Averaging Over Varieties of Variance; Evidence from Survey Experiments on Immigration

Mohit Negi *
NM College of Commerce & Economics, Mumbai
and
Emma Pendl-Robinson
Junior Data Scientist, Mathematica Policy Research
and
Marco Scipioni
Project Officer, European Commission

December 14, 2020

Abstract

Recent literature on immigration attitudes has focused on estimating via survey experiments the determinants of such attitudes regarding immigration. Considering the degree to which opinions are polarized on immigration, while knowing the overall public opinions on specific questions is of substantive interest, so is the appreciation of the amount of variation that we can observe. We expand upon a recently published survey experiment by first devising a multilevel model accounting for variation between both respondents and countries. Simulating from such multilevel model fails to retrieve significant differences between high and low skilled immigrants as in the original analysis of the survey. We model the spread at the country level and find that skill level of the immigrant has a small but significant effect on the spread in several countries and that the difference in such a spread between countries depends on the features of the immigrant.

Keywords:

*The authors would like to thank the students in our class who generously provided feeback to our paper (Kacie Dragan and Yi Ning Chang), the teaching staff of GOV2001 (Professor Gary King and TFs Georgina Evans, Shusei Eshima, Sascha Riaz), and the authors of the original paper who kindly answered to our requests for clarifications.

1 Introduction

Research on attitudes towards immigration and their role in shaping contemporary politics and society has greatly expanded over the past decades. Past studies have captured preferences in public opinions regarding immigration (Hainmueller & Hopkins (2015); Bansak et al. (2016); Ford & Jennings (2020)), tested the stability of attitudes over time (Daniel et al. (2019); Kustov et al. (n.d.)), and investigated their role in shaping political outcomes (Grande et al. (2018); Ford & Jennings (2020); Stockemer et al. (2018); Green-Pedersen & Otjes (2017)). In this area, experimental methods have been applied mainly in the form of survey experiments to capture the effect of treatments on respondents' attitudes. A recent article titled, "Economic and cultural drivers of immigrant support worldwide," by Valentino et al. (2019) is remarkable because of the numerous treatments deployed in the survey and its cross-national scope. The paper aims to explain the causal determinants of public opposition to immigration in 11 countries with a series of survey treatments relative to type of employment, racial composition, family status, and country of origin.

We extend the original paper by Valentino et al. (2019) in two main ways. First, we explore the possibility of fully exploiting the nested structure of the data. In their analysis, after having clustered the repeated observations for each survey respondent, the authors either analysed the entire sample together or they analyse each country separately. Such strategy prevents them to explore how information relative to each grouping variable is captured in the model. Considering that the article's cross-national scope is one of its strength, being able to simultaneously analyse effects between and within countries is preferable. We provide a rough description of the importance to include each of the clustering levels in a single multilevel model by calculating the intraclass correlation. We also propagate uncertainty during estimation via simulating from our multilevel model. The results of the simulation - while subject to assumptions and dependent upon several methodological choices - point towards a more cautionary approach towards one of the main result in the original paper, namely higher support towards highly skilled immigrants.

Second, we expand the paper by shifting the question of interest from *relative to measures of central tendency* to *dispersion*. Much of the literature on attitudes towards immigration tends to present central tendency measures - e.g. means or median - of specific

opinions regarding immigrants, such as views on skills levels of migrants, or towards family reunification, or whether immigrants represent a challenge to welfare states. We believe that the analysis should also concentrate on how dispersed the opinions within countries are. While it is informative to know that, for instance, views on immigrants contributions to the welfare state may be more positive in country A rather than country B, and what could be the reasons for that, it is equally important to appreciate what is the relative degree of dispersion (or consensus) within a country. If measurements of central tendencies are relatively close, but variance in the data is large, this should alert the reader about the substantive implications of differences-in-means estimators.

We study this dispersion between respondents within countries by measuring the "spread" as the standard deviation in the data. Our results indicate that while the spread between countries is smaller than the spread among respondents, the former is non-negligible. This suggests that a model that includes all nesting levels - in this case, observations nested within countries - is more informative than a model with either no country identifier or individual regressions for each state.

Further, following best practices in the literature (Gelman et al. (2006)), we propagate uncertainty during estimation by simulating from our multilevel model. As a consequence, we were not able to retrieve significant changes within countries depending on immigrants' skills level, and between countries among identical treatment assignment.

Modeling the spread shows that the countries in our sample can be roughly split into two groups. A first, South Korea (KR) and Japan (JP) show a relative smaller spread in our simulations. A second, larger group of Western countries display larger but similar levels of spread. Substantively, our simulated standard deviations range from about 0.26 to about 0.35, on a dependent variable ranging from 0 to 1. This means the spread conditional on the treatments in the paper is rather large, suggesting that the public opinions across all countries in the sample are not extremely cohesive. This makes intuitively sense, considering how polarised public opinions are on immigration, and thus the expectations of a firmly held opinions on such divisive issue is perhaps unrealistic. However, this should be better regarded as a comparative matter, in as much as judgements on how cohesive public opinions should be checked across domains. Thus, future research could consider

applying an explicit modeling of variance across public policy domains to test for significant differences.

2 The Quantities Of Interest (QOI)

In the literature on attitudes towards immigration, several experiments have striven to identify the types of immigrants that a given population may prefer, usually through conjoint experiments. These types of studies were conducted in the United States (US), and the European Union (EU). In the US Hainmueller & Hopkins (2015) investigate, 'whether various subgroups of Americans respond differently to specific immigrant attributes.' They report that, to their surprise, 'our results uncover a sweeping consensus across different groups about which types of immigrants to admit.' This 'hidden American immigration consensus' has important policy consequences according to Hainmueller & Hopkins (2015), in as much as 'mass-level consensus indicates that any disagreements among the mass public are more likely to be over how many immigrants to admit or how to handle immigrants already here than over whom to admit.' In the EU, Bansak et al. (2016) observe that the probability of accepting asylum seekers depends on the applicants' features, such as whether they are male or female, or where do they come from, or their previous occupation. In particular, both studies use conjoint experiments, which 'ask subjects to evaluate hypothetical profiles with multiple, randomly varied attributes and are widely used [...] to measure preferences and the relative importance of structural determinants of multidimensional decision-making' (Bansak et al. (2016)). With an experiment nested in the most recent ESS wave, Ford & Jennings (2020) test whether opposition to immigration is more consistently connected with immigrants' skills rather than their origin. They find systematic evidence for a skills premium, meaning that immigrants with higher skills are generally more accepted than those with lower ones, but do not find the same widespread opposition based on a differentiation of immigrants as coming from a European or non-European country.

The key point of the present paper is that all these survey experiments-based studies use measures of central tendencies - such as average effects of treatments - as their key quantities of interest. Our proposal in this paper is to shift the focus towards a differ-

ent quantity of interest, namely the variance. In the context of mounting divisions and polarisation regarding immigration in many countries (see *Around the World, More Say Immigrants Are a Strength Than a Burden;2019 IIS 2910-S1.10514* (2019)), knowing the average opinion regarding immigration is both informative and of substantive interest, as it is testing whether being more informed regarding immigration would soften opposition, or whether specific immigrants' traits are connected with more scepticism. Besides all this, if opinions are widely different and fragmented within countries, public policies would benefit from also appreciating the degree to which opinions vary, and whether such variation dwindles when conditioned upon a (series of) treatment.

3 Data

The *average causal effects* [1] in Valentino et al. (2019) article refer to the effects of the following treatments: 'varying the skill level (high and low), family status (single or married with children), national origin (Middle Eastern versus Latin American/Asian) and racial phenotype (Afrocentric, non-Afrocentric) of an individual immigrant [shown in the vignettes]'.

The authors estimate such causal effects in two main ways: one model including all countries in the sample and then 11 separate country-specific regressions [2].

The effect of the treatments is measured on a dependent variable - *support* - which is a composite indicator of three different questions: *admission*, *length of stay*, and *naturalisation*. The composite variable ranges from 0 to 1.

The coding and aggregation decisions made by Valentino et al. (2019) may be regarded

[1] According to Imai (2017), 'Researchers want to estimate the unknown value of a *quantity of interest* using observed data. We refer to the quantity of interest as a *parameter*, and the method to compute its estimate as an *estimator*'. Imai provides a concrete example which is close to our paper. He states that, 'in randomised controlled trials, the average outcome difference between the treatment and control groups represents the estimator for the average causal effect, which is our parameter' - in other words, the QOI (Imai (2017)).

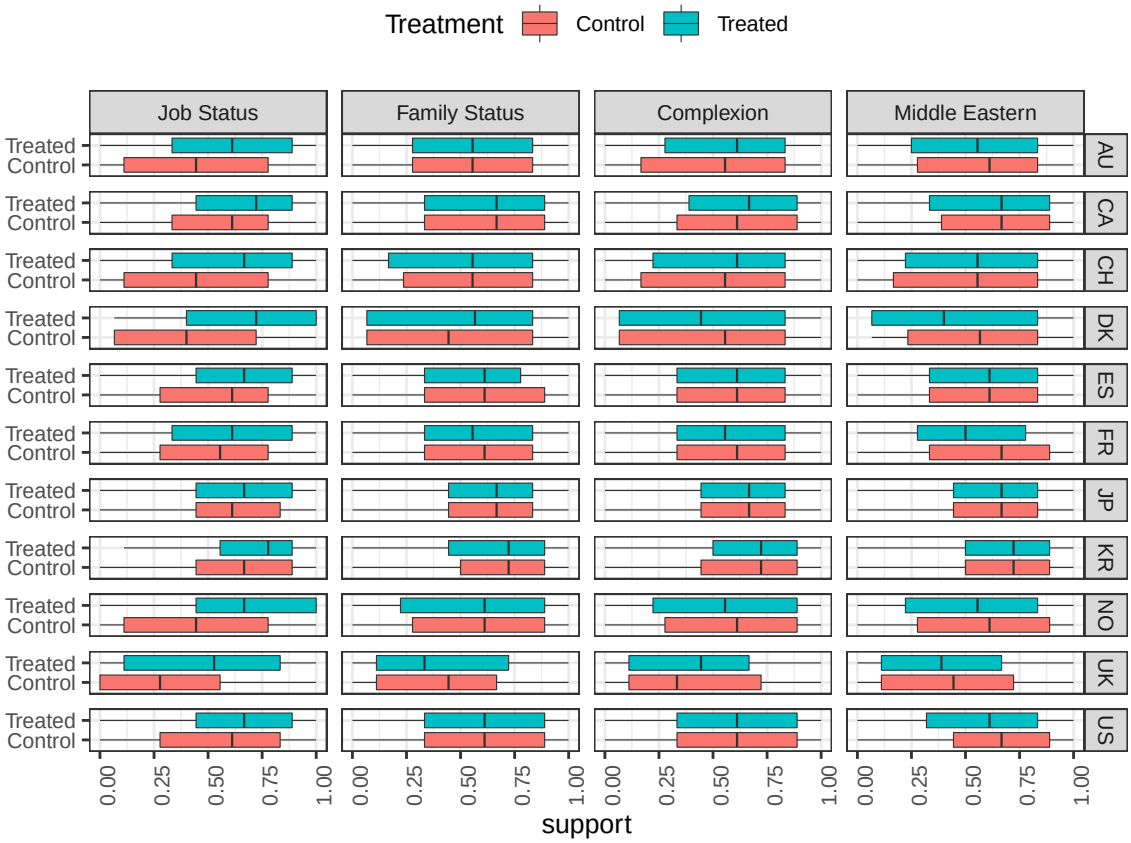
[2] Following the same presentation as the original authors of the paper, we abbreviate the 11 countries - Australia, Canada, Switzerland, Denmark, Spain, France, Japan, S.Korea, Norway, UK, USA - as following: AU, CA, CH, DK, ES, FR, JP, KR, NO, UK, US)

as problematic by some. Having coded *'can't say'* and *'undecided'* as 1 on a 0-2 scale (respectively, for the *admission* and *naturalisation* variables), and then aggregating them into a composite indicator, may skew the results in unpredictable directions. Perhaps it would have been better to check the results by having the same model run on these variables without these values (i.e. treat them as *don't know* and drop them, which would have turned all models into logistic regressions). Alternatively, if one consider these answers - again, the *'can't say'* and *'undecided'* - to be substantive answers (for perfectly justifiable reasons), it would have been better to treat these values as baselines within three multinomial logits (one for each variable), not assuming that they can be scaled in the middle as 1s. This is particularly concerning because then they are summed up and rescaled, thus treating them as ratio data/variables (which they are not). Moreover, we have no clue as to the distribution of these values in the original variables - in other words, how pervasive the *'can't say'* and *'undecided'* answers were. We asked the authors for the original variables and they denied access to the data[3].

Figure 1 below offer a quick glimpse at the distribution of the dependent variable in Valentino et al. (2019) original article. There is ample variation in the data when we unpack support by treatment and control groups, besides the type of treatment and country. For the skill-level treatment we see substantial differences between the medians of treated and control group for all states, whereas same does not apply to family-related treatments. At first sight, variance seems to be more country related rather than treatment-based. In other words, countries such as South Korea (KR) and Japan (JP) seem to have consistently low variance across treatments, whereas Denmark (DK) has relatively high variance disregarding the specific treatment being applied.

[3]Citing directly from their reply: 'They also are part of the country-specific files, which were gathered independently by the different researchers (using different grants). We don't have permission to redistribute the individual surveys'.

Figure 1 : Distribution of outcome variable conditional upon treatment, by country



4 Multilevel Model - Taking into account the nested structure of the data

A *mixed model* - as termed by the authors - accounts for multiple observations per subject. These multiple observations per subjects are likely to be correlated, as belonging to the same subject. This makes one of the assumptions of an OLS model - namely independent observations - likely to be violated (Agresti (2018)).

Indeed, the underlying idea of a multi-level model is that we take out of the error term another constant, which is connected to the clustering in the data (in the case at hand, repeated measurements of the same individual - so, it captures an additional source of variance at the individual level).

The research design of the article foresees not only repeated observations for each in-

terviewees, but also each interviewees nested within 11 countries. One of the authors' model specification envisages all countries pooled together (which the authors termed *all countries* in the graphs, and *all* in the regression tables). A possible explanation for this is that Valentino et al. (2019) wanted to understand if there were global trends that persisted regardless of country-specific features.

In this section, we couple the replication of the original model by Valentino et al. (2019) (termed *all* in the table below) with another model (labelled *MLM* for multilevel model, see Gelman et al. (2006)) which acknowledges the fact that we not only have repeated measurements for each subject, but also these subjects are nested within different countries (this is captured in the table below as the difference between the total number of *observations*, compared to the *group level* observations).

The two models are almost identical in both coefficients and standard errors, with the exception of the intercept, which has a much larger standard error in the second specification. In the *MLM* model, we have the same slopes for the two levels (individuals and country), as these are posited by the model to be fixed effects, but we see that - if needed - we are able to calculate the regression line at both levels. What is perhaps more interesting is that we are able to look at the differences in the estimated fixed effects in each of the two models with predictors. But we are now able to appreciate that what's included in the intercept of the nested model is actually considerable variation among states. Indeed, the intercept goes from a minimum of 0.44 in United Kingdom (UK), to a maximum of 0.71 in South Korea (KR). In short, we do not have to model a separate regression country for each country, and we can keep in a single model variation due to the respondents and variation due to country.

Besides changes in coefficients and standard errors, there are drops in both the Akaike information criterion (AIC) and Bayesian information criterion (BIC), a signal of the improvements in model fit of the second specification.

The *residual standard deviation* and the *level 1 standard deviation* barely change between the two specifications.

However, the different model specification enables us to capture the variance related to country level grouping, which is the the *Level 2 sd* in the table.

Table 1: Model comparison

	Dependent variable:	
	support	
	Valentino's all country	MLM
Intercept	0.52^{***} (0.004)	0.53^{***} (0.02)
Job Status	0.12^{***} (0.004)	0.12^{***} (0.004)
Family Status	-0.002 (0.002)	-0.002 (0.002)
Middle Eastern	-0.02^{***} (0.002)	-0.02^{***} (0.002)
Complexion	0.001 (0.004)	0.002 (0.004)
Vignette Position	-0.04^{***} (0.002)	-0.04^{***} (0.002)
Residual sd	0.14	0.14
Level 1 sd	0.29	0.28
Level 2 sd		0.07
Groups level 1	19734	19734
Groups level 2		11
Observations	37,579	37,579
Log Likelihood	$-3{,}688.43$	$-3{,}212.61$
Akaike Inf. Crit.	7,392.86	6,443.22
Bayesian Inf. Crit.	7,461.13	6,520.02
Note:		*p<0.1; **p<0.05; ***p<0.01

A formal notation for both the intercept and error term may clarify where do the fixed and the random effects components of the model come from. Indeed, one can regard both the systematic components β_j and the error term ε_{ij} as being random variables assumed to have normal distribution. The intercept in the systematic component can be written as $\beta_{0j} = \gamma_{00} + \delta_{0j}$, wherein γ_{00} is the mean intercept across all j clusters, and the random effect δ_{0j}, which stands for the deviation of a given cluster j from the mean intercept across all j clusters γ_{00} (*Statistical Methods for the Social & Behavioural Sciences: A Model-Based Approach* (2018)). At an abstract level, they can be written as $\beta_j \sim \mathcal{N}(\gamma, \sigma^2_\alpha)$, and $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2_y)$, respectively. This perspective enables us to understand that the original model by Valentino et al. (2019) had two main sources of stochastic variation

in the dependent variable, one related to the *level 1 observations* (i.e., the vignettes), and another component stemming from the variation among the *level 2 grouping variables* (i.e., the respondents in the original paper), The *intraclass correlation (ICC)* [4] captures the extent to which the variance can thus be apportioned between the various levels in a multilevel model.

Our results below indicate that while the authors' original model was able to account for about 0.68 of the variation in the data, our *MLM* model is able to apportion variation partly to the respondents (0.578278) and partly to countries (0.138261).

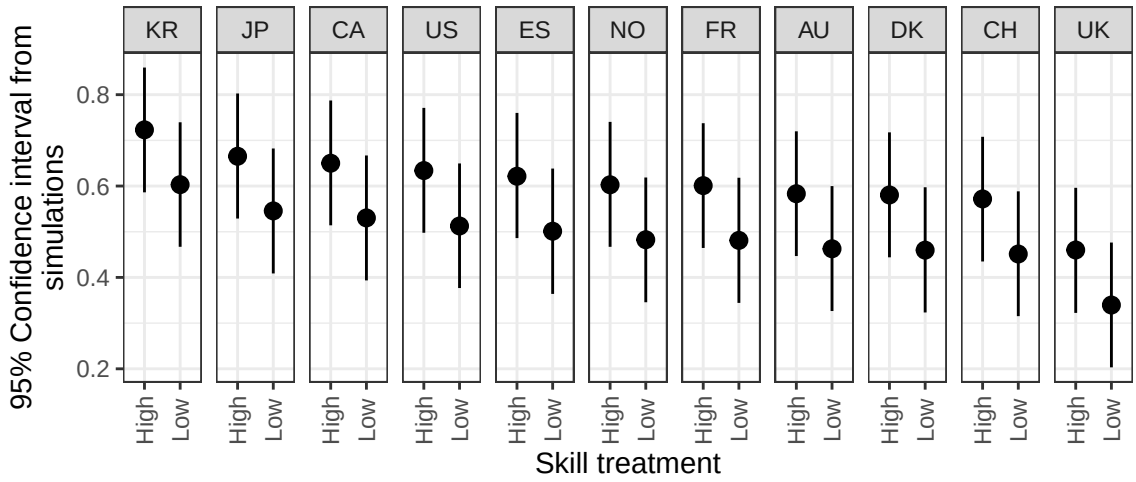
4.1 Prediction from simulation

If we want to accurately portray both the stochastic and systematic uncertainty from our *MLM* model, we can simulate each of these two components. One of the choices when simulating from a multilevel model is the selection of the very level we are predicting. Gelman et al. (2006) offer a way to think about this choice in terms of '*Prediction for a new observation in an existing group*' or '*Prediction for a new observation in a new group*'. Such choice entails simulating from distributions with different means and standard deviations. For instance, in the case of the former - i.e. 'Prediction for a new observation in an existing group' - Gelman et al. (2006) use the residual standard deviation, whereas in the case of the latter - namely, 'prediction for a new observation in a new group' - they use standard deviation of the level 1 grouping variable. In the case of our *MLM* model, we have a third standard deviation, which is related to the level 2 grouping variables, namely states. We focus on this source of variation as the added value of this model compared to the original by Valentino et al. (2019) is that we are able to include in a single model both variation within and between countries. To obtain confidence intervals from the simulations-based predicted probabilities, we set the covariates at median level, and pay attention to the difference between being shown a low-skilled versus high-skilled immigrants in the vignettes. This is similar to what Valentino et al. (2019) do in their

[4] 'the fraction of total variation in the data that is accounted for by between-group variation. The intraclass correlation can also be thought of as the correlation among units within the same group. [...]. The greater the correlation among units within a group (that is, the bigger ICC is) the greater the impact on the standard error' Gelman et al. (2006).

paper, where they plot the difference in the average treatment effect for their skill-related treatment.

Figure 2: Average treatment effect of job status by country, simulating from multilevel model



As a consequence of the simulations, we notice that - within each country - the confidence intervals for treatment and control tend to overlap, casting some doubts on the actual differences between the treatment effects. Further, we also see that - within each treatment - confidence intervals also overlap (even in the case of the two cases at the opposite side of the spectrum, i.e. South Korea (KR) and United Kingdom (UK)), again raising questions on the difference between countries. Overall, we also see that the confidence interval are much larger compared to those in Valentino et al. (2019) as a consequence of the propagation of uncertainty in both the systematic and stochastic components.

Importantly, the reader should bear in mind that these results are conditional upon keeping the fixed slopes (something that could be changed in a different specification) and simulating from normal distribution with standard deviation equal to the standard deviation of our level 2 grouping variable. Future research may wish to assess whether changing these model specification and assumptions lead to different results.

5 Modeling the Spread in Immigration Support

The second way we add to existing literature is by modelling the standard deviation of immigrant support as a measure of "spread in attitudes towards immigrants" between the

surveyed individuals within each country. Our model provides us with two interesting insights - When looking at heterogeneity in support for immigrants, the features of the immigrant matters a lot more than the features of the respondents. Likewise, the spreads vary a lot more between different countries for certain kinds of immigrants.

5.1 Motivation and Model description

The multi-level structure of our dataset motivates us to use a multiple equations approach where we model the responses to the two vignettes separately which helps us to account for the correlation between the two responses from each individual in the survey. Each pair of responses is modeled as a draw from a bivariate normal distribution. We then reparametrize not only the mean, but also the standard deviations of the distribution as a function of our treatment covariates. The parameters to both the mean and standard deviation equations are then jointly estimated using Maximum Likelihood.

This is an alternative approach to the Panel Linear model estimated using generalized least squares (GLS) in the original paper. We make the additional assumption of normality to enable us to run simulations for the purpose of calculating interesting quantities of interest in the subsequent sections. We work with the following model:

$$\text{Support for Immigrant in first vignette by respondent i} = Y_{1i}$$
$$\text{Support for Immigrant in second vignette by respondent i} = Y_{2i} \; \mathbf{Y_i} \sim \mathcal{N}(\mu_\mathbf{i}, \mathbf{\Sigma_i})$$
$$\mu_\mathbf{i} = \begin{bmatrix} \mu_{1i} \\ \mu_{2i} \end{bmatrix}$$

Where,

$\mu_{1i} = \beta_0 + \beta_1 JobStatus_{1i} + \beta_2 FamilyStatus_{1i} + \beta_3 MiddleEastern_{1i} + \beta_4 Complexion_{1i} + \beta_5 VignettePosition_{1i}$

$\mu_{2i} = \beta_0 + \beta_1 JobStatus_{2i} + \beta_2 FamilyStatus_{2i} + \beta_3 MiddleEastern_{2i} + \beta_4 Complexion_{2i} + \beta_5 VignettePosition_{2i}$

and,

$$\mathbf{\Sigma_i} = \begin{bmatrix} \sigma_{1i}^2 & \sigma_{12i} \\ \sigma_{12i} & \sigma_{2i}^2 \end{bmatrix}$$

We reparametrize the standard deviations,

$\sigma_{1i} = \alpha_0 + \alpha_1 JobStatus_{1i} + \alpha_2 FamilyStatus_{1i} + \alpha_3 MiddleEastern_{1i} + \alpha_4 Complexion_{1i}$

$\sigma_{2i} = \alpha_0 + \alpha_1 JobStatus_{2i} + \alpha_2 FamilyStatus_{2i} + \alpha_3 MiddleEastern_{2i} + \alpha_4 Complexion_{2i}$

Within-respondent Correlation $= \rho_i = \frac{\sigma_{12i}}{\sigma_{1i}\sigma_{2i}}$

Our model successfully replicates the estimates from the authors' original model and produces the correct standard errors thereby reassuring us that our model is not misspecified. Our approach also provides a flexible, software package-free framework to model various features of the distribution as functions of a wide variety of variables that might be available to the researcher. This might be very useful for testing certain hypotheses and calculating certain quantities of interest.

In our paper, we focus on two questions. Firstly, do these immigrant features (our treatments) impact the "spread" in attitudes towards immigrants amongst the individuals in a country? This spread in attitudes, measured as the standard deviation in immigrant support, is a key indicator of dispersion of opinions on certain kinds of immigrants in a population where a small "spread" might indicate a sort of consensus amongst the people about the desirability of the immigrant. Secondly, we shift focus to the question of how this spread as defined above differs between the different countries in our sample.

5.2 Results

Our model shows that the Job Status of the immigrant has a negative effect on the spread of support in five out of our ten country sample. Having a Middle Eastern background has a positive effect on spread in the US. That said, Figure 3 clearly illustrates that even though these effects are statistically significant, they are not very substantial. When compared to the effects of these treatments on the mean support, the effects of those same treatments on the spread are quite small.

The large intercept in the standard deviation (SD) equation implies that the treatments do not have much explanatory power for not only the spread, but also the mean support for the immigrants. It tells us that a large part of the variation in responses of different individuals is absorbed in the intercept, which remains unexplained by our treatment variables. It is clear that more data and more variables are needed to make more substantive

arguments and inform better policy.

Our approach, which we believe extends beyond just this paper in its usefulness, can be a very useful way to evaluate proposed policy changes. A well-informed policy decision requires that attention is paid not only to how immigration attitudes will change, but also how the overall distribution of attitudes in a population changes.

Table 2: Model Results : Standard Deviation Equation

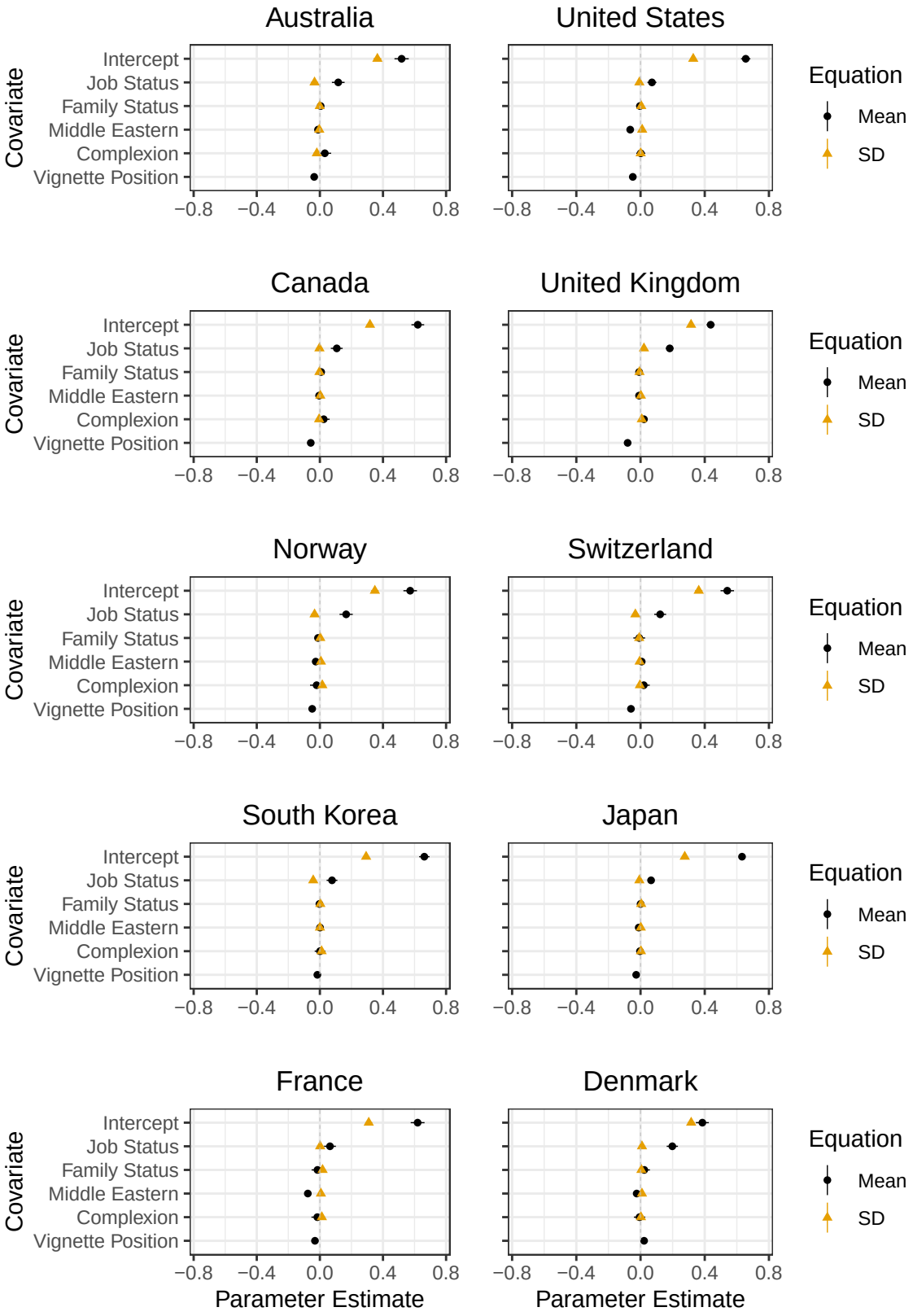
	Dependent variable:				
	SD of Support for Immigrant				
	AU	US	CA	UK	NO
Intercept	0.364^{***} (0.012)	0.328^{***} (0.008)	0.318^{***} (0.010)	0.315^{***} (0.007)	0.349^{***} (0.012)
Job Status	-0.034^{***} (0.011)	-0.007 (0.007)	-0.003 (0.010)	0.021^{***} (0.006)	-0.034^{***} (0.011)
Family Status	-0.001 (0.007)	0.004 (0.005)	-0.003 (0.007)	-0.006 (0.004)	0.002 (0.006)
Middle Eastern	-0.004 (0.007)	0.010^{**} (0.005)	0.003 (0.007)	0.002 (0.004)	0.007 (0.006)
Complexion	-0.019^{*} (0.011)	0.001 (0.007)	-0.005 (0.010)	0.008 (0.006)	0.015 (0.011)

Table 3: Model Results : Standard Deviation Equation (Continued)

	Dependent variable:				
	ImmigrantSupport				
	CH	KR	JP	FR	DK
Intercept	0.363^{***} (0.011)	0.293^{***} (0.009)	0.276^{***} (0.004)	0.310^{***} (0.011)	0.316^{***} (0.010)
Job Status	-0.032^{***} (0.010)	-0.042^{***} (0.009)	-0.007^{*} (0.004)	0.002 (0.010)	0.009 (0.009)
Family Status	-0.008 (0.010)	0.003 (0.004)	0.004^{*} (0.003)	0.017^{*} (0.010)	0.006 (0.009)
Middle Eastern	-0.005 (0.005)	0.0004 (0.004)	0.002 (0.003)	0.007 (0.007)	0.009 (0.006)
Complexion	-0.004 (0.010)	0.011 (0.009)	0.003 (0.004)	0.012 (0.010)	0.003 (0.009)

Note: *p<0.1; **p<0.05; ***p<0.01

Figure 3 : Parameter Estimates (Maximum Likelihood)



The horizontal bars represent 95% confidence intervals

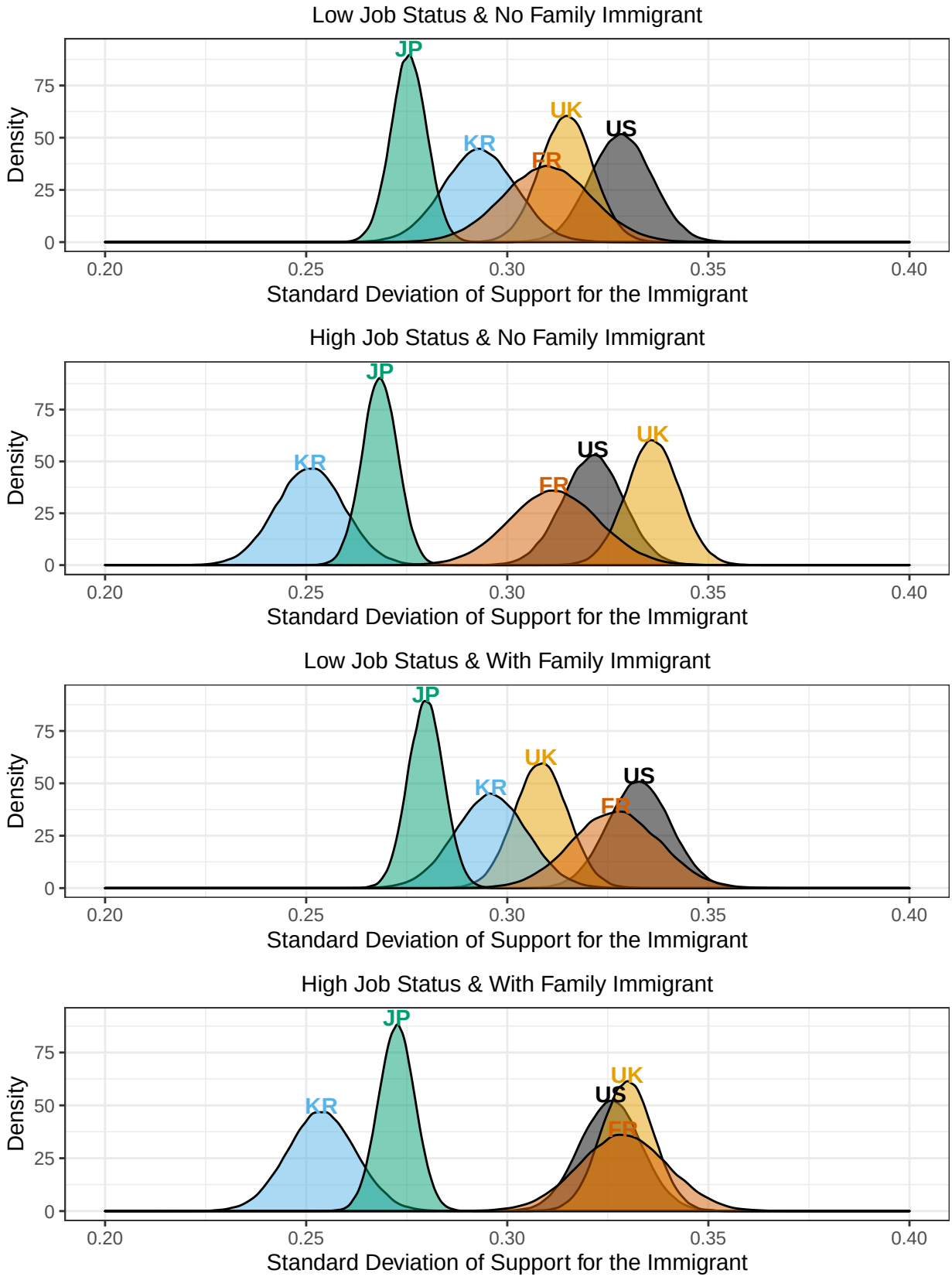
5.3 Differences in the spread of Immigrant Support across countries.

Now we look at how this spread in attitudes differs across countries. We simulate the between-respondent spread in support for immigrants of a specific kind (For example, High Job Status and No Family) for each country. We plot our results in Figure 4. Only the United States (US), United Kingdom (UK), France (FR), Japan (JP) and South Korea (KR) are retained to avoid overcrowding in the figure. For each plot, Middle Eastern and Complexion are set to 0, which means the immigrant is not middle eastern and their complexion has not been darkened by the researchers. The figure plots the distribution of the results from $100,000$ simulations, which helps us visualize not only the point estimates of between-respondent standard deviation for each country, but also the estimation uncertainty of those estimates.

The figure shows an interesting pattern between the Western countries (US, UK and France) and the Eastern ones (Japan and South Korea). For a low job status immigrant with or without family, there is not a large difference between the countries. However, for immigrants with a high job status, the difference between Eastern and Western nations increases. A large part is due to the spread in South Korea becoming much smaller. For high job status immigrants with families, the spread is very similar amongst the western nations and is larger than than the spread in Japan and South Korea.

Even though our model helps us simulate how the spreads vary amongst countries, a more careful study into the reasons for such patterns might be warranted.

Figure 4 : Between−respondent Spread in Support by Country and Immigrant type



6 Conclusions

The purpose of this paper is to expand upon the current trend in research on immigration attitudes by switching from a focus on difference in means and measures of central tendencies, to an analysis of the variance in attitudes between and within countries. In the context of division and fragmentation connected with opinions on immigration in many countries, it is of substantive interest to understand the degree to which such opinions actually vary.

The literature on immigration attitudes is expected to grow in the future. Our paper illustrates how future analyses in this domain can benefit by looking at measures of spread such as the standard deviation in addition to typical measures of central tendencies. By leveraging the nested structure of most datasets in this domain, simulating first differences for treatment variables, explicitly modeling the spread as a function of explanatory variables, and using simulations to calculate quantities of interest, a researcher can obtain interesting insights that might otherwise have been unexplored.

After accounting for the nested structure and simulating first differences, we did not retrieve significant effects of the skill-based treatment within or between countries for each treatment group. We model the spread in immigration support between respondents and find that the skill-based treatment has a negligible negative effect on spread in many countries. This implies that features of an immigrant do not play a large role in building consensus or division within a population regarding immigration attitudes. The countries in our sample have different amounts of spread among respondents. We used simulations to look at how the features of an immigrant affect this difference between the countries. Our main recommendations concern both methods and substance. In terms of methods, we adopt a simulation based approach to estimate quantities of interest also in the case of survey experiments with complex modelling structure. Even if the simulations are subject to assumptions and subjective decisions in terms of quantities of interest, they have the advantage of propagating uncertainties during estimation. We encourage researchers to match our analysis and estimation of spread in public opinion in other policy domains. On that comparative basis, we would be better positioned to understand whether our results are peculiar to attitudes towards immigration, or if they are in common with other domains.

References

Agresti, A. (2018), *Statistical methods for the social sciences*, fifth edition. edn, Pearson, Boston.

Around the World, More Say Immigrants Are a Strength Than a Burden;2019 IIS 2910-S1.10514 (2019).

Bansak, K., Hainmueller, J. & Hangartner, D. (2016), 'How economic, humanitarian, and religious concerns shape european attitudes toward asylum seekers', *Science (American Association for the Advancement of Science)* **354**(6309), 217–222.

Daniel, S., Arne, N., Doris, U. & Johanna, S. (2019), 'The "refugee crisis," immigration attitudes, and euroscepticism', *The International migration review* **54**(3), 019791831987992–912.

Ford, R. & Jennings, W. (2020), 'The changing cleavage politics of western europe', *Annual review of political science* **23**(1), 295–314.

Gelman, A., Hill, J., Alvarez, R. M., Beck, N. L. & Wu, L. L. (2006), *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Analytical methods for social research, Cambridge University Press, Cambridge.

Grande, E., Schwarzbözl, T. & Fatke, M. (2018), 'Politicizing immigration in western europe', *Journal of European public policy* **26**(10), 1–20.

Green-Pedersen, C. & Otjes, S. (2017), 'A hot topic? immigration on the agenda in western europe', *Party politics* **25**(3), 424–434.

Hainmueller, J. & Hopkins, D. J. (2015), 'The hidden american immigration consensus: A conjoint analysis of attitudes toward immigrants', *American journal of political science* **59**(3), 529–548.

Imai, K. (2017), *Quantitative social science : an introduction*, Princeton University Press, Princeton.

Kustov, A., Laaker, D. & Reller, C. (n.d.), 'The stability of immigration attitudes: Evidence and implications', *SSRN Electronic Journal* .

Statistical Methods for the Social & Behavioural Sciences: A Model-Based Approach (2018), **2018**(22).

Stockemer, D., Lentz, T. & Mayer, D. (2018), 'Individual predictors of the radical right-wing vote in europe: A meta-analysis of articles in peer-reviewed journals (1995–2016)', *Government and opposition (London)* **53**(3), 569–593.

Valentino, N. A., Soroka, S. N., Iyengar, S., Aalberg, T., Duch, R., Fraile, M., Hahn, K. S., Hansen, K. M., Harell, A., Helbling, M., Jackman, S. D. & Kobayashi, T. (2019), 'Economic and cultural drivers of immigrant support worldwide', *British journal of political science* **49**(4), 1201–1226.