

Supporting information for “The effects of perspective taking on multiple dimensions of discrimination”

Contents

1	Model specification	2
1.1	Equation	2
1.2	Software implementation	3
2	Sample description	4
2.1	Time taken	6
2.2	Vignettes assigned	6
3	Full model results	7
4	Robustness checks and sensitivity analysis	14
4.1	Sensitivity to priors	14
4.2	Low quality individuals - Quick replies	15
4.3	Low quality individuals - Same profile order	16
4.4	Compliance	17
4.5	NLP coding of open responses	18
5	Certainty of the response	19
6	Convergence tests	20

1 Model specification

1.1 Equation

Model description:

$Y_o \sim$	$\mathcal{B}(\pi_o)$	Main data component
$\pi_o =$	$\text{logit}(\beta_{i,f} F_f)$	Linear relationship
$\beta_{i,f} \sim$	$\mathcal{MN}(\mu_{i,f}, \Sigma_\beta)$	
$\Sigma_\beta \sim$	$\mathcal{IW}(0, 9)$	VCov for individual effects
$\mu_{i,f} =$	$(\theta_{f,v} C_{i,v})$	Explanation of individual behavior
$\theta_{f,v} \sim$	$\mathcal{N}(\omega_f, \sigma_\theta)$	Priors for explanatory variables of individual behavior
$\omega_f \sim$	$\mathcal{N}(0, 2.5)$	Prior for the effects shared by outcome
$\sigma_{\theta_{f,v}} \sim$	$\mathcal{U}(0, 1)$	Prior for SD of individual behavior

Where:

- Y : Outcome variable capturing whether a profile has been prioritized (1) or not (0).
- o : Observation
- i : Individual
- F : Matrix with the observations of features (the discrimination sources, plus intercept and first shown profile), for each experimental data point.
- C : Matrix with the characteristics v of the respondents, including their population p and the treatment t .
- f : Feature
- $\theta_{f,v}$: Main parameters of interest capturing the individual variables affecting discrimination effects by profiles' feature.
- ω_f : Hyper-parameters capturing the shared effect of individual characteristics on features.
- Σ_β : Variance-covariance matrix of the individual effects prioritizing the different profiles.
- σ_θ : Within feature/individual characteristic's standard deviations.

1.2 Software implementation

The JAGS code for the model is the following:

```
1 model {
2   for (o in 1:nO) {
3     Y[o] ~ dbern(p[o])
4     logit(p[o]) <- inprod(beta[id[o],1:nF], X[o,1:nF])
5   }
6   #
7   # Priors for effects
8   #
9   for (id in 1:nId) {
10    for (f in 1:nF) {
11      beta[id,f] ~ dnorm(mu[id,f], 2.5^-2)
12    }
13  }
14  for (f in 1:nF) {
15    for (id in 1:nId) {
16      mu[id,f] <- inprod(theta[f,1:nCov], C[id,1:nCov])
17    }
18    for (cov in 1:nCov) {
19      theta[f,cov] ~ dnorm(omega[f,cov], tau_theta[f,cov])
20    }
21  }
22  for (cov in 1:nCov) {
23    omega[f,cov] ~ dnorm(0, 2.5^-2)
24    tau_theta[f,cov] <- pow(sigma_theta[f,cov], -2)
25    sigma_theta[f,cov] ~ dt(0, 0.5^-2, 3)I(0,)
26  }
27  #
28  # Missing data
29  #
30  for (id in 1:nId) {
31    for (v in cov.missing) {
32      C[id,v] ~ dnorm(0, 1^-2)
33    }
34  }
35 }
```

2 Sample description

Table 1: Summary descriptives table by groups of 'Treatment'

	No N=534	Complier (Language) N=375	Complier (Bureaucracy) N=73	Non-complier (No idea) N=62	Non-complier (Bad quality) N=32	p.overall
Gender:						0.681
Female	278 (52.1%)	193 (51.5%)	33 (45.2%)	36 (58.1%)	17 (53.1%)	
Male	256 (47.9%)	182 (48.5%)	40 (54.8%)	26 (41.9%)	15 (46.9%)	
Age	50.0 (16.0)	50.8 (16.1)	50.1 (17.8)	50.0 (15.1)	43.7 (15.4)	0.212
Education:						
Primary	148 (27.7%)	90 (24.0%)	17 (23.3%)	25 (40.3%)	11 (34.4%)	
Secondary	193 (36.1%)	140 (37.3%)	27 (37.0%)	23 (37.1%)	10 (31.2%)	
Still studying	11 (2.06%)	9 (2.40%)	1 (1.37%)	0 (0.00%)	2 (6.25%)	
Tertiary	182 (34.1%)	136 (36.3%)	28 (38.4%)	14 (22.6%)	9 (28.1%)	
Ideology	4.72 (1.95)	4.71 (1.86)	4.42 (1.65)	4.68 (1.47)	5.34 (1.66)	0.246
Migration:						0.188
No	466 (87.3%)	338 (90.1%)	62 (84.9%)	51 (82.3%)	26 (81.2%)	
Yes	68 (12.7%)	37 (9.87%)	11 (15.1%)	11 (17.7%)	6 (18.8%)	
Income availability:						0.040
Not reported	427 (80.0%)	307 (81.9%)	66 (90.4%)	48 (77.4%)	21 (65.6%)	
Reported	107 (20.0%)	68 (18.1%)	7 (9.59%)	14 (22.6%)	11 (34.4%)	
Income level (low):						0.348
Above €1,500	429 (80.3%)	308 (82.1%)	64 (87.7%)	46 (74.2%)	26 (81.2%)	
Below €1,500	105 (19.7%)	67 (17.9%)	9 (12.3%)	16 (25.8%)	6 (18.8%)	
Income level (high):						0.181
Above €4,000	85 (15.9%)	66 (17.6%)	15 (20.5%)	4 (6.45%)	4 (12.5%)	
Below €4,000	449 (84.1%)	309 (82.4%)	58 (79.5%)	58 (93.5%)	28 (87.5%)	

2.1 Time taken

Table 2: (#tab:tab-time-completion-median.tex)Median time to complete the survey, by Treatment.

Treatment	Median time	Median time (seconds)
No	5.733	344.0
Yes	6.042	362.5

2.2 Vignettes assigned

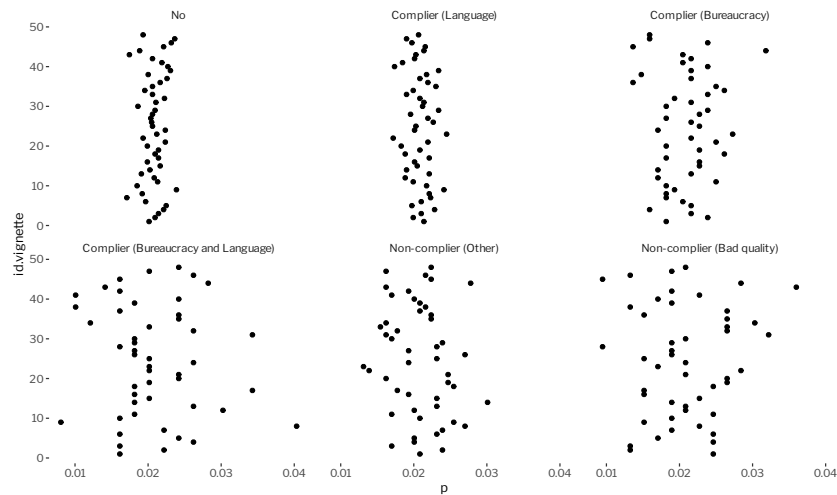


Figure 1: Distribution of vignette ids by treatment. Each unique vignette id (with up to 48 unique combinations) is shown against the proportion of individuals in each of the treatment groups that has received them. Bigger groups (the control group, which is half the sample) have less variable percentages due to its size, and each vignette is shown about two percent of the time (which corresponds to the 48 vignette combinations).

3 Full model results

Table 3 presents the model fit in terms of percent correctly predicted.

Treatment	Median PCP
No	78%
Yes	79%

Table 3: Posterior median percent correctly predicted.

The Hierarchical Bayes model estimated (see equation in Section 1.1) produces utilities for each individual (1,075) and each of the features. These are the β parameters that have subscripts over i individuals and f features. The point estimates of such $\beta_{i,f}$ parameters are shown in Figure 2 in this document (Online Appendix).

The main Figures in the text come from simply taking means of the two treatment groups. The grey densities show the distribution of the average utility for untreated individuals, whereas the black line represents the distribution of the average utility for treated individuals. Hence, the average treatment effect reported in the figures is the difference between the two distributions.

For practical reasons, we can not provide a tabular form of the $1,076 * 8 = 8,608$ parameters estimated, but their point estimate distribution is shown in Figure 2. In addition, the tables with the summaries of the posterior distributions of the effects are shown in Tables 4 to 6

Figure 3 in this document (Online Appendix) shows the parameters of the model reported in the article.

Table 4: Summary of the posterior distributions of the utilities. Median and standard deviations (Figure 2 in the main article).

Feature	Treatment	Median	SD
Nationality Netherlands	No	0.642	0.078
Nationality Netherlands	Yes	0.428	0.079
Gender Female	No	0.524	0.079
Gender Female	Yes	0.795	0.079
Profession Nurse	No	0.706	0.078
Profession Nurse	Yes	0.772	0.079
Language Full	No	1.447	0.080
Language Full	Yes	1.501	0.079
Age 40	No	0.305	0.093
Age 40	Yes	0.356	0.094
Age 55	No	0.227	0.094
Age 55	Yes	0.159	0.097
First shown	No	0.389	0.078
First shown	Yes	0.361	0.080

Table 5: Summary of the posterior distributions of the utilities. Median and standard deviations (Figure 3 in the main article).

Feature	Treatment	Median	SD
Nationality Netherlands	Complier (Bureaucracy)	-0.033	0.193
Nationality Netherlands	Complier (Language)	0.507	0.101
Nationality Netherlands	No	0.642	0.078
Nationality Netherlands	Non-complier (Bad quality)	1.027	0.311
Nationality Netherlands	Non-complier (Other)	0.327	0.203
Gender Female	Complier (Bureaucracy)	0.737	0.196
Gender Female	Complier (Language)	0.858	0.099
Gender Female	No	0.524	0.079
Gender Female	Non-complier (Bad quality)	0.298	0.301
Gender Female	Non-complier (Other)	0.798	0.202
Profession Nurse	Complier (Bureaucracy)	1.234	0.191
Profession Nurse	Complier (Language)	0.995	0.102
Profession Nurse	No	0.706	0.078
Profession Nurse	Non-complier (Bad quality)	-0.636	0.307
Profession Nurse	Non-complier (Other)	-0.090	0.205
Language Full	Complier (Bureaucracy)	0.989	0.193
Language Full	Complier (Language)	1.710	0.101
Language Full	No	1.447	0.080
Language Full	Non-complier (Bad quality)	0.629	0.301
Language Full	Non-complier (Other)	1.525	0.202
Age 40	Complier (Bureaucracy)	0.085	0.234
Age 40	Complier (Language)	0.472	0.119
Age 40	No	0.305	0.093
Age 40	Non-complier (Bad quality)	0.024	0.369
Age 40	Non-complier (Other)	0.273	0.238
Age 55	Complier (Bureaucracy)	0.901	0.237
Age 55	Complier (Language)	-0.062	0.123
Age 55	No	0.227	0.094
Age 55	Non-complier (Bad quality)	0.325	0.376
Age 55	Non-complier (Other)	0.235	0.250
First shown	Complier (Bureaucracy)	-0.098	0.201
First shown	Complier (Language)	0.329	0.101
First shown	No	0.389	0.078
First shown	Non-complier (Bad quality)	0.599	0.314
First shown	Non-complier (Other)	0.892	0.207

Table 6: Summary of the posterior distributions of the utilities. Median and standard deviations (Figure 4 in the article).

Feature	Treatment	Variable	Group	Median	SD
Nationality Netherlands	No	Ideology (Right)	Remaining sample	0.552	0.084
Nationality Netherlands	No	Ideology (Right)	Subgroup	0.962	0.139
Nationality Netherlands	No	Income (Low)	Remaining sample	0.666	0.080
Nationality Netherlands	No	Income (Low)	Subgroup	0.414	0.212
Nationality Netherlands	No	Migrant	Remaining sample	0.698	0.082
Nationality Netherlands	No	Migrant	Subgroup	0.262	0.179
Nationality Netherlands	No	Not favourable to immigration	Remaining sample	0.544	0.100

Nationality Netherlands	No	Not favourable to immigration	Subgroup	0.736	0.099
Nationality Netherlands	Yes	Ideology (Right)	Remaining sample	0.329	0.085
Nationality Netherlands	Yes	Ideology (Right)	Subgroup	0.777	0.143
Nationality Netherlands	Yes	Income (Low)	Remaining sample	0.427	0.081
Nationality Netherlands	Yes	Income (Low)	Subgroup	0.439	0.235
Nationality Netherlands	Yes	Migrant	Remaining sample	0.474	0.082
Nationality Netherlands	Yes	Migrant	Subgroup	0.085	0.190
Nationality Netherlands	Yes	Not favourable to immigration	Remaining sample	0.149	0.099
Nationality Netherlands	Yes	Not favourable to immigration	Subgroup	0.714	0.100
Gender Female	No	Ideology (Right)	Remaining sample	0.571	0.085
Gender Female	No	Ideology (Right)	Subgroup	0.361	0.140
Gender Female	No	Income (Low)	Remaining sample	0.555	0.081
Gender Female	No	Income (Low)	Subgroup	0.221	0.212
Gender Female	No	Migrant	Remaining sample	0.558	0.082
Gender Female	No	Migrant	Subgroup	0.293	0.179
Gender Female	No	Not favourable to immigration	Remaining sample	0.463	0.101
Gender Female	No	Not favourable to immigration	Subgroup	0.584	0.100
Gender Female	Yes	Ideology (Right)	Remaining sample	0.774	0.086
Gender Female	Yes	Ideology (Right)	Subgroup	0.877	0.141
Gender Female	Yes	Income (Low)	Remaining sample	0.786	0.081
Gender Female	Yes	Income (Low)	Subgroup	0.931	0.245
Gender Female	Yes	Migrant	Remaining sample	0.825	0.083
Gender Female	Yes	Migrant	Subgroup	0.581	0.187
Gender Female	Yes	Not favourable to immigration	Remaining sample	0.870	0.100
Gender Female	Yes	Not favourable to immigration	Subgroup	0.721	0.101
Profession Nurse	No	Ideology (Right)	Remaining sample	0.676	0.085
Profession Nurse	No	Ideology (Right)	Subgroup	0.803	0.138
Profession Nurse	No	Income (Low)	Remaining sample	0.749	0.080
Profession Nurse	No	Income (Low)	Subgroup	0.256	0.211
Profession Nurse	No	Migrant	Remaining sample	0.730	0.082
Profession Nurse	No	Migrant	Subgroup	0.540	0.178
Profession Nurse	No	Not favourable to immigration	Remaining sample	0.744	0.100
Profession Nurse	No	Not favourable to immigration	Subgroup	0.672	0.100
Profession Nurse	Yes	Ideology (Right)	Remaining sample	0.778	0.085
Profession Nurse	Yes	Ideology (Right)	Subgroup	0.753	0.142
Profession Nurse	Yes	Income (Low)	Remaining sample	0.792	0.081
Profession Nurse	Yes	Income (Low)	Subgroup	0.509	0.245
Profession Nurse	Yes	Migrant	Remaining sample	0.805	0.082
Profession Nurse	Yes	Migrant	Subgroup	0.526	0.187
Profession Nurse	Yes	Not favourable to immigration	Remaining sample	0.716	0.100
Profession Nurse	Yes	Not favourable to immigration	Subgroup	0.827	0.101
Language Full	No	Ideology (Right)	Remaining sample	1.441	0.087
Language Full	No	Ideology (Right)	Subgroup	1.467	0.141
Language Full	No	Income (Low)	Remaining sample	1.408	0.083
Language Full	No	Income (Low)	Subgroup	1.846	0.214
Language Full	No	Migrant	Remaining sample	1.456	0.083
Language Full	No	Migrant	Subgroup	1.380	0.180
Language Full	No	Not favourable to immigration	Remaining sample	1.210	0.103
Language Full	No	Not favourable to immigration	Subgroup	1.668	0.101

Language Full	Yes	Ideology (Right)	Remaining sample	1.478	0.086
Language Full	Yes	Ideology (Right)	Subgroup	1.582	0.141
Language Full	Yes	Income (Low)	Remaining sample	1.511	0.081
Language Full	Yes	Income (Low)	Subgroup	1.375	0.241
Language Full	Yes	Migrant	Remaining sample	1.530	0.083
Language Full	Yes	Migrant	Subgroup	1.289	0.190
Language Full	Yes	Not favourable to immigration	Remaining sample	1.284	0.100
Language Full	Yes	Not favourable to immigration	Subgroup	1.722	0.104
Age 40	No	Ideology (Right)	Remaining sample	0.253	0.099
Age 40	No	Ideology (Right)	Subgroup	0.488	0.161
Age 40	No	Income (Low)	Remaining sample	0.348	0.095
Age 40	No	Income (Low)	Subgroup	-0.136	0.241
Age 40	No	Migrant	Remaining sample	0.271	0.096
Age 40	No	Migrant	Subgroup	0.536	0.200
Age 40	No	Not favourable to immigration	Remaining sample	0.294	0.116
Age 40	No	Not favourable to immigration	Subgroup	0.316	0.114
Age 40	Yes	Ideology (Right)	Remaining sample	0.322	0.101
Age 40	Yes	Ideology (Right)	Subgroup	0.474	0.158
Age 40	Yes	Income (Low)	Remaining sample	0.377	0.096
Age 40	Yes	Income (Low)	Subgroup	0.084	0.265
Age 40	Yes	Migrant	Remaining sample	0.366	0.098
Age 40	Yes	Migrant	Subgroup	0.272	0.207
Age 40	Yes	Not favourable to immigration	Remaining sample	0.289	0.116
Age 40	Yes	Not favourable to immigration	Subgroup	0.423	0.116
Age 55	No	Ideology (Right)	Remaining sample	0.180	0.100
Age 55	No	Ideology (Right)	Subgroup	0.393	0.157
Age 55	No	Income (Low)	Remaining sample	0.215	0.097
Age 55	No	Income (Low)	Subgroup	0.344	0.230
Age 55	No	Migrant	Remaining sample	0.242	0.097
Age 55	No	Migrant	Subgroup	0.133	0.204
Age 55	No	Not favourable to immigration	Remaining sample	0.232	0.117
Age 55	No	Not favourable to immigration	Subgroup	0.223	0.115
Age 55	Yes	Ideology (Right)	Remaining sample	0.179	0.103
Age 55	Yes	Ideology (Right)	Subgroup	0.087	0.167
Age 55	Yes	Income (Low)	Remaining sample	0.160	0.098
Age 55	Yes	Income (Low)	Subgroup	0.155	0.272
Age 55	Yes	Migrant	Remaining sample	0.109	0.100
Age 55	Yes	Migrant	Subgroup	0.527	0.212
Age 55	Yes	Not favourable to immigration	Remaining sample	0.111	0.117
Age 55	Yes	Not favourable to immigration	Subgroup	0.207	0.120
First shown	No	Ideology (Right)	Remaining sample	0.507	0.086
First shown	No	Ideology (Right)	Subgroup	-0.030	0.140
First shown	No	Income (Low)	Remaining sample	0.349	0.081
First shown	No	Income (Low)	Subgroup	0.791	0.213
First shown	No	Migrant	Remaining sample	0.373	0.082
First shown	No	Migrant	Subgroup	0.495	0.184
First shown	No	Not favourable to immigration	Remaining sample	0.529	0.102
First shown	No	Not favourable to immigration	Subgroup	0.257	0.099
First shown	Yes	Ideology (Right)	Remaining sample	0.383	0.086

First shown	Yes	Ideology (Right)	Subgroup	0.286	0.142
First shown	Yes	Income (Low)	Remaining sample	0.288	0.081
First shown	Yes	Income (Low)	Subgroup	1.313	0.246
First shown	Yes	Migrant	Remaining sample	0.386	0.083
First shown	Yes	Migrant	Subgroup	0.191	0.189
First shown	Yes	Not favourable to immigration	Remaining sample	0.432	0.100
First shown	Yes	Not favourable to immigration	Subgroup	0.289	0.102

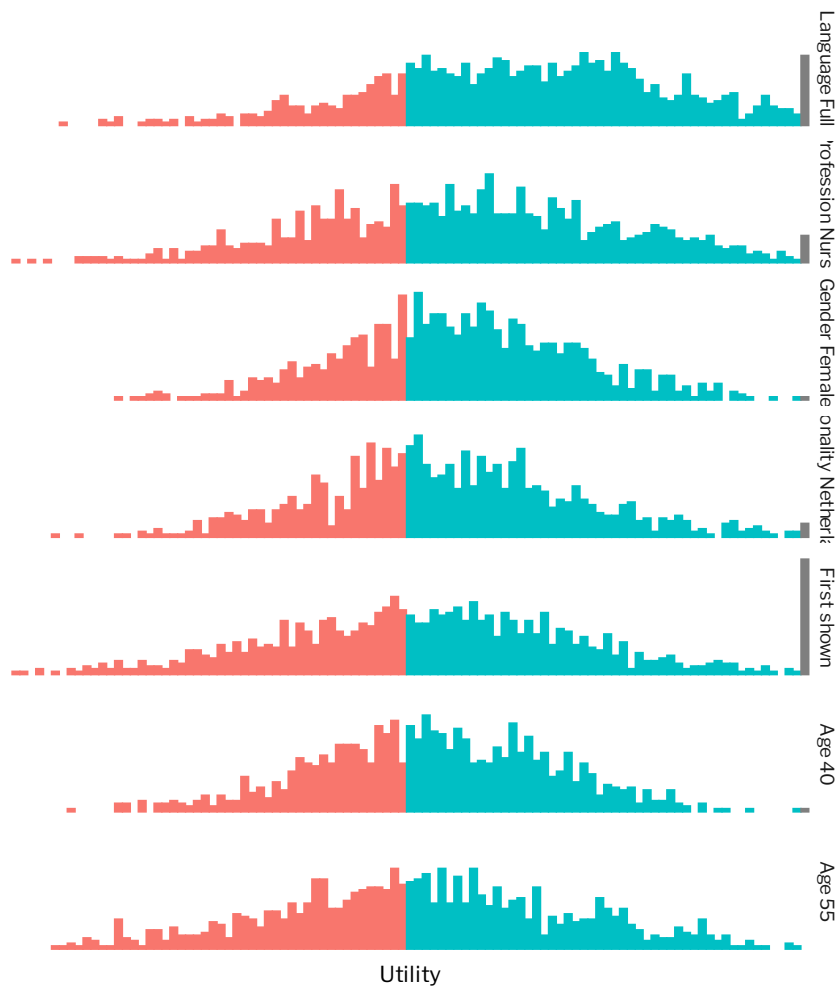


Figure 2: Individual utilities for each conjoint feature (β), as summarized by the median of the posterior distribution.

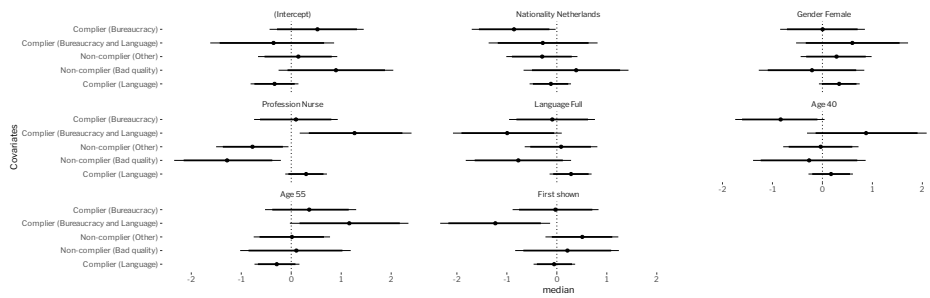


Figure 3: Posterior distributions for the model parameters that capture the differences for the sociodemographical variables (θ). Dots represent the median effect, while thick and thin lines cover the 90 and 95 percent of the credibility interval of the parameter.

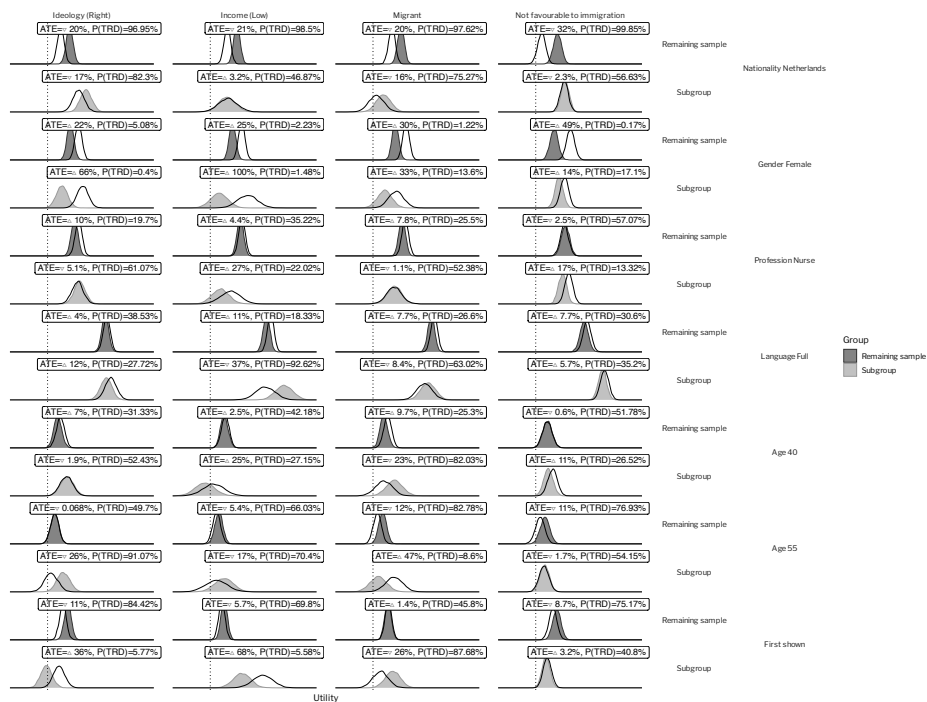


Figure 4: Average treatment effects. Comparison showing the average effect for treated individuals (curve with the line), and the effect for non-treated individuals (filled density curve). The shadows of grey compare subgroups of the population against the remaining sample (those with right ideology against the rest of individuals; those with low income against the rest of individuals, and so on). All individual features are displayed.

4 Robustness checks and sensitivity analysis

4.1 Sensitivity to priors

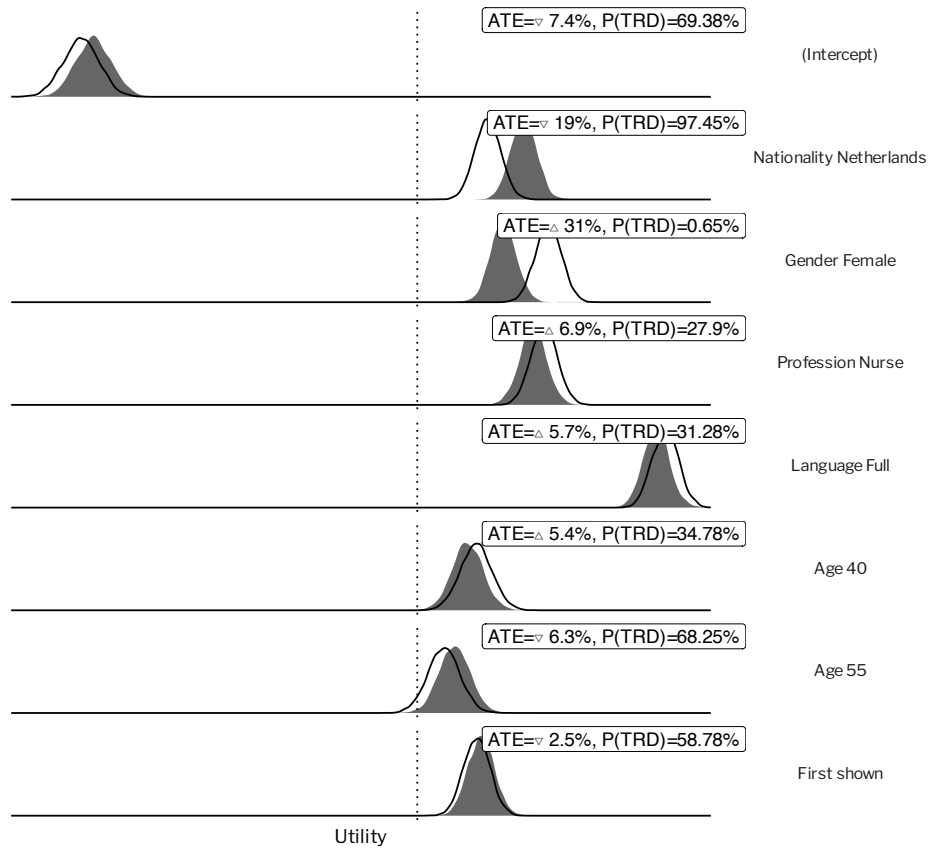


Figure 5: Average treatment effects. These results correspond to a model like the one shown in the main text, but the priors for the individual utilities are even more vague ($\sigma = 10$). Shown here for sensitivity.

4.2 Low quality individuals - Quick replies

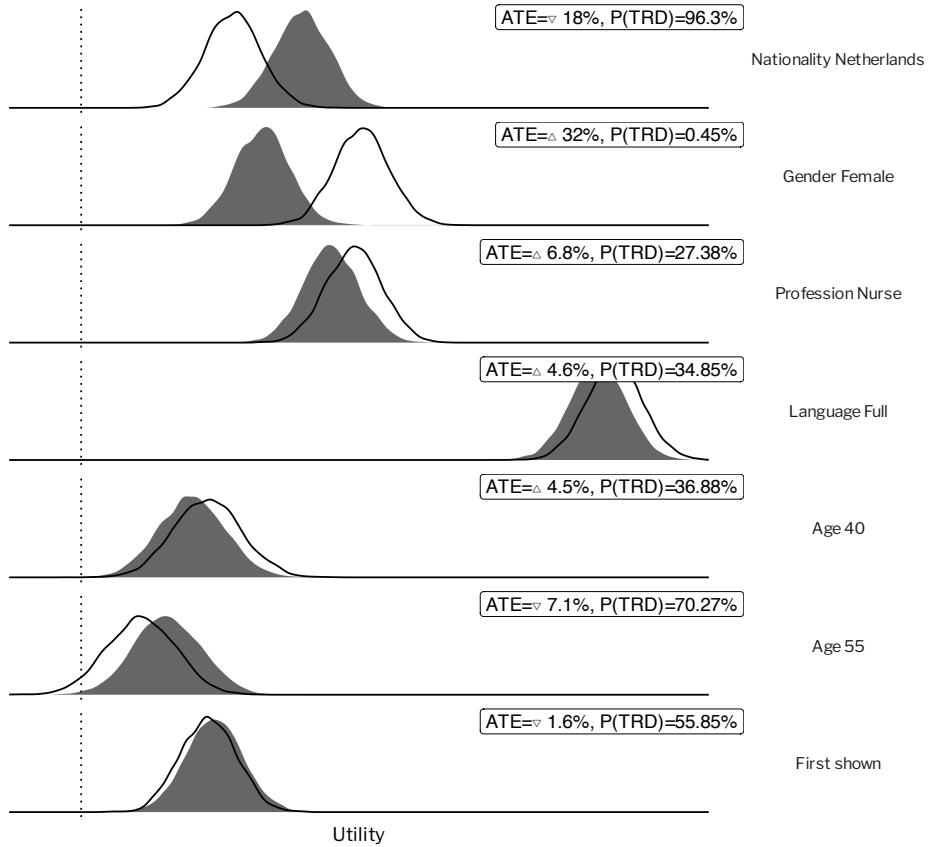


Figure 6: Average treatment effects. These results correspond to a model like the one shown in the main text, but without individuals that take the survey very quick (below 2 standard deviations of the geometric mean).

4.3 Low quality individuals - Same profile order

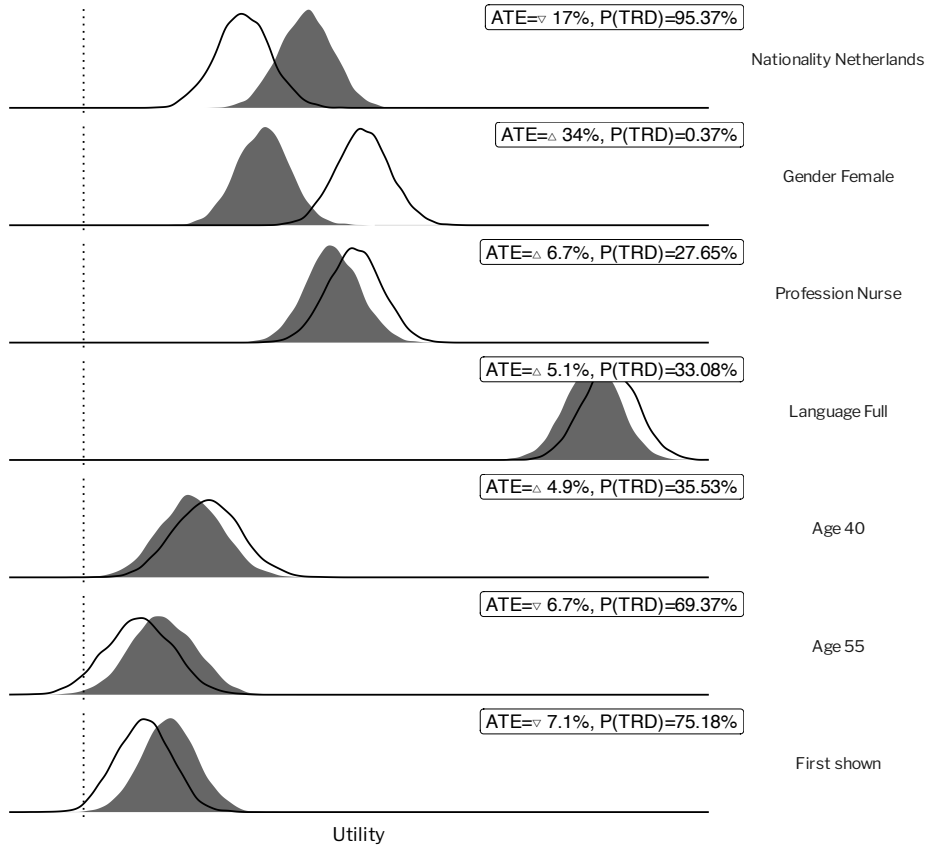


Figure 7: Average treatment effects. These results correspond to a model like the one shown in the main text, but without individuals that always respond to the same profile (first or second) throughout the survey.

4.4 Compliance

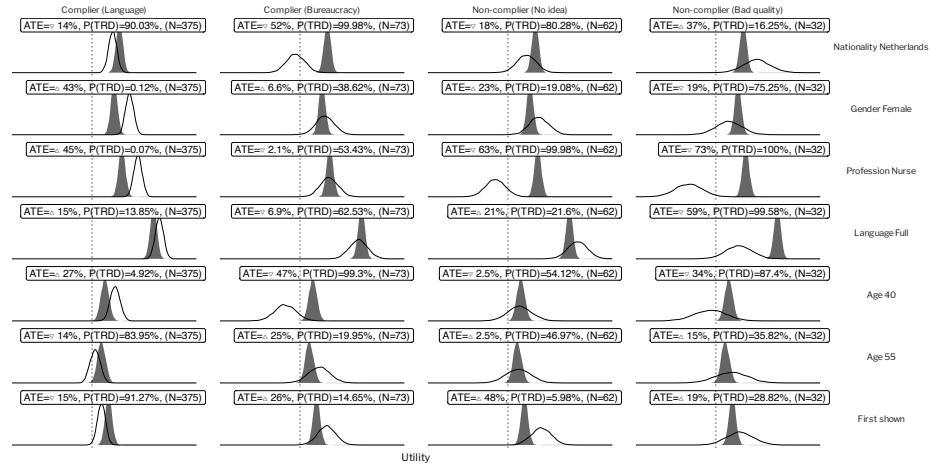


Figure 8: Average treatment effects. All details about compliers (as shown in the main text), as well as non-compliers, divided between those who reply that they have no idea, and those who provide an erratic text.

4.5 NLP coding of open responses

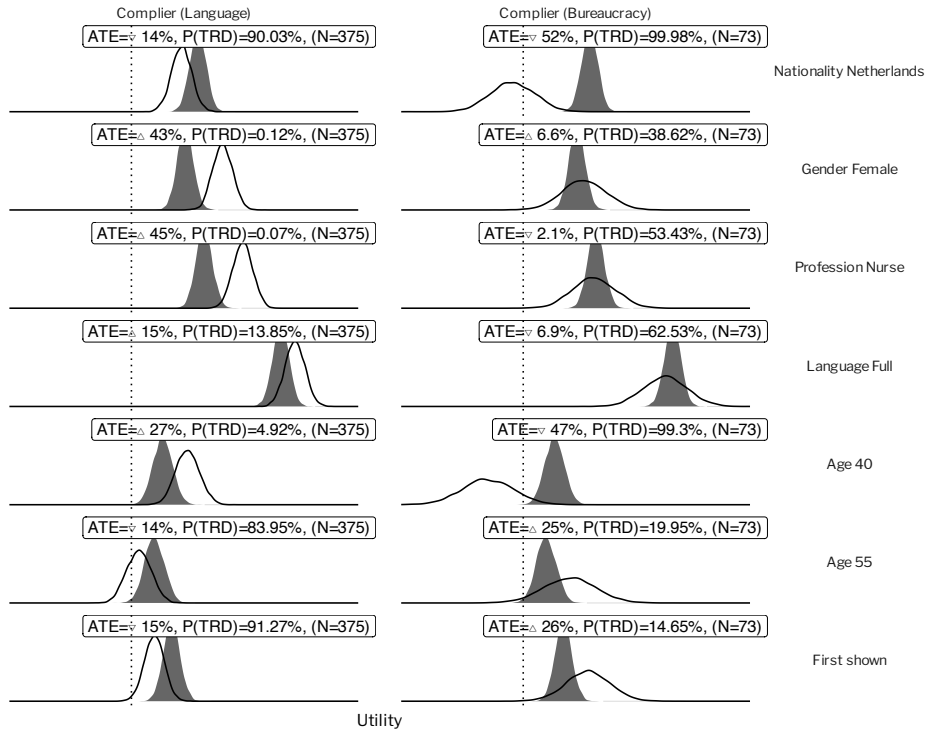


Figure 9: Average treatment effects. Same as Figure 2 in the main text, but with natural language processing for coding the compliance group.

5 Certainty of the response

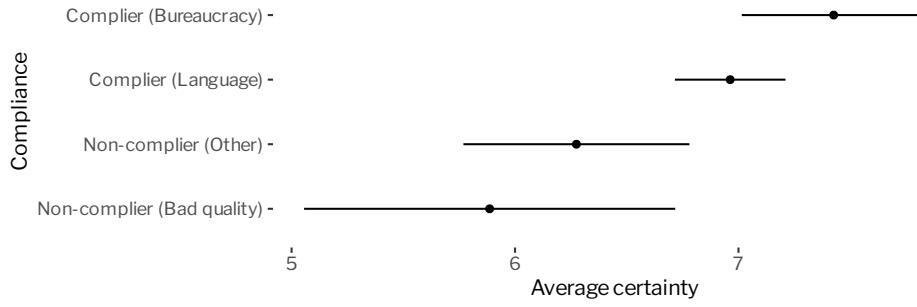


Figure 10: Average certainty (and confidence interval) by treatment / compliance.

Treatment	Mean	low	high
Non-complier (Bad quality)	5.89	5.06	6.72
Non-complier (Other)	6.27	5.77	6.78
Complier (Language)	6.96	6.72	7.21
No	7.28	7.10	7.45
Complier (Bureaucracy)	7.43	7.01	7.84

Table 7: Average certainty of each respondent, by treatment and compliance.

6 Convergence tests

This section includes Bayesian convergence tests for the prioritizations, the main parameters of interest (β). The rest of the parameters can be provided by request.

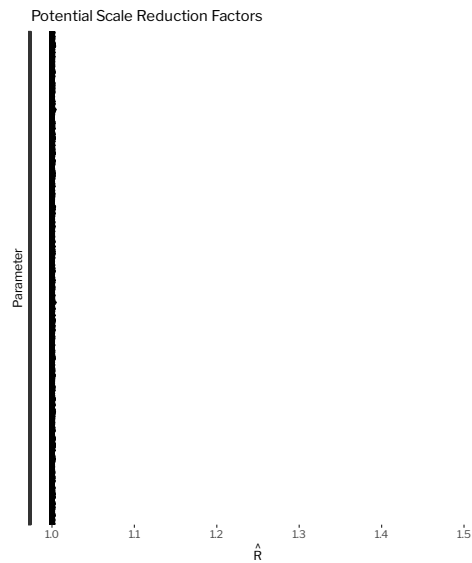


Figure 11: Potential Scale Reduction Factor (\hat{R}).

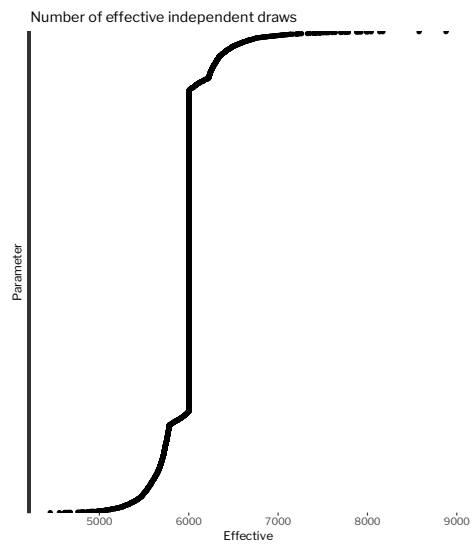


Figure 12: Number of effective parameters (\hat{R}).