

## Appendix B: Measurement Model

We aggregate our scores using an item-response model (IRT). This approach allows respondents to assign different scores given the same underlying subjective rating of a face. Thus, one respondent may grade a given MP's face as 4 on the physically dominant scale, whilst another who perceives it the same way might grade it as a 5. Given respondents graded faces on a 7-point scale, we use what is termed as a graded response model (GRT) in which the latent underlying judgements of an individual are related to the scores they given each MP via a cumulative logistic link function:

$$P(y = c) = \frac{\exp(\sum_{j=1}^{c-1}(\psi - \tau_j))}{\sum_{r=1}^c \exp(\sum_{j=1}^{r-1}(\psi - \tau_j))}$$

Here,  $P(y = c)$  is the probability that a given respondent assigns a given MP a score of  $c$  where  $c = 1, 2, \dots, 7$ .  $\psi$  is continuous latent variable, that captures a respondent's perception of an MP. Here,  $\tau$  is a  $1 \times 6$  vector of thresholds, describing which values of  $\psi$  correspond to which reported scores.

Because each respondent may map perceptions to scores differently, we allow for the  $\tau$  to vary by respondent. Note, that since respondents are only assessing one trait for multiple MPs, we perform this analysis separately by trait thus making no assumptions about how scores in one relate to another. We do assume, however, that  $\tau$  is fixed within respondents. That is, that the range of underlying perceptions that leads to a given grade is the same for all MPs.

Given the large number of parameters we estimate  $P(y = c)$  for each value of  $c$ , for each politician, for each trait, and the vector  $\tau$  for each respondent), we employ a Bayesian approach. This requires us to specify priors on these parameters. We use weakly informative priors such that our estimates will coincide with those of a classical maximum likelihood estimator. In particular, following Bruckner (2019) we set half-Normal(0,3) priors for both the politician and respondent parameters.

This model is fitted using Hamiltonian Markov-Chain Monte-Carlo (HMC) using STAN (Carpenter et al., 2017) and the brms R package (Bruckner, 2017).

Having obtained the estimated posterior distribution we then compute the expected value for each respondent-politician pair  $\sum_{c=1}^C P(c = y) \times C$  and then average these over all respondents to obtain the face ratings.

### Additional References:

Bürkner, P. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1 - 28. doi:<http://dx.doi.org/10.18637/jss.v080.i01>

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1), 1 - 32. doi:<http://dx.doi.org/10.18637/jss.v076.i01>