SUPPLEMENTARY MATERIALS

Supplementary note 1: Summary statistics

Study 1: Risky choice dataset. Subjects chose the safe option in 66.1% of the trials (s.d. = 20%). The mean estimate of the loss aversion coefficient λ was 2.68 (s.d. = 1.65), ranging from 0.75 to 7.56 across subjects. The mean RT when choosing the safe option was 1.57 s (s.d. = 0.47), when choosing the risky option: 2.38 s (s.d. = 1.42) (RT in the histogram below truncated to 5 s for presentation purposes).

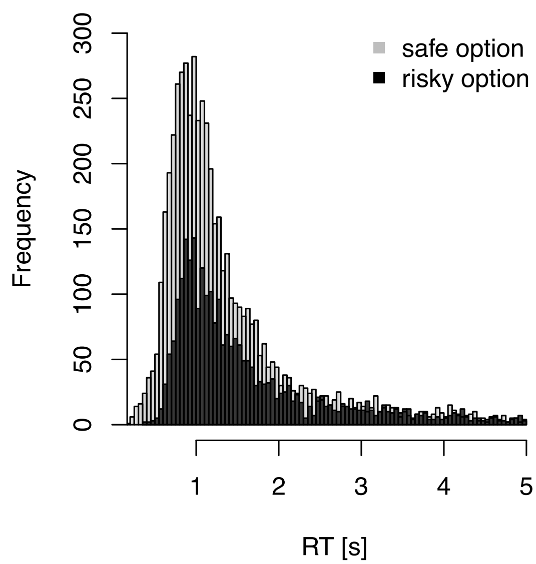


Figure S1. Choice and RT distributions in the risky choice dataset.

Study 2: Intertemporal choice. Subjects chose the later option in 49% of the trials (s.d. = 24%). The mean estimate of the discount coefficient k was 0.01 (s.d. = 0.008), ranging from 0.0007 to 0.34 across subjects. The mean RT when choosing the later option was 1.29 s (s.d. = 0.28), when choosing the sooner option was 1.27 s (s.d. = 0.24).

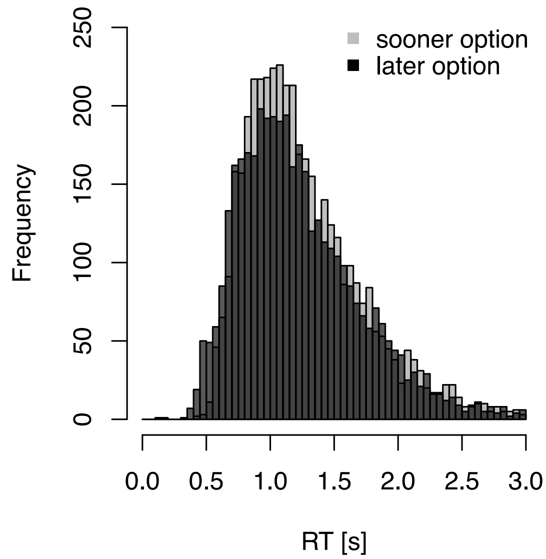


Figure S2. Choice and RT distributions in the intertemporal choice dataset.

Study 3: Social preference. (a) Disadvantageous inequality. Subjects chose the altruistic option in 15% of the trials (s.d. = 9%). The mean estimate of the preference parameter α was -0.09 (s.d. = 0.19), ranging from -0.46 to 0.64 across subjects. The mean RT when choosing the altruistic option was 3.08 s (s.d. = 1.1), and the selfish option was 2.63 s (s.d. = 0.97).

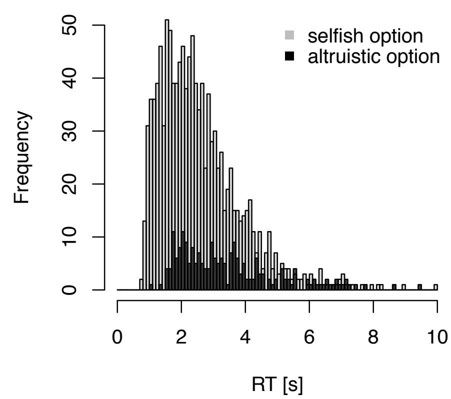


Figure S3. Choice and RT distributions in the social preference dataset (disadvantageous inequality).

(b) Advantageous inequality. Subjects chose the altruistic option in 33% of the trials (s.d. = 20%). The mean estimate of the preference parameter β was 0.36 (s.d. = 0.32), ranging from -0.34 to 1.07 across subjects. The mean RT when choosing the altruistic option was 2.81 s (s.d. = 0.9), when choosing the selfish option was 2.5 s (s.d. = 0.9).

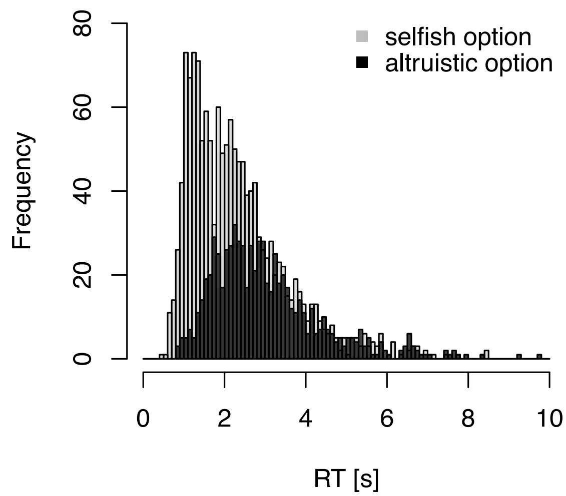


Figure S4. Choice and RT distributions in the social preference dataset (advantageous inequality).

Supplementary Note 2: Estimation methods

Choice-based method

We estimate each individual preference function $u(\cdot\,|\theta)$, where θ is a subject-specific parameter (temporal discount factor, loss aversion, or inequity aversion), in the standard way as follows. We assume that for each pair of options and choice $a = 1{,}2$ the error terms in utilities follow the type I extreme value distribution, so the probability of choosing option 1 is a logistic function

$$p(a_i = 1) = \frac{1}{1+e^{-\mu(u_1(\cdot|\theta)-u_2(\cdot|\theta))}}, \tag{1}$$

where μ and θ are free parameters that can be estimated for each subject individually maximizing a likelihood function

$$LL = \sum_n \left(\log(p(a_n = 1)) \cdot 1(a_t = 1) + \log\big(1 - p(a_n = 1)\big) \cdot 1(a_t = 2)\right), \tag{2}$$

where n is the trial number, a_n is the choice made by the subject on that trial, and $1(\cdot)$ is the indicator function. Figure S1 shows the subject level correlations between the predicted and the actual choices.

Top RT decile method

For each decision problem on each trial n, we calculate the indifference parameter value θ_n^{ind} as a solution to the equation

$$u_1(\cdot\,|\theta_n) = u_2(\cdot\,|\theta_n). \tag{3}$$

Then we average the indifference values on the trials in the top RT decile to obtain the final parameter estimate:

$$\hat{\theta} = \frac{\sum_n(\theta_n^{ind} \cdot 1(F(RT_n) \geq 0.9))}{\sum_n 1(F(RT_n) \geq 0.9)}, \tag{4}$$

where RT_n is the response time on trial n, $F(\cdot)$ is the empirical RT distribution for the specific subject, and $1(\cdot)$ is the indicator function.

Local regression (LOWESS) method

As the top 10% approach only utilizes part of the data, we developed another approach based on the peaks in RTs, this time using all the available RT data. The local regression (LOWESS) method also uses the "revealed indifference" approach: for each individual subject, we run a local polynomial regression of RTs on the indifference parameter values and use that regression to identify the indifference value that produces the highest predicted RT (see **Figure 4f**). As the peak of this line is typically close to the choice-based estimate and the observations with the highest RTs, this method is quite similar in its predictions to the top-RT-decile method. This method requires the researcher to choose the local regression smoothing parameter. Here we use a value of 0.5 as it produces the best results across all of the data sets (though other values can work better for specific datasets; see **Figure S7**).

As in the previous method, for each decision problem on each trial n, we estimate the indifference parameter value θ_n^{ind} solving the equation

$$u_1(\cdot\,|\theta_n) = u_2(\cdot\,|\theta_n). \tag{5}$$

For each individual subject, we regress response time $\log(RT)$ in every trial n on the corresponding indifference parameter value θ_n^{ind} using a local polynomial regression (LOWESS, (Cleveland, 1979)) in the R package *stats*:

$$RT = g\left(\theta_n^{ind}\right) + \varepsilon_n. \tag{6}$$

Then we obtain the parameter estimate $\hat{\theta}$ by inverting the fitted regression line at the maximum predicted response time $\widehat{RT}$:

$$\hat{\theta} = \hat{g}^{-1}\left(\max\left(\widehat{RT}\right)\right). \tag{7}$$

Although this approach generally produces results similar to the top decile approach, it can be affected by outliers (e.g. sparse data and unusually high RTs around extreme indifference points) and thus sometimes misestimates individual subject parameters, producing predictions that are in some cases worse than those produced by the top RT decile approach: social choice α: r = 0.12, p = 0.52; social choice β: r = 0.6, p = 0.001; intertemporal choice k: r = 0.44, p = 0.01; risky choice λ: r = 0.88, p < 0.001; Pearson correlations. This can be mitigated by using a specific smoothing parameter for each data set, but our goal was to identify a method that works well across all the datasets.

Drift-diffusion model (DDM) method

In the DDM a latent decision variable evolves over time with an average drift rate plus Gaussian noise (the Wiener diffusion) until it reaches one of two pre-determined boundaries, which correspond to the two choice options. Given the boundary separation parameter, the drift rate, the non-decision time (the component of RT not attributable to the decision process itself, e.g. moving one's hand to indicate the choice), and the variance of the Gaussian noise, it is possible to calculate choice probabilities and choice-contingent RT distributions.

In the model, we assume that a subject observes a set of alternatives $j \in \{1,2\}$. The choice process involves two components: a constant boundary threshold b and a decision variable $y(t)$ that evolves over time according to the following differential equation:

$$dy(t) = v \cdot dt + \sigma \cdot dW \tag{8}$$

where $y(t)$ is accumulated evidence towards option 1 (with $y(0) = 0$), v is the drift rate, which is assumed to be a linear function of the subjective-value difference:

$$v \equiv z \cdot (u_1(\theta) - u_2(\theta)), \tag{9}$$

where $z \in R^+$ is a scaling parameter, $u_j(\cdot)$ is the subjective value of the given alternative, and θ is the subject-specific parameter. Finally, $\sigma \cdot dW$ is a Wiener process (i.e. Brownian motion) that represents Gaussian white noise with variance σ^2. Without loss of generality, we normalize $\sigma = 1$ as it can only be identified up to scale (due to the arbitrary units on y).

We define the response time RT as the first time that the absolute value of the decision variable reaches a boundary $b \in R^+$, plus a non-stochastic component known as non-decision time ($\tau \in R^+$, typically interpreted as the time that a subject needs to process the information on the screen):

$$RT = \min\{t: |y(t)| \geq b\} + \tau. \tag{10}$$

The choice outcome $a \in \{1,2\}$ is defined as follows:

$$a = \begin{cases} 1 \; if \; y(RT) = b \\ 2 \; if \; y(RT) = -b \end{cases} \tag{11}$$

Now, assuming without loss of generality $v \geq 0$, we can calculate the choice probabilities $p(a = j)$, the expected RT, and an approximate probability density function (PDF) for the RTs (minus τ) as follows (Wabersich & Vandekerckhove, 2014):

$$p(a = 1) = \begin{cases} \frac{e^{2vb}-1}{e^{2vb}-e^{-2vb}} & if \; v > 0 \\ \frac{1}{2} & if \; v = 0 \end{cases} \tag{12}$$

$$E[RT] = \begin{cases} \frac{b}{v}\left(1 - \frac{2}{e^{2vb}-e^{-2vb}}\right) + \tau & if \; v > 0 \\ b^2 + \tau & if \; v = 0 \end{cases} \tag{13}$$

$$f(t) = \frac{\pi}{4b^2} e^{-\frac{v^2 t}{2}} \sum_{m=1}^{\infty} (-1)^{m-1} m \cdot e^{-\frac{m^2 \pi^2 t}{8b^2}} \sin\left(\frac{m\pi}{2}\right) (e^{vb} + e^{-vb}) \tag{14}$$

We use a density function of the Wiener distribution from the RWiener R package (Wabersich and Vandekerckhove 2014) to estimate the likelihood (14) for the observed RT on every given trial assuming a set of parameters $(b, \tau, z, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of individual subjects' parameters. Essentially, the identification of the individual parameters is possible due to the fact that RTs are predicted to vary as the subjective value difference v varies across trials and subjects.

One alternative way to assess the goodness-of-fit is to examine the number of choices that are consistent with the estimated parameter values. To do so, we used the RT-estimated parameters to identify the "preferred" alternatives in every trial and compared those to the actual choice outcomes. RT-estimated parameters were able to explain a high proportion of choices in the datasets (social choice α: 79% respectively ($p < 0.001$); social choice β: 80% ($p < 0.001$); intertemporal choice: 76% ($p < 0.001$); risky choice: 77% ($p < 0.001$); p-values denote two-sided Wilcoxon signed rank test significance at the subject level, comparing these proportions to chance). For a stricter test, we calculated an average of all indifference points for each experiment (which roughly corresponds to the mean of the experimenter's prior parameter distribution) and made choice predictions for each subject using this single value. The DDM accuracy rates beat this baseline in two out of four cases (for the intertemporal choice and social choice β, $p < 0.05$, two-sided Wilcoxon signed rank test).

Chabris et al. (2009) method

Here we follow the method suggested by Chabris et al. (2009), which uses the full RT distribution to estimate the preference parameters.

Let the difference in the two utilities on trial *n* be

$$\Delta_n \equiv |u_{1n}(\cdot\,|\theta) - u_{2n}(\cdot\,|\theta)|. \tag{15}$$

Assume that the decision difficulty is a convex and decreasing function of this difference:

$$\Gamma(\Delta_n) \equiv \frac{2}{1+e^{\omega \Delta_n}}, \tag{16}$$

where ω is a free parameter.

The response times are then modeled as a function of the trial number and the choice difficulty:

$$RT_n = \beta_0 + \beta_1 n + \beta_2 \Gamma(\Delta_n) + \varepsilon_n. \tag{17}$$

To estimate the set of parameters $(\hat{\theta}, \hat{\omega}, \hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$ we follow the original paper and minimize the error function

$$\sum_n \left(RT_n - \beta_0 - \beta_1 n - \beta_2 \Gamma(\Delta_n)\right)^2. \tag{18}$$

Unlike Chabris et al. (2009), we estimate this model for every individual subject (see Table S1).

SUPPLEMENTARY TABLES

Table S1

	DDM	Chabris et al (2009)	Top 10% RT	LOWESS (0.5)
Social choice (α)	0.39	0.22	**0.44**	0.12
Social choice (β)	0.52	0.53	0.56	**0.6**
Intertemporal choice	0.57	0.59	**0.71**	0.44
Risk choice	0.36	0.16	0.64	**0.88**

Pearson correlation between parameters estimated from choices and RTs, at the subject level. The best performing method for each data set is shown in bold. For estimation methods details see Supplementary Note 2.

SUPPLEMENTARY FIGURES

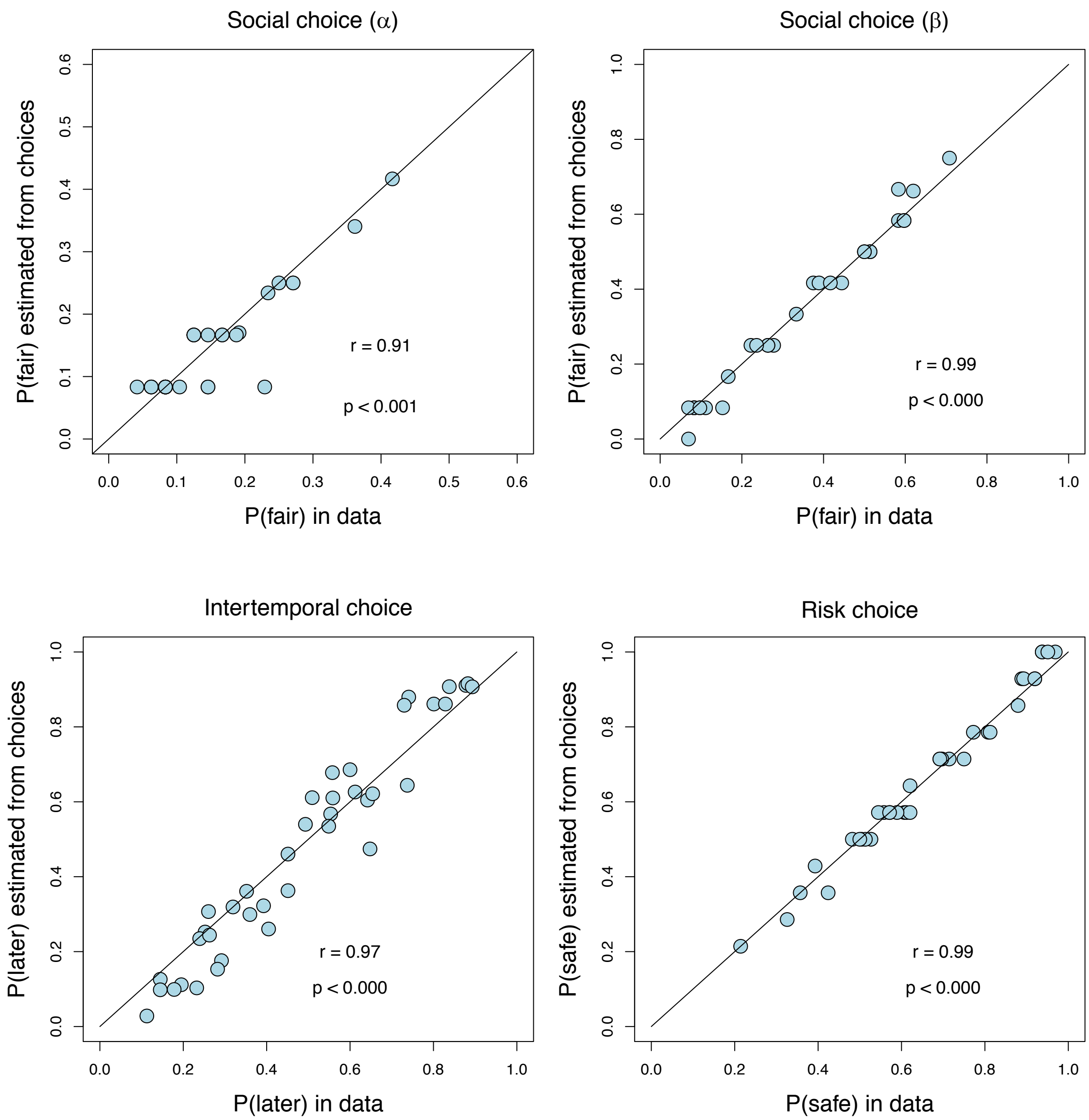
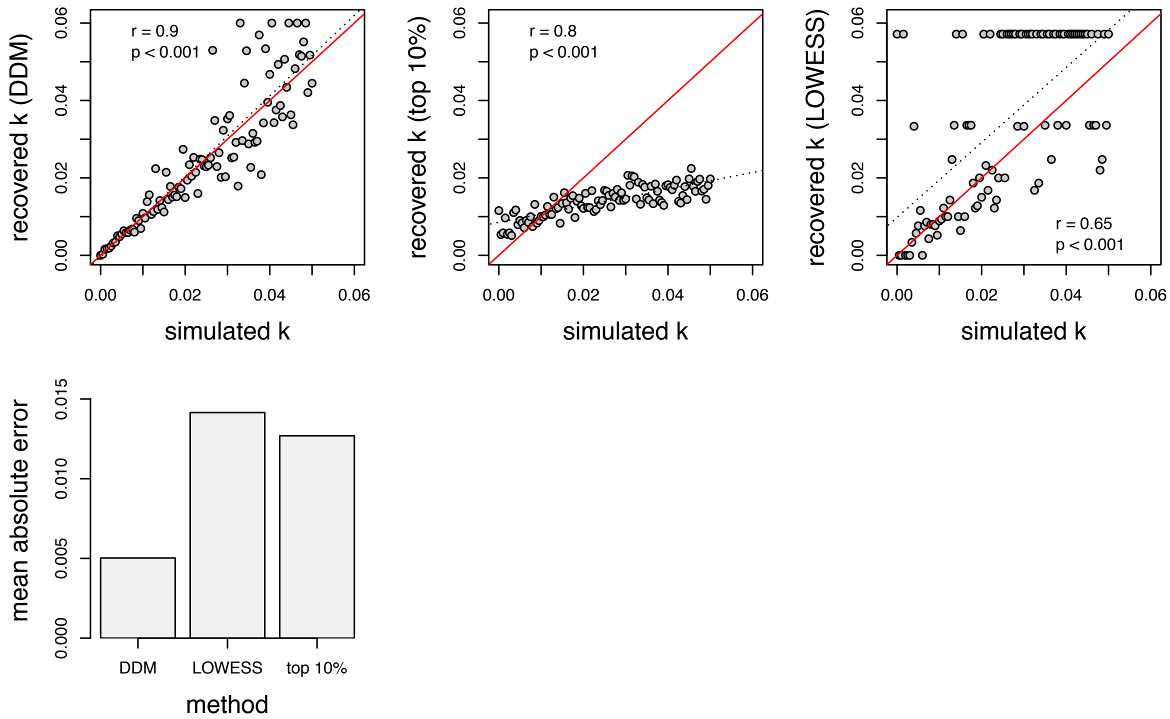
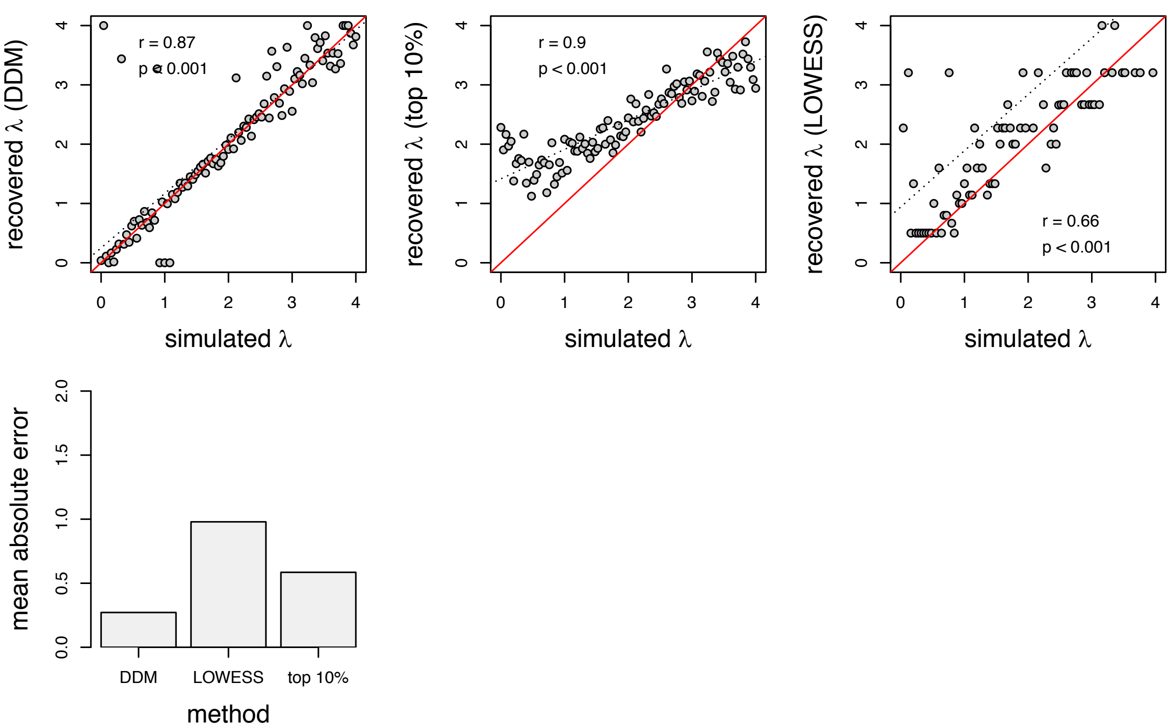


Figure S5. Subject-level correlation (Pearson) between choice proportions in the data and as predicted by choice-estimated preference functions (see Supplementary Note 2 for details). The solid lines are 45 degrees.

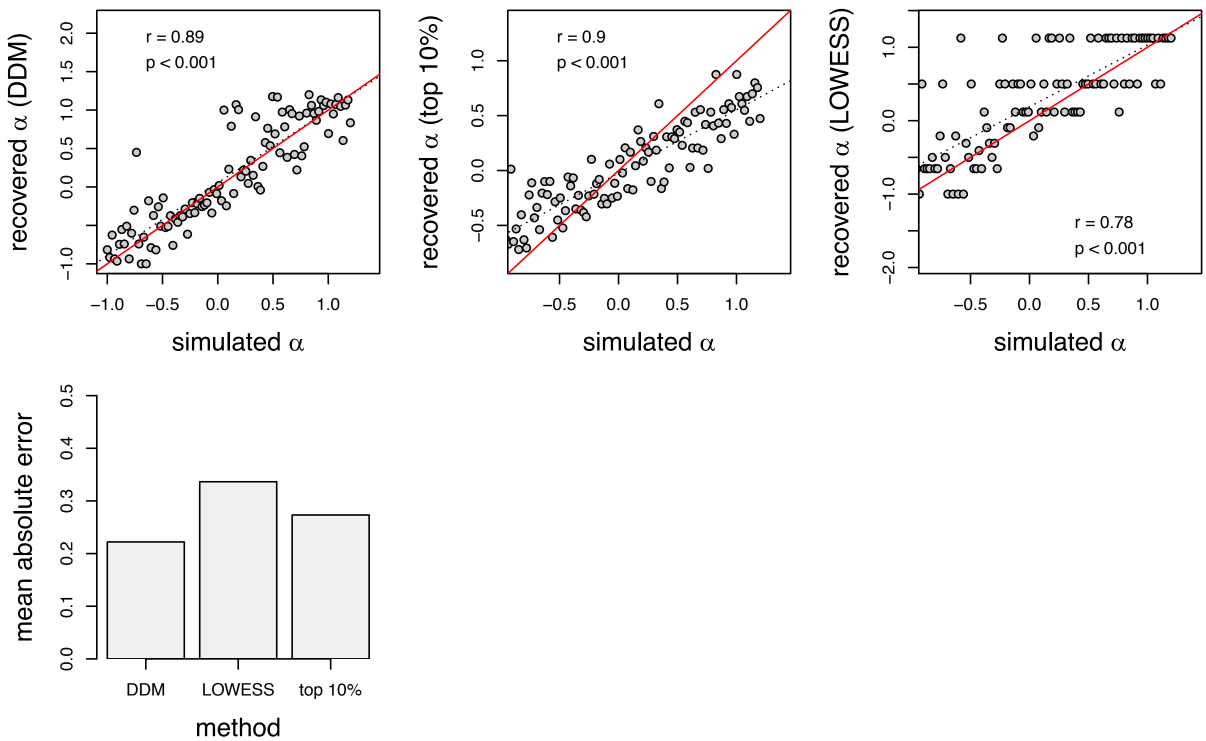
Intertemporal choice



Risky choice



Social choice (α)



Social choice (β)

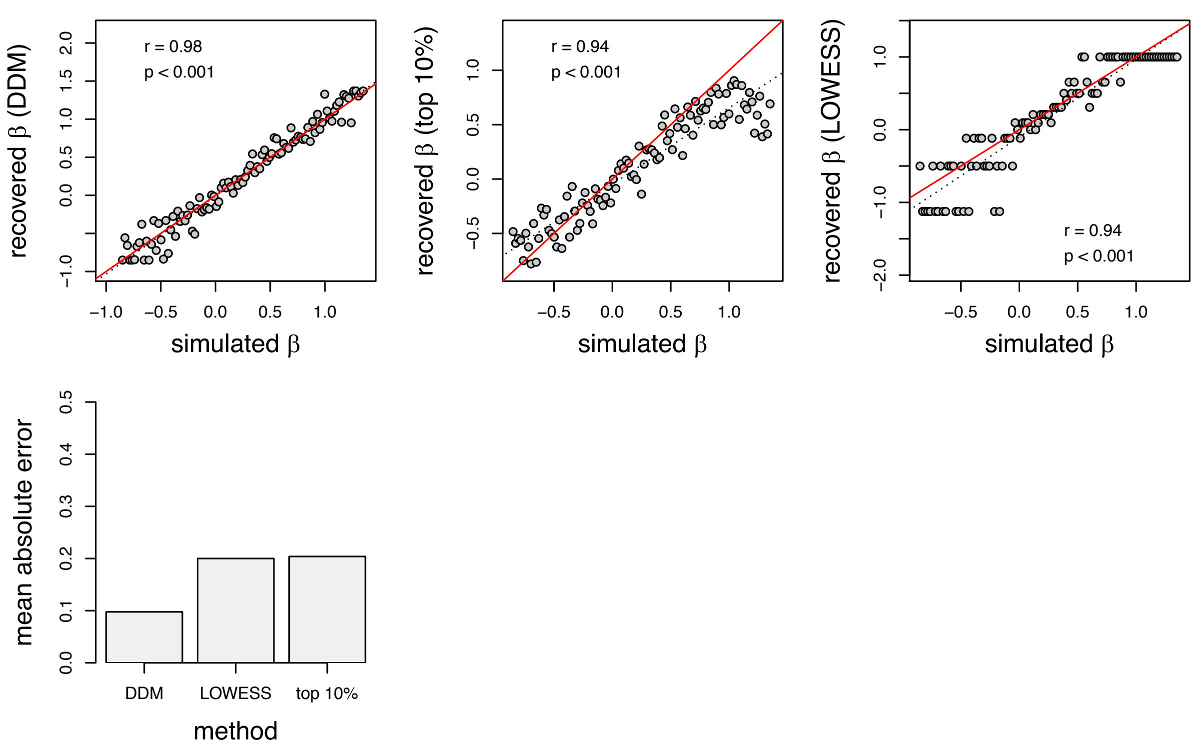


Figure S6. Parameter recovery for three parameter estimation methods: DDM using only RT (top left panel), top 10% RT trials (top middle panel), and LOWESS regression (top right panel). For each dataset we generated 101 simulated subjects with the preference parameter varying from the lowest to the highest indifference value used in the experiment, using the same set of decision problems, with the other parameters of the DDM fixed at the median values across real subjects. For each simulated subject (each having 216 (intertemporal), 224 (risky), 48 (social α), and 72 (social β) trials) we recovered the value of the preference parameter using the three methods described in Supplementary Note 2. The bottom panel in each plot shows the mean absolute error for all three methods. The scatter plots report Pearson correlations between the simulated and the recovered values of the preference parameter.

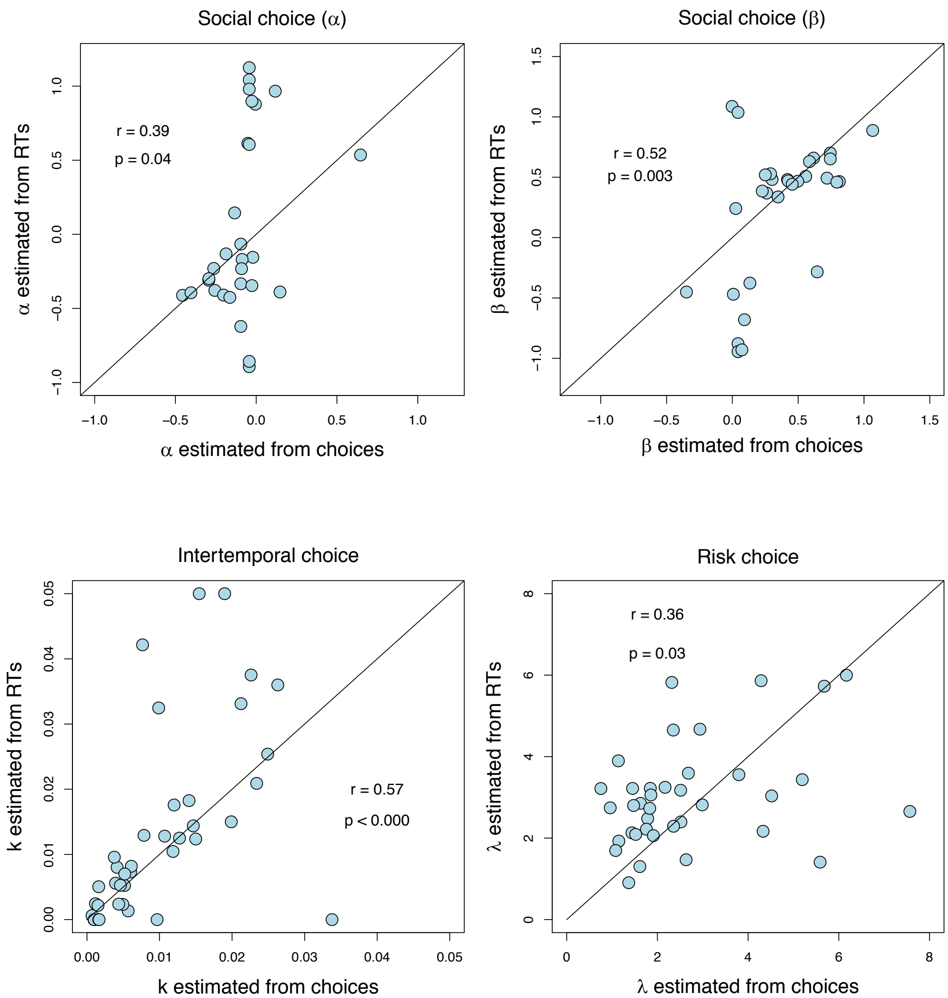


Figure S7. The DDM estimates of subjects' preference function parameters, estimating DDM parameters at the group level. Subject-level correlation (Pearson) between parameters estimated from choice data and RT data using the drift-diffusion model (DDM). The solid lines are 45 degree lines.

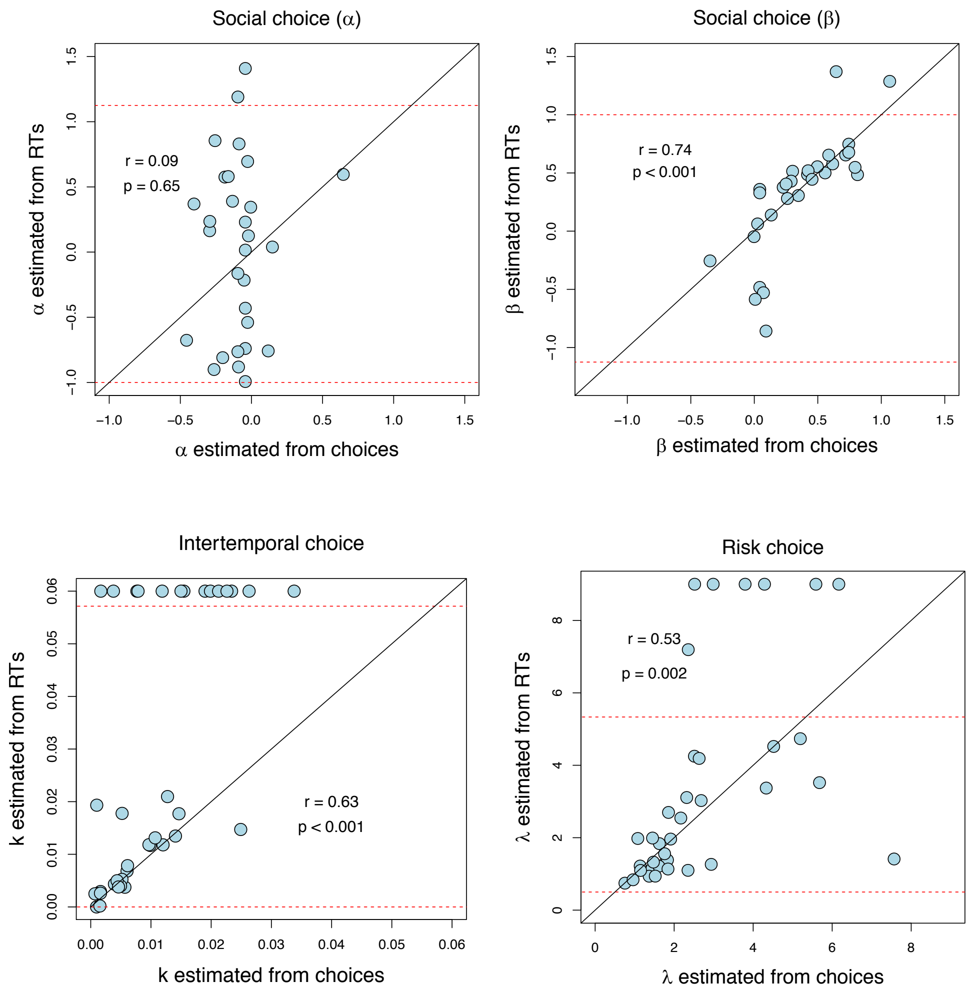


Figure S8. The DDM estimates of subjects' preference function parameters, fitting the model to each subject individually. Subject-level correlation (Pearson) between parameters estimated from choice data and RT data using the drift-diffusion model (DDM), including only the subjects with parameters estimated within the range of indifference points in the experiment (red dotted lines). The solid lines are 45 degree lines.

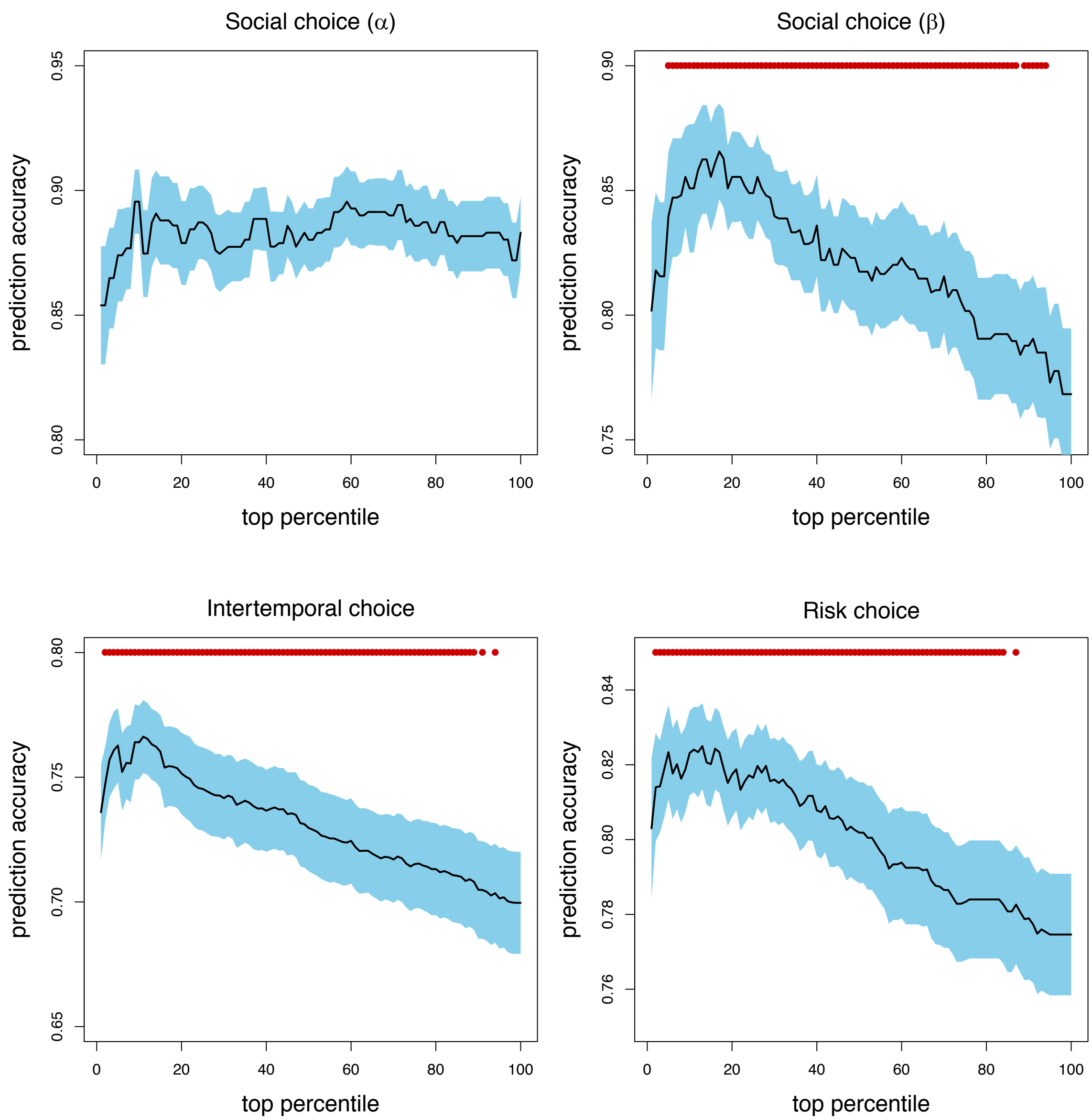


Figure S9. Choice prediction accuracy as a function of the percentage of slowest trials used in the parameter estimation from 1 to 100%. Averaging the longest 10-20% RT trials provides the best choice prediction accuracy. The solid black lines denote mean prediction accuracy across subjects, the shaded areas show standard errors at the subject level, the red dots above the graphs indicate significant difference from the baseline at the p = 0.05 level for each corresponding percentile (Wilcoxon signed rank test). The baseline is the average of all indifference points across trials. It is important to emphasize that the baseline to which we compare the predictive power is not chance (50%) as almost any experimenter uses some prior knowledge of the parameter distribution in the population to select their choice problems. For example, an experimenter studying intertemporal choice might select a set of choice problems so that the average subject would choose the immediate option half of the time and the delayed option the other half of the time. So if you were to average the indifference points from the trials in such an experiment, you would be able to predict behavior quite accurately, on average. In such an experiment, behavior in trials with extreme indifference points will be very predictable. That is, on a trial designed to make a very patient subject indifferent, most subjects will have a strong preference for the immediate option. Similarly, on a trial designed to make an impatient person indifferent, most subjects will have a strong preference for the delayed option. Thus behavior in many of an experiment's trials is quite easy to predict because those trials are only included to identify parameter values for extreme subjects. For instance, a single loss-aversion coefficient of $\lambda = 2$ can predict about 75% of choices in our risky-choice dataset.

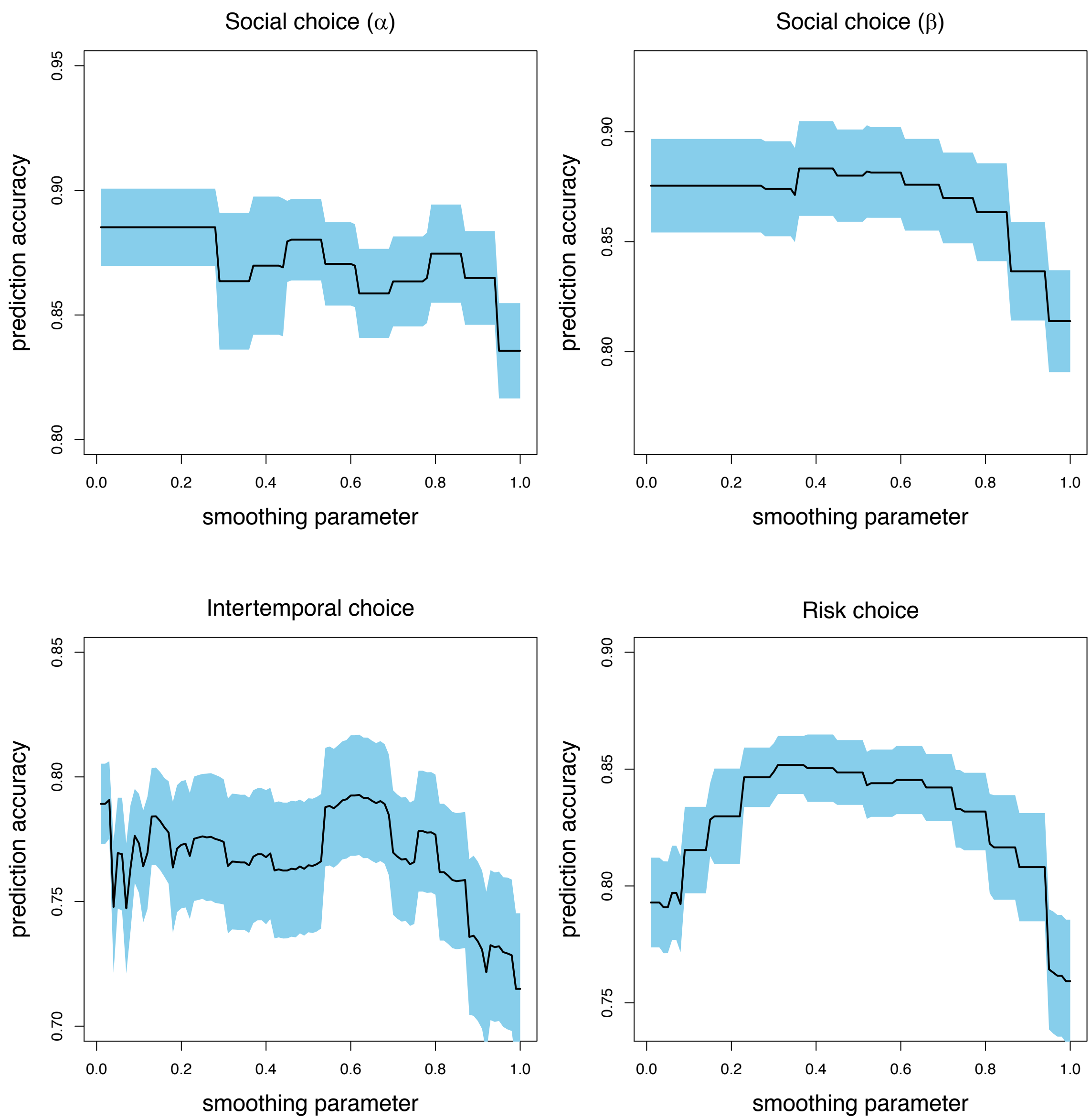


Figure S10. Choice prediction accuracy as a function of the smoothing parameter of the LOWESS regression model. The solid black lines denote mean prediction accuracy across subjects, the shaded areas show standard errors at the subject level.