

Heterogeneity in Effort Provision: Evidence from a lab-in-the-field experiment

Zvonimir Bašić¹, Stefania Bortolotti², Daniel Salicath¹,
Stefan Schmidt¹³, Sebastian O. Schneider¹, Matthias Sutter¹³⁴

¹Max-Planck-Institute for Research on Collective Goods, Bonn

²University of Bologna

³University of Cologne

⁴University of Innsbruck

August 31, 2023

Introduction

Incentives are supposed to increase effort, yet individuals react differently to incentives

- intrinsic \leftrightarrow extrinsic motivation
- thriving \leftrightarrow choking under competitive environments¹

This study: We examine this heterogeneity:

- How do personal characteristics, ability, preferences and one's socio-economic background shape one's sorting decisions and performance under various incentive schemes?
- Results from a lab-in-the-field experiment in German high-schools

¹[Ariely et al., 2009], [Dohmen, 2008]

Overview

- **Pre-Treatment and Treatment session** (2-6 weeks apart):
 - 1.) 5-min Real Effort Task (RET) & survey on socio-demographics, personality characteristics, economic preferences, ...
 - 2.) 20-min RET under three different payment schemes
 - ▶ RET
- **2x3 between-subject design:**
(Fixed, Piece Rate, Tournament) × (Exogenous, Endogenous)
- **Treatment assignment:**
After first session based on min MSE method²
 - ▶ Balance Table

²[Schneider and Schlather, 2017]

Payment Schemes

Fixed

- Flat payment (6.5€)
- Independent of the number of correctly solved tables

Piece Rate

- Piece rate (0.06€) per correctly solved table

Tournament

- Either high (0.08€) or low (0.04€) piece rate per correctly solved table
- High rate if own performance was higher than performance of other randomly chosen participant in “Tournament”

Part 1 – socio-demographics, traits and preferences	Exogenous treatment	Endogenous treatment
Personal ID	✓	✓
RET instructions	✓	✓
Ability (RET 5 min)	✓	✓
IQ test (Raven's matrices, 5min)	✓	✓
SES questionnaire	✓	✓
Big Five (BFI-44)	✓	✓
Competitiveness (14-item)	✓	✓
Preference module	✓	✓
Positive parenting (6-item)	✓	✓
Grit (8-item)	✓	✓
Average payment (€)	€4 + RET (5 mins)	€4 + RET (5 mins)
Average time	45 min	45 min

Part 2 – Effort provision and incentives	Exogenous treatment	Endogenous treatment
Personal ID	✓	✓
Belief elicitation	✓	✓
RET instructions	✓	✓
Instructions (incentive schemes)	1 incentive scheme (within session randomization) [§]	all 3 incentive schemes
Choice (incentive scheme)	-	✓
RET (20 min)	✓	✓
Average payment (€)	€1 + RET (20 min) + belief	€1 + RET (20 min) + belief
Average time	45 min	45 min

Research Questions

Sorting Decisions & Performance:

- 1.) Which characteristics predict sorting into different incentive schemes?
- 2.) Which characteristics predict performance across incentive schemes and treatments?
- 3.) Do subjects sort in a “performance-maximizing” way?

Research Questions

Sorting Decisions & Performance:

- 1.) Which characteristics predict sorting into different incentive schemes?
- 2.) Which characteristics predict performance across incentive schemes and treatments?
- 3.) Do subjects sort in a “performance-maximizing” way?

Research Questions

Sorting Decisions & Performance:

- 1.) Which characteristics predict sorting into different incentive schemes?
- 2.) Which characteristics predict performance across incentive schemes and treatments?
- 3.) Do subjects sort in a “performance-maximizing” way?

Procedures

- 1914 adolescents between 15-21 in German highschools [▶ Map](#)
- Highly heterogeneous sample close to joining the labor market
- Data collection:
 - 03/2019 to 08/2022
 - o-Tree³ running on tablets in lecture halls
 - Earnings: 4€/1€ fixed payment + earnings from RET (+ belief incentivization in part 2)
 - On average: 5.65€ in part 1 and 8.71€ in part 2,
- Pre-Registration: AEARCTR-0008360

³[Chen et al., 2016]

Sorting Decisions

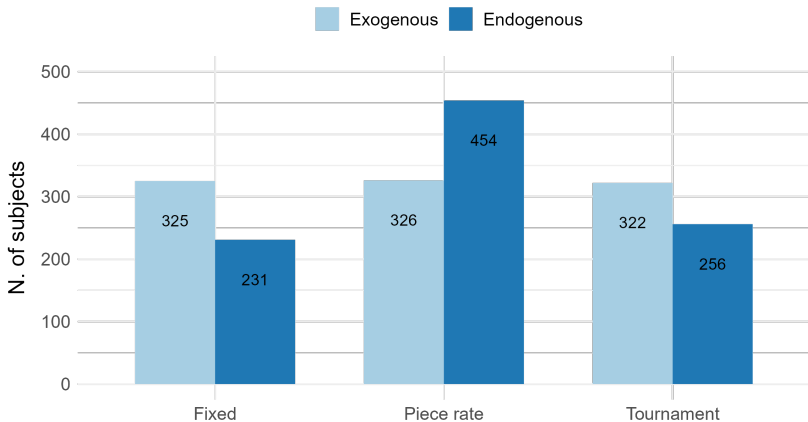


Figure: Number of subjects across treatments

LASSO: Sorting Decisions - Determinants

	Endogenous			
	Logit			Multin. Logit
	Fixed Rate	Piece Rate	Tournament Rate	FR - PR - TR
	(I)	(II)	(III)	(IV)
Female (=1)		2.014*** (0.273)	0.532*** (0.109)	0.809 (0.132)
Grade Math			1.336*** (0.126)	
Grade German	0.737*** (0.059)			
Risk Taking			1.340** (0.168)	
Enjoy Competition (1-5)	0.697*** (0.071)	0.837** (0.070)	1.837*** (0.159)	1.644*** (0.120)
Belief on rel. performance (0-1)			3.048*** (0.913)	2.699*** (0.738)
Productivity (resid.)	0.908*** (0.016)		1.100*** (0.027)	1.101*** (0.016)
Constant	3.355** (1.584)	1.108 (0.343)	0.010*** (0.006)	
Num.Obs.	941	941	941	941
RMSE	0.41	0.49	0.40	1.85

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Clustered standard errors on the session level. Coefficients represent odds ratios. Feature selection is based on a LASSO regression. Remaining predictors are used for predicting sorting decisions via OLS.

Performance (Part 2)

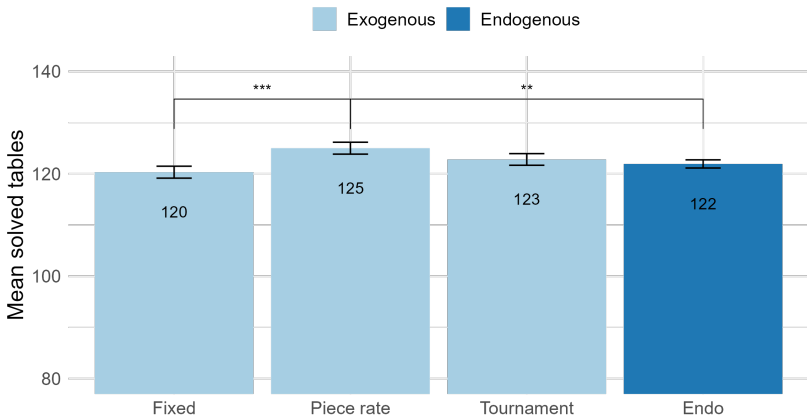


Figure: Mean completed tasks in part 2, by treatment and payment scheme in part 2

[▶ CDF](#)[▶ Part 1](#)[▶ 2x3: Sorting](#)

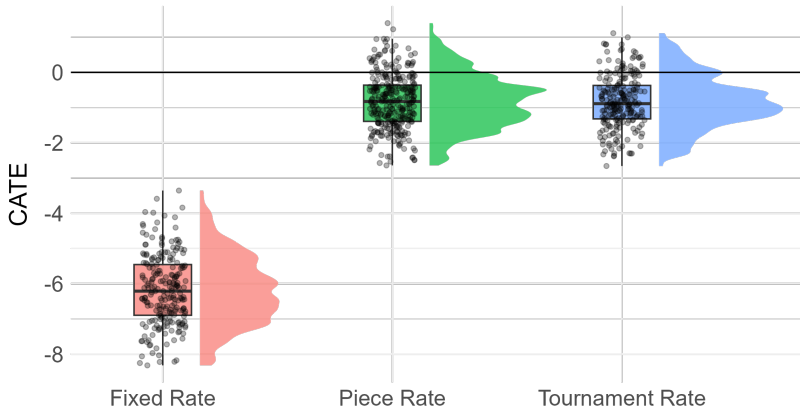
LASSO: Performance (Part 2)

	Exogenous			Endogenous		
	Fixed Rate	Piece Rate	Tournament Rate	Fixed Rate	Piece Rate	Tournament Rate
	(I)	(II)	(III)	(IV)	(V)	(VI)
Female (=1)	3.978** (1.626)				2.994** (1.419)	
Age (rel. to grade mean)	-3.007*** (0.737)	-3.874*** (1.141)				
Grade (9-13)	4.046*** (0.779)	3.551** (1.452)	4.110*** (1.077)		1.621* (0.864)	1.824 (1.363)
IQ (Raven 0-10)					0.781 (0.475)	1.425* (0.790)
Grade Math	4.009*** (0.849)	2.225*** (0.761)	1.764** (0.896)		1.803*** (0.670)	
Grade German						2.039* (1.126)
SES Index			-1.506 (1.130)		-1.003* (0.580)	
Agreeableness (0-5)					1.612 (1.402)	2.044 (1.423)
Conscientiousness (0-5)					0.898 (2.004)	
Positive Parenting (0-5)		-1.026 (0.939)	-2.552** (1.001)	-3.024* (1.787)		
Belief on rel. performance (0-1)	23.344*** (4.681)	24.141*** (4.392)	23.013*** (3.645)	29.347*** (7.471)	20.483*** (3.698)	19.460*** (3.259)
Productivity (resid.)	1.696*** (0.221)	1.582*** (0.276)	1.857*** (0.308)	1.160** (0.556)	2.448*** (0.212)	2.282*** (0.329)
Constant	43.209*** (11.029)	66.109*** (17.872)	65.390*** (14.447)	101.349*** (8.016)	73.556*** (12.943)	73.416*** (17.317)
Num.Obs.	325	326	322	231	454	256
R2 Adj.	0.359	0.355	0.345	0.136	0.470	0.478

* p < 0.1, ** p < 0.05, *** p < 0.01

Note: Clustered standard errors on the session level

Treatment Effects (CRF)



▶ Restricted Sample

▶ Individual Side

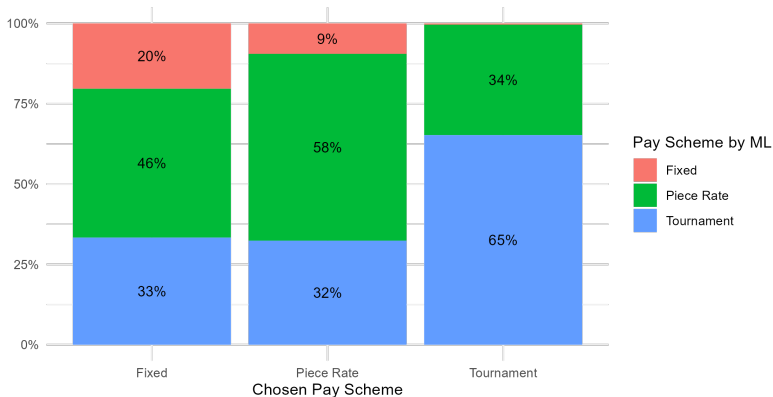
LASSO: Heterogeneity in CATE (CRF)

	CATE (pred.)
	(1)
Productivity (resid.)	0.049*** (0.004)
Age (rel. to grade mean)	0.403*** (0.031)
SES Index	0.203*** (0.023)
Extraversion (0-5)	-0.160*** (0.030)
Positive Parenting (0-5)	-0.137*** (0.027)
Piece Rate	5.295*** (0.057)
Tournament Rate	5.296*** (0.061)
Constant	-6.436*** (0.171)
Num.Obs.	766
R2 Adj.	0.937

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

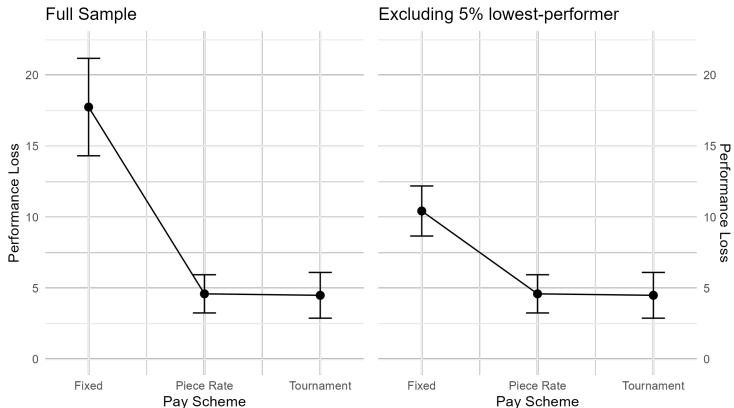
Note: Clustered standard errors on the session level

Treatment Assignment by ML



Treatment Assignment by ML

Predicted - Actual Performance



Bars show confidence intervals; SE's clustered on the session level

Thank you!

stschmidt@coll.mpg.de

[🐦@Stefan__Schmidt](https://twitter.com/Stefan__Schmidt)





Abeler, J., Falk, A., Goette, L., and Huffman, D. (2011).
Reference points and effort provision.
American Economic Review, 101(2):470–92.



Ariely, D., Gneezy, U., Loewenstein, G., and Mazar, N. (2009).
Large stakes and big mistakes.
The Review of Economic Studies, 76(2):451–469.



Chen, D. L., Schonger, M., and Wickens, C. (2016).
otree- an open-source platform for laboratory, online, and field
experiments.
Journal of Behavioral and Experimental Finance, 9:88–97.



Dohmen, T. J. (2008).
Do professionals choke under pressure?
Journal of Economic Behavior & Organization,
65(3-4):636–653.



Schneider, S. O. and Schlather, M. (2017).

A new approach to treatment assignment for one and multiple treatment groups.

Technical Report 228.

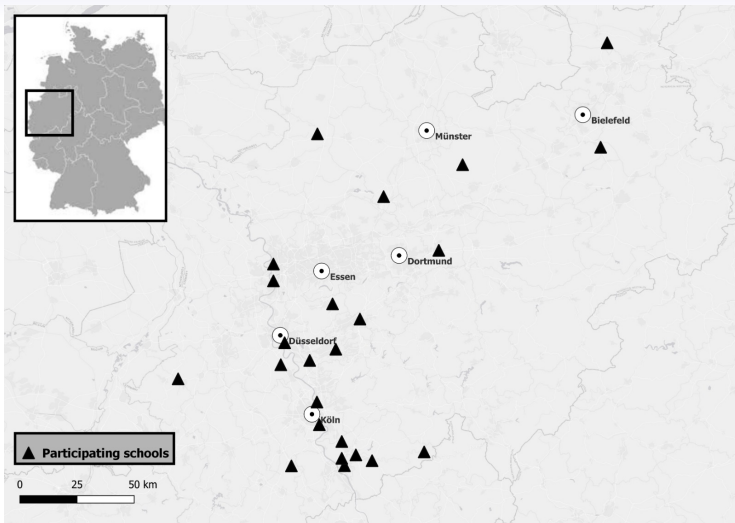


Figure: Map of participating schools

▶ back

Real-effort task – [Abeler et al., 2011]

Time remaining: **20:00**

Number of attempts remaining: **3**

Correctly solved tables: **0**

Table 1

0	1	1	0	0	0	0	0
0	0	0	0	1	0	1	1
0	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32

Next

Characteristic	Overall	Fixed rate	Piece rate	Tournament	p-value
Productivity part 1 (RET 5 min)	26.84 (6.19)	26.93 (6.27)	26.80 (6.51)	26.80 (5.80)	0.84
Performance Rank (Belief; 0-1)	0.55 (0.24)	0.53 (0.25)	0.56 (0.24)	0.55 (0.24)	0.28
Female (=1)	0.53 (0.49)	0.52 (0.50)	0.53 (0.49)	0.54 (0.49)	0.95
Age	17.14 (1.11)	17.19 (1.12)	17.09 (1.06)	17.13 (1.13)	0.54
Grade (9-13)	11.28 (0.97)	11.32 (0.98)	11.27 (0.95)	11.26 (0.99)	0.71
IQ (Raven 0-10)	5.12 (1.43)	5.12 (1.35)	5.16 (1.45)	5.09 (1.49)	0.80
Grade Math	2.82 (1.13)	2.81 (1.07)	2.83 (1.18)	2.82 (1.14)	0.98
Grade German	2.72 (0.92)	2.72 (0.96)	2.72 (0.88)	2.72 (0.92)	0.94
Born Germany (=1)	0.94 (0.24)	0.94 (0.23)	0.94 (0.25)	0.93 (0.26)	0.70
Speak German at home (=1)	0.94 (0.24)	0.95 (0.22)	0.95 (0.22)	0.92 (0.27)	0.15
Parents German (=1)	0.60 (0.49)	0.58 (0.49)	0.61 (0.49)	0.60 (0.49)	0.69
Mother univ. diploma (=1)	0.13 (0.33)	0.12 (0.33)	0.13 (0.34)	0.12 (0.33)	0.97
Father univ. diploma (=1)	0.18 (0.38)	0.18 (0.38)	0.18 (0.38)	0.17 (0.38)	>0.99
One parent univ. diploma (=1)	0.23 (0.42)	0.23 (0.42)	0.23 (0.42)	0.24 (0.43)	0.96
Single parent (=1)	0.21 (0.41)	0.21 (0.41)	0.21 (0.41)	0.21 (0.41)	0.99
Number of siblings	1.69 (1.13)	1.62 (1.10)	1.71 (1.12)	1.73 (1.18)	0.50
Books at home (1-6)	2.38 (1.31)	2.39 (1.43)	2.39 (1.26)	2.35 (1.24)	0.96
Pocket money (0-95)	27.87 (25.37)	29.35 (26.19)	26.76 (26.03)	27.51 (23.80)	0.27
Number of cars (0-3)	1.90 (0.82)	1.83 (0.80)	1.92 (0.84)	1.94 (0.83)	0.20
Number of holidays (0-3)	1.78 (1.05)	1.74 (1.05)	1.88 (1.06)	1.73 (1.04)	0.093
PISA wealth index (0-17)	12.95 (2.28)	12.82 (2.33)	13.06 (2.22)	12.98 (2.28)	0.44
FAS index (0-10)	6.75 (1.91)	6.61 (1.93)	6.89 (1.96)	6.74 (1.84)	0.21
low SES (=1)	0.21 (0.41)	0.21 (0.41)	0.23 (0.42)	0.20 (0.40)	0.68
Patience (1-32)	18.25 (11.77)	18.71 (11.49)	18.29 (11.94)	17.77 (11.90)	0.59
Patience survey (0-10)	7.19 (1.97)	7.23 (2.01)	7.09 (1.97)	7.26 (1.93)	0.47
Risk (1-32)	10.41 (6.56)	10.35 (6.51)	10.44 (6.38)	10.45 (6.80)	0.93
Risk survey (0-10)	5.83 (1.93)	5.85 (1.95)	5.84 (1.97)	5.80 (1.88)	>0.99
Altruism (0-10)	7.52 (2.20)	7.46 (2.28)	7.54 (2.20)	7.55 (2.12)	0.96
Extraversion (1-5)	3.47 (0.78)	3.46 (0.78)	3.46 (0.77)	3.50 (0.80)	0.70
Agreeableness (1-5)	3.56 (0.57)	3.54 (0.56)	3.55 (0.57)	3.58 (0.57)	0.83
Conscientiousness (1-5)	3.36 (0.42)	3.34 (0.43)	3.35 (0.40)	3.37 (0.44)	0.81
Neuroticism (1-5)	2.91 (0.73)	2.95 (0.76)	2.92 (0.70)	2.88 (0.72)	0.57
Openness (1-5)	3.41 (0.66)	3.39 (0.68)	3.43 (0.61)	3.40 (0.68)	0.61
Enjoy competition (1-5)	3.09 (0.98)	3.07 (1.00)	3.13 (0.95)	3.07 (0.99)	0.72
Positive parenting (1-5)	3.40 (0.91)	3.39 (0.91)	3.40 (0.91)	3.41 (0.92)	0.90
Grit (1-5)	3.23 (0.46)	3.23 (0.50)	3.23 (0.46)	3.24 (0.42)	0.84

Note:

The p-values report tests of difference in means across the three different payment schemes.

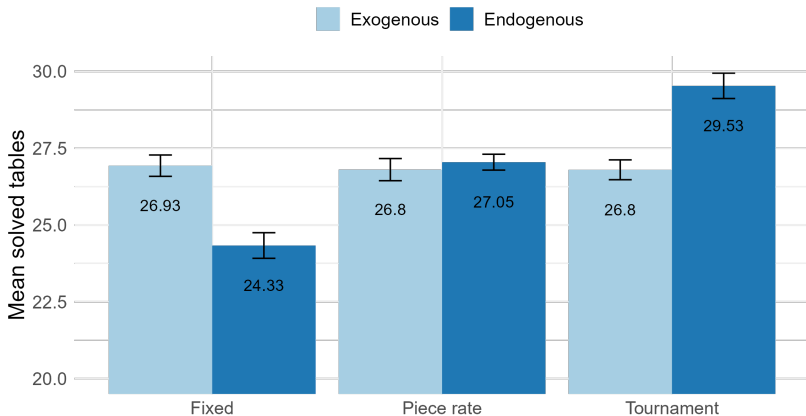


Figure: Mean completed tasks in part 1, by treatment and payment scheme in part 2

▶ CDF

▶ Back

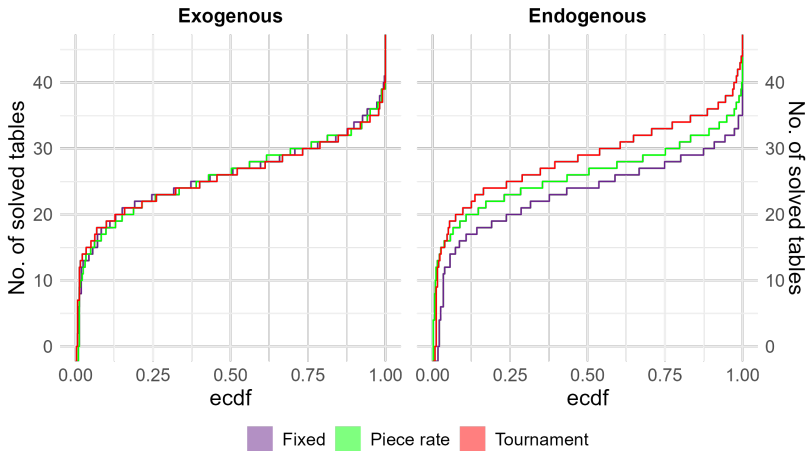


Figure: Mean completed tasks in part 1, by treatment and payment scheme in part 2 [▶ Back](#)

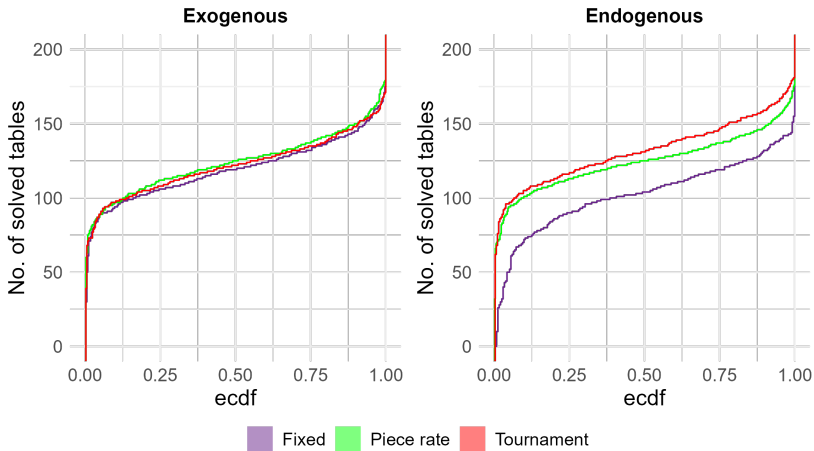


Figure: Mean completed tasks in part 2, by treatment and payment scheme in part 2 [▶ Back](#)

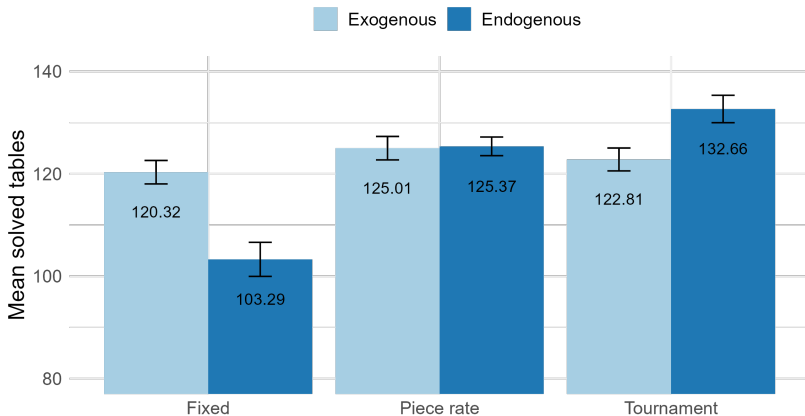
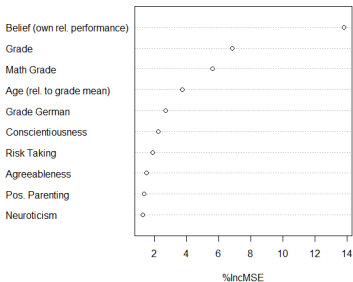


Figure: Mean completed tasks in part 2, by treatment and payment scheme in part 2 [▶ Back](#)

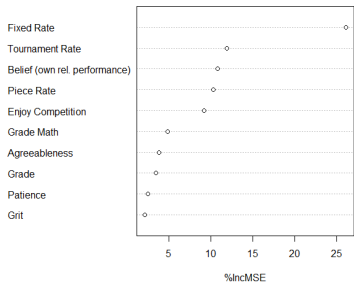
Determinants of Performance

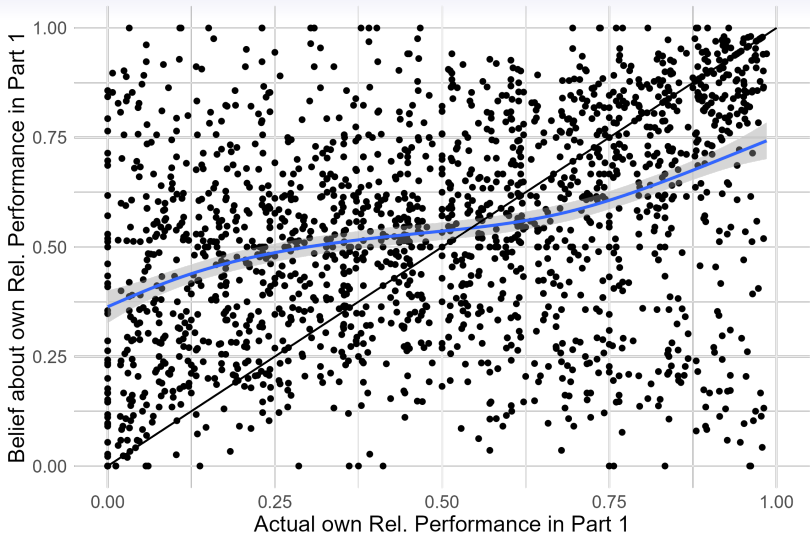
Prediction via Random Forests:

Top 10 - Variable Importance (Exogenous)



Top 10 - Variable Importance (Endogenous)





LASSO: Beliefs (own rel. Performance)

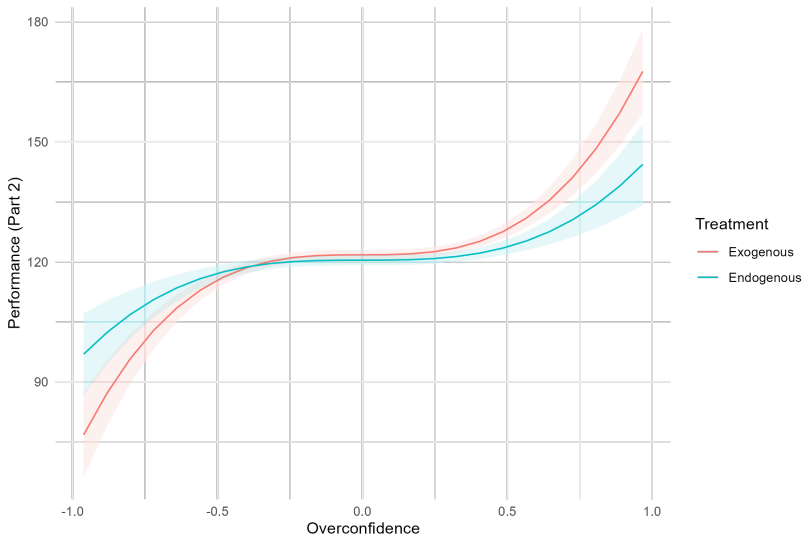
	Belief on rel. performance (0-1)	Overconfidence (-1 - 1)
	(I)	(II)
Female (=1)	-0.026** (0.011)	-0.038*** (0.009)
Age (rel. to grade mean)		0.028*** (0.006)
IQ (Raven 0-10)		-0.004 (0.003)
Grade Math		-0.007* (0.004)
Grade German		-0.018*** (0.005)
SES Index	0.016*** (0.006)	0.017*** (0.005)
Patience Index		-0.015*** (0.006)
Risk Taking	0.019*** (0.006)	
Extraversion (0-5)		0.020*** (0.005)
Agreeableness (0-5)		0.016** (0.008)
Neuroticism (0-5)	-0.017** (0.009)	
Openness (0-5)		-0.013* (0.007)
Enjoy Competition (1-5)		-0.017*** (0.005)
Positive Parenting (0-5)		0.039*** (0.006)
Productivity (resid.)		-0.042*** (0.002)
Constant	0.605*** (0.022)	0.043 (0.054)
Num.Obs.	1914	1914
R2 Adj.	0.014	0.637

* p < 0.1, ** p < 0.05, *** p < 0.01

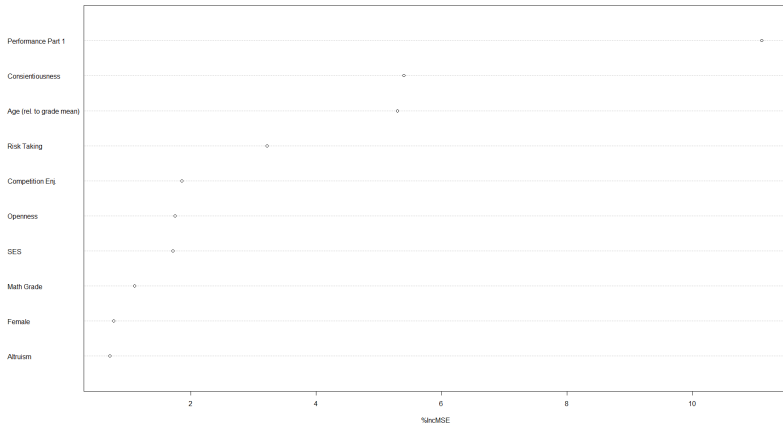
Note:

Clustered standard errors on the session level

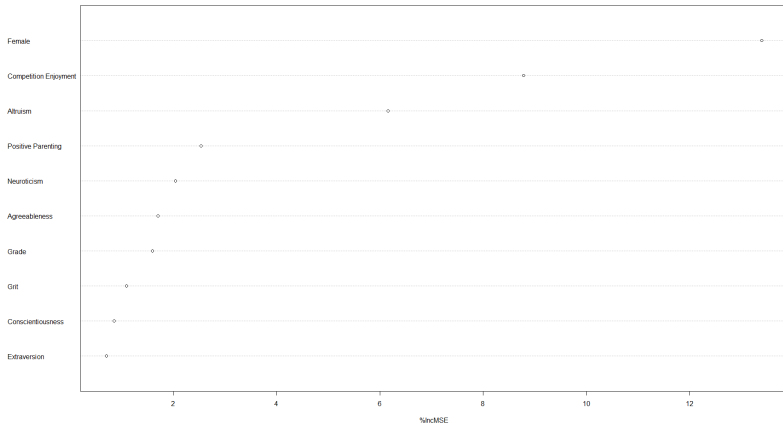
Overconfidence in own rel. Performance



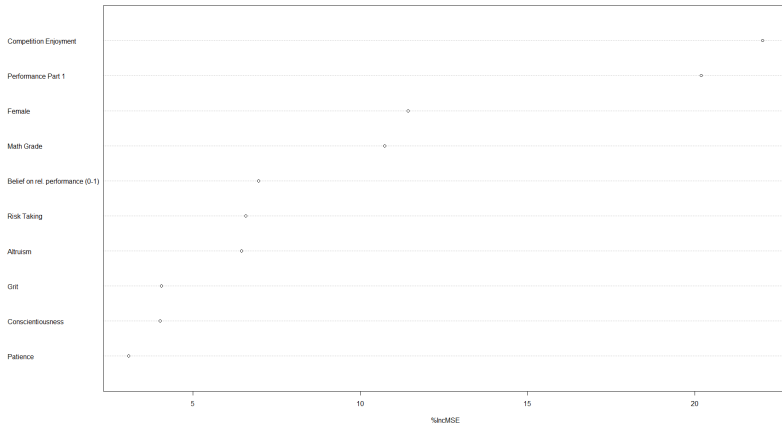
Top 10 - Variable Importance (Endo - Fixed Rate)

[▶ back](#)

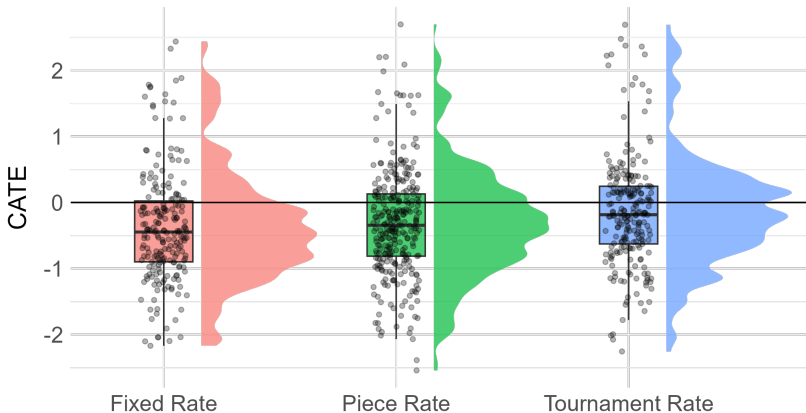
Top 10 - Variable Importance (Endo - Piece Rate)

[▶ back](#)

Top 10 - Variable Importance (Endo - Tournament Rate)

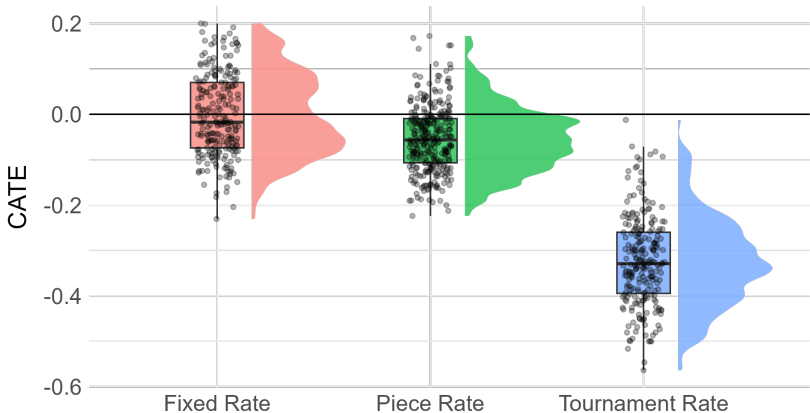
[▶ back](#)

Treatment Effects (CRF) - Excluding bottom 5% (low-performers)



Treatment Effects (CRF) - Individual Side

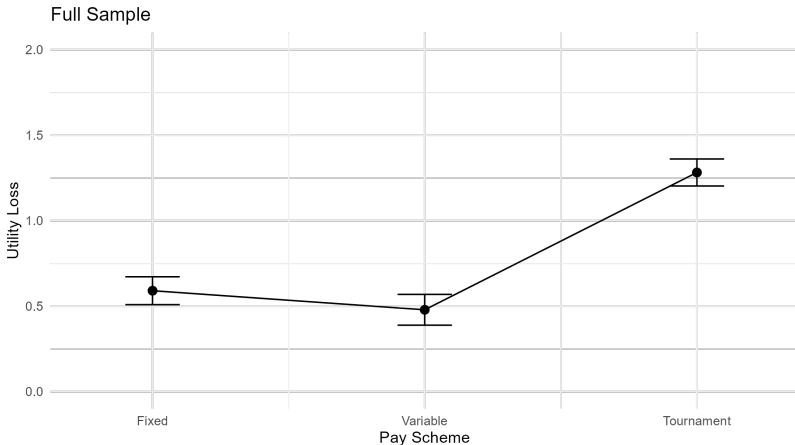
▶ Utility Function



▶ back

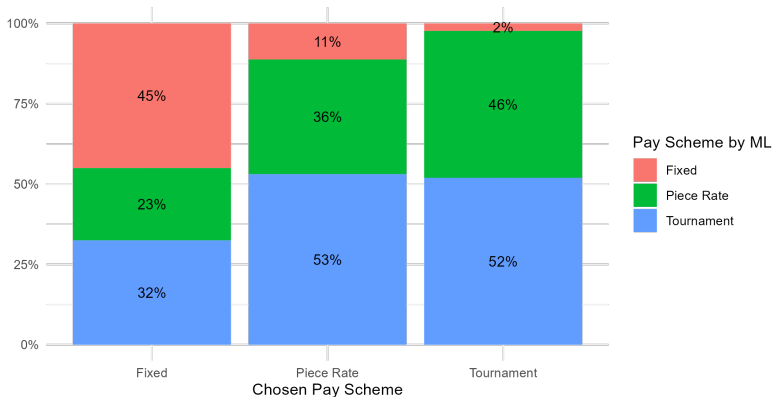
Treatment Assignment by ML

Predicted - Actual Utility



Bars show confidence intervals; SE's clustered on the session level

Treatment Assignment by ML



LASSO: Heterogeneity in CATE (CRF) - Individual Side

	CATE (pred.)
	(1)
Female (=1)	0.018*** (0.006)
SES Index	0.014*** (0.003)
Patience Index	-0.018*** (0.003)
Agreeableness (0-5)	0.081*** (0.005)
Openness (0-5)	0.019*** (0.004)
Enjoy Competition (1-5)	-0.008** (0.003)
Positive Parenting (0-5)	0.013*** (0.003)
Pay Scheme	-0.153*** (0.004)
Constant	-0.063** (0.030)
Num.Obs.	766
R2 Adj.	0.769

* p < 0.1, ** p < 0.05, *** p < 0.01

[▶ back](#)

Utility Function

$$\text{Utility} = \text{payoff} - \text{Effort Costs (survey)}$$

- “How much effort did you exert?” (Likert 1-7)
- “How stressed did you feel?” (Likert 1-7)
- “How exhausted did you get?” (Likert 1-7)

Effort costs = Mean response (normalized by population SD)