PNAS
www.pnas.org

Supplementary Information for

Preferences predict who commits crime among young men

Thomas Epper, Ernst Fehr, Kristoffer B. Hvidberg, Claus T. Kreiner, Søren Leth-Petersen, Gregers N. Rasmussen

Corresponding Author: Ernst Fehr
E-mail: ernst.fehr@econ.uzh.ch

This PDF file includes:

Supplementary text
Fig. S1 (not allowed for Brief Reports)
Tables S1 to S8 (not allowed for Brief Reports)
SI References

Supporting Information Text

A Basic Choice Model of Criminal Behavior. Here we illustrate the role of risk preferences and time preferences in a simple two-period model of criminal behavior. We assume the preferences of an individual can be represented by the objective function

$$\Omega\left(a\right)=u\left(c_1\left(a\right)\right)+\beta\left[pu\left(c_2^A\left(a\right)\right)+\left(1-p\right)u\left(c_2^B\left(a\right)\right)\right],$$

where c_1 denotes consumption in period 1, depending on whether the individual engage in crime ($a = 1$) or not ($a = 0$), whereas p is the probability of detection in which case consumption in period 2 becomes c_2^A. Otherwise consumption in period 2 equals c_2^B. The parameter β is the weight of an individual on future utility (patience) and $u\left(\cdot\right)$ is a concave utility function.

The consumption levels equal

$$\begin{aligned} c_1\left(a\right) &= \left(1+a\gamma\right)y, \\ c_2^A\left(a\right) &= \left(1-a\eta\right)y, \\ c_2^B\left(a\right) &= y, \end{aligned}$$

where y is a fixed income, γ is the gain from criminal activity measured in proportion to income, and η is the loss if detected measured in proportion to income. Gains and losses can be money/consumption but, more broadly it can also include non-pecuniary effects that correspond to a given gain or loss in consumption. Crime enforcement policy can change the probability of detection p and the loss if detected η.

The incentive to commit crime equals $\Psi = \Omega\left(1\right) - \Omega\left(0\right)$, which gives

$$\Psi=u\left(\left(1+\gamma\right)y\right)+\beta\left[pu\left(\left(1-\eta\right)y\right)+\left(1-p\right)u\left(y\right)\right]-\left(1+\beta\right)u\left(y\right). \qquad [1]$$

If it is assumed that the consumption levels associated with criminal activity are well approximated by second-order Taylor expansions around the non-crime consumption level, y, then we can write

$$\begin{aligned} u\left(\left(1+\gamma\right)y\right) &\approx u\left(y\right)+u'\left(y\right)\gamma y+\frac{1}{2}u''\left(y\right)\left(\gamma y\right)^2=u\left(y\right)+u'\left(y\right)y\gamma\left(1-\frac{1}{2}\theta\gamma\right), \\ u\left(\left(1-\eta\right)y\right) &\approx u\left(y\right)-u'\left(y\right)\eta y+\frac{1}{2}u''\left(y\right)\left(\eta y\right)^2=u\left(y\right)-u'\left(y\right)y\eta\left(1+\frac{1}{2}\theta\eta\right), \end{aligned}$$

where $\theta \equiv -\frac{u''(y)y}{u'(y)}$ is the coefficient of relative risk aversion. By substituting these expressions into (1), the incentive to commit crime becomes

$$\Psi=u'\left(y\right)y\left[\gamma\left(1-\frac{1}{2}\theta\gamma\right)-\beta p\eta\left(1+\frac{1}{2}\theta\eta\right)\right]. \qquad [2]$$

The first term in the bracket is the benefit of engaging in crime, which is decreasing in the risk aversion parameter θ because marginal utility is decreasing. The second term in the bracket is the expected loss, which is increasing in the risk aversion parameter and in the patience parameter β. It follows from the equation that the incentive to commit crime is decreasing in both the risk aversion parameter and the patience parameter, i.e., $\partial\Psi/\partial\theta < 0$ and $\partial\Psi/\partial\beta < 0$. It also follows that the incentive to commit crime is decreasing in the enforcement parameters, i.e., $\partial\Psi/\partial p < 0$ and $\partial\Psi/\partial\eta < 0$, and that the effect of enforcement on the incentive to commit crime is (numerically) larger for more risk averse and more patient people, i.e., $\frac{\partial^2\Psi}{\partial p\partial\theta} < 0$, $\frac{\partial^2\Psi}{\partial p\partial\beta} < 0$, $\frac{\partial^2\Psi}{\partial\eta\partial\theta} < 0$, and $\frac{\partial^2\Psi}{\partial\eta\partial\beta} < 0$.

If we assume that the propensity to commit crime Π is an increasing, concave function of the incentive to commit crime, i.e., $\Pi : \Psi \to (0, 1)$ where $\Pi'\left(\Psi\right) > 0$ and $\Pi''\left(\Psi\right) \leq 0$, then we have

Proposition (i) *The propensity to commit crime is decreasing in the degree of risk aversion θ and patience β, i.e., $\partial\Pi/\partial\theta < 0$ and $\partial\Pi/\partial\beta < 0$.* ($ii$) *Stricter crime enforcement, which increases p or η, reduce crime propensities most for risk averse and patient individuals, $\frac{\partial^2\Pi}{\partial p\partial\theta} < 0$, $\frac{\partial^2\Pi}{\partial p\partial\beta} < 0$, $\frac{\partial^2\Pi}{\partial\eta\partial\theta} < 0$, and $\frac{\partial^2\Pi}{\partial\eta\partial\beta} < 0$.*

Proof (i) By differentiating, we find $\partial\Pi/\partial\theta = \Pi'\left(\Psi\right)\frac{\partial\Psi}{\partial\theta} < 0$ and $\Pi'\left(\Psi\right)\frac{\partial\Psi}{\partial\beta} < 0$. ($ii$) By differentiating, we find $\frac{\partial^2\Pi}{\partial p\partial\theta} = \Pi'\left(\Psi\right)\frac{\partial^2\Psi}{\partial p\partial\theta} + \Pi''\left(\Psi\right)\frac{\partial\Psi}{\partial p}\frac{\partial\Psi}{\partial\theta} < 0$, $\frac{\partial^2\Pi}{\partial p\partial\beta} = \Pi'\left(\Psi\right)\frac{\partial^2\Psi}{\partial p\partial\beta} + \Pi''\left(\Psi\right)\frac{\partial\Psi}{\partial p}\frac{\partial\Psi}{\partial\beta} < 0$, $\frac{\partial^2\Pi}{\partial\eta\partial\theta} = \Pi'\left(\Psi\right)\frac{\partial^2\Psi}{\partial\eta\partial\theta} + \Pi''\left(\Psi\right)\frac{\partial\Psi}{\partial\eta}\frac{\partial\Psi}{\partial\theta} < 0$, and $\frac{\partial^2\Pi}{\partial\eta\partial\beta} = \Pi'\left(\Psi\right)\frac{\partial^2\Psi}{\partial\eta\partial\beta} + \Pi''\left(\Psi\right)\frac{\partial\Psi}{\partial\eta}\frac{\partial\Psi}{\partial\beta} < 0$. ■

The proposition shows that people with a lower degree of risk aversion and a lower degree of patience are more prone to commit crime and are less responsive to crime enforcement.

Administrative data. From Statistics Denmark we use the crime registers for charges (KRSI) and convictions (KRAF). Entries in the two registers can be linked using a file number (`journr`[*]). From KRSI we use the date of the committed offence (`sig_ger1dto`). Not all convictions have a related charge. In these cases we assume that the offence was committed on the date of the conviction. The convictions in the KRAF register are categorized using a seven digit code in the variable `afg_ger7`. The first digit indicates whether the conviction is related to the criminal code (1), the traffic code (2) or other special laws (3). We discard convictions under the traffic code. The two first digits are used two divide the convictions into sexual offences (11), violent offences (12), property offences (13) and drug offences (32). We use this information to generate indicators for having been convicted of different types of crime. From the register UDFK we have information on grades in lower secondary school. We compute the grade point average across subjects in the final exams. For 63 participants we do not have information on any final exams in 9th or 10th grade. For 43 of these participants, we use self-reported GPA in the survey or their mark for general proficiency where possible. For the remaining 20 participants we assume that they have not passed any final exams and assign them a GPA of 0. Based on the grade point average, we compute the rank/percentile position of each participant in the grade point average distribution of all participants, `gpa_r`. We use ordinal ranking, which ensures a uniform distribution of percentile positions, also when some of the underlying observations are identical. We also compute the GPA only based on the subject math, `gpa_math_r`, which we use in the sensitivity analyses in column 2 of Table S5.

We use the population register (BEF) to identify participants' gender, birth year, immigrant and descendant status, region of residence, and whether their municipality is characterized as a city (according to the "Kommunegrupper" characterization provided by Statistics Denmark). We also use information in the BEF register to identify the participants' parents and the variable `fm_mark` to identify whether they live with both parents or not, e.g. if the parents are divorced. Using municipality (`kom`) and address (`bopikom`) in the BEF register we link participants to the household register (HUST). Here we obtain information on the number of children in the household when the participant was 0 years old as an indicator of whether the participant is the first born child, and the number of children in the household when the participant was 10 years old as an indicator for having siblings.

For the parents, we also use the population register to compute their age when the participant was born and the crime register to compute an indicator of whether they have committed a criminal offence, excl. traffic offences. In the analyses, we use whether the parents are convicted of a crime committed from age 15 to 20 (as for the young men in the sample), but the results are the same if we use an indicator for whether the parents have been convicted at any age. Furthermore, we use the income register (IND) to compute the parents' average income in 2015 prices across the years when the participant was 17 to 19 years old. As for the grade point average, we compute the percentile position in the parent income distribution, `income_parents_r`. We use the employment register (IDAP) to compute indicators of parents employment status in 2017 (employed, self-employed, unemployed and not in the work force) and unemployment between 2008 and 2017. We fist compute the average share of time they have been unemployed (based on `arledgr`) and compute indicators of not having been unemployed, having been unemployed from 0 to 10% of the time and from 10% to 100% of the time. Finally, we use the education register (UDDA) to compute indicators for the parents' levels of education based on the Danish ISCED classification (primary or lower secondary, upper secondary, short cycle tertiary, bachelor or equivalent, and master, doctoral or equivalent).

For each participant, we predict the probability of participation, i.e. the propensity score, based on a probit model where we include the register data that is available for both participants and non-participants as explanatory variables. We use this for the propensity score weighting in column 6 of Table S5.

Experimental data. Based on the data from our incentivized behavioral experiments, we compute each individual's time, risk and social preferences. In addition, the experimental data enable us to identify individuals' present bias and their behindness and aheadness aversion which we use for sensitivity and robustness checks in table S4.

We use the time experiment to compute the impatience measure, which is the mean share kept across choice situations in Table S6. Based on this, we compute the percentile positions in the impatience distribution for all participants, `impatience_r`, in the same way as we do for grade point average and parental income. Since participants were randomly assigned to either a high or low stake condition, we compute the percentiles separately for the two conditions. We also use the data from the time experiment to compute percentile positions within the distribution of present bias, `present_bi_r`. Here we first compute the difference in the share kept between two situations with the same interest rate but with different timing of payout, e.g. situation 1 and 9 in Tables S7. We then compute the mean across the 8 computed differences and rank the participants accordingly. Similarly, we use the risk experiment to compute each participant's risk tolerance percentile position, `risk_toler_r`. Finally, we use the social experiment to compute each participant's altruism percentile position, `altruism_r`. As seen in Table S8, the cost of giving varies across situations, and in particular it is positive in some cases, i.e. it cost money for the participant to increase the other's payoff, and negative in others, i.e. the participant gets money for increasing the other's payoff. We compute a measure of aheadness aversion from the situations with a positive cost of giving (i.e. negative slope in the trade-off) and a measure of behindness aversion from the situations with a negative cost of giving (i.e. a positive slope in the trade-off). Again, we compute each participant's positions in the distributions, `altruism_pos_r` and `altruism_neg_r`.

As an alternative to the non-parametric preference measures described above, we also estimate structural preference parameters for the three choice domains and use these in the sensitivity analysis in column 5 of Table S4. To estimate the discrete choice models we assume random utility with choice-domain-specific error parameters. More specifically, we back out individual-level parameters from mixed logit models with normally distributed behavioral parameters. As for the

[*] Names written with monospaced typewriter typeface refer to variable names. For the variables from Statistics Denmark's records, these are the names Statistic Denmark has assigned (see definitions here: www.dst.dk/da/Statistik/dokumentation/Times).

non-parametric preference measures, we then compute percentile positions based on the estimated parameter distribution. We implement a hierarchical Bayesian procedure to estimate the mixed logit models. This procedure is described in detail elsewhere (1, 2).[†] In the following models, i denotes the individual and j denotes the choice situation.

Based on the risk task, we define R_{ij} as the binary lottery the individual faces as a consequence of the chosen allocation (see *Material and Methods*) and estimate an expected utility model (3), which assigns the value V to the alternative R_{ij}:

$$V(R_{ij}) = p_j u_i(w_{1ij}) + (1 - p_j)u_i(w_{2ij})\,,$$

with $u_i(w_{\cdot ij}) = w_{\cdot ij}^{1-\rho_i}$ where p_j is the probability that the good state occurs, w_{1ij} and w_{2ij} are the payoffs in the good and the bad state, and ρ_i denotes the (Arrow-Pratt) coefficient of relative risk aversion. Higher values of ρ_i indicates comparatively more risk aversion.

Based on the time task, we define T_{ij} as the income stream the individual faces as a consequence of the chosen allocation and estimate a quasi-hyperbolic discounted utility model (4),[‡] which assigns the value V to the alternative T_{ij}:

$$V(T_{ij}) = d(t_{1j})v_i(w_{1ij}) + d(t_{2j})v_i(w_{2ij})\,,$$

with $v_i(w_{\cdot ij}) = w_{\cdot ij}^{1-\gamma_i}$, where t_{1j} and t_{2j} denote the payment delays in months relative to the time of the experiment $(0 \leq t_{1j} < t_{2j})$, w_{1ij} and w_{2ij} are the payoffs at the earlier and later point in time, and γ denotes an (Arrow-Pratt-type) coefficient of relative aversion towards income fluctuations over time. Finally, $d(t_{\cdot j}) = 1$ if $t_{\cdot j} = 0$ and $d(t_{\cdot j}) = \beta_i e^{-\eta_i \frac{t_{\cdot j}}{12}}$ otherwise, where $\eta_i \geq 0$ denotes the (annualized) rate of time preference and β_i present bias (for $\beta_i < 1$) or future bias (for $\beta_i > 1$).

Based on the social task, we define S_{ij} as the interpersonal distribution the individual faces as a consequence of the chosen allocation and estimate an inequality aversion model (6), which assigns the value V to the alternative S_{ij}:

$$V(S_{ij}) = w_{\text{own}j} - \alpha_i(w_{\text{other}j} - w_{\text{own}j})\mathbb{1}[w_{\text{own}j} < w_{\text{other}j}] - \beta_i(w_{\text{own}j} - w_{\text{other}j})\mathbb{1}[w_{\text{own}j} > w_{\text{other}j}]\,,$$

where $w_{\text{own}j}$ is the payoff to the participant, $w_{\text{other}j}$ is the payoff to the other person and α_i and β_i denote behindness aversion and aheadness aversion, respectively.

Participants in the experiment also responded to a few survey questions: self-reported gender, self-reported birth year, self-reported GPA and self-reported self-control. For the last variable the question asked was "I am good at exercising self-control in my actions and decisions" and participants answered this question on a 7 point Likert scale where 1 was "strongly disagree" and 7 was "strongly agree". Using the same procedure as for the other explanatory variables, we compute the percentile position in the self-control distribution, `selfcontrol_r`. For all the percentile position variables we also compute the z-scores based on the raw variables where we subtract the mean across all participant and divide by the standard deviation. We use this in Table S5.

[†] We assume uninformative prior distributions. In addition, we adopt appropriate transformations of the distributions to ensure that the estimated parameters lie within the range of their theoretical support.

[‡] Since the individuals were assigned to a high or low stake condition, we estimate separate models for the two conditions and obtain condition-specific rankings of individuals to avoid the ordering being confounded by the magnitude effect (see e.g. (5)).

Figures

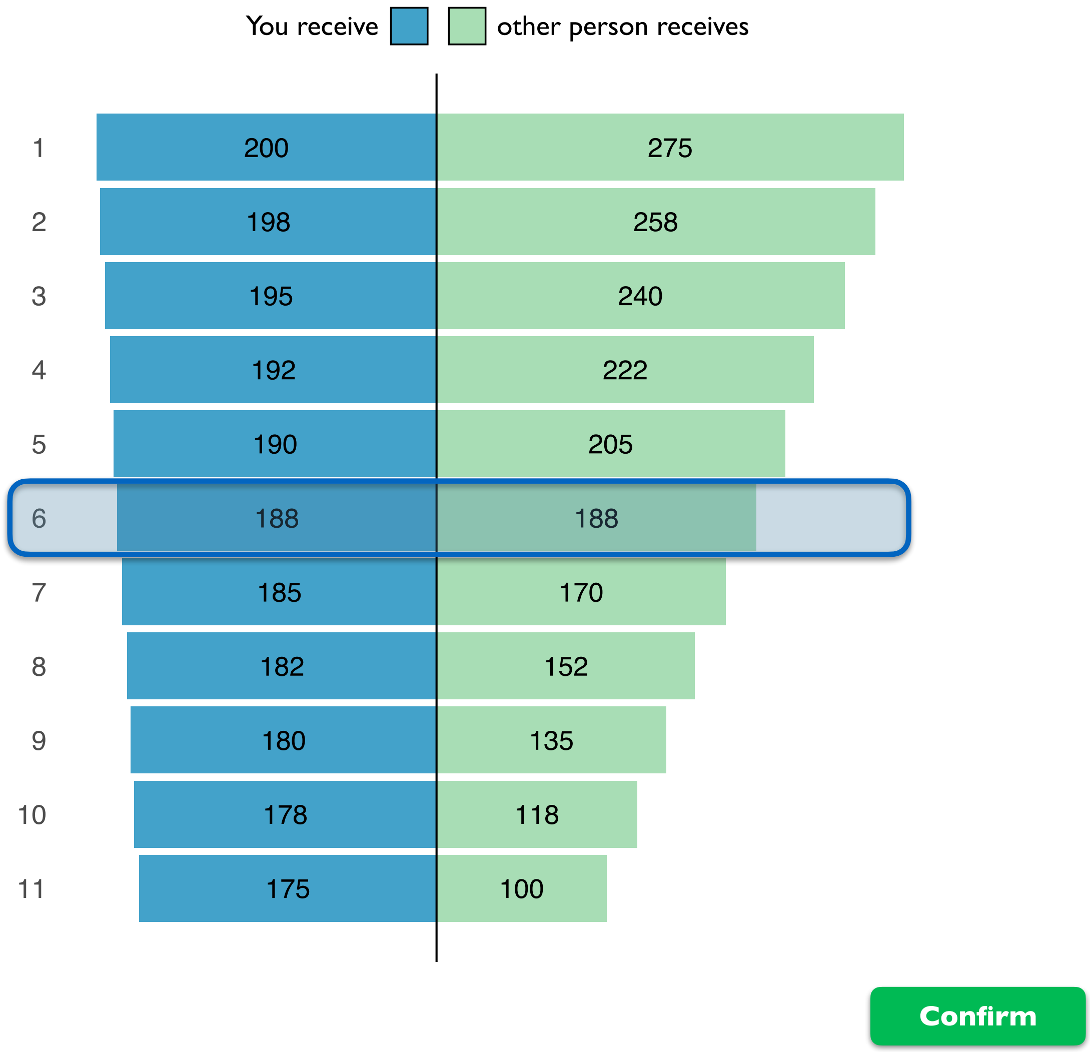

Fig. S1. Screenshot of online experiment eliciting social preferences

Notes: This is an example of one of the choice situations (Situation 9 in Table S8). The subject can choose one out of eleven payoff allocations where the blue bars to the left show the money received by the subject, while the green bars to the right show the money received by a randomly assigned other person. The individual has chosen option number six, which gives DKK 188 to each person.

Tables

Table S1. Summary statistics

	Sample (1)	Logins (2)	Invited (3)	Difference (1) - (3) (4)	P-value (%) (5)
Crime age 15-20 (%)	13.4	14.0	19.0	-5.6	0.0
GPA	7.5	7.4	6.6	0.9	0.0
Immigrant (%)	3.2	3.4	5.2	-2.0	0.0
Descendant (%)	6.6	7.0	7.9	-1.3	0.3
Northern Jutland (%)	9.8	9.9	10.5	-0.7	16.3
Middle Jutland (%)	24.6	24.3	23.7	0.9	21.0
Southern Denmark (%)	25.4	25.2	23.1	2.3	0.2
Copenhagen (%)	26.4	26.2	28.0	-1.7	3.1
Sealand (%)	13.8	14.4	14.7	-0.9	13.1
Capitol municipality (%)	21.9	21.7	23.9	-2.0	0.6
Large city municipality (%)	11.9	11.8	11.0	0.9	12.8
Small city municipality (%)	25.7	25.2	24.3	1.4	6.4
Hinterland municipality (%)	19.0	19.1	17.8	1.2	8.1
Rural municipality (%)	21.5	22.2	23.0	-1.5	3.9
First born (%)	37.9	37.7	36.4	1.5	7.6
Only child (%)	10.8	10.3	10.2	0.6	26.4
Lives w. both parents (%)	60.1	58.9	55.0	5.1	0.0
Father's inc. (1.000 DKK)	565.9	552.4	542.3	23.6	5.0
Mother's inc. (1.000 DKK)	399.4	394.1	386.8	12.6	0.7
Mother's age at birth	30.2	30.1	29.8	0.4	0.0
Father's age at birth	32.7	32.6	32.5	0.2	8.0
Missing father info (%)	2.3	2.6	3.3	-1.0	0.0
Missing mother info (%)	0.6	0.8	1.7	-1.1	0.0
Observations	2254	2650	7054		

Notes: In column (1), the *Sample* consists of the respondents who completed all the experiments on the online platform and are used in the analysis. In column (2), *Logins* are everyone who logged into the online platform. Column (3) shows descriptives for the random sample of 18 year old men who were invited to participate. Column (4) shows the difference between participants used in the analysis and everyone who was invited (incl. participants). Column (5) shows the P-values of the differences in column (4). The P-values are calculated using partially overlapping samples t-test with Welch's degrees of freedom (7, 8).

Table S2. Pairwise correlations of regressors and outcome

	Convicted	Risk	Impatience	Altruism	Self-control	GPA	Income	Conv. parent
Convicted (=1)	1.00	0.09	0.07	-0.07	-0.12	-0.17	-0.10	0.10
Risk tolerance	0.09	1.00	-0.12	-0.11	-0.04	-0.13	-0.05	0.04
Impatience	0.07	-0.12	1.00	-0.07	-0.11	-0.12	-0.07	0.01
Altruism	-0.07	-0.11	-0.07	1.00	0.01	0.27	0.08	-0.06
Self-control	-0.12	-0.04	-0.11	0.01	1.00	0.12	0.08	-0.04
GPA	-0.17	-0.13	-0.12	0.27	0.12	1.00	0.34	-0.12
Parental income	-0.10	-0.05	-0.07	0.08	0.08	0.34	1.00	-0.08
Convicted parent (=1)	0.10	0.04	0.01	-0.06	-0.04	-0.12	-0.08	1.00

Notes: The table shows pairwise correlations between the different key regressors in the analysis and between each regressor and the outcome (*Convicted*).

Table S3. Economic preferences and probability of having been convicted of an offence committed at age 15 to 20

	Probability of having been convicted	
Risk tolerance	7.87***	(2.37)
Impatience	4.97*	(2.44)
Altruism	-2.46	(2.42)
Self-control	-10.06***	(2.43)
GPA	-13.75***	(2.76)
Parental income	-1.57	(3.37)
Convicted parent (=1)	5.69**	(2.21)
Geography (=1)		
Middle Jutland	1.33	(2.39)
Southern Denmark	0.21	(2.35)
Copenhagen	3.26	(2.64)
Sealand	1.19	(2.74)
Urban area	-1.42	(1.75)
Family and background (=1)		
Immigrant	-2.50	(5.06)
Descendant	5.29*	(2.69)
Lives with both parents	-3.59*	(1.49)
Only child	1.67	(2.23)
First born	-4.62**	(1.60)
Misreport age or gender (=1)	3.53	(4.14)
Missing father information	-8.85	(9.05)
Father's age at birth (=1)		
24-25	-2.06	(4.87)
26-27	-2.30	(4.40)
28-29	-4.95	(4.46)
30-31	-2.35	(4.38)
32-33	-4.37	(4.43)
34-35	-6.65	(4.58)
36-37	-8.51	(4.76)
38-39	-6.96	(4.98)
>39	-8.70	(4.91)
Father's educational level (=1)		
Upper secondary	-0.30	(2.04)
Short cycle tertiary	0.64	(3.33)
Bachelor or equivalent	2.75	(2.71)
Master, Doctoral or equivalent	2.19	(3.02)
Missing	-3.82	(5.66)
Father's employment status (=1)		
Self-employed	2.23	(2.64)
Unemployed	-12.13	(6.90)
Not in the workforce	4.46*	(2.27)
Missing	1.49	(3.30)
Father's avg. unemployment 10 yr. (=1)		
1-10%	0.06	(1.90)
11-100%	-0.24	(2.44)
Missing mother information	1.89	(11.44)
Mother's age at birth (=1)		
24-25	1.78	(3.55)
26-27	2.49	(3.23)
28-29	4.70	(3.25)
30-31	3.97	(3.41)
32-33	2.59	(3.52)
34-35	2.76	(3.63)
36-37	1.82	(4.18)
38-39	5.19	(4.94)
>39	8.51	(5.62)
Upper secondary	-1.58	(2.18)
Mother's educational level (=1)		
Short cycle tertiary	-4.25	(4.12)
Bachelor or equivalent	-0.40	(2.49)
Master, Doctoral or equivalent	2.44	(3.17)
Missing	4.27	(5.93)
Mother's employment status (=1)		
Self-employed	3.32	(3.30)
Unemployed	-10.22	(5.89)
Not in the workforce	1.73	(2.06)
Missing	0.59	(5.43)
Mother's avg. unemployment 10 yr. (=1)		
1-10%	1.42	(1.84)
11-100%	4.83*	(2.08)
Observations	2254	

Notes: The table reports the marginal effects in p.p. from the same probit model as in column (6) of Table 1. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table S4. Economic preferences and probability of having been convicted of an offence committed at age 15 to 20

	Probability of having been convicted				
	(1)	(2)	(3)	(4)	(5)
Present bias	2.0			0.8	-3.5
	(2.4)			(2.3)	(2.7)
Aheadness aversion		-1.8		-1.8	-2.3
		(2.4)		(2.3)	(2.6)
Behindness aversion			10.9***	1.4	3.0
			(2.5)	(2.7)	(2.7)
Risk tolerance				7.9***	7.9**
				(2.4)	(2.4)
Impatience				4.9*	6.2*
				(2.4)	(2.7)
Self-control				-10.0***	-10.3***
				(2.4)	(2.4)
GPA				-13.9***	-13.4***
				(2.8)	(2.8)
Parental income				-1.5	-2.1
				(3.4)	(3.4)
Convicted parent (=1)				5.7**	5.5*
				(2.2)	(2.2)
Observations	2254	2254	2254	2254	2254
Individual controls				✓	✓
Parental controls				✓	✓
Structural prefs.					✓

Notes: The table reports the marginal effects in p.p. from estimated probit models. *Present bias*, *Behindness aversion*, *Aheadness aversion*, *Risk tolerance*, *Impatience*, *Self-control*, *GPA*, and *Parental income* are all within cohort in sample ranks. *Convicted parent* is an indicator. *Individual controls* include regional FE, an urban area indicator, immigrant and descendant status, a living with both parents indicator, an only child indicator, a first born indicator and an indicator for misreported age or gender in survey. *Parental controls* include educational level, age at child's birth, employment status and unemployment history. *Structural prefs.* indicates that we use structurally estimated preference measures in column (5) instead of the non-parametric measures. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table S5. Economic preferences and probability of having been convicted of an offence committed at age 15 to 20

	Baseline	Math GPA	LPM	Bias-adjusted	z-scores	Weighted	Incl. women	All information
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Risk tolerance	7.9***	7.9***	8.1**	7.3	2.3***	9.1**	4.7***	6.8**
	(2.4)	(2.4)	(2.5)		(0.7)	(3.0)	(1.4)	(2.5)
Impatience	5.0*	5.1*	4.7*	3.7	1.6*	7.0*	5.4***	4.5
	(2.4)	(2.4)	(2.4)		(0.7)	(3.0)	(1.5)	(2.5)
Altruism	-2.5	-2.3	-2.4	-0.1	-0.9	-1.0	0.2	-2.8
	(2.4)	(2.4)	(2.5)		(0.7)	(3.0)	(1.4)	(2.5)
Self-control	-10.1***	-9.9***	-10.2***	-8.5	-2.9***	-11.3***	-7.0***	-8.0**
	(2.4)	(2.5)	(2.6)		(0.6)	(3.1)	(1.4)	(2.6)
GPA	-13.7***	-12.5***	-13.1***	-8.0	-3.8***	-17.2***	-12.2***	-12.0***
	(2.8)	(2.6)	(2.7)		(0.7)	(3.7)	(1.6)	(2.9)
Parental income	-1.6	-1.7	-2.3	6.9	-0.5	-3.5	0.0	-1.5
	(3.4)	(3.4)	(3.5)		(0.5)	(4.0)	(1.9)	(3.6)
Convicted parent (=1)	5.7**	5.6*	7.8*	6.4	5.5*	7.2**	3.4**	5.0*
	(2.2)	(2.2)	(3.1)		(2.2)	(2.7)	(1.2)	(2.3)
Observations	2254	2254	2254	2254	2254	2254	4503	2011
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Parental controls	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Column (1) is the baseline result from column (6) of Table 1 in the main text. Column (2) uses GPA only for math. Column (3) shows OLS estimates in p.p. from a linear probability model. Column (4) shows biased-adjusted OLS estimates using the Oster bounding approach (9) with $\delta = 1$ and $\Pi = 1.3$. Column (5) uses z-scores for the explanatory variables instead of ranks. Column (6) weighs the observations with the inverse probability of being in the sample using all explanatory variables that are also available for non-participants. Column (7) includes women born in 1999. Column (8) excludes participants for whom we do not observe all information in the administrative data, most importantly GPA and parental information. *Risk tolerance*, *Impatience*, *Altruism*, *Self-control*, *GPA*, and *Parental income* are all within cohort in sample ranks. *Convicted parent* is an indicator. *Controls* include regional FE, an urban area indicator, immigrant and descendant status, a living with both parents indicator, an only child indicator, a first born indicator, an indicator for misreported age or gender in survey, parents' educational level, parents' age at child's birth, parents' employment status and unemployment history. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table S6. Choice situations in the time experiment

	Low stakes		High stakes				
Situation	x_1	x_2	x_1	x_2	t_1	t_2	Rate
1	250	251	7500	7530	0	2	0.024
2	250	256	7500	7680	0	2	0.153
3	250	261	7500	7830	0	2	0.295
4	250	266	7500	7980	0	2	0.451
5	250	271	7500	8130	0	2	0.622
6	250	276	7500	8280	0	2	0.811
7	250	281	7500	8430	0	2	1.016
8	250	286	7500	8580	0	2	1.242
9	250	251	7500	7530	2	4	0.024
10	250	256	7500	7680	2	4	0.153
11	250	261	7500	7830	2	4	0.295
12	250	266	7500	7980	2	4	0.451
13	250	271	7500	8130	2	4	0.622
14	250	276	7500	8280	2	4	0.811
15	250	281	7500	8430	2	4	1.016
16	250	286	7500	8580	2	4	1.242

Notes: x_1 is the amount the participant can get paid out sooner and x_2 is the amount the participant can get paid out later. They differ by whether the participant was assigned to the low or high stake treatment. t_1 indicates the sooner payout time (either within 24 hours (0) or in 2 months) while t_2 indicates the later payout time (either in 2 months or in 4 months). The user interface displayed the delays in weeks to avoid confounds by payments at different weekdays. Rate is the annualized interest rate the participant gets on the amount saved for two months. For instance, in Situation 1, Rate=0.024 refers to a yearly interest rate of 2.4%. Panel (a) of Figure 1 in the main text illustrates situation 10 with low stakes.

Table S7. Choice situations in the risk experiment

Situation	DKK	p	Good	Bad
1	250	0.5	1.21	0.81
2	250	0.2	1.41	0.91
3	250	0.8	1.11	0.61
4	250	0.5	1.31	0.71
5	250	0.2	1.61	0.86
6	250	0.8	1.16	0.41
7	250	0.5	1.35	0.75
8	250	0.2	1.65	0.90
9	250	0.8	1.20	0.45
10	250	0.6	1.50	0.40
11	250	0.4	1.72	0.62
12	250	0.6	1.45	0.35
13	250	0.4	1.67	0.57
14	250	0.5	1.51	0.50
15	250	0.5	1.61	0.60

Notes: *DKK* is the amount the participant can keep or invest in the lottery. p is the probability that the lottery will give the good state. *Good* is the multiplier of the investment in the good state while *Bad* is the multiplier of the investment in the bad state. Panel (b) of Figure 1 in the main text illustrates situation 10.

Table S8. Choice situations in the social experiment

Situation	own_1	$other_1$	own_2	$other_2$	Cost of giving
1	262.5	137.5	112.5	237.5	1.500
2	250.0	125.0	125.0	250.0	1.000
3	237.5	112.5	137.5	262.5	0.667
4	225.0	112.5	150.0	262.5	0.500
5	212.5	112.5	162.5	262.5	0.333
6	212.5	100.0	162.5	275.0	0.286
7	200.0	100.0	175.0	275.0	0.143
8	187.5	100.0	187.5	275.0	-0.000
9	175.0	100.0	200.0	275.0	-0.143
10	175.0	112.5	200.0	262.5	-0.167
11	162.5	100.0	212.5	275.0	-0.286
12	162.5	112.5	212.5	262.5	-0.333
13	187.5	187.5	200.0	250.0	-0.200
14	212.5	112.5	187.5	187.5	0.333
15	187.5	187.5	212.5	262.5	-0.333
16	250.0	125.0	187.5	187.5	1.000
17	187.5	187.5	225.0	275.0	-0.429
18	262.5	162.5	187.5	187.5	3.000
19	187.5	187.5	192.5	292.5	-0.048
20	192.5	92.5	187.5	187.5	0.053

Notes: own_1 is the amount the participant gets if he/she gives the smallest possible amount to the other person. $other_1$ is the smallest possible amount to give. own_2 is the amount the participant gets if he/she gives the largest possible amount to the other person. $other_2$ is the largest possible amount to give. Thus, for a given choice situation ($other_1$, $other_1$) and ($other_2$, $other_2$) represent the most extreme allocations in the set of feasible payoff allocations. In every choice situation there were 9 further feasible payoff allocations located between the extremes in an equi-distant way. *Cost of giving* denotes the cost of the participant per DKK given to the other person. A negative value means that the participant benefit per DKK given to the other person.

References

1. G Allenby, An Introduction to Hierarchical Bayesian Modeling, Tutorial Notes, Advanced Research Technique Forum, American Marketing Association (1997).
2. K Train, A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit, Working Paper, Department of Economics, University of California, Berkeley (2001).
3. O Morgenstern, JV Neumann, *Theory of Games and Economic Behavior.* (Princeton university press), (1953).
4. D Laibson, Golden Eggs and Hyperbolic Discounting. *The Q. J. Econ.* **112**, 443–478 (1997).
5. S Frederick, G Loewenstein, T O'Donoghue, Time Discounting and Time Preference: A Critical Review. *J. Econ. Lit.* **40**, 351–401 (2002).
6. E Fehr, KM Schmidt, A Theory of Fairness, Competition, and Cooperation. *The Q. J. Econ.* **114**, 817–868 (1999).
7. B Derrick, B Russ, D Toher, P White, Test statistics for the comparison of means for two samples that include both paired and independent observations. *J. Mod. Appl. Stat. Methods* **16**, 137–157 (2017).
8. B Derrick, D Toher, P White, How to compare the means of two samples that include paired observations and independent observations: A companion to Derrick, Russ, Toher and White (2017). *The Quant. Methods for Psychol.* **13**, 120–126 (2017).
9. E Oster, Unobservable Selection and Coefficient Stability: Theory and Evidence. *J. Bus. & Econ. Stat.* **37**, 187–204 (2019).