants	1,489	989	1,489	989	1,489	989

Notes: Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for apartment, order, and subject pool fixed effects.

TABLE A.5: OLS regressions of estimation error across different periods

	Task 1		Tasks 2-5		Tasks 6-10	
	(1)	(2)	(3)	(4)	(5)	(6)
PerformancePay	-1.703	-0.668	-2.112	-5.710	0.063	0.602
	(2.522)	(6.939)	(1.801)	(5.173)	(1.587)	(4.563)
Tournament	-4.125*	-5.675	-4.196**	-10.417**	-1.241	-2.914
	(2.438)	(6.382)	(1.704)	(4.941)	(1.285)	(3.754)
AIAdvice	-17.672***	-17.006***	-23.348***	-28.861***	-24.069***	-24.403***
	(2.707)	(5.555)	(1.954)	(4.330)	(1.809)	(2.644)
HumanAIAdvice	-18.858***	-19.768***	-23.514***	-27.425***	-24.631***	-25.258***
	(2.682)	(5.436)	(2.004)	(4.358)	(1.786)	(2.638)
AIAdvice × PerfPay		-2.945		5.812		-1.313
		(7.540)		(5.360)		(4.641)
AIAdvice × Tourn		1.066		10.709**		2.317
		(6.884)		(5.005)		(3.820)
HumanAIAdvice × PerfPay		-0.726		4.316		-0.854
		(7.352)		(5.355)		(4.622)
HumanAIAdvice × Tourn		2.976		7.644		2.398
		(6.764)		(5.073)		(3.828)
IsFemale		-4.314*		-1.888		-1.175
		(2.339)		(1.524)		(1.420)
Age		-0.316		-0.176		-0.238**
		(0.210)		(0.145)		(0.112)
HasUnivDegree		-1.261		0.255		-0.347
		(2.205)		(1.509)		(1.346)
IsAustrian		-0.956		2.409*		1.268
		(2.045)		(1.378)		(1.147)
Constant	30.640***	42.169***	33.477***	40.378***	30.634***	36.675***
	(4.609)	(10.593)	(4.133)	(7.999)	(1.934)	(3.930)
Observations	1,489	1,489	5,956	5,956	7,445	7,445
R-squared	0.115	0.120	0.156	0.161	0.166	0.168
N Participants	1,489	1,489	1,489	1,489	1,489	1,489

Notes: Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for apartment, order, and subject pool fixed effects.

TABLE A.6: OLS Regressions of uncensored weight of advice on treatments

	(1)	(2)	(3)	(4)
PerformancePay	0.104***	0.107***	0.112***	0.112***
	(0.026)	(0.026)	(0.032)	(0.032)
Tournament	0.072***	0.074***	0.109***	0.111***
	(0.023)	(0.023)	(0.032)	(0.031)
HumanAIAdvice		0.040**	0.066**	0.065**
		(0.020)	(0.032)	(0.031)
HumanAIAdvice × PerfPay			-0.008	-0.007
			(0.052)	(0.052)
HumanAIAdvice × Tourn			-0.071	-0.074*
			(0.045)	(0.044)
IsFemale				0.006
				(0.020)
Age				0.003
				(0.003)
HasUnivDegree				-0.034
				(0.021)
IsAustrian				-0.036*
				(0.022)
Constant	0.588***	0.567***	0.553***	0.522***
	(0.044)	(0.044)	(0.046)	(0.068)
Observations	9,711	9,711	9,711	9,711
R-squared	0.012	0.013	0.014	0.015
N Participants	989	989	989	989

Notes: Dependent is the weight of advice, but not censored between 0 and 1. Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for apartment, order, and subject pool fixed effects.

TABLE A.7: OLS regressions of weight of advice across split-samples related to unfamiliarity and overconfidence

	Vienna knowledge		Airbnb experience		Task familiarity		Overconfident	
	Low	High	Low	High	Low	High	No	Yes
PerformancePay	0.097***	0.082**	0.142***	0.080***	0.061**	0.126***	0.148***	0.081***
	(0.027)	(0.034)	(0.045)	(0.024)	(0.027)	(0.035)	(0.038)	(0.026)
Tournament	0.051**	0.083**	0.079*	0.050**	0.022	0.126***	0.053	0.078***
	(0.026)	(0.032)	(0.042)	(0.023)	(0.023)	(0.036)	(0.035)	(0.025)
HumanAIAdvice	0.049*	0.042	0.047	0.049**	0.014	0.100***	0.056*	0.047*
	(0.026)	(0.034)	(0.043)	(0.024)	(0.026)	(0.034)	(0.032)	(0.026)
HumanAIAdvice × PerfPay	-0.066*	-0.002	-0.112*	-0.030	0.002	-0.112**	-0.085	-0.033
	(0.037)	(0.050)	(0.063)	(0.034)	(0.037)	(0.048)	(0.054)	(0.036)
HumanAIAdvice × Tourn	-0.026	-0.068	-0.059	-0.032	0.016	-0.117**	-0.067	-0.036
	(0.035)	(0.046)	(0.058)	(0.032)	(0.034)	(0.047)	(0.048)	(0.034)
IsFemale	0.000	-0.032	0.006	-0.018	-0.028*	0.012	-0.032	-0.007
	(0.015)	(0.021)	(0.026)	(0.014)	(0.015)	(0.020)	(0.021)	(0.014)
Age	0.000	0.001	-0.002	0.001	0.002	0.000	0.004	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
HasUnivDegree	-0.019	-0.038*	-0.030	-0.019	-0.025	-0.033	-0.044*	-0.020
	(0.016)	(0.023)	(0.028)	(0.015)	(0.016)	(0.021)	(0.025)	(0.016)
IsAustrian	-0.011	-0.026	-0.006	-0.029**	-0.034**	0.007	-0.035	-0.019
	(0.016)	(0.021)	(0.026)	(0.014)	(0.015)	(0.021)	(0.022)	(0.015)
Constant	0.449***	0.425***	0.524***	0.420***	0.431***	0.449***	0.458***	0.405***
	(0.052)	(0.070)	(0.070)	(0.049)	(0.049)	(0.075)	(0.077)	(0.049)
Observations	6,546	3,165	2,370	7,341	5,512	4,199	3,422	6,289
R-squared	0.029	0.039	0.040	0.031	0.031	0.040	0.037	0.035
N Participants	666	323	241	748	562	427	348	641

Notes: Dependent in all regressions is weight of advice. Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for apartment, order, and subject pool fixed effects.

TABLE A.8: OLS REGRESSIONS OF TIME SPENT ON TASK, WITH DEMOGRAPHICS

		Time spent on				
Dependent	Both estimates		Initial estimate			Final estimate
Conditions	All	Advice only	All	No advice	Advice only	Advice
Model	(1)	(2)	(3)	(4)	(5)	(6)
PerformancePay	5.224***	7.274***	5.036***	3.551	6.400***	0.874
	(1.566)	(2.792)	(1.417)	(2.454)	(2.374)	(0.672)
Tournament	4.861***	6.040**	4.325***	3.638	5.610***	0.430
	(1.538)	(2.548)	(1.376)	(2.383)	(2.134)	(0.732)
HumanAIAdvice		-0.361			0.361	-0.722
		(2.170)			(1.826)	(0.543)
HumanAIAdvice		-1.925			-1.906	-0.019
× PerfPay		(3.880)			(3.380)	(0.837)
HumanAIAdvice		-0.940			-1.881	0.941
× Tourn		(3.959)			(3.392)	(0.923)
IsFemale	2.889**	2.332	2.637**	3.497*	2.181	0.151
	(1.354)	(1.723)	(1.217)	(2.111)	(1.476)	(0.403)
Age	0.265*	0.412**	0.203	-0.030	0.304*	0.108***
	(0.149)	(0.189)	(0.129)	(0.200)	(0.166)	(0.037)
HasUnivDegree	-2.758*	-3.652**	-1.994	0.070	-3.188**	-0.464
	(1.452)	(1.835)	(1.304)	(2.348)	(1.592)	(0.394)
IsAustrian	-2.785**	-2.897	-2.201*	-2.766	-1.884	-1.013**
	(1.394)	(1.821)	(1.227)	(2.043)	(1.538)	(0.456)
Constant	85.420***	84.626***	69.623***	87.228***	60.900***	23.726***
	(4.121)	(5.460)	(3.637)	(6.070)	(4.675)	(1.379)
Observations	14,890	9,890	14,890	5,000	9,890	9,890
R-squared	0.181	0.183	0.138	0.192	0.118	0.243
N Participants	1,489	989	1,489	500	989	989

Notes: Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for apartment, order, and subject pool fixed effects.

TABLE A.9: OLS regressions of time spent on initial estimate

	(1)	(2)	(3)	(4)
Financial Incentives	3.437*		3.437*	
	(2.023)		(2.023)	
Algorithmic Advice	-5.774***	-5.771***		
	(1.777)	(1.777)		
Financial Incentives × Algorithmic Advice	1.643			
	(2.433)			
PerformancePay		3.229		3.229
		(2.442)		(2.442)
Tournament		3.665		3.665
		(2.396)		(2.396)
PerformancePay × Algorithmic Advice		2.366		
		(2.971)		
Tournament × Algorithmic Advice		0.919		
		(2.922)		
AIAdvice			-6.026***	-6.023***
			(1.951)	(1.951)
HumanAIAdvice			-5.547***	-5.544***
			(2.030)	(2.030)
Financial Incentives × AIAdvice			2.574	
			(2.716)	
Financial Incentives × HumanAIAdvice			0.646	
			(2.850)	
PerformancePay × AIAdvice				3.351
				(3.401)
Tournament × AIAdvice				1.772
				(3.208)
PerformancePay × HumanAIAdvice				1.263
				(3.431)
Tournament × HumanAIAdvice				0.046
				(3.522)
Constant	77.261***	77.311***	77.265***	77.316***
	(2.368)	(2.366)	(2.370)	(2.366)
Observations	14,890	14,890	14,890	14,890
R-squared	0.139	0.139	0.139	0.139
N Participants	1,489	1,489	1,489	1,489
Post-hoc F tests p-values				
FinInc + FinInc × Alg. Advice = 0	0.0002			
PerfPay + PerfPay × Alg. Advice = 0		0.0010		
Tourn + Tourn × Alg. Advice = 0		0.0061		
FinInc + FinInc × AIAdvice = 0			0.0009	
FinInc + FinInc × HumanAIAdvice = 0			0.0419	
PerfPay + PerfPay × AIAdvice = 0				0.0057
PerfPay + PerfPay × HumanAIAdvice = 0				0.0627
Tourn + Tourn × AIAdvice = 0				0.0109
Tourn + Tourn × HumanAIAdvice = 0				0.1499

Notes: Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for apartment, order, and subject pool fixed effects.

TABLE A.10: OLS REGRESSIONS OF ESTIMATION ERROR IN ALGORITHMIC ADVICE TREATMENTS, INCLUDING WEIGHT OF ADVICE AS AN INDEPENDENT VARIABLE

	(1)	(2)	(3)	(4)	(5)
PerformancePay	-0.323		-0.344	-0.100	-0.047
	(0.513)		(0.510)	(0.648)	(0.648)
Tournament	-0.781		-0.795	-0.291	-0.337
	(0.509)		(0.506)	(0.683)	(0.681)
HumanAIAdvice		-0.250	-0.275	0.197	0.161
		(0.413)	(0.409)	(0.748)	(0.754)
HumanAIAdvice × PerfPay				-0.458	-0.469
				(1.008)	(1.005)
HumanAIAdvice × Tourn				-1.000	-0.892
				(1.014)	(1.018)
Weight of Advice	-3.337***	-3.370***	-3.324***	-3.339***	-3.321***
	(0.481)	(0.486)	(0.481)	(0.486)	(0.487)
IsFemale = 1					0.511
					(0.443)
Age					-0.008
					(0.046)
HasUnivDegree = 1					0.261
					(0.439)
IsAustrian = 1					-0.172
					(0.484)
Constant	15.295***	15.038***	15.432***	15.190***	15.096***
	(0.978)	(0.953)	(0.984)	(1.018)	(1.560)
Observations	9,711	9,711	9,711	9,711	9,711
R-squared	0.204	0.204	0.204	0.204	0.205
N Participants	989	989	989	989	989

Notes: Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for apartment, order, and subject pool fixed effects.

TABLE A.11: OLS regressions of questionnaire answers (Table 6) with treatment interaction terms

	(1) Task enjoyment	(2) Algorithm is credible	(3) Use of algorithm	(4) Use of context info
PerformancePay	-0.579***	0.201	7.754***	0.528
	(0.218)	(0.209)	(2.652)	(1.576)
Tournament	-0.881***	0.400*	8.170***	-0.163
	(0.234)	(0.204)	(2.450)	(1.625)
AIAdvice	0.236			
	(0.205)			
AIAdvice × PerfPay	0.252			
	(0.298)			
AIAdvice × Tourn	1.044***			
	(0.300)			
HumanAIAdvice	0.557***	0.284	3.630	-0.066
	(0.187)	(0.202)	(2.499)	(1.611)
HumanAIAdvice × PerfPay	0.156	0.022	0.168	-2.402
	(0.300)	(0.279)	(3.639)	(2.435)
HumanAIAdvice × Tourn	0.457	-0.221	-2.384	1.289
	(0.298)	(0.274)	(3.488)	(2.357)
Constant	7.700***	5.988***	46.681***	82.166***
	(0.180)	(0.217)	(2.497)	(1.541)
Observations	1,489	989	989	989
R-squared	0.059	0.010	0.030	0.003

Notes: Robust standard errors in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for subject pool fixed effects.

B Exploring the relationship between weight of advice and final estimate error

In this appendix we aim to explore why a significant increase in weight of advice due to financial incentives did not translate into a statistically significant reduction in the final estimation error, despite higher advice utilization being beneficial to performance (see also the regressions reported in Table A.10). Our suspicion is that since the final estimation error depends on several factors such as initial estimates and weight of advice, heterogeneity in these may be responsible for the (mostly) non-significance of our results on the effects of financial incentives on final estimation errors.

Our regression model 1 below estimates treatment effects on initial estimates, to verify whether incentives (and thus increased time spent on the task) have negative or positive effects on the error of participants' own initial estimate. As a second step, our approach is to separately remove heterogeneity in either initial estimates or in weight of advice, and to explore how our analysis results respond. Namely, in models 2 and 3 below we reduce heterogeneity in weight of advice (by either assuming that all participants use the same average weight of advice in their treatment, or that they employ the estimated weight of advice, respectively) while in model 4 we reduce the heterogeneity in initial estimates (by assuming all participants use the same average initial estimate for that apartment).

In Table B.1 we report results from these four regressions using data from the advice conditions. All models use treatment indicators and their interactions as independents. In order to also single out the effect of *performance pay* and *tournament incentives* within the *Human AI advice* conditions, we also report results from post-estimation f-tests on whether the joint effects $PerformancePay$ + $HumanAIAdvice \times PerformancePay$ and $Tournament$ + $HumanAIAdvice \times Tournament$ are estimated to be significantly different from zero.

Model 1 regresses the error in the initial estimate (before receiving advice) on treatment indicators. The results show that there is no evidence that financial incentives would affect the error in initial estimates. The effects for *tournament incentives* are always statistically not significant, the effect of *performance pay* is not significant with regular AI advice and even statistically significantly negative in the *Human AI* condition. This is inconsistent with a potential mechanism whereby financial incentives at the same time increase error in initial estimates (before advice) and weight of advice, such that both would cancel out each other.

The other three models use simulated final estimates as dependent variables. For the "Simulated Final Estimate 1" (regressed in model 2), we take the individual estimates of participants, but assume that all participants in a treatment condition use a weight of advice equal to the average weight of advice in their treatment condition. That is, we take out any within-treatment heterogeneity in weight of advice. We find that under this simulation, *performance pay* reduces error in both advice conditions while the effects for *tournament pay* are negative but not significant.

The "Simulated Final Estimate 2" (regressed in model 3) reduces noise in weight of advice in a different way. Instead of using the "raw" weight of advice calculated from comparing the individual initial and final estimate, we use the estimated weight of advice for this observation (based on model 4 presented in Table 2) in order to augment the initial estimate and arrive at the final estimate. The results are almost the same as with the first simulation: *performance pay* reduces participants' estimate error but *tournament pay* does not.

Finally, the "Simulated Final Estimate 3" (regressed in model 4) removes a different kind of heterogeneity: the one in initial estimates. When calculating the final estimate for each participant and apartment, it uses the average initial estimate across all participants for this treatment condition and apartment, and augments it with the given advice using the participant's individual weight of advice for this task. Under this simulation of the final estimate, *performance pay* significantly reduces error in the *AI advice* condition but has no significant effect in the *human-AI condition*, while it is the other way around for *tournament pay*: it reduces error when the AI is human-augmented but not when it is not.

The three simulation models show that heterogeneity both in terms of weight of advice and initial estimates may be responsible for the statistical null result of financial incentives on estimation error, since removing these types of heterogeneity leads to significant effects (albeit not consistently).

TABLE B.1: OLS REGRESSIONS EXPLORING EFFECTS OF WEIGHT OF ADVICE ON ERROR IN FINAL ESTIMATE ERROR

	Error Initial Estimate	Error Simulated Final Estimate 1	Error Simulated Final Estimate 2	Error Simulated Final Estimate 3
	(1)	(2)	(3)	(4)
PerformancePay	-0.595	-1.459**	-1.335**	-0.335**
	(1.119)	(0.624)	(0.585)	(0.138)
Tournament	0.662	-0.546	-0.394	-0.056
	(1.275)	(0.728)	(0.692)	(0.142)
HumanAIAdvice	1.721	0.060	0.014	-0.274*
	(1.345)	(0.764)	(0.705)	(0.141)
HumanAIAdvice × PerfPay	-2.163	-0.265	-0.241	0.536***
	(1.714)	(0.935)	(0.872)	(0.199)
HumanAIAdvice × Tourn	-2.772	-0.767	-0.856	-0.631***
	(2.003)	(1.108)	(1.032)	(0.204)
Constant	33.634***	19.714***	17.600***	3.457***
	(2.555)	(1.423)	(1.268)	(0.197)
Observations	9,890	9,890	9,890	9,711
R-squared	0.103	0.166	0.176	0.792
N Participants	989	989	989	989

Notes: Robust standard errors clustered by participant. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. All regressions control for apartment, order, and subject pool fixed effects.

C Additional survey study on credibility perceptions of different forms of human-AI collaborative advice

C.1 Study Design

To provide additional insights into how participants respond to the human-AI framing of the algorithmic advice, we conducted a follow-up study. Specifically, in an experimental vignette study we examine whether participants' perceived credibility of the algorithmic advice is influenced by how the human-centered AI advice is generated. We contrast the description of the human-centered AI algorithm from our main experiment (our baseline) with five additional treatments. Each treatment describes a different mechanism of how the human experts interacted with the AI algorithm to generate the eventual human-centered AI advice.

In order to be able to freely confront participants with different mechanisms of how the human experts interacted with the AI algorithm to generate the eventual human-centered AI advice, while at the same time avoiding any deception, we employ a vignette scenario study. We first familiarized our survey participants with the Airbnb price estimation task by presenting them with an apartment listing (which as in the main experiment includes a photo, a description, the average review scores, and a map with the location) and asking them to estimate the price per night of this apartment. Every participant saw the same apartment, and the estimate was not incentivized. The main purpose of this exercise was to have participants immerse themselves into the task in order to get a feeling for the challenges encountered.

Then, we presented participants with a vignette scenario. They were asked to imagine that they no longer have to do the task on their own, but that they receive advice from a human-centered AI algorithm. We provide participants with the following description of the human-centered AI advice, which is identical to our main experiment:

Human-centered AI considers not just numbers but also human experience and advice. In fact, the human-centered AI advice in our study consists of two elements: (1) a random forest model trained on a large data set and (2) the expertise of 5 individuals who are familiar with the real-estate sector in Vienna.

(1) The random forest model is trained on a real dataset of approximately 12,000 Airbnbs in Vienna. In particular, the random forest model generates estimates of Airbnb listing prices based on an ensemble learning method for regression that operates by constructing a multitude of decision trees and returns the average estimate of the individual trees. The model takes into account the following input variables: room type (apartment, private room), number of bedrooms, number of beds, number of accommodated guests, district of Vienna, number of reviews, average review rating, and whether host is a superhost or not.

(2) In addition to this, the price estimate of the human-centered AI incorporates advice from 5 experts. The 5 experts have substantial experience in the pricing of Airbnb apartments and are familiar with the housing and accommodation sector in Vienna.

Our baseline scenario uses the same human-centered advice description as in our main experiment, such we do not provide any other information to participants. In five additional treatments we add another sentence to detail how the price prediction by the human experts is combined with the price prediction by the AI algorithm. The design of our new treatments is based on the hybrid intelligence concepts developed by Dellermann et al. (2019). Between-participants, we randomly assign the following additional explanations as treatment conditions:

- **Treatment 0 (baseline):** No additional text.
- **Treatment 1 (50/50 weighting):** *The eventual advice of the human-centered AI algorithm is the weighted average of the prediction of the random forest model (50% weight) and the average prediction of the 5 human experts (50% weight).*
- **Treatment 2 (80/20 weighting):** *The eventual advice of the human-centered AI algorithm is the weighted average of the prediction of the random forest model (80% weight) and the average prediction of the 5 human experts (20% weight).*
- **Treatment 3 (Human adjusts AI):** *The eventual advice of the human-centered AI algorithm is the result of a process where the random forest model first made an initial prediction which was provided to the 5 human experts, who then made the final prediction.*
- **Treatment 4 (AI adjusts human):** *The eventual advice of the human-centered AI algorithm is the result of a process where the 5 human experts first made their predictions which were then provided to the algorithm, which made the final prediction.*
- **Treatment 5 (Human-AI collaboration):** *The eventual advice of the human-centered AI algorithm is the result of a process where the 5 human experts and the algorithm interacted with each other to make the final prediction.*

As our main dependent variable, we ask participants to what extent they would find the advice of such a human-centered AI algorithm to be a credible source for estimating the price of an Airbnb apartment. This question on source credibility, which is based on Chen et al. (2022), is identical to an item in our post-experiment questionnaire of the main experiment. Participants respond on a scale from 0 (to no extent) to 10 (to a very large extent). After the measurement of our dependent variable, on the next survey page we employed an ex-post recall check by asking participants to remember the exact description of the humanAI algorithm. The survey ended with a few brief questions on demographics.

We randomized the presentation format since it may affect how rather complex information is processed by participants. Half of the participants received all the information on the human-centered AI algorithm on one single page, while the remaining half received the information sequentially on multiple screens. We control for this information provision format in our subsequent analyses.

We recruited participants from a large public university in Austria via a university-wide survey mailing list. In total, we received 1601 responses, from which we excluded 72 participants for suspected double participation from the same IP address, and 13 participants who did not provide their contact details upon the study's completion. Thus, our final dataset includes observations from 1,516 participants. In terms of demographics, the average participant age is 22.3 years and 52% identify as female. About one third are masters degree students, and the remaining two thirds are undergraduate students in business, economics, or law. 54% are Austrian nationals, while more than 70 different nationalities are present in the remaining half. As a "thank you" for participating in our survey study, participants could enter a lottery, with five winners each receiving EUR 100.

C.2 Results

In Table C.1, we present descriptive statistics on participants' perceived credibility in the human-centered AI advice. Columns 1-3 present descriptives for the full sample, while columns 4-6 report aggregates only for those participants who passed the ex-post recall check. Due to the complex nature and lengthy description of the human-centered AI algorithms, a rather large fraction of participants failed the ex-post recall check in our study. However, we do not observe any significant differences in our study results between the two sample groups.

Table C.2 reports results from OLS regressions of the perceived credibility of the human-centered AI advice on our five treatments conditions. Columns 1-3 refer to the whole sample, columns 4-6 only consider participants who passed the ex-post recall check. Columns 1 and 4 are simple OLS models, columns 2 and 5 control for the presentation format (sequential vs. simultaneous), and columns 3 and 6 controls for participant demographics. Across all specifications, we do not observe any statistically significant differences between our baseline treatment and the five additional treatments. Also across the five treatments, we do not observe a clear pattern that some human-AI collaborations would be rated higher or lower on credibility than others. In fact, across all treatments, participants rate the advice credibility with the same value of around 6 out of 10.

TABLE C.1: Average credibility across treatments

	Full sample			Passed recall check		
	N	Avg.	(StdDev)	N	Avg.	(StdDev)
Treatment 0: Baseline	259	6.205	(1.765)	259	6.205	(1.765)
Treatment 1: 50/50 weighting	247	6.170	(1.690)	185	6.249	(1.682)
Treatment 2: 80/20 weighting	255	6.157	(1.716)	210	6.181	(1.738)
Treatment 3: Human adjusts AI	262	6.137	(1.750)	132	6.136	(1.716)
Treatment 4: AI adjusts human	261	6.061	(1.729)	94	5.989	(1.569)
Treatment 5: Human-AI coll.	232	6.091	(1.806)	87	6.023	(1.874)
All	1,516	6.137	(1.740)	967	6.161	(1.726)

This suggests that the way in which the human expert predictions and the AI predictions are combined does not meaningfully affect the perceived credibility of the human-centered AI advice. Thus, we conclude that our inferences from the main experiment are likely generalizable to many different settings in practice, regardless of how specifically human and machine expertise is combined to generate a human-centered AI advice.

TABLE C.2: Average credibility across treatments

	Full sample			Passed recall check		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment = 1	-0.035	-0.034	-0.030	0.044	0.039	0.045
	(0.155)	(0.155)	(0.154)	(0.166)	(0.167)	(0.166)
Treatment = 2	-0.048	-0.054	-0.048	-0.024	-0.031	-0.033
	(0.154)	(0.154)	(0.153)	(0.161)	(0.161)	(0.160)
Treatment = 3	-0.067	-0.066	-0.068	-0.068	-0.075	-0.087
	(0.153)	(0.153)	(0.152)	(0.185)	(0.185)	(0.185)
Treatment = 4	-0.143	-0.140	-0.158	-0.215	-0.213	-0.232
	(0.153)	(0.153)	(0.152)	(0.208)	(0.208)	(0.208)
Treatment = 5	-0.114	-0.114	-0.115	-0.182	-0.177	-0.184
	(0.157)	(0.157)	(0.157)	(0.214)	(0.214)	(0.214)
Sequential		0.098	0.089		0.089	0.072
		(0.090)	(0.089)		(0.112)	(0.112)
Age			-0.017			-0.006
			(0.014)			(0.018)
Austrian			-0.058			-0.032
			(0.091)			(0.113)
Female			-0.274***			-0.254**
			(0.090)			(0.111)
MA degree			0.257***			0.251**
			(0.099)			(0.126)
Constant	6.205***	6.155***	6.364***	6.205***	6.160***	6.104***
	(0.108)	(0.117)	(0.276)	(0.107)	(0.121)	(0.342)
Observations	1,516	1,516	1,516	967	967	967
R-squared	0.001	0.002	0.013	0.002	0.003	0.015

Notes: The dependent is the credibility rating [1-10] of the humanAI algorithm. Columns 1-3 report results from the full sample, columns 4-6 only include participants who passed the recall check. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

D The algorithm underlying the algorithmic advice in the experiment

The algorithm used for the experimental stimuli (the algorithmic advice) is based on a random forest model which is trained on a real dataset obtained from *www.insideairbnb.com*. The raw data set of Vienna listings was originally scraped on 09.06.2021 and contained 11,567 observations. The model generates estimates of Airbnb listing prices based on a random forest model – an ensemble learning method for regression that operates by constructing a multitude of decision trees and returns the average estimate of the individual trees. We include the full code of the model in the replication package.

D.1 Data cleaning and variables

The dataset contains the following information: numerical ID (id), weblink to actual Airbnb listing (listing_url), listing title (name), description of Apartment (description), title picture of the apartment (picture_url), yes/no indication if host is an "Airbnb Superhost" (host_is_superhost), yes/no indication if host's identity was verified (host_identity_verified), the district in Vienna where the apartment is located (neighbourhood_cleansed), an indicator whether an entire home/apartment or just a room was offered (room_type), the number of guests than can be accommodated (accommodates), the number of bathrooms (bathrooms_text), the number of bedrooms (bedrooms), the number of beds (beds), the apartment base price per night (price), the total number of reviews (number_of_reviews), the number of reviews last three months (number_of_reviews_ltm), the number of reviews last 30 days (number_of_reviews_l30d), the date of the first review (first_review), the date of the last review (last_review), the overall review score (review_scores_rating), and the review scores on sub-categories accuracy, cleanliness, check-in, communication, location, and value (review_scores_XX).

Before the dataset is used in the model, it is cleaned and reduced. In particular, observations of hotel rooms or shared rooms, where superhost status was not known, where prices were outliers (prices smaller than 30 USD and above 300 USD), or where hosts were very inexperienced (less than 5 reviews) were dropped. The final dataset for training the algorithm consists of 5,426 observations.

The algorithm takes into account the following input variables: room type (apartment, private room), number of bedrooms, number of beds, number of accommodated guests, district of Vienna, number of reviews, average review rating, and whether host is a superhost or not. The algorithm deploys an 80/20 stratified sampling split, so 80% of the data was used as a training set and 20% as a test set.

D.2 Model specifications and selection

Our algorithm compares decision-tree models and random forest models. It is implemented in *R*. First, we perform grid-based hyperparameter search for a decision tree model. We use a grid of potential values for the hyperparameter(s) that we want to try – we specify potential values for the hyperparameters, and the *tune_grid()* function builds separate models using these values. We compare how well the individual values work on the cross-validated data set and select the "best" set of hyperparameters to predict the separate test data set. To include grid search in our pipeline, we specify three things: a) in the *decision_tree()* function we specify which hyperparameters we want to tune, b) we define a grid of hyperparameter values and provide it to the *fit* function, c) we use the function *tune_grid()* to tell *tidymodels* that we want to tune the previously defined hyperparameters. We then select the hyperparameters based on the best Root Mean Squared Errors, and predict the test set.

Second, we use random search as an efficient alternative to grid search in decision tree modeling. In random search, we switch our grid of hyperparameter values to *grid_random*, and *R* creates random values for the hyperparameters to try. We specify a random grid using 10 different combinations of values. We then compare the results of the random search with the results of the grid search. We select the best hyperparameters according to the metric Root Mean Squared Error and predict the test set. Then we compare our decision-tree models based on grid and random search.

Third, we specify a random forest model. Random Forests are a type of bagging approach where multiple, independent decision trees are built on separate, independent, bootstrapped data sets. Random forests improve upon standard bagging approaches by de-correlating the individual trees. If we have a set of very strong predictors in the data set, these predictors will be used over and over by the separate trees. To prevent this, random Forests randomly select a subset of predictors at each step that are considered for determining the next split in the tree. This guarantees that "less important" predictors are chosen sometimes. In *R*, this hyperparameter is called *mtry*. We try different values for this parameter and select the one that works best for our dataset. Additional parameters that we tune are *trees* (the number of trees) and *min_n* (minimum number of data points). To specify cross-validation, we use the *vfold_cv* function. We predict the test set.

Finally, we can compare the predictions of decision tree and random forest models on the test set. We a) sort the results according to a specific metric Root Mean Squared Error, b) extract the best model type, c) extract the best hyperparameters for this model type, d) re-train the final model using the best hyperparameters, and e) predict the test data. Ultimately the random forest model is chosen based on this metric.

E Screenshots of experimental instructions and decision screens

Screenshot 1: Welcome screen, payment information, and consent form

WELCOME!

Thank you for participating. In this study, we try to better understand how people form economic judgements and make decisions.

The study will take around **15-20 minutes**. Your main task will be to submit price estimates for past Airbnb listings in Vienna. This will be followed by a questionnaire.

100 individuals out of all participants in this study will be selected for payment. If you are one of these 100 randomly selected individuals, you can earn up to **EUR 100**. (The average payment for paid participants will be around EUR 50.) The study is funded by WU Vienna.

If you are one of the 100 participants chosen for payment, we will inform you by email about your payment. In order to do so, we will ask you to provide us with your name and email address at the end of the study. The randomly selected participants will then be asked for their IBAN account number, and paid by WU via IBAN bank transfer. Once this is completed, any identifying data will be removed from the dataset, including names, email addresses, and IBAN numbers. Only the WU Financial Accounting department (Finanzbuchhaltung) will keep records of the transfers as required by law.

Irrespective of you being chosen for payment or not, all your answers and decisions in this task will be treated **confidentially**. Results of this study will only be presented in aggregated form.

Your participation is **completely voluntary** and you can stop participating at any time and for any reason or for no reason at all. (However, you will not be eligible for payment if you do not complete the study.)

Contact: If you have any questions or comments about this study, please feel free to contact Georg Lintner at georg.lintner@wu.ac.at.

Please indicate below that you agree to participate in our study:

- ○ I agree to participate.
- ○ I do not agree to participate.

Screenshot 2: Filtering question

In the next 15-20 minutes, I am able to fully concentrate on this study.

- ○ Yes.
- ○ No.

Screenshot 3: Ground rules of the study

Before you start, please consider the following ground rules:

- **Participating multiple times in the study is strictly forbidden!** Doing so will lead to an immediate disqualification and you will not receive any payment!
- Please **do not communicate** with other individuals while participating in the study!
- **We promise not to misrepresent facts or deceive you in any way!** All the information provided in the study is 100% true and your performance as well as your payments will be determined exactly as described.

Screenshot 4: Basic task description

Your task

Your task in this experiment will be to **estimate the price per night** of **10 different Airbnbs in Vienna**. The price per night in our task is defined as the **base price excluding any additional fees** (i.e. excl. Airbnb service fee and/or cleaning fees).

“Airbnbs” are apartments or private rooms rented for short-term tourism or business stays via the online platform Airbnb.com.

The 10 different Airbnb listings were **real listings in Vienna in June 2021**, but they are **currently not active** anymore. Hence, looking for the prices online will not be effective, and we ask you not to do so for time reasons. We are interested in **your** estimates.

Screenshot 5: Detailed task description and comprehension questions

Airbnb price estimation

To make your estimate, we will provide you with the actual Airbnb listing in Vienna, which includes a photo, a brief description of the apartment, the approximate location, the average rating by previous guests, and some information on the host. You will see an example for a typical Airbnb listing in a moment.

The label "entire apartment" refers to renting the whole apartment - no facilities and rooms are shared with the host or other guests. The "private room" label refers to renting a room in a shared apartment. Kitchen and bathroom are usually shared with the host or other guests.

Please note that there might be some **attention checks** included in the task. An attention check will look similar to the 10 apartment listings, but will ask you in the text to submit a particular number. In case you fail an attention check, you will not receive any payment.

For our research purposes we rely on you that you try to make as accurate price estimates as possible.

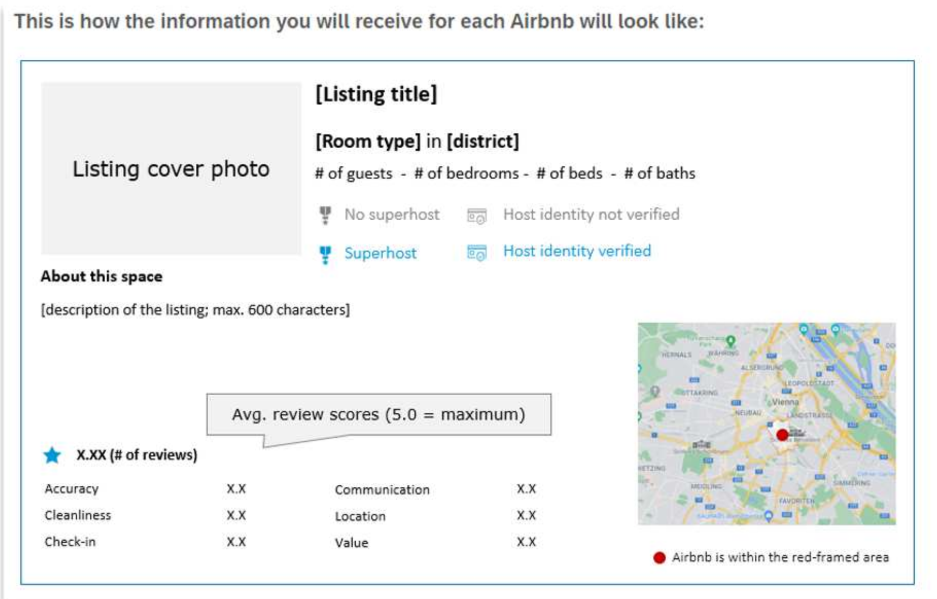
Your task is to estimate the price per night of 10 Airbnb listings in Vienna.

- True.
- False.
- I don't know.

Your task is to make as accurate price estimates as possible.

- True.
- False.
- I don't know.

Screenshot 6: Example visualization of an Airbnb listing



Screenshot 7: Detailed instructions on the AI Algorithm (only AI treatments)

Decision Support by an Artificial Intelligence (AI) Algorithm

To support you in making your estimates, we will also provide you with an **advice that was generated by an artificial intelligence algorithm.**

The algorithm itself is based on a **random forest model** which is trained on a real dataset of approximately 12,000 Airbnbs in Vienna (as of June 2021). In particular, the random forest model generates estimates of Airbnb listing prices based on an ensemble learning method for regression that operates by constructing a multitude of decision trees and returns the average estimate of the individual trees. The algorithm takes into account the following input variables: room type (apartment, private room), number of bed rooms, number of beds, number of accommodated guests, district of Vienna, number of reviews, average review rating, and whether host is a superhost or not.

However, please be advised that **the algorithm is not perfect. Its estimates can be above or below the actual listing price.** For some Airbnbs, the algorithm produces relatively accurate estimates, for some others slightly deviating estimates, and for some Airbnbs its estimates can be far off the actual listing price. Based on previous analysis, on average the **algorithm is about 30% off the true price**.

For each of the 10 Airbnb listings, you will first be asked to make an **initial price estimate without having the algorithmic advice.** Only then you will receive the advice from the AI algorithm, and you are asked to make a **second price estimate.** The second estimate can be equal or different to your first estimate. So for each Airbnb listing you will be asked to make two estimates.

The AI algorithm is trained on a real dataset of Airbnb listings in Vienna.

- ○ True.
- ○ False.
- ○ I don't know.

The AI algorithm is based on a

- ○ Random forest algorithm.
- ○ Support vector machine.
- ○ Simple linear regression.

The average estimation error of the AI algorithm is

- ○ 0%
- ○ 30%
- ○ 50%

Screenshot 8: Detailed instructions on the Human-framed AI Algorithm

Decision Support by a Human-Centered Artificial Intelligence (AI) Algorithm

To support you in making your estimates, we will also provide you with an **advice that was generated by a human-centered artificial intelligence algorithm**. Human-centered AI takes into account not just numbers but also human experience and advice.

The algorithm itself is based on a **random forest model** which is trained on a real dataset of approximately 12,000 Airbnbs in Vienna (as of June 2021). In particular, the random forest model generates estimates of Airbnb listing prices based on an ensemble learning method for regression that operates by constructing a multitude of decision trees and returns the average estimate of the individual trees. The algorithm takes into account the following input variables: room type (apartment, private room), number of bed rooms, number of beds, number of accommodated guests, district of Vienna, number of reviews, average review rating, and whether host is a superhost or not.

In addition to this, the price estimate incorporates expert advice from 5 individuals. The **5 experts have substantial experience** in the pricing of Airbnb apartments and are familiar with the housing and accommodation sector in Vienna.

However, please be advised that **the human-centered algorithm is not perfect. Its estimates can be above or below the actual listing price.** For some Airbnbs, the algorithm produces relatively accurate estimates, for some others slightly deviating estimates, and for some Airbnbs its estimates can be far off the actual listing price. Based on previous analysis, on average the **algorithm is about 30% off the true price.**

For each of the 10 Airbnb listings, you will first be asked to make an **initial price estimate without having the algorithmic advice**. Only then you will receive the advice from the AI algorithm, and you are asked to make a **second price estimate**. The second estimate can be equal or different to your first estimate. So for each Airbnb listing you will be asked to make two estimates.

The estimate by the decision support system is generated by a human-centered AI algorithm.

- ○ True.
- ○ False.
- ○ I don't know.

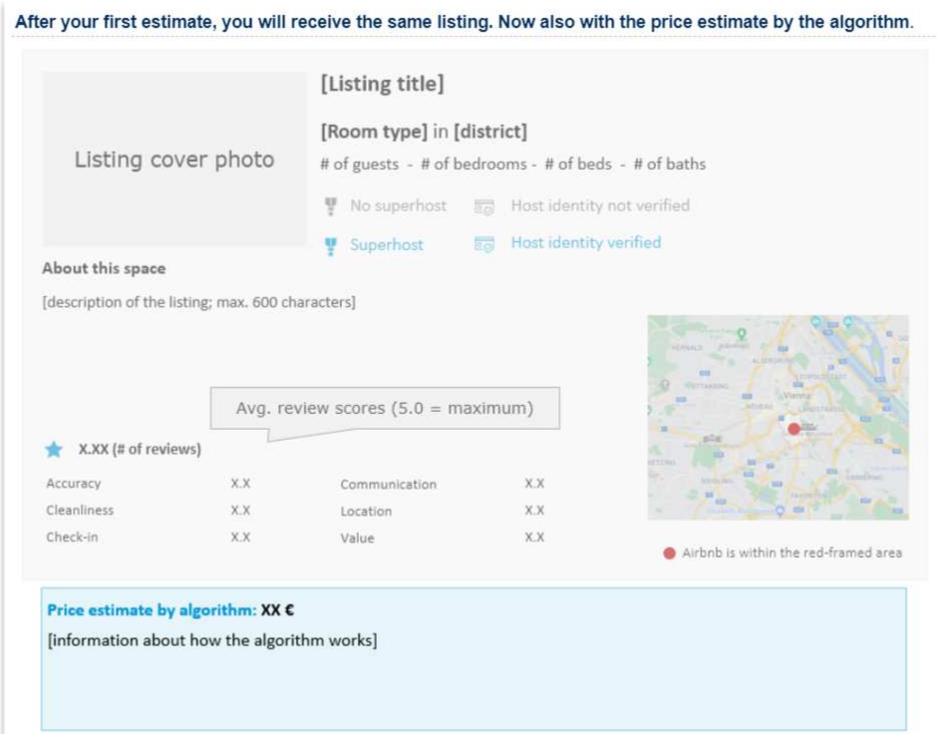
The human-centered AI algorithm is trained on

- ○ A dataset of 12,000 Airbnb listings.
- ○ Advice of 5 experts.
- ○ Both: a dataset of 12,000 Airbnb listings and the advice of 5 experts.

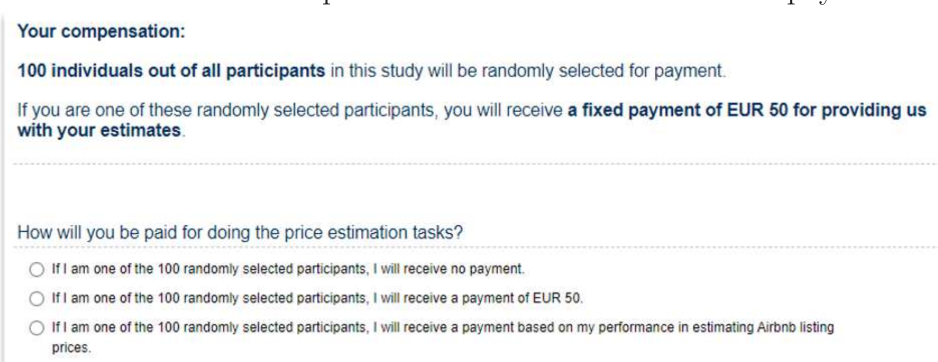
The average estimation error of the AI is

- ○ 0%
- ○ 30%
- ○ 50%

Screenshot 9: Example of Airbnb listing with algorithmic advice



Screenshot 10: Compensation contract details – fixed payment



Your compensation:

100 individuals out of all participants in this study will be randomly selected for payment.

If you are one of these randomly selected participants, you will receive **a fixed payment of EUR 50 for providing us with your estimates**.

How will you be paid for doing the price estimation tasks?

- ○ If I am one of the 100 randomly selected participants, I will receive no payment.
- ○ If I am one of the 100 randomly selected participants, I will receive a payment of EUR 50.
- ○ If I am one of the 100 randomly selected participants, I will receive a payment based on my performance in estimating Airbnb listing prices.

Screenshot 11: Compensation contract details – tournament incentives (as shown to treatment groups without algorithmic advice)

Your Compensation:

100 individuals out of all participants in this study will be randomly selected for payment.

If you are one of these randomly selected participants, you can earn **EUR 100 conditional on your performance compared to another participant**. We will randomly select one of your 10 price estimates. You will be randomly matched with another participant who is also selected for payment, and your price estimate is compared to the other participant's price estimate. **If your estimate is closer to the true price, you receive EUR 100** and the other participant receives EUR 0. Vice versa, if the other participant's estimate is closer to the true price, then the other participant receives EUR 100 and you receive EUR 0. If you submitted exactly the same price estimate, then it will be randomly determined who receives EUR 100 and who receives EUR 0.

How will you be paid for doing the price estimation tasks?

- ○ If I am one of the 100 randomly selected participants, I will receive no payment.
- ○ If I am one of the 100 randomly selected participants, I will receive a payment of EUR 50.
- ○ If I am one of the 100 randomly selected participants, I will receive a payment of EUR 100 if my price estimate is more accurate than the price estimate of another randomly matched participant.

Screenshot 12: Compensation contract details – performance-based incentives (as shown to treatment groups with algorithmic advice)

Your Compensation:

100 individuals out of all participants in this study will be randomly selected for payment.

If you are one of these randomly selected participants, **you can earn EUR 100 for your price estimation**. The computer will randomly select one of your 20 price estimates (you will make 2 estimates for each listing), and your likelihood to receive a payment of **EUR 100** will depend on how close your price estimate is to the true price. **The better your price estimate, the higher are your chances to receive the EUR 100**, such that your optimal action is to submit your best estimate of the price.

In particular, we will calculate a score between 0 and 100 as follows:

$$Score = \max\left(100 - 0.2 \times (Estimate - TruePrice)^2, 0\right)$$

Based on this function, you can approximately expect the following scores contingent on how much your estimate deviates from the true price:

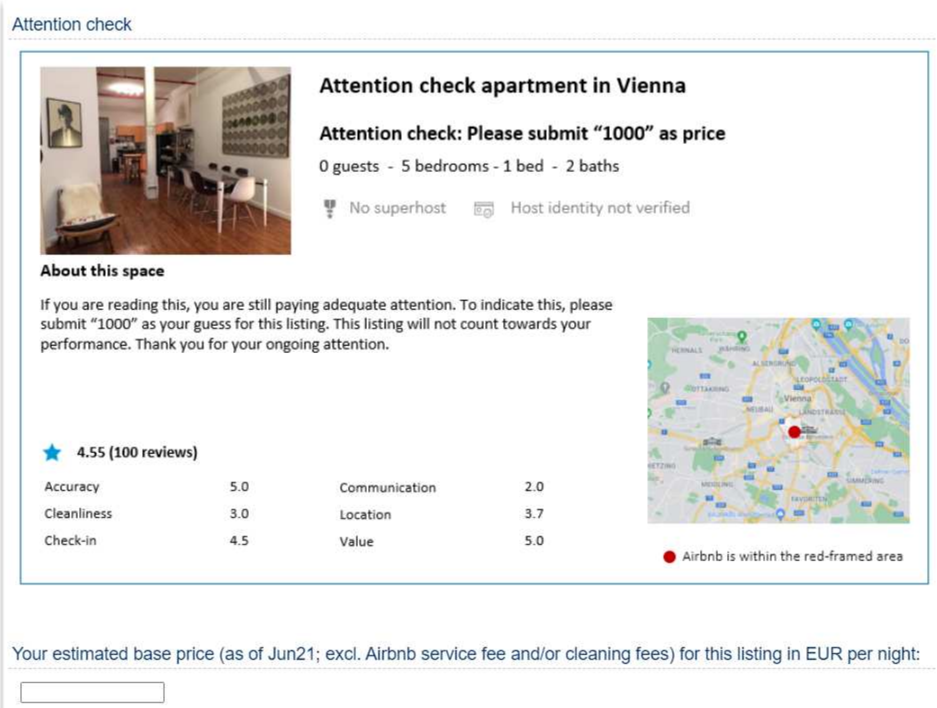
Estimate deviation	Score (=likelihood to receive EUR 100)
EUR 0	100
EUR 5	95
EUR 10	80
EUR 15	55
EUR 20	20
EUR 22	12
EUR 24	3
>=EUR 25	0

The computer will randomly draw a number between 0 and 100 (with each number being equally likely). If your score is higher than or equal to that random number, you will receive the EUR 100. If your score is lower than the random number, you will receive EUR 0 (nothing). Thus, your score directly represents your likelihood (in percent) to receive EUR 100, and the higher your score, the higher are your chances to receive the EUR 100.

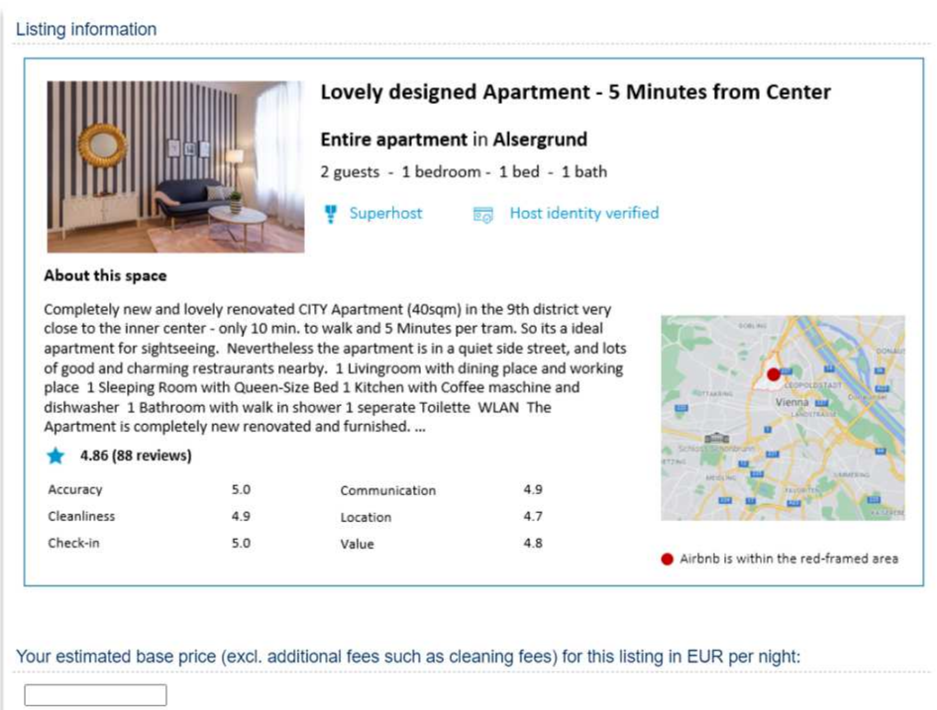
How will you be paid for doing the price estimation tasks?

- ○ If I am one of the 100 randomly selected participants, I will receive no payment.
- ○ If I am one of the 100 randomly selected participants, I will receive a payment of EUR 50.
- ○ If I am one of the 100 randomly selected participants, I will receive a payment based on my performance in estimating Airbnb listing prices.

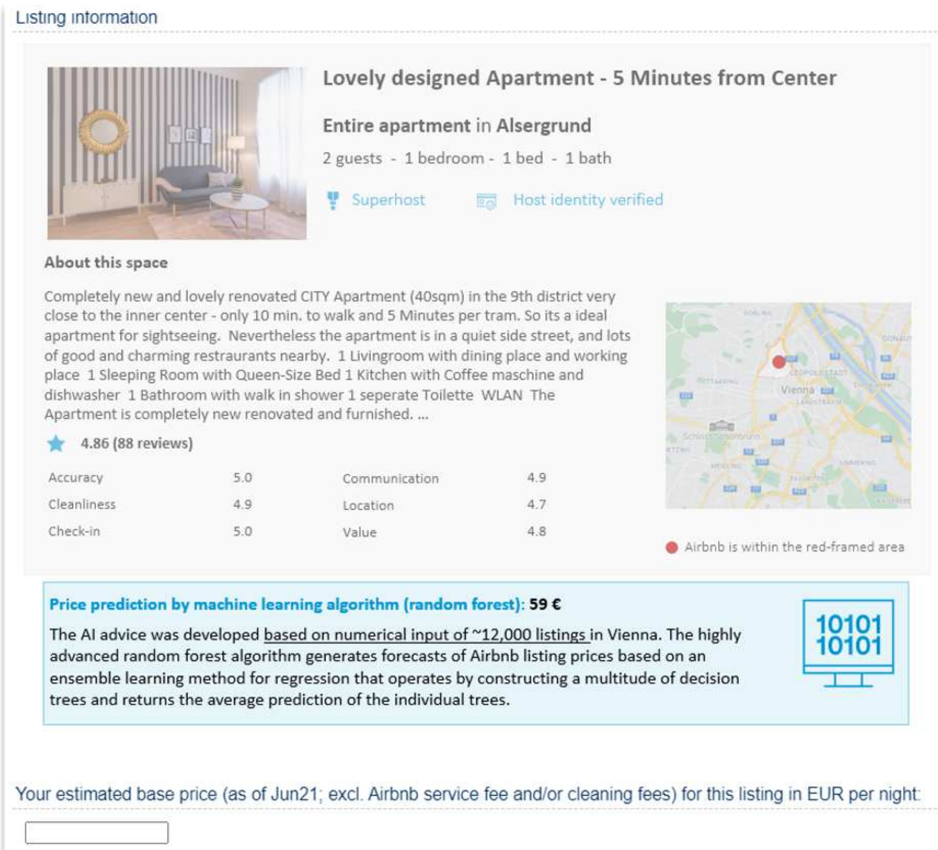
Screenshot 13: Attention check (same for all treatment groups)



Screenshot 14: Example Airbnb listing – without algorithmic advice (treatment groups)



Screenshot 15: Example Airbnb listing – with AI advice



Screenshot 16: Example Airbnb listing – with human-framed AI advice

Listing information

Lovely designed Apartment - 5 Minutes from Center

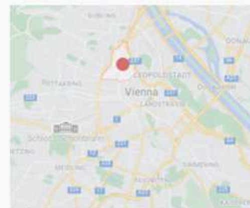
Entire apartment in **Alsergrund**

2 guests - 1 bedroom - 1 bed - 1 bath

Superhost   Host identity verified

About this space

Completely new and lovely renovated CITY Apartment (40sqm) in the 9th district very close to the inner center - only 10 min. to walk and 5 Minutes per tram. So its a ideal apartment for sightseeing. Nevertheless the apartment is in a quiet side street, and lots of good and charming restraurants nearby. 1 Livingroom with dining place and working place 1 Sleeping Room with Queen-Size Bed 1 Kitchen with Coffee maschine and dishwasher 1 Bathroom with walk in shower 1 seperate Toilette WLAN The Apartment is completely new renovated and furnished. ...



Airbnb is within the red-framed area

4.86 (88 reviews)

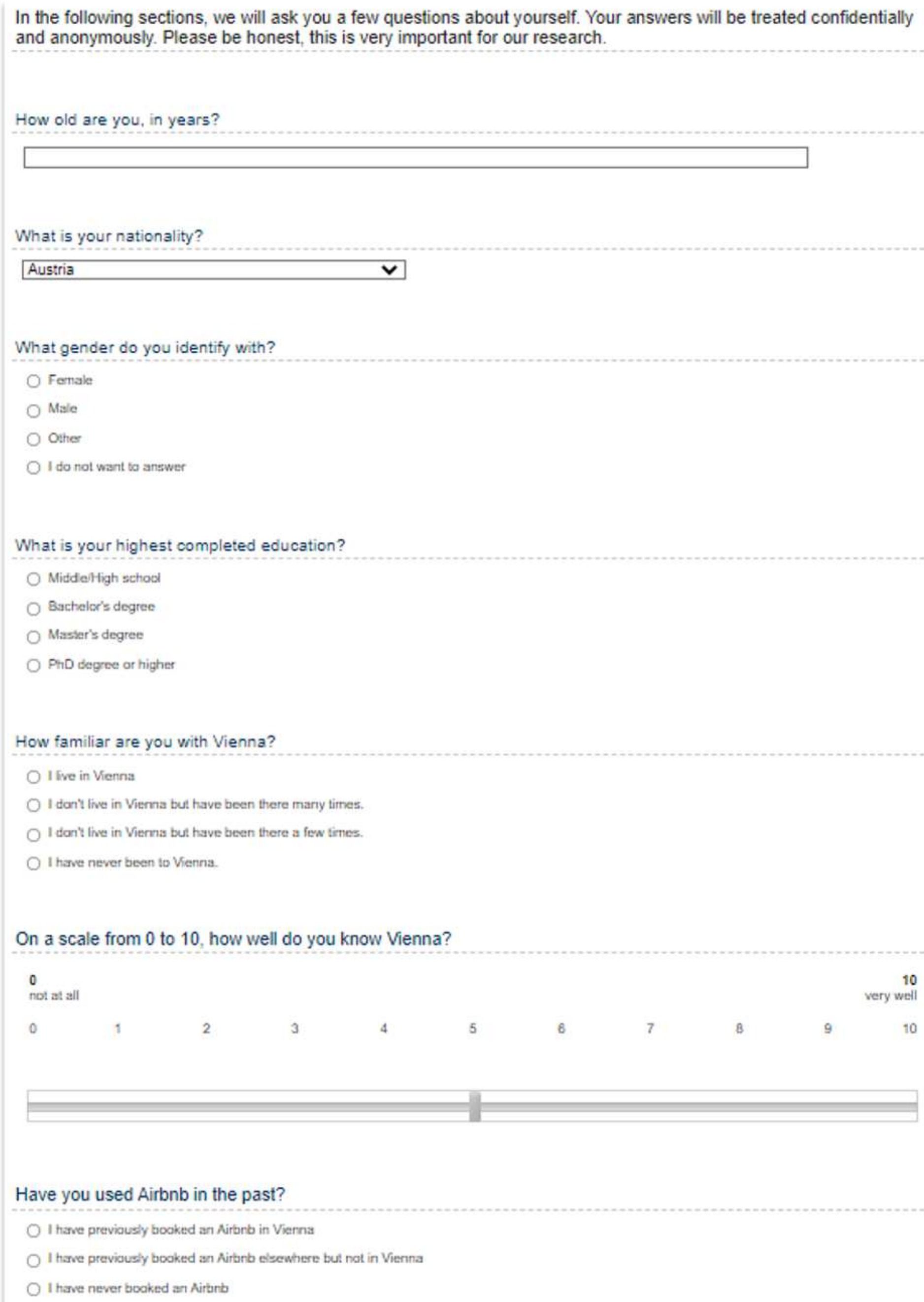
Accuracy	5.0	Communication	4.9
Cleanliness	4.9	Location	4.7
Check-in	5.0	Value	4.8

Price prediction by human-centered AI system: 59 €

The AI advice was developed based on numerical input of ~12,000 listings in Vienna as well as predictions by human experts. Human-centered AI is defined by systems that are continuously improving because of human input while providing an effective experience between human and algorithm.

Your estimated base price (as of Jun21; excl. Airbnb service fee and/or cleaning fees) for this listing in EUR per night:

Screenshot 17: Post experiment questionnaire – demographics



In the following sections, we will ask you a few questions about yourself. Your answers will be treated confidentially and anonymously. Please be honest, this is very important for our research.

How old are you, in years?

What is your nationality?

Austria

What gender do you identify with?

- Female
- Male
- Other
- I do not want to answer

What is your highest completed education?

- Middle/High school
- Bachelor's degree
- Master's degree
- PhD degree or higher

How familiar are you with Vienna?

- I live in Vienna
- I don't live in Vienna but have been there many times.
- I don't live in Vienna but have been there a few times.
- I have never been to Vienna.

On a scale from 0 to 10, how well do you know Vienna?

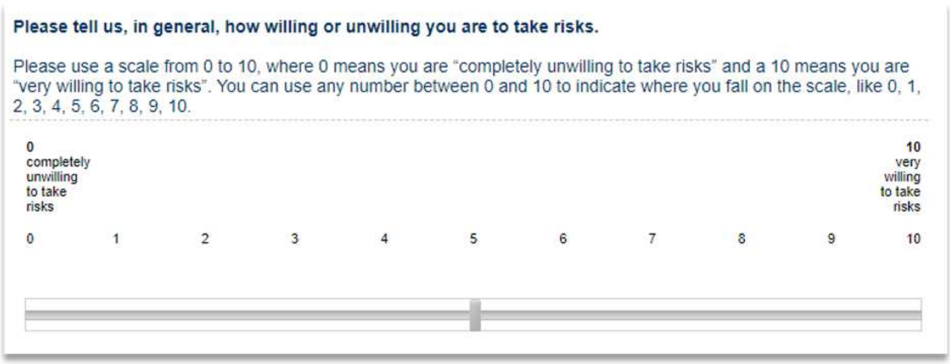
0 not at all — 10 very well

0 1 2 3 4 5 6 7 8 9 10

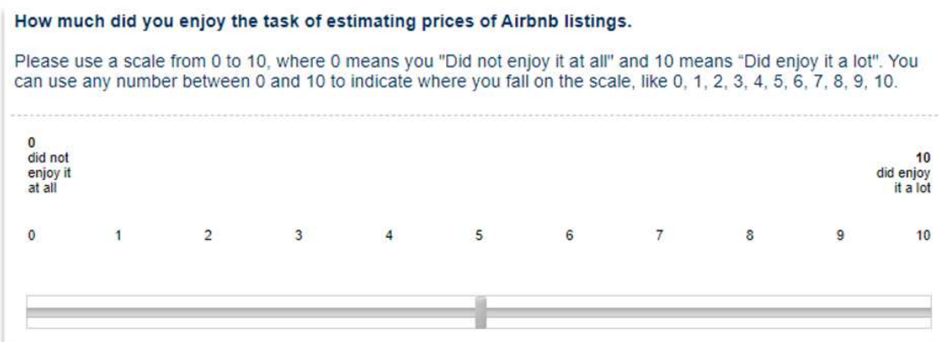
Have you used Airbnb in the past?

- I have previously booked an Airbnb in Vienna
- I have previously booked an Airbnb elsewhere but not in Vienna
- I have never booked an Airbnb

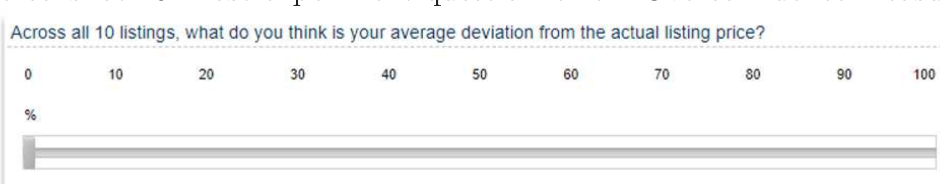
Screenshot 18: Post experiment questionnaire – Risk taking measure



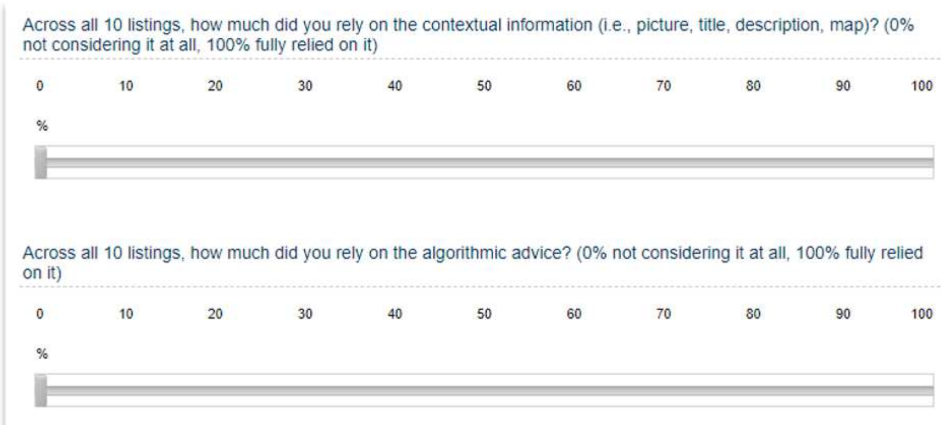
Screenshot 19: Post experiment questionnaire – Task enjoyment measure



Screenshot 20: Post experiment questionnaire – Overconfidence measure



Screenshot 21: Post experiment questionnaire – Information reliance measures (bottom question only shown in algorithmic advice treatments)



Across all 10 listings, how much did you rely on the contextual information (i.e., picture, title, description, map)? (0% not considering it at all, 100% fully relied on it)

0 10 20 30 40 50 60 70 80 90 100

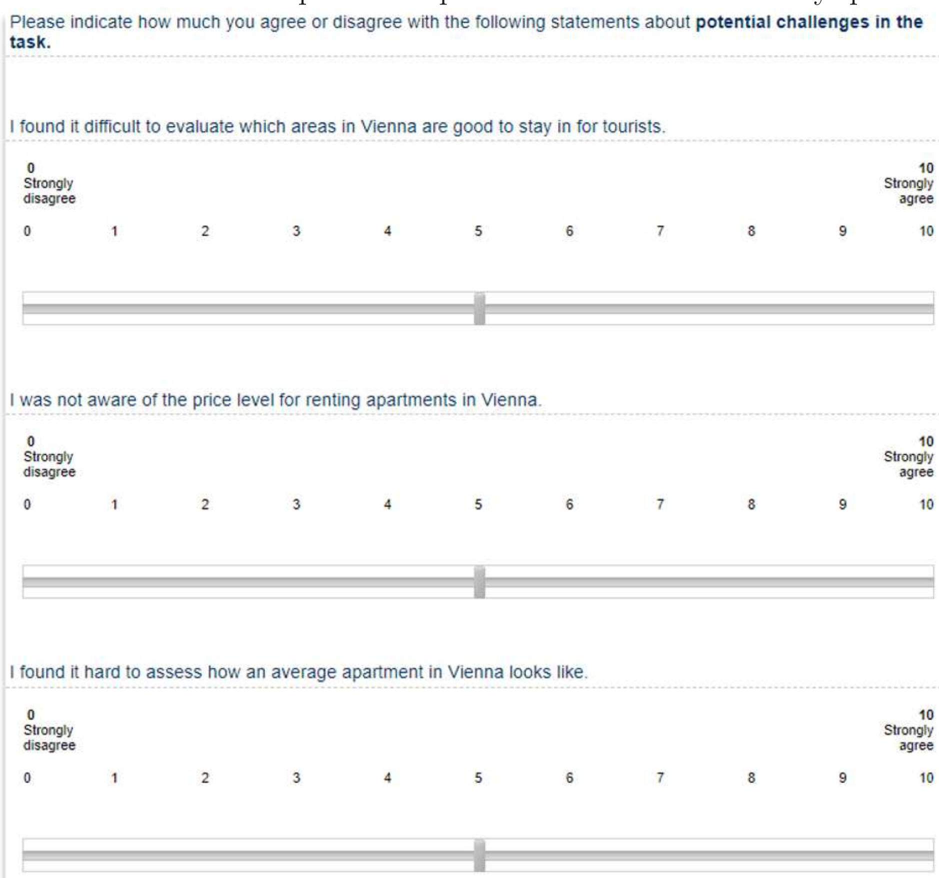
%

Across all 10 listings, how much did you rely on the algorithmic advice? (0% not considering it at all, 100% fully relied on it)

0 10 20 30 40 50 60 70 80 90 100

%

Screenshot 22: Post experiment questionnaire – Unfamiliarity questions



Please indicate how much you agree or disagree with the following statements about **potential challenges in the task.**

I found it difficult to evaluate which areas in Vienna are good to stay in for tourists.

0 Strongly disagree — 10 Strongly agree

0 1 2 3 4 5 6 7 8 9 10

I was not aware of the price level for renting apartments in Vienna.

0 Strongly disagree — 10 Strongly agree

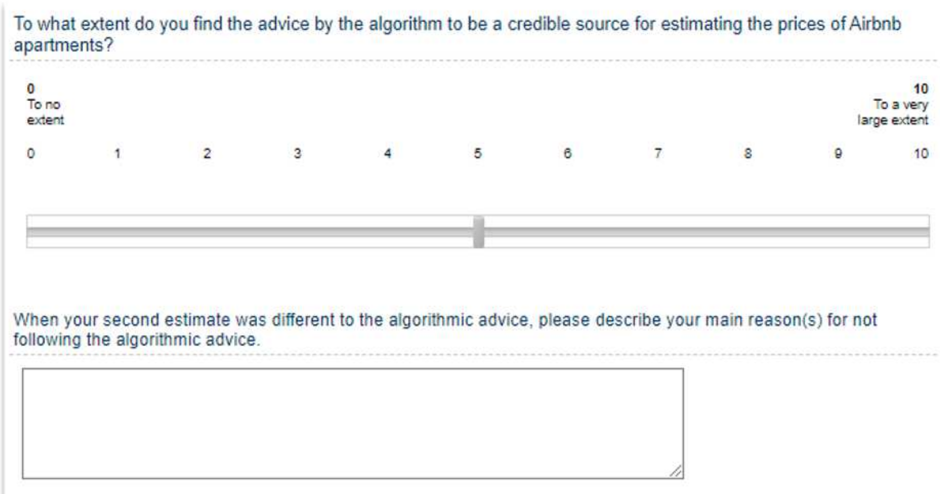
0 1 2 3 4 5 6 7 8 9 10

I found it hard to assess how an average apartment in Vienna looks like.

0 Strongly disagree — 10 Strongly agree

0 1 2 3 4 5 6 7 8 9 10

Screenshot 23: Post experiment questionnaire – Source credibility measure and free text (only shown in algorithmic advice treatments)



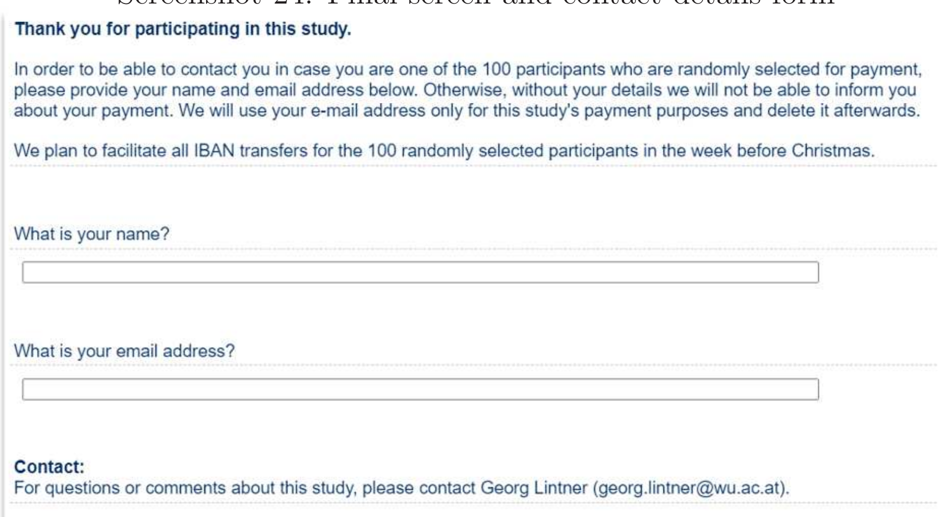
To what extent do you find the advice by the algorithm to be a credible source for estimating the prices of Airbnb apartments?

0 To no extent — 10 To a very large extent

0 1 2 3 4 5 6 7 8 9 10

When your second estimate was different to the algorithmic advice, please describe your main reason(s) for not following the algorithmic advice.

Screenshot 24: Final screen and contact details form



Thank you for participating in this study.

In order to be able to contact you in case you are one of the 100 participants who are randomly selected for payment, please provide your name and email address below. Otherwise, without your details we will not be able to inform you about your payment. We will use your e-mail address only for this study's payment purposes and delete it afterwards.

We plan to facilitate all IBAN transfers for the 100 randomly selected participants in the week before Christmas.

What is your name?

What is your email address?

Contact:
For questions or comments about this study, please contact Georg Lintner (georg.lintner@wu.ac.at).