

Supplementary Appendices

Sub-national forecasts of COVID-19 vaccine acceptance across the UK: a large-scale cross-sectional spatial modelling study

A. de Figueiredo, PhD^{1,2†}

¹ Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK

² Department of Mathematics, Imperial College London, London, UK

† alex.defigueiredo@lshtm.ac.uk

These appendices contain a number of additional details and figures in support of Forecasted trends in COVID-19 vaccine acceptance across the UK: a large-scale cross-sectional spatial modelling study. The contents of these three appendices are detailed below.

Appendix 1: Respondents' socio-econo-demographic background	2
Appendix 2: Supplementary figures	3
<i>Mapping COVID-19 vaccine acceptance: unsure about accepting a COVID-19 vaccine</i>	3
<i>Mapping COVID-19 vaccine acceptance: would not accept a COVID-19 vaccine</i>	3
Appendix 3: Model details	5
<i>Model: Multilevel regression</i>	6
<i>Model: Poststratification</i>	7
<i>Model: Implementation and output</i>	7
<i>Model: Sensitivity analysis on hierarchical variance priors</i>	Error! Bookmark not defined.

Appendix 1: Respondents' socio-econo-demographic background

The number (and percentage) of respondents within each socio-econo-demographic group (sex, age, highest education level, religious affiliation, ethnicity, employment status, and primary language) is shown in table A1. The number (and percentage) of each group replying to each of the four response options to “If a new coronavirus (COVID-19) vaccine became available, would you accept it?” is also shown.

		If a new coronavirus (COVID-19) vaccine became available, would you accept the vaccine for yourself?									
Socio-econo-demographic	N	%	No, definitely not (N)	No, definitely not (%)	Unsure, but leaning towards no (N)	Unsure, but leaning towards no (%)	Unsure, but leaning towards yes (N)	Unsure, but leaning towards yes (%)	Yes, definitely (N)	Yes, definitely (%)	
SEX	Female	8682	51.6%	788	9.1	1120	12.9	3071	35.4	3703	42.6
	Male	8138	48.4%	560	6.9	699	8.6	2429	29.9	4450	54.7
AGE	18-24	1978	11.8%	180	9.1	251	12.7	753	38.1	794	40.1
	25-34	2934	17.4%	352	12	414	14.1	1022	34.8	1146	39.1
	35-44	3027	18.0%	298	9.8	391	12.9	1042	34.4	1296	42.8
	45-54	3015	17.9%	231	7.7	351	11.6	1025	34	1408	46.7
	55-64	2528	15.0%	164	6.5	226	8.9	789	31.2	1349	53.4
	65-79	3150	18.7%	117	3.7	176	5.6	826	26.2	2031	64.5
	80+	188	1.1%	6	3.2	10	5.3	43	22.9	129	68.6
highest EDUCATION	Level 1-3	7291	43.3%	629	8.6	823	11.3	2476	34	3363	46.1
	Level 4	7057	42.0%	453	6.4	712	10.1	2221	31.5	3671	52
	None/Other	2472	14.7%	266	10.8	284	11.5	803	32.5	1119	45.3
RELIGIOUS affiliation	Atheist or agnostic	5076	30.2%	362	7.1	521	10.3	1692	33.3	2501	49.3
	Christian	8279	49.2%	598	7.2	831	10	2549	30.8	4301	52
	Hindu	177	1.1%	6	3.4	13	7.3	62	35	96	54.2
	Jewish	143	0.9%	7	4.9	11	7.7	38	26.6	87	60.8
	Muslim	534	3.2%	66	12.4	80	15	209	39.1	179	33.5
	Not given	1159	6.9%	134	11.6	164	14.2	439	37.9	422	36.4
	other religion	1452	8.6%	175	12.1	199	13.7	511	35.2	567	39
ETHNICITY	Asian/Asian British	845	5.0%	65	7.7	103	12.2	336	39.8	341	40.4
	Black/Black British	413	2.5%	81	19.6	88	21.3	113	27.4	131	31.7
	Mixed	321	1.9%	41	12.8	40	12.5	107	33.3	133	41.4
	other ethnicity	315	1.9%	52	16.5	51	16.2	112	35.6	100	31.8
	White	14926	88.7%	1109	7.4	1537	10.3	4832	32.4	7448	49.9
EMPLOYMENT status	full-time	7326	43.6%	641	8.8	801	10.9	2371	32.4	3513	48
	looking after home / family	925	5.5%	89	9.6	133	14.4	345	37.3	358	38.7
	other	105	0.6%	23	21.9	12	11.4	37	35.2	33	31.4
	part-time	2942	17.5%	238	8.1	394	13.4	1079	36.7	1231	41.8
	retired / disabled	3834	22.8%	192	5	285	7.4	1033	26.9	2324	60.6
	student	844	5.0%	54	6.4	92	10.9	340	40.3	358	42.4
	unemployed	844	5.0%	111	13.2	102	12.1	295	35	336	39.8
primary LANGUAGE	English or Welsh	15656	93.1%	1179	7.5	1637	10.5	5090	32.5	7750	49.5
	Other	1026	6.1%	138	13.4	163	15.9	360	35.1	365	35.6
	Polish	138	0.8%	31	22.5	19	13.8	50	36.2	38	27.5

Table A1. Socio-econo-demographic breakdown of the 16,820 survey participants and their intent to accept a COVID-19 vaccine

Appendix 2: Supplementary figures

Mapping COVID-19 vaccine acceptance: unsure about accepting a COVID-19 vaccine

The estimated proportion of respondents in each of the UK's 174 NUTS3 regions who are unsure about whether they would accept a COVID-19 vaccine ('unsure, but leaning towards yes' or 'unsure, but leaning towards no') is shown in figure A1 below. Estimates are derived via multilevel regression and poststratification as described in the main text and Appendix 3.

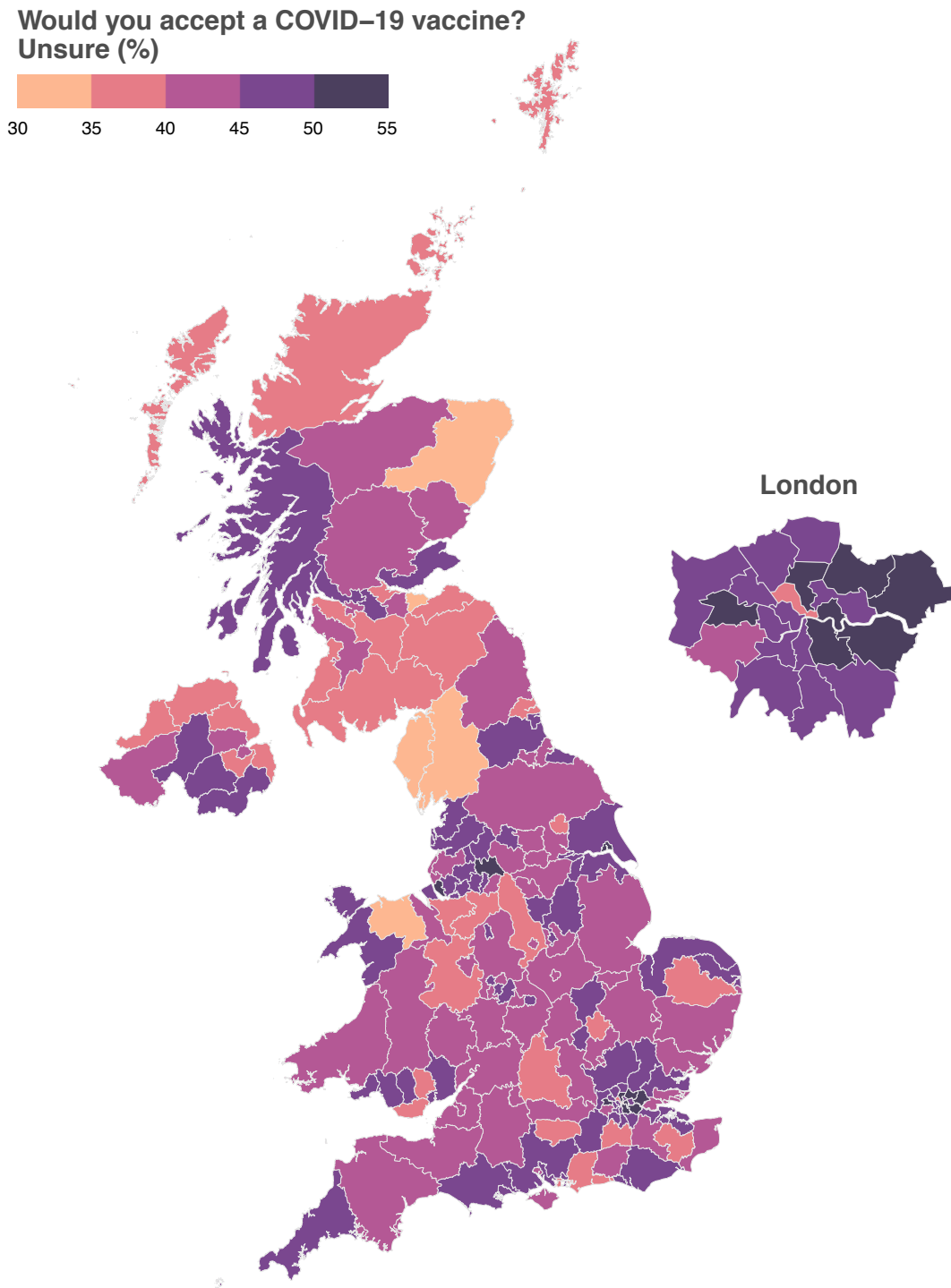


Figure A1. Regional level MRP estimates of the proportion of each UK region who are unsure about accepting a COVID-19 vaccine

Mapping COVID-19 vaccine acceptance: would not accept a COVID-19 vaccine

The estimated proportion of respondents in each of the UK's 174 NUTS3 regions who state that they would not accept a COVID-19 vaccine ('no, definitely not') is shown in figure A2 below. Estimates are derived via multilevel regression and poststratification as described in the main text and Appendix 3.

Would you accept a COVID-19 vaccine?
no, definitely not (%)

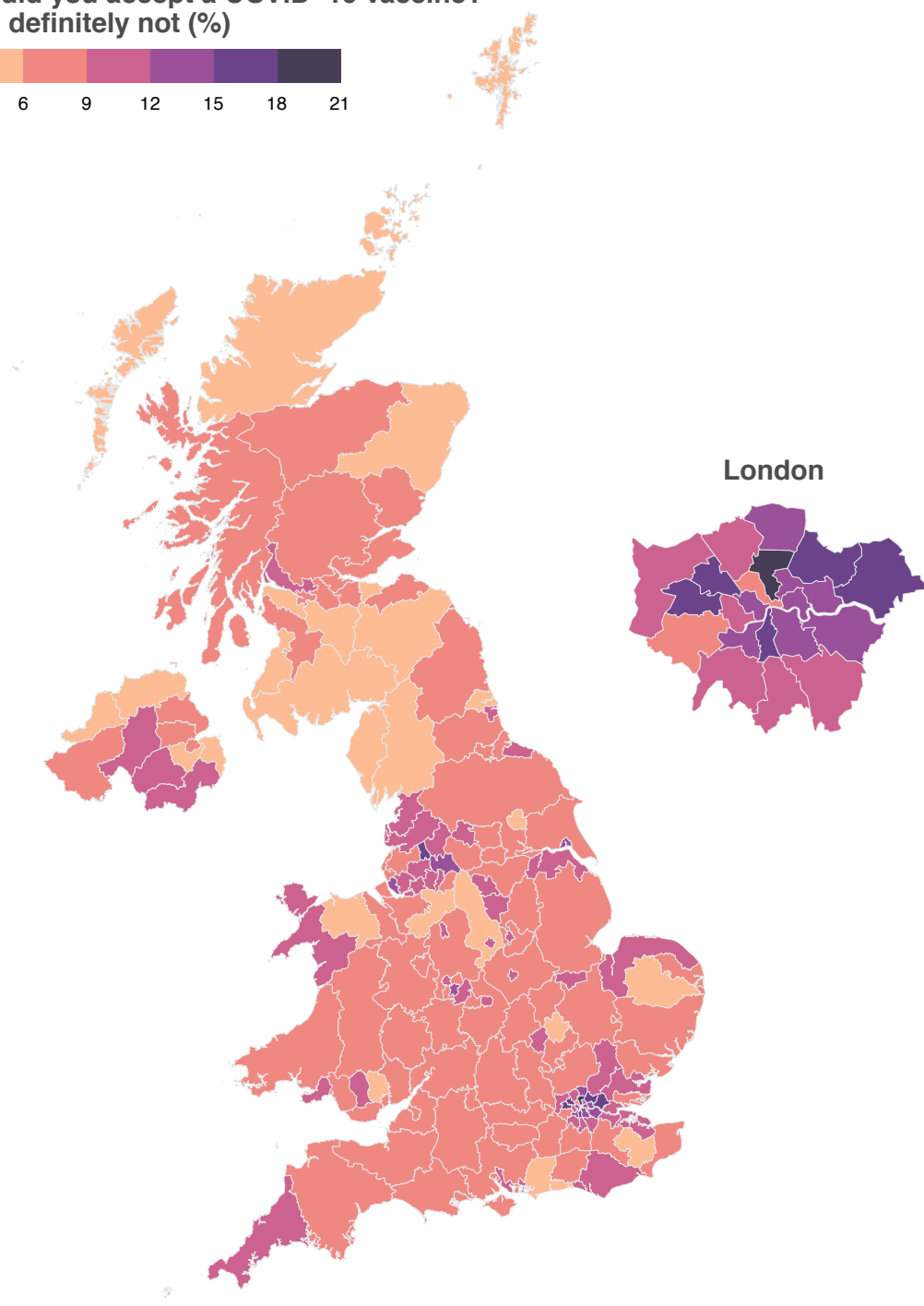
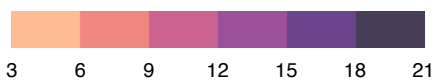


Figure A2. Regional level MRP estimates of the proportion of each UK region who would definitely not accept a COVID-19 vaccine

socio–econo–demographic determinants of intent to accept a COVID–19 vaccine (random–effects)

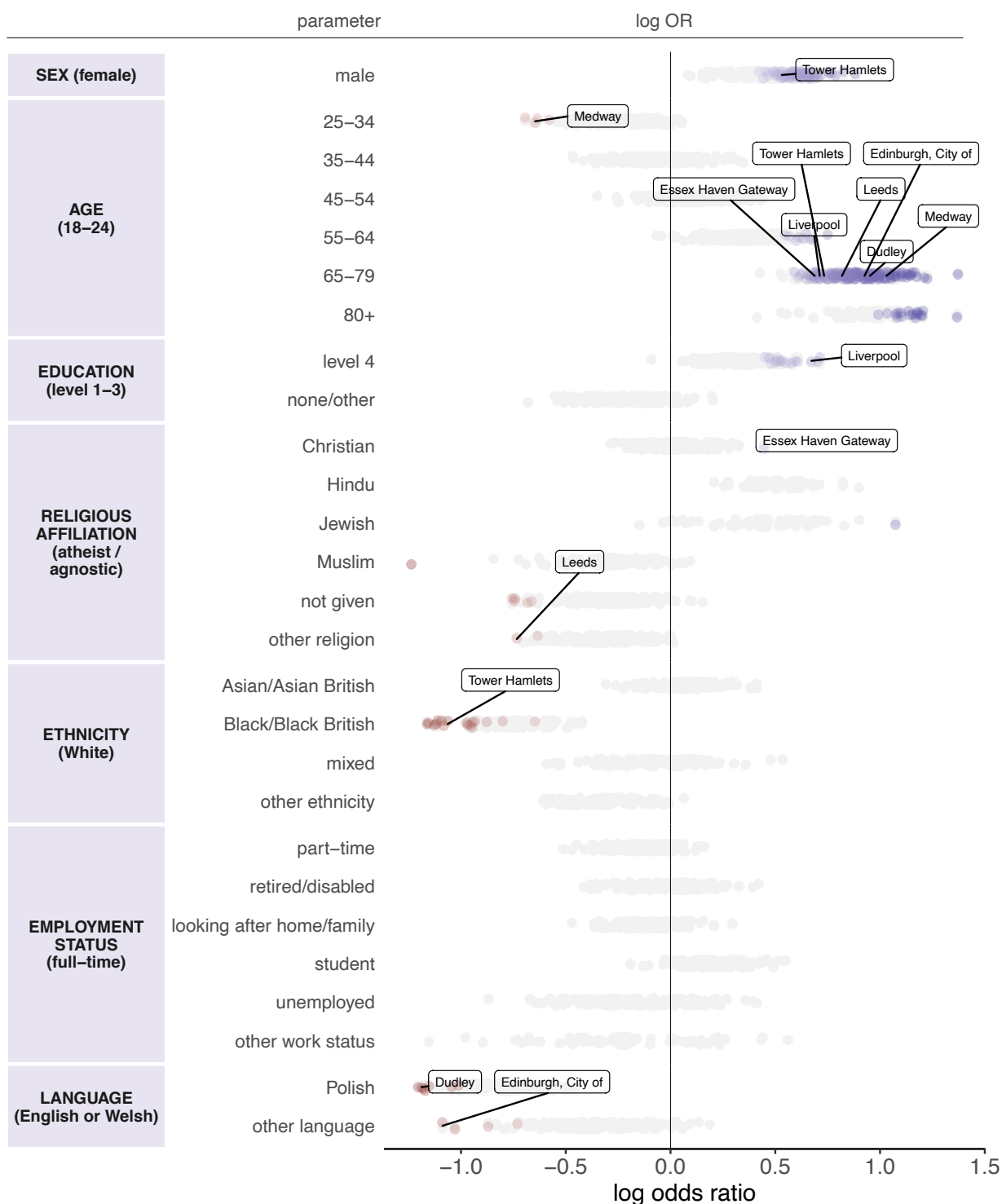


Figure A3. Multilevel regression random effects. All random effect parameters from the multilevel regression are shown in grey. Parameter log odds ratios are plotted with corresponding 95% HPDIs and coloured by effect size, where blues (reds) denote a positive (negative) association between the factor and intent to accept a vaccine relative to the baseline group (which is provided in parentheses on the left column). The darker the colour the stronger the association. All parameters whose 95% HPDI excludes zero are coloured. A selection of sub-national regions are shown, see supplementary data file (random effects) for all inferred random-effect parameters with HPDIs for all 174 regions.

Appendix 3: Model details

Multilevel regression and poststratification (MRP) is a tool that can be used to estimate opinions aggregated at sub-national regions from survey data collected at the national level, via partial pooling of information between these national and sub-national scales¹. This pooling of information between the two levels is a compromise between estimates derived via a total aggregation of data (to estimate national trends only) and estimates via complete disaggregation (that is, estimating regional trends only). The former suffers from a loss of information at the regional level while the latter suffers from possible low data counts and the loss of statistical power. More pooling of information will occur in regions with low relative numbers of surveyed individuals and less pooling in regions with high relative counts.

In brief (and relating specifically to this study), the first step of MRP is to conduct a multilevel regression to estimate, for each stratum (that is, a possible combination of individual characteristics) and for each region, the probability of COVID-19 vaccine acceptance. The second step is to reweight (poststratify) these strata probabilities by the frequency with which a given strata appears in a population. In this study individual-level UK census data is used to perform the reweighting.

Model: Multilevel regression

Individual intent to accept a COVID-19 vaccine is specified as $y_{ij} \in \{1,2,3,4\}$, where 1 = “no, definitely not”, 2 = “unsure, but leaning towards no”, 3 = “unsure, but leaning towards yes”, and 4 = “yes, definitely” and $1 < 2 < 3 < 4$. Here, $j = 1, \dots, 174$ is one of the $J = 174$ third National Territorial Units for Statistics (NUTS3) regions in the UK, and $i = 1, \dots, n_j$, where n_j is the number of individuals surveyed in region j . $\sum_j n_j = 16,820$ is the total number of respondents in the survey. A breakdown of the number of respondents in each region and a summary of their socio-econo-demographic status is given in the supplementary data file.

Intent to accept a COVID-19 vaccine is modelled as a multilevel ordinal regression with the proportional odds assumption²,

$$Y_{ij} | \mathbf{p}_{ij}, n_j, \sim \text{Multi}(\mathbf{p}_{ij}, 1)$$

$$\log \frac{\Pr(Y_{ij} \leq k | X_{ij})}{\Pr(Y_{ij} > k | X_{ij})} = \rho_k + \beta_j^{SEX[i]} + \beta_j^{AGE[i]} + \beta_j^{EDU[i]} + \beta_j^{REL[i]} + \beta_j^{ETH[i]} + \beta_j^{EMP[i]} + \beta_j^{LAN[i]}$$

for $k = 1, \dots, 3$,

where $\beta_j^{SEX[i]}$, $\beta_j^{AGE[i]}$, $\beta_j^{EDU[i]}$, $\beta_j^{REL[i]}$, $\beta_j^{ETH[i]}$, $\beta_j^{EMP[i]}$, and $\beta_j^{LAN[i]}$ are the random-effect varying intercepts for sex, age, highest education level, religious affiliation, ethnicity, employment status, and primary language (respectively); ρ_k are probability threshold parameters; $k \in \{1,2,3,4\}$ is the ordinal response category; $\mathbf{p}_{ij} = [\Pr(Y_{ij} = 1), \Pr(Y_{ij} = 2), \Pr(Y_{ij} = 3), \Pr(Y_{ij} = 4)]$; and X_{ij} is the covariate data for individual i in region j . The baseline group for the regression corresponds to an individual who is male, aged 18-24, has an education level 1-3, is an atheist or agnostic, is White, works full-time, and speaks English or Welsh as their primary language.

In line with prior recommendations for variance components in hierarchical models ^{1,3}, default weakly informative priors are chosen for the random-effects regression coefficients β^j ,

$$\begin{aligned}\beta_j^c | \sigma^c &\sim N(\gamma^c, (\sigma^c)^2), & \text{for } c = 1, \dots, 27 \\ 1/(\sigma^c)^2 = \tau^c &\sim N_+(0,1), & \text{for } c = 1, \dots, 27 \\ \beta^c &\sim N(0,1), & \text{for } c = 1, \dots, 27,\end{aligned}$$

where c indexes the regression coefficients: excluding the threshold parameters, there are 27 fixed-effect parameters: $\gamma_j^{SEX=female}$, $\gamma_j^{AGE=25-34}$, $\gamma_j^{AGE=35-44}$, $\gamma_j^{AGE=45-54}$, $\gamma_j^{AGE=55-64}$, $\gamma_j^{AGE=65-79}$, $\gamma_j^{AGE=80+}$, $\gamma_j^{EDU=level 4}$, $\gamma_j^{EDU=none}$, $\gamma_j^{REL=Christian}$, $\gamma_j^{REL=Hindu}$, $\gamma_j^{REL=Jewish}$, $\gamma_j^{REL=Muslim}$, $\gamma_j^{REL=Not given}$, $\gamma_j^{REL=Other}$, $\gamma_j^{ETH=Asian/Asian British}$, $\gamma_j^{ETH=Black/Black British}$, $\gamma_j^{ETH=Other}$, $\gamma_j^{EMP=part-time}$, $\gamma_j^{EMP=retired/disabled}$, $\gamma_j^{EMP=student}$, $\gamma_j^{EMP=other}$, $\gamma_j^{EMP=looking after home}$, $\gamma_j^{EMP=unemployed}$, $\gamma_j^{LAN=Polish}$, and $\gamma_j^{LAN=other}$. (And, thus, a total of $27 \times 174 = 4,698$ random effect parameters.)

Model: Poststratification

There are $S = 30,870$ socio-econo-demographic strata (two sexes \times seven age groups \times three education levels \times seven affiliations for religion \times five ethnicity groupings \times seven employment statuses \times three languages). Denoting the posterior probabilities of COVID-19 vaccination intent for each stratum $s = 1, \dots, S$ and NUTS3 region $j = 1, \dots, J$ as θ_{sjk} (where, as a reminder, $k \in \{1,2,3,4\}$ denotes the response), then the MRP estimate for the intent to vaccinate within each of the UK's 174 NUTS3 regions is,

$$\Phi_{jk} = \frac{\sum_s N_{sj} \theta_{sjk}}{\sum_s N_{sj}},$$

where N_{js} is the number of individuals in stratum s in region j according to the UK census microdata sample. In the main text, this quantity Φ_{jk} is computed for $k = 4$ (“yes, [I] definitely [would accept a COVID-19 vaccine]”). Estimates are computed for the remaining response options and are shown in figures A1 and A2 (note that options for “unsure, but leaning towards yes” and “unsure, but leaning towards no” have been combined).

Model: Implementation and output

The multilevel regression model detailed above is implemented using JAGS version 4.3.0 (implemented via `rjags`⁴) and R version 4.0.3. 25,000 posterior samples (excluding the first 5,000 for model burn-in) was sufficient for successful convergence and all posterior draws were well-mixed. The posterior draws for the fixed effects are shown in figure A4 and all look visibly well-mixed and all except “other work status” ($p = 0.04$) have Geweke p -values above 0.05. There are too many posterior draws to plot for all random-effects, but we show posterior draws for the first UK NUTS3 region alphabetically (Hartlepool and Stockton-on-Tees) in figure A5 with a histogram of Geweke p -values for

¹ Instead of an noninformative $N_+(0,100)$ distribution over the standard deviation of hierarchical variance parameters ³, a weakly-informative $N_+(0,1)$ prior is placed over the *precision* of these parameters, which places 95% of σ^c 's prior mass between 0.54 and 4.05.

all model parameters (fixed effects, random effects, and variance components) to demonstrate universally good mixing and convergence in figure A3. In the computation of the Geweke statistic, the first 10% and final 50% of the posterior samples used for computation are used. A slightly larger fraction of Geweke p -values fall below 0.05 than is expected by chance (0.082 compared to 0.05 by chance). Manual inspection of these chains revealed no cause for concern: chains showed no ill-mixing or convergence issues.

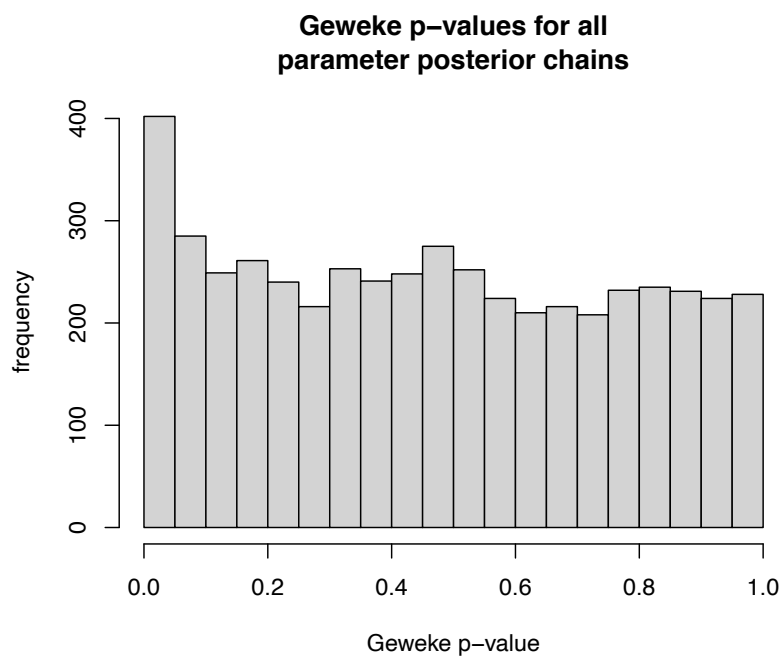


Figure A3. Distribution of Geweke p -values for all parameter posterior chains.

Mixing of posterior samples for fixed-effect parameters

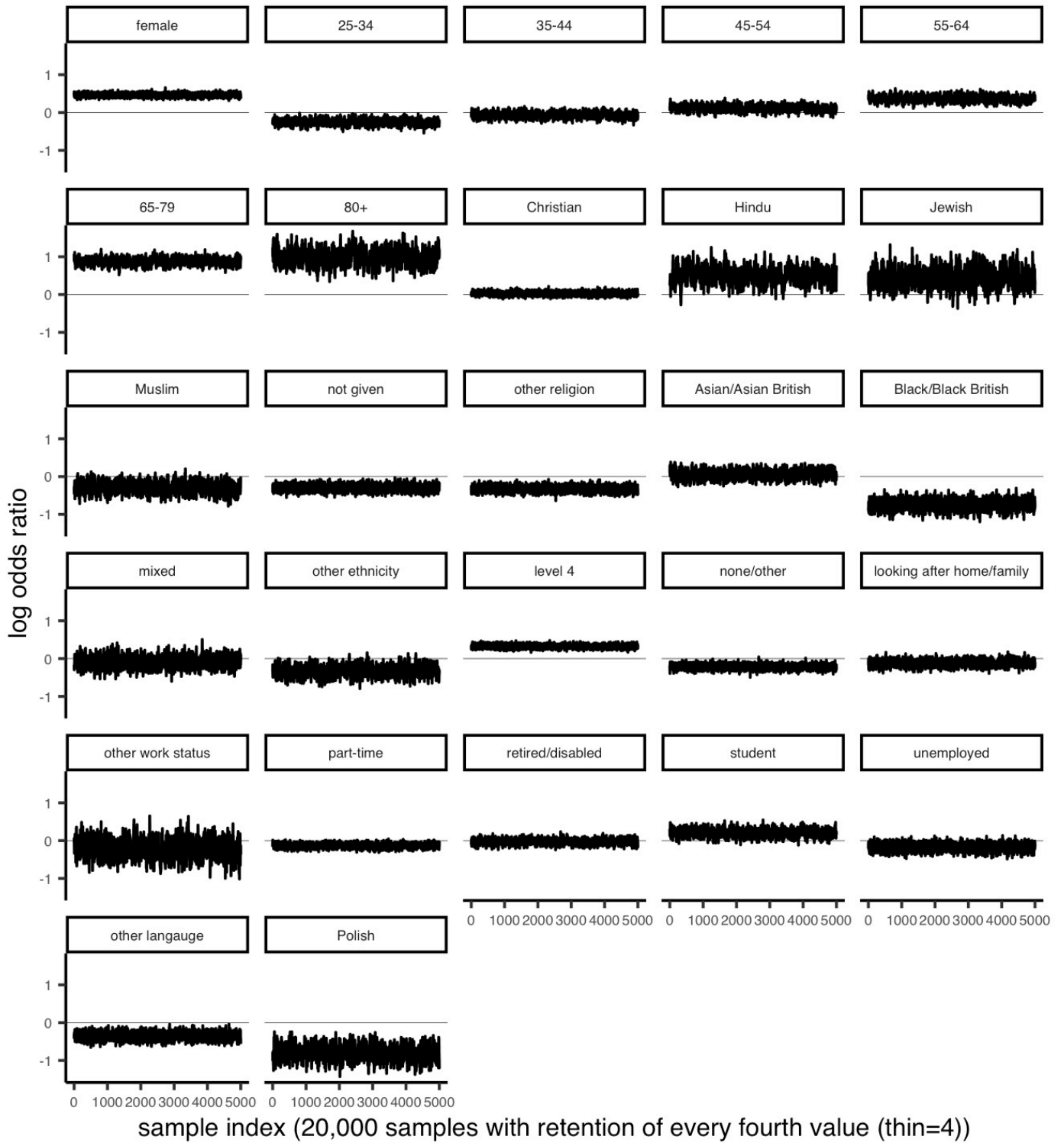


Figure A4 Mixing and convergence of the fixed-effect parameters in the multilevel regression model

Mixing of posterior samples for random-effect parameters
Hartlepool and Stockton-on-Tees

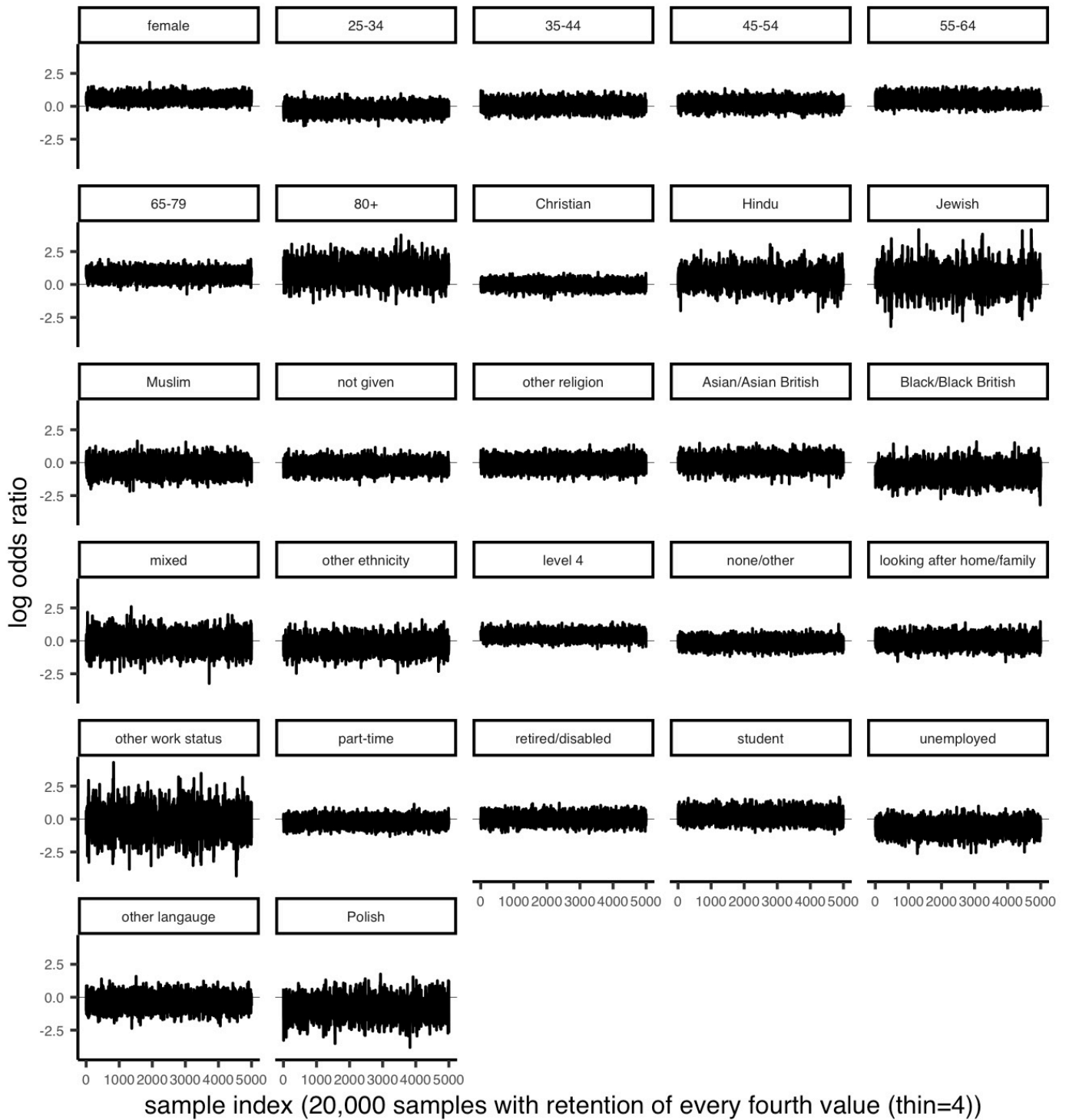


Figure A5 Mixing and convergence of the random-effect parameters for Hartlepool and Stockton-on-Tees in the multilevel regression model.

Mixing of posterior samples for variance parameters

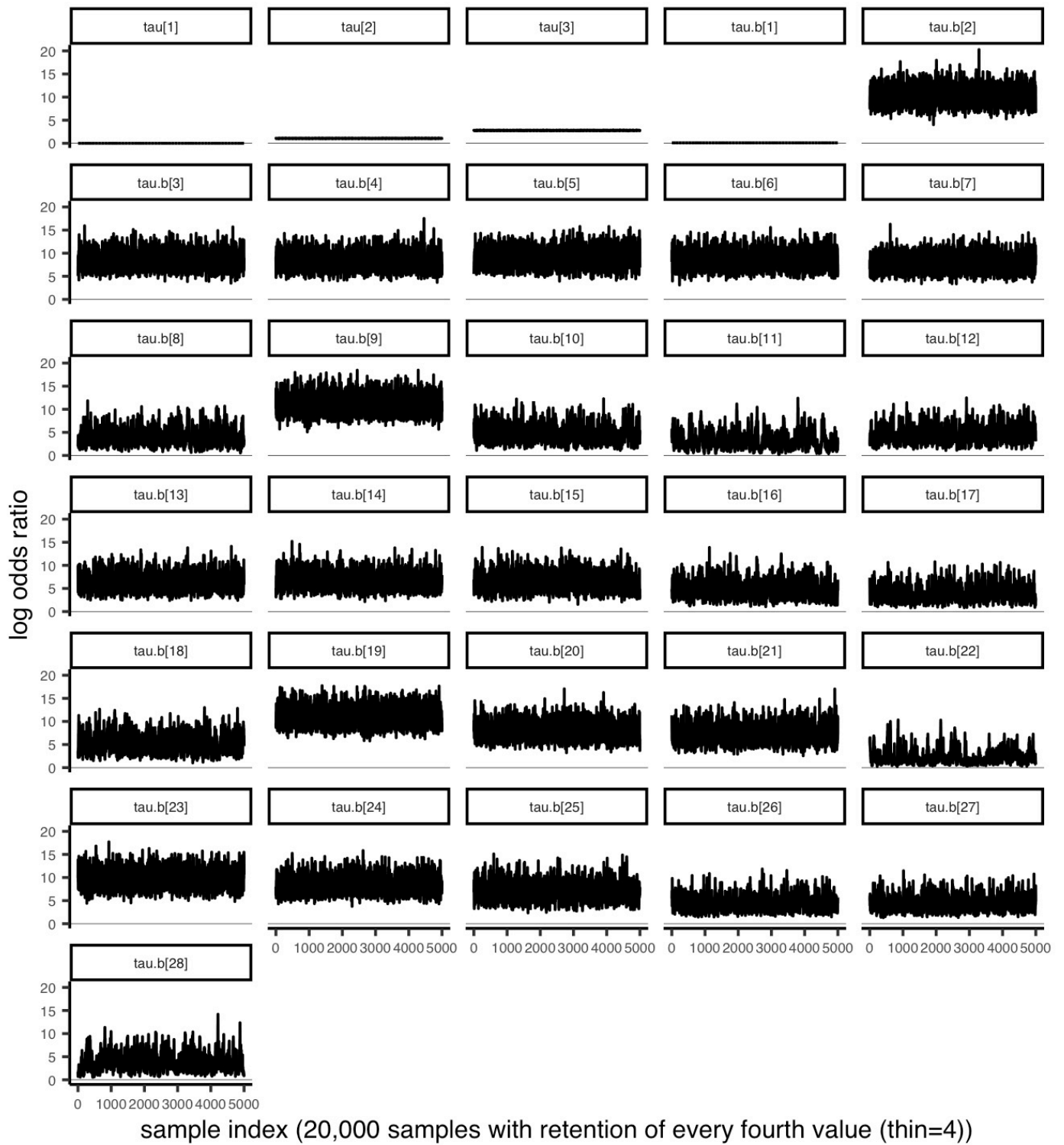


Figure A6 Mixing and convergence of the variance ($\tau.b[1:28]$) and threshold ($\tau[1:4]$) parameters for in the multilevel regression model. Note that for model identifiability, $\tau[1] = 0$ is fixed.

References

1. Gao, Y., Kennedy, L., Simpson, D. & Gelman, A. Improving multilevel regression and poststratification with structured priors. *arXiv* (2019) doi:10.1214/20-ba1223.
2. Hedeker, D. Multilevel models for ordinal and nominal variables. in *Handbook of Multilevel Analysis* (2008). doi:10.1007/978-0-387-73186-5_6.
3. Gelman, A. Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). *Bayesian Analysis* (2006) doi:10.1214/06-BA117A.
4. Plummer, M. rjags: Bayesian graphical models using MCMC. *R package version 3-13* (2016).