Bayesian Multilevel Modeling for the Intersections of Race, Gender, and Class

Melina Much *

January 29, 2024

Abstract

Intersectionality is widely recognized as one of the largest contributions to the study of race, gender, and class across the academy. However, the quantitative operationalization of intersectionality within political science is often unsatisfactory. I provide evidence that the Bayesian Multilevel Model is an accessible and flexible tool for understanding intersectional dynamics in political behavior. Using both a synthetic simulation and a real-world example with the American National Election Survey (ANES), I show how Bayesian Multilevel Models increase our inferential understanding of group-based heterogeneity in public opinion and political behavior. Conventional techniques, such as interaction terms, frequently yield estimates that are obscured by considerable noise, making it challenging to discern meaningful patterns. In contrast, the Bayesian Multilevel Model excels at revealing underlying patterns in small sample-size environments. In doing so, the model better captures the interwoven nature of race, gender, and class that often goes unnoticed in political science research.

*This paper is available at PRQ. Melina thanks the following individuals for their feedback: Ines Levin, Danielle Thomsen, Michael Tesler, Jan Box-Steffensmeier, Jane Junn, Amanda Bittner, Amber Spry, Christine Slaughter, Rebecca Kreitzer, and Kira Sanbonmatsu. Additionally, she would like to thank the Empirical Study of Gender Network (EGEN) 2022 conference, their NSF grant #2215500, and her EGEN mentor, Dawn Teele, for their generous feedback and support. Lastly, she thanks the Political Methodology Section of APSA at the 2020 and 2021 conferences and the Gender and Political Psychology Conference.

Introduction

Scholars have long been interested in the influence of social identities on American politics (Berelson et al. 1954; Converse et al. 1961; Dawson 1994; Kinder et al. 1996; Mason 2018). The salience of race during Barack Obama's presidency and the racialized and gendered politics of the Trump era has made understanding the complex nature of intersecting identities all the more important for researchers (Tesler 2016; Sides et al. 2018; Jardina 2019).

American politics research, however, has not effectively updated the methods used to study increasingly salient intersecting identities. The leading quantitative methods used to account for race, gender, and class in politics often struggle to convey their compounding nature. In recent years, the discipline has made strides to overcome this methodological divide by using intersectional theory (Crenshaw 1989, 1991). However, the bulk of work on political behavior and attitudes in the United States omits intersectionality–particularly as it relates to the unique experiences of women of color (WOC) in the United States. This issue is exacerbated by the discipline's documented gender blindness (Forman-Rabinovici and Mandel 2023). As a result, crucial identity patterns in political life are overlooked.

Some of these failures of operationalization and omission stem from the difficulty in operationalizing intersectionality as it requires rich contextual understandings of lived experience and working against the hegemonic norm that race, gender, and class can be studied in isolation from each other (Simien 2007). However, there are also major methodological impediments to understanding the complex nature of intersecting identities in political behavior. In particular, there are real data limitations for quantitative scholars who study the intersection of already underrepresented subgroups (Frasure-Yokley 2018; Barreto et al. 2018). In addition, terms such as gender and sex are often conflated despite being separate concepts, contributing to the challenge of precisely studying social identities (Bittner and Goodyear-Grant 2017).

A segment of political science scholars such as McCall (2005), Weldon (2006), Junn (2007), Simien (2007), Hancock (2007b,a, 2019), and Spry (2018) theorize how to use intersectionality within quantitative methods and challenge these norms. These scholars have created frameworks for applying intersectionality quantitatively, outlined the limitations of current approaches, and rethought our survey methodology to understand multidimensional identity better. I update this literature by demonstrating that quantitative methods are already available to capture these relationships. Scholars have illustrated in fields such as sociology and epidemiology that intersectionality can be examined using multilevel modeling (MLM) (Jones et al. 2016; Evans et al. 2018; Merlo 2018; Johnston et al. 2018; Bell et al. 2019; Evans et al. 2020).[1] I aim to bridge the existing body of research on American politics and intersectional models while considering the discipline's specific data constraints and areas of substantive focus.

This piece solidifies Bayesian Multilevel Models (BMLMs) as a better option than frequentist regression using indicators or interaction terms in studying multidimensional identities.[2] The contemporary quantitative methods used to study multidimensional identities have shortcomings that can be addressed by BMLMs, namely, sample size limitations and false discovery rates. BMLMs can account for such group-based heterogeneity in data-scarce environments and reduce multiple testing issues by incorporating priors and specifying group-level effects via partial pooling of groups.

BMLMs have been shown to have utility for intersectionality in other fields but have yet to be incorporated into intersectional political science. I adapt an existing intersectional modeling approach, Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy (MAIHDA), with a simpler model more suited to

Political Science sample sizes and contexts. I also offer deeper engagement with Bayesian methodologies by incorporating informative priors not yet explored in the intersectional modeling literature. Using Bayesian methods can also show statistical uncertainty more candidly as an alternative to the p-value (Wasserstein and Lazar 2016; Wasserstein et al. 2019). This piece shows how informative priors built on previous intersectional work can bolster the model's estimation process and theoretical utility. Building on these methodological opportunities, this piece shows that BMLMs are a better option than standard regression for studying multidimensional identities on many fronts.

First, I outline commonly used methods to quantitatively measure identity, including indicator/dummy variables, interaction terms, and subgroup regressions (no pooling). Second, I present why the BMLM is a more suitable method when modeling the structure of the multidimensional lived experience with privileges and/or oppression. Third, I apply the BMLM to both a synthetic data simulation and a real-world example with the 2020 American National Election Study Survey (ANES). These applications showcase the advantages of BMLMs and offer new substantive insights into how identity mediates the relationship between ideology and partisanship. I conclude by showing that researchers can use this model to recalibrate their conventional approaches to understanding the interwoven nature of identity.[3]

Operationalizing Intersectionality

Intersectionality was pioneered by Kimberlé Crenshaw in 1989 as she critiqued identity literature and White feminism for a lack of understanding of the interwoven nature of race/ethnicity, gender, and class. In particular, Crenshaw highlighted how the lived experience of racism, sexism, and classism do not line up with the societal understanding of them being separate concepts and left the experiences of lowerincome Black women in the shadows. Crenshaw's original intention was to explain this in terms of the legal system as an interventionist and practitioner-oriented approach. The intellectual lineage of scholars of color describing these intersections is long; however, Crenshaw coined the term and approach in a way that has been widely disseminated in academic circles and politics alike (Davis 2008; Cho et al. 2013).[4]

This work is heavily intertwined with feminist and black feminist methodologies in the discipline and beyond. Feminist researchers have long debated the utility of quantitative methods for feminist work as they are rooted in positivist traditions which are seemingly at odds with feminist approaches (Hesse-Biber and Leavy 2007; Ramazanoglu and Holland 2007; Ackerly and True 2010). Many of these scholars operate outside the positivist realm, through interpretive or ethnographic methods, as these methods provide richer context (Jordan-Zachery 2007; Alexander-Floyd 2012). The nuance and context needed for a detailed description of intersections of oppression were often more suited to methods outside of quantitative methods (McCall 2005).

However, many feminist scholars have embraced quantitative methods while rejecting positivist notions of objective truth, instead offering a more holistic understanding of bias and the situated nature of research (Stauffer and O'Brien 2018). My work is aligned with the Stauffer and O'Brien (2018) approach in arguing that a feminist methodology uses the best available method to answer a given question and can reject problematic positivist notions of truth and objectivity. The BMLM is a tool for all researchers, especially feminist researchers, who have questions involving identity as measured by survey studies.[5] The method proposed in this piece is derived from taking the charges of contextual richness and theoretical robustness from interpretive scholars seriously by applying Bayesian methods. By using informative priors, researchers explicitly account for the interwoven and dynamic nature of social identities from the outset, producing results that are more contextually grounded and reflective of situated knowledge.

By integrating an intersectional perspective with feminist methodological commitments, the BMLM approach ensures that the complexities of social identities are not just acknowledged but systematically incorporated into the analytical process, making it especially well-suited for studying U.S. politics and beyond.

Using the Intersectional Research Paradigm

To situate this work within the broader intersectional literature, I first distinguish between intersectionality and multidimensionality. Intersectionality's intellectual trajectory has oscillated between broader interpretations that employ it as a research paradigm (Hancock 2007a,b), and more narrow interpretations that intersectionality should solely focus on the experiences of Black women and social justice projects (Alexander-Floyd 2012). Therefore, it is important to employ the concept with specificity with how it is being used, and who is centered in intersectional analysis. I use Hancock's intersectional research paradigm to frame modeling choices and analysis, which poses that intersectionality is an empirical worldview about how identity operates. In terms of terminology within that paradigm, multidimensionality typically refers to the larger scope of how race, gender, and class operate interdependently for all individuals, not just those with multiple marginalized identities (Simien 2007; García Bedolla 2007; Spry 2018). In this paper, I use multidimensionality to explain the effects on individuals with at least one axis of privilege in the model. I use intersectionality to describe the experiences of women of color (WOC). This understanding of intersectionality and multidimensionality guides the research questions in a way that centers WOC as the focus of intersectionality (Nash 2018).

There is also a rich lineage of Americanists using intersectionality have used it to explain: the race-gendering of WOC in political institutions such as Congress (Hawkesworth 2003; Smooth 2011; Brown 2012), the political behavior of WOC (Junn and Masuoka 2008; Junn 2017; Junn and Masuoka 2020; Brown 2014; Ojeda and Slaughter 2019), voting rights (Montoya 2020), interrogating U.S. democracy (García Bedolla 2007), and political attitudes of different racial groups of women (Frasure-Yokley 2018; Gershon et al. 2019). Researchers have also sought to apply intersectionality in the comparative context (Weldon 2006), and have created better datasets for intersectional analysis (Barreto et al. 2018). This paper focuses on building on this intersectional work to further show intersectionality's importance to understanding politics.

The following sections examine the most commonly used methods to incorporate identity into quantitative work and how these methods can better incorporate context when faced with data limitations. Scholars utilize three main approaches for identity, none of which has the same utility as the BMLM. Finally, I do not claim to evaluate intersectionality's "existence" in any of my applications, especially when a regression coefficient is insignificant, as some research does. This work is meant to build on the long academic lineage of intersectionality, even in contexts lacking quantitative evidence.

Indicator Approach

The most commonly used approach to account for identity is indicator variables. Additive understandings of identity lead to a reliance solely on binary indicator variables to account for race, gender, and class. Binary dummy variables in regression indicate the influence of a single identity on the outcome and do not consider any combined impact of these identities. Intersectional research shows us that WOCs in particular do not operate solely through their gender, racial, or class identity at a given moment, but are constantly influenced by the three structures (Junn 2007; Hancock 2007b). Indicator methods cannot be used to study intersectionality as they privilege one aspect of identity and do not consider the holistic impact of all identities. This is similar but distinct from complete pooling, which would be not including identity groups as variables at all.[6] Both complete pooling and indicator variables are troubling if the main point of the study is to investigate intersectional identity, as these tactics "pool" or dilute multiple identity group-based heterogeneity (Gelman and Hill 2006).[7]

Interaction Approach

To address the shortcomings of the indicator approach, methods scholars use race and gender interaction terms (accounting for the shared effect of race and gender) to model intersectional theory (Weldon 2006; Block et al. 2023). This does indeed allow for estimating the impact of intersectional or multidimensional race/gender groups. However, it is common practice to drop class from the analyses for sample size or simplicity. This may be from small sample size issues as referenced in the Appendix of the Block et al. (2023) paper, but the authors do not directly address in the main text why class is not included. It is also possible that these pieces focus on two identities for simplicity and ease of interpretation as they explore early innovations in quantitative intersectional methods. However, neglecting class is problematic as much of the intersectionality literature focuses on the influence of all three identities; class modifies both raced and gendered experiences. In addition, interaction effects can also have noisy estimates for small sample sizes and are not ideal for estimating grouped effects (Gelman and Hill 2006).

In the synthetic simulation section, I demonstrate that insignificant interaction terms often result from data scarcity in certain sub-groups rather than a true absence of substantive intersectional patterns. This limitation can lead some scholars to mistakenly conclude that intersectionality is not empirically present when, in reality, the data structure obscures its impact. Lastly, interaction terms require interpretation in terms of a reference level. This is problematic for a handful of reasons, namely that, by and large, the racial group chosen to compare against is Whites. Choosing Whites as the baseline without careful consideration can reinforce biases that Whites are the norm, and people of color (POC) are deviations from that norm. I will show how MLMs allow interpretation without considering a baseline or reference group.

Scholars have also documented that researchers frequently misinterpret high-order interactions, particularly regarding uncertainty around coefficients (Brambor et al. 2006). Brambor et al. (2006) show that computing marginal effects is a more reliable way to assess interaction effects, as focusing solely on coefficient significance can obscure essential findings. Yet, nearly twenty years later, the discipline still largely adheres to norms that prioritize coefficient significance (and do not go further) over more nuanced interpretations like marginal effects. Given the importance of intersectionality in centering marginalized groups, I argue for a modeling approach that avoids these pitfalls.

There is merit in arguing for the use of interaction terms for their accessibility (many scholars are familiar with them) and the fact that interaction terms account for both the additive and multiplicative forms of race, gender, and class intersections. However, I argue that the MLM provides a more robust framework for understanding intersectionality because it aligns with the group-based structure of the data. Intersectionality is not just about examining identities in isolation but about understanding how structural power shapes the differences between identity groups. Unlike interaction terms, which may statistically capture these dynamics without explicitly accounting for the underlying structures, the MLM better reflects how structural forces produce group-based differences. The MLM, by modeling these overarching structures, offers a more theoretically coherent approach to studying intersectionality.

Sub-Group Regression Approach

Subgroup regressions, (or a separate regression for each race/ethnic and gender intersection), are another method used by intersectional scholars (Frasure-Yokley 2018; Hancock 2019). This tactic lacks modeling parsimony but accounts for the grouped nature of race and gender.[8] Within the methods community, this is called the "no pooling" approach. No pooling methods separate all groups, and regression is run separately for each. This can lead to overfitting issues, and it is problematic for groups with small amounts of data in each group, as the researcher is likely to get more extreme estimates (Gelman and Hill 2006).

Most datasets do not often allow for one to subset by race, gender, and class and still maintain large enough sample sizes for robust statistical analysis. This issue was outlined in the original discussions of intersectionality within Crenshaw's analysis of defendant experiences with the legal system. The defendants outlined in "Demarginalizing the Intersections" had trouble proving their point as data on Black women were few and far between (Crenshaw 1989). Researchers also cannot directly compare coefficients across multiple models, so direct comparisons of effects across race, gender, and class subgroups are lost. Moreover, interpretation is problematic for multiple dependent variables of interest and various subgroups, as this requires many models to be run for each subset and the effect being measured. Lastly, running many regression tests on the same dataset increases the risk of multiple testing issues, inflating the false discovery rate. As a result, a researcher may find statistically significant coefficients by chance rather than due to genuine patterns in the data. I will not consider this an optimal intersectional method in most cases when sample sizes are small, and a researcher needs to compare different identity groups directly.[9]

BMLMs provide an intuitive way to solve the problems posed by the aforementioned methods. In a single model, we can produce an intercept and slope relative to the intersections of race, gender, and class. This approach reduces noise and the risk of overfitting while maintaining meaningful comparisons between groups. It also provides more information about the influence of intersectionality.

Bayesian Multilevel Models for Intersectional Analysis

Multilevel models (MLMs) are designed to account for related groups within data and can be implemented using either Bayesian or frequentist approaches. While both frameworks offer advantages, the Bayesian approach provides additional benefits that make it the preferred choice when feasible. In particular, the BMLM allows us to use informative priors and performs better in small sample-size environments. This section highlights the well-documented advantages of MLMs in general while emphasizing the unique strengths of BMLMs. A classic example of multilevel models is analyzing student data without accounting for classroom effects. In such studies, failing to consider the influence of a specific classroom overlooks a crucial factor in a student's academic experience (Gelman and Hill 2006; Peugh 2010). Without accounting for the effect of a given classroom (or another grouping factor like identity), important trends shared by students within that classroom may be missed.[10] The classrooms can be thought of as the structural impact of intersections of racism, sexism, and classism on identities. For intersectionality, structural oppression is interwoven on many fronts, which creates unique group-level effects for WOC, specifically in many political outcomes. Multidimensionality also demonstrates that identities with some level of privilege (via race, gender, or class) also have unique lived experiences, given the impact of different axes of privilege and oppression.

To account for these related structures in the data, the MLM calculates individual-level effects (the traditional main effects of a regression model), and group-level effects (also called random or mixed effects) which vary across individuals in the sample according to the group. These group-level effects capture the naturally occurring patterns that result from the unique impact of the classrooms, or in this case, race, gender, and class combinations. While this modeling approach has not been approached in political science for intersectionality specifically, the BMLM has a proven track record of success in modeling intersectionality in other disciplines (Jones et al. 2016; Evans et al. 2018; Merlo 2018; Johnston et al. 2018; Evans et al. 2020; Bell et al. 2019). These scholars are more often in data-rich environments and specify their models slightly differently (sample sizes in the tens of thousands); however, those approaches show that political science as a discipline can also benefit from updating the functional form of intersectionality to a BMLM. An outline of the recommended BMLM formula for political science use is featured below.

The general form of the individual-level effects is featured below, where i represents a given individual, and j represents the identity groups.

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \epsilon_{ij}$$

Where the MLM differs from traditional fixed effects regression is the estimation of group-specific effects for both the intercept β_{0j} and the coefficients β_{1j} through partial pooling, which weights the group estimates with the individual level estimates. In other words, partial pooling uses the overall average for the individual level to inform group-level estimates (Gelman and Hill 2006). The MM also measures residual error terms at the group-level u_{oj}. Those formulas are shown below. In this example, we can estimate the group effect (j) across different coefficients of interest β (slope β_{0j} and coefficients β_{1j}), using the grand mean γ_{00}, deviation from the grand mean in the second level γ_{10}, and group-specific means γ_{01}, γ_{11}. The group-level intercept is β_{0j} and the slope is β_{1j}.

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(X_{ij}) + u_{oj}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(X_{ij}) + u_{oj}$$

I specify the models to have individual-level effects, and a grouping variable for each race, gender, and class combination to find specific group-level effects (by the method of partial pooling).[11] There has been some debate about using these identities in MLMs and whether MAIHDA authors argue that demographics should be accounted for at both the individual level and then again at the second level based on race, gender, and class combinations. This runs into issues with collinearity because similar concepts are used at both the individual and group levels (Wilkes and Karimi 2023; Evans et al. 2024). Collinearity inflates standard errors and reduces precision, making it difficult to determine whether an effect stems from individual-level differences or broader group-based patterns. Evans et al. (2024) attempts to lay rest to this claim by stating identities in MAIHDA should be used at the group level and not redundantly in the individual. I also do not recommend accounting for identities at the individual level if one is specifying group-level effects. I argue that they should be incorporated at the second level as groups.

I again propose using Bayesian methods for MLMs as they allow for further context to be built into modeling practice and have better small sample size estimation (Rupp et al. 2004; Kruschke et al. 2012; Zyphur and Oswald 2015). Bayes' theorem is as follows (Clark 2018):

$$posterior \propto prior * data$$

$$updated\,belief = prior * current\,evidence$$

Each portion of the Bayesian equation is posed in the context of distributions of values. This is a critical differentiation of Bayesian estimation from frequentist estimation. The BMLM estimates individual level and group effects as distributions rather than just point estimates (as in classical regression) through Stan programming and Hamiltonian Monte Carlo (HMC) sampling in the *brms* package (Bürkner 2017).

An in-depth explanation of Bayesian methods will not be given here.[12] While BMLM cannot solve all historical data scarcity issues, it can allow researchers to make better use of long-standing datasets like the ANES, or their original surveys which have small sample sizes. In the following subsections, this article details the benefits of moving to a Bayesian framework, the benefits of a multilevel model, and the benefits of combining the two.

Bayesian Benefits

The literature has well documented that Bayesian methods are better suited for small sample group size environments because they utilize a prior distribution during the estimation process (Rupp et al. 2004; Kruschke et al. 2012; Zyphur and Oswald 2015). Sample sizes have long been an obstacle for researchers who do quantitative intersectional work. Crenshaw clearly documented issues of sample sizes of Black women being too small for robust statistical analyses (Crenshaw 1989). Surveys such as the ANES, the longest-standing and arguably most well-known American election survey, have very small sample sizes of intersectional race, gender, and class groups because it is set up to be representative to the American public. While representative sampling isn't inherently problematic, researchers of intersectionality are often left with lackluster sample sizes for disaggregation. It will likely be the case that identity sample sizes will continue to be small in the foreseeable future in these representative datasets, and intersectional research will always need to disaggregate. Therefore, different estimation tactics need to be explored, such as Bayesian frameworks, which operate better with small sample sizes.

This sample size advantage does come at a cost, however, as it requires that careful attention be paid to the prior distribution. As scholars such as McNeish (2016) and Smid et al. (2020) point out, estimates can be biased if certain software defaults or flat priors are used without caution. As outlined in the ANES application, I argue for informative priors based on previous years of the ANES with some adjustments based on extant understandings of the intersectional literature in that given topic.

This use of priors can be leveraged by intersectional scholars to create estimates that take into account intersectional literature that would not be included in frequentist estimation. Bayesian methods with informative intersectional priors leveraging the spectrum of qualitative, quantitative, and interpretive knowledge will better inform multidimensional group heterogeneities than frequentist frameworks. Critical race and feminist scholars have long acknowledged the persistent salience of racism, sexism, and classism's influence on political behavior and attitudes, but much of this work exists in qualitative and interpretive spaces. Therefore, I argue that building these works into model specifications will lead to more realistic results, especially in a data-scarce environment where it is logical to lean on decades of previous research on a topic. Incorporating literature into the estimation process helps contextualize the model and its estimates, making Bayesian methods more compatible with feminist methodologies that emphasize the creation of situated knowledge (Stauffer and O'Brien 2018).

Informative priors in this paper are garnered by running a comparable model in a previous year of the dataset (2016), and using those estimates as the basis for the prior of the year in question. Then, I make adjustments based on the state of intersectional research on the topic. I also recommend utilizing a Lewandowski-Kurowicka-Joe (LKJ) prior on the correlation matrix to account for the inter-group correlation effects. Defining the LKJ prior (which allows for correlations between the levels of the grouping variable) helps build in the context of the interrelated nature of race, gender, and class as implied by multidimensional identity research. An LKJ prior of less than 1 as used by this piece denotes a prior which allows for higher between-group correlations, and those greater than 1 lead to less.[13] A researcher might prefer an LKJ prior with a shape parameter less than 1 because it induces a prior distribution that favors stronger correlations among the random effects. This can be beneficial when modeling hierarchical data where groups are expected to share similar patterns due to common attributes such as race or gender within the grouping variable. A more technical account of these priors is explained in the ANES section and the Appendix.

Bayesian methods also offer a more comprehensive perspective on uncertainty compared to traditional frequentist regression. Instead of providing a single point estimate for the "Truth," Bayesian methods estimate it as a distribution of values. This distribution, obtained through thousands of draws, not only yields a mean or median that can act like a point estimate but also provides a range of uncertainty to assess other statistical moments. Consequently, Bayesian estimation enhances our understanding of the true relationship under investigation, presenting a nuanced alternative to the frequentist paradigm (Gelman et al. 2013). By moving away from reliance on p-values, Bayesian methods articulate uncertainty through these distributions, avoiding the potential misinterpretations associated with p-values. This approach aligns with contemporary critiques of frequentist statistical significance which emphasize the use of Bayesian alternatives that offer a more transparent representation of estimate uncertainty (Wasserstein and Lazar 2016).

Bayesian estimation is now more accessible to scholars than ever before. Previously, they required lengthy package and syntax knowledge of (but not limited to): BUGS, JAGS, or Stan programming.[14] These skills were needed beyond learning Bayesian statistical theory, which are not as widely accessible as frequentist pedagogy. The "brms" package is one of a few R-based programs that interface with R to create multilevel model Stan code for you without needing to learn Stan (Bürkner 2017).[15] The brms package enables researchers to fit a wide range of Bayesian multilevel models with greater ease, utilizing weakly informative default priors that aim to improve convergence while exerting minimal influence on parameter estimation and being computationally stable (Bürkner 2017). Scholars of race, gender, class, and intersectionality broadly can utilize this step towards accessibility in using Bayesian methods.

Multilevel Benefits

Additionally, some benefits come from using a multilevel model alone in a frequentist setting. Partial pooling gives more precise results than other pooling methods (smaller error bars), particularly when there is a small sample size in a given group Gelman and Hill (2006). This approach can reveal statistically significant effects that may be obscured by the noise in unpooled analyses, especially for smaller groups. The benefit of small sample size precision comes with a caveat that the within-group sample size can only be so small. Research has shown that small amounts of groups in frequentist settings (in this case, the number of identities) can lead to inflated second-level standard errors and unreliable coefficient estimates, which is inherently problematic since intersectional effects operate at the second level (Maas and Hox 2005; Paccagnella 2011; Stegmueller 2013; McNeish and Stapleton 2016; Park and Yu 2018). This drawback is part of the benefit of using a Bayesian approach, as scholars have documented it ameliorating these small sample size concerns in groups (Maas and Hox 2005). I follow recent recommendations that there should be at least 20 groups and at least 10 people per group in the Bayesian context (Park and Yu 2018). This does indeed limit the amount of structural identities that can be included in analyses. A researcher might want to add sexuality or disability status as an additional grouping identity that is commonly referenced in qualitative or interpretive literature; however, as we will see with the later ANES application, further splitting would lead to unusable subgroup sizes in many cases. Adding additional identities must be weighed by both theoretical reasoning for inclusion and data limitations.

Partial pooling is theoretically more compatible with intersectionality logic as it shares information across all the race, gender, and class groups. A fully unpooled model (sub-group modeling approach) assumes that all race, gender, and class combinations are fully distinct from each other. This would mean that we could not learn from Black women in the lowest class about Black women in the middle class, or could not learn about Black women in the lowest class from White women in the lowest class. A completely pooled model would say something even more extreme, that all Americans are the same regardless of race, gender, or class (and their intersections). In a partially pooled model, we obtain a distinct estimate for each group (e.g., Black women in the lowest class), but that estimate is informed by related groups through shared identity dimensions. Specifically, the estimate is influenced by other individuals who share gender and class (e.g., White women in the lowest class) as well as those who share race across different class categories like Black men. Therefore, partial pooling allows the researcher to recognize that there is shared lived experience on any one of these shared dimensions that informs the overall combined identity. I argue that this approach more accurately captures identity-based experiences, as entire subfields—such as Black Politics and Gender and Politics—have documented the significance of lived experience within these groups. By leveraging partial pooling, we incorporate these shared experiences into the estimation process while still recognizing that identity is fundamentally intersectional.

Multilevel models also allow one to estimate new quantities such as the intra-class correlation coefficient (ICC) or, in Bayesian contexts, variance partition coefficients (VPC), which are helpful to scholars who care about group-based heterogeneity (i.e., identity scholars).[16] The ICC/VPC is a percentage that shows how much of the total variation explained by the model can be attributed to the grouping variable specification. It "ranges from 0 if the grouping conveys no information to 1 if all members of a group are identical (Gelman and Hill 2006)." In the frequentist context, the ICC is the ratio of the between-cluster variance to the total variance. The VPC in the Bayesian context involves computing a variance decomposition using the posterior predictive distribution. Initially, it generates samples from the posterior predictive distribution without conditioning on group-level terms, and also obtains samples conditioned on all random effects. Subsequently, it calculates variances for each of these samples. This can be thought of as analogous to the ICC (Lüdecke 2024).[17] This measure provides utility to intersectional scholars as they now have insight into how influential race, gender, and class are in accounting for explained variation in the model.

Below I feature a table that provides an overview of the conventional methods in comparison to the BMLM. While the process is more complicated, it is a small price to pay for increasing small sample size performance, minimizing noise and overfitting issues, providing new variation terms, and incorporating prior context.

Method	**Functional Form**	**Benefits**	**Drawbacks**
Single Regressions (Complete Pooling)	$Y_i = \beta_0 + \epsilon_i$	- Accessibility to the academic community	- Does not address intersectionality - Not considered an intersectional approach
Indicator Variables	$Y_i = \beta_0 + \beta_1 \text{Race}_i + \beta_2 \text{Gender}_i$ $+ \beta_3 \text{Class}_i + \epsilon_i$	- Accessible to the academic community - Explicitly includes race, gender, and class	- Assumes additive effects - Ignores intersectional differences
Interaction Terms	$Y_i = \beta_0 + \beta_1 \text{Race}_i + \beta_2 \text{Gender}_i$ $+ \beta_3 \text{Class}_i + \beta_4 (\text{Race} \times \text{Gender})$ $+ \beta_5 (\text{Race} \times \text{Class})$ $+ \beta_6 (\text{Gender} \times \text{Class})$ $+ \beta_7 (\text{Race} \times \text{Gender} \times \text{Class}) + \epsilon_i$	- Captures identity interactions - Maintains accessibility	- Requires large sample sizes - Becomes noisy with third-order interactions
Sub-Group Regressions	*Black Working-Class Women:* $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ *Working-Class Latinas:* $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ *White Working-Class Women:* $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$	- Directly models intersectional differences - Provides clear subgroup comparisons	- Difficult to interpret across groups - Increases false discovery risks
Multilevel Modeling	*Level 1:* $Y_{ijk} = \beta_{0jk} + \beta_{1jk} X_{ijk} + \epsilon_{ijk}$ *Level 2 Intercept:* $\beta_{0jk} = \gamma_{00} + \gamma_{01}(\text{Race})$ $+\gamma_{02}(\text{Gender}) + \gamma_{03}(\text{Class}) + u_{0jk}$ *Level 2 Slope:* $\beta_{1jk} = \gamma_{10} + \gamma_{11}(\text{Race})$ $+\gamma_{12}(\text{Gender}) + \gamma_{13}(\text{Class}) + u_{1jk}$	- Accounts for group-based heterogeneity - Reduces overfitting via partial pooling - In Bayesian context has priors	- Less accessible - Minimal gains in data-rich environments

Table 1: Model Comparisons Including Race, Gender, and Class

Adding More Identities

Before applying the method, it is important to consider not only data constraints but, theoretically, how many additional identities can and should be incorporated into the analysis. I interpret intersectionality as such that the identities most essential to address are those with a clear privilege-oppression dynamic embedded within societal structures. This is why race, gender, and class are the most commonly studied, as they are directly tied to systemic racism, patriarchy, and classism. Beyond these, sexuality and disability would logically be the next identities to consider. However, certain identities that hold significance in American politics—such as religion, geographic divides (rural/suburban/urban), and partisanship—do not, in my view, function in the same way as structural intersectional identities.

These factors should be incorporated into the model differently than race, gender, and class. The key advantage of MLMs is their ability to capture structurally embedded patterns of oppression, so any additional identities added at the grouping level should also be clearly structural in nature. Maintaining a stricter definition of what constitutes an intersectional identity, i.e., focusing on those with precise structural privilege-oppression dynamics, ultimately keeps the model theoretically grounded and empirically manageable. This approach ensures that the focus stays on those most marginalized within societal power structures rather than diluting intersectionality's core purpose by incorporating identities that, while politically relevant, do not operate through the same embedded systems of oppression.

Valid arguments could be made to include identities beyond race, gender, and class (especially for rural/urban/suburban divides), but which ones are chosen is ultimately dependent on a researcher's interpretation of intersectionality. In my approach, however, commonly relevant identities such as religion, geographic divides, and partisanship are best incorporated at the first level of the model rather than as structural grouping variables. This means that rather than treating them as intersectional identities at the second level, we would include them at the first level and allow the structural identities to vary around them.[18]

Data and Method

In this section, I detail a synthetic data simulation of BMLM performance and a real-world example in the American National Election Study in 2020. The synthetic approach is meant to show first how BMLM outperforms the next best method for capturing intersectionality, the interaction term approach, when we know what ground truth is. In this case, data was created to mimic a scenario in the ANES where a researcher is analyzing a given year and modeling an outcome that has race, gender, and class-based effects. Since the data was created with exact parameters for the size and nature of the effects, we can compare the performance of each model in uncovering the identity effect built into the data. I compare the performance of a frequentist interaction approach and a BMLM to demonstrate that the proposed tactic estimates with more precision and reduces risk of Type II error, or a false negative from small sample size complications. I support this with a Monte Carlo simulation detailed in Appendix A.

This is followed by an example from the 2020 ANES which investigates a burgeoning area of partisanship research—the ideology/partisanship paradox (White and Laird 2020; Jefferson 2020)— to demonstrate the utility of the BMLM. This body of research challenges the conventional understanding that ideology (as an independent variable) consistently predicts and correlates with partisanship (as a dependent variable). Instead, it demonstrates that this relationship primarily holds for White Americans, but not for Black Americans. I argue that this finding should be substantively extended to certain subsets of the Latino(a) population. More specifically, this phenomenon appears to also vary at the intersection of race, gender, and class for Latino(a) communities. Methodologically, I display findings that show the utility of the BMLM in uncovering results that would be insignificant in an interaction framework because of the small sample size environment. These findings are bolstered by cross-validation and out-of-sample performance metrics to demonstrate BMLM performance.

Synthetic Data Simulation

The synthetic simulation is meant to show the performance of different modeling choices in a situation where the researcher knows the true relationship between our inputs and output variables and can compare the results from each test. First, a synthetic dataset was created with an output that resembles a standard feeling thermometer (FT), as seen in surveys like the ANES, where attitudes towards a group or person are measured on a scale from 0-100, with zero being cold/negative, and 100 being warm/positive. Variables were also created to mimic race, gender, and class in sample sizes found in a representative sample of the American public with a total sample size of 3,000. Lastly, the output was created to have dependent effects on race, gender, and class for Black and White women in different class groups.[19][20] The size of the impact is shown in Table 2. A model that does well will be able to uncover these values with high certainty and accuracy. In this example, I chose not to specify informative priors to closely compare the estimates from the frequentist methods to the Bayesian ones.[21] Due to the data being generated solely using the influence of identity features, I will not feature the ICC/VPC because the utility is gained when we don't know the sources of variation.

Three different modeling approaches are presented. The first is a three-way interaction estimated in a frequentist paradigm which is argued by Block et al. (2023).

Size of Effect		
Class	**White Women**	**Black Women**
Lowest	-16	16
Working	-12	12
Middle	-8	8
Highest	-4	4

Table 2: Synthetic FT Effects

The second is a frequentist multilevel model (FMLM), and the third is the BMLM. Figure 1 demonstrates that small sample size environments for these subgroups do create noisy estimates (large error bars) as documented in the literature (Gelman and Hill 2006).[22]

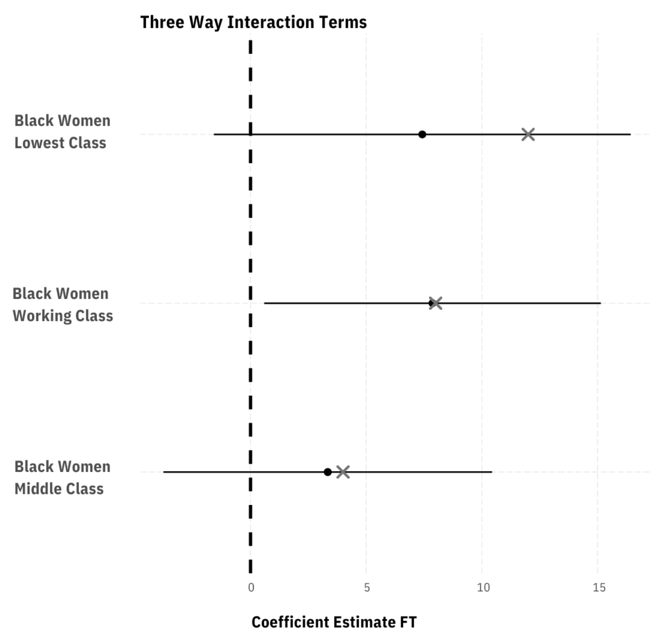

Figure 1: Interaction Terms
The dependent variable is the synthetic feeling thermometer, and these coefficients should be assessed as the group identity interaction term between race, gender, and class. X is the true effect shown in Table 2.

In this case, we can see the drawback of these higher-order interaction terms. They generally are accurate (i.e. get close to the specified value of the group on average, shown by the proximity of the X and dot), but are not precise, and have large error bars that lead to “insignificant” substantive takeaways. Due to discipline-wide norms of interpreting frequentist p-values as direct evidence, a researcher could come away from this investigation and find that intersectionality does not “exist” in this case despite a sizeable intersectional pattern in the underlying data. This interpretation is based on a modeling issue under data-scarce environments rather than what is happening in reality. I support this finding with Monte Carlo simulation results in Appendix A, which show that insignificant p-values are more common in the interaction context. Also, the average standard error across estimates is much greater with three-way interactions. Below, I feature two MLMs demonstrating the importance of moving to new modeling approaches.

In Figure 2, I feature a FMLM. These models can be interpreted without a baseline (or reference level), which is a clear benefit in terms of straightforward interpretation. We can directly look for the values featured in Table 2 in the results.[23] In this case, we can see that not only is the MLM accurate, but it is also more precise than its interaction counterpart (smaller error bars). This is because the MLM makes better use of sparse data than high-order interaction terms via partial pooling.

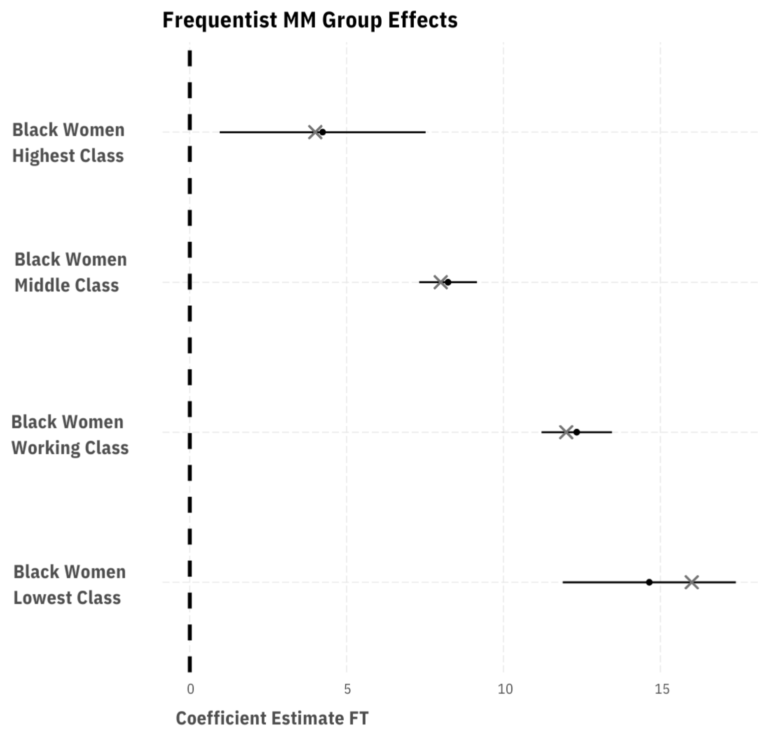

Figure 2: Frequentist Multilevel Model
The dependent variable is the synthetic feeling thermometer and these coefficients should be assessed as the group level random intercept or the change in FT score by group identity. X is the true effect shown in Table 2.

Extending this to a BMLM, the interpretation, and gain remain similar when comparing Figure 2 and Figure 3. However, since the BMLM uses an updated estimation technique, the group-level estimates now have a clearer picture of uncertainty. Figure 3 demonstrates the precision of the point estimates and the consistency of these estimates across multiple iterations. The narrow and symmetrical shape of the uncertainty distribution indicates a high degree of confidence in the central tendency of the estimates. This concentrated distribution suggests that the model produces stable and reliable results, with minimal dispersion around the point estimate. For the smallest of subsamples (groups in the highest and lowest class categories), we are

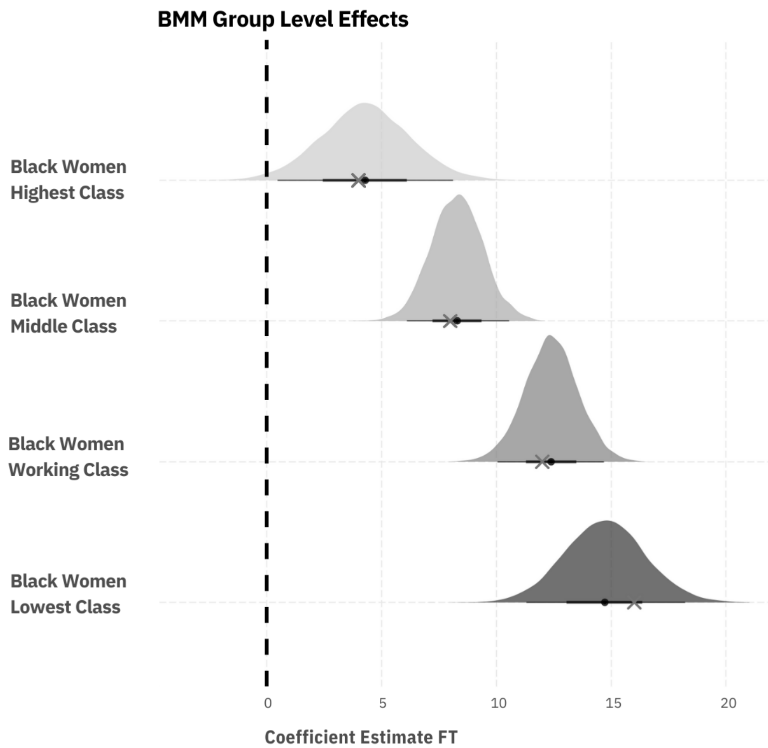

Figure 3: Bayesian Multilevel Model
The dependent variable is the synthetic feeling thermometer and these coefficients should be assessed as the group level random intercept or the change in FT score by group identity. X is the true effect shown in Table 2.

also able to see concretely that the distribution of estimates is more spread out showing that we are less certain about the point estimate for smaller sample size groups. However, all subgroups are squarely centered away from zero, showing that in either the case of the FMLM or BMLM the researcher would not come away with an understanding that a pattern did not exist when it did. Results from comparing the root mean squared error (RMSE) (which is a metric for out-of-sample performance) support the idea that the BMLM has more accurate findings than other methods with the lowest RMSE at 4.70 showing the best performance. This is compared to the RMSE for both frequentist MLM and interaction terms at 5.79.[24] Both frequentist models, FMLM and the interaction model, yield the same RMSE because they are similarly accurate. Nevertheless, it's important to note that the frequentist interaction term, as discussed in the literature, exhibits a lack of precision compared to the FMLM. The choice of model significantly influences the interpretation of statistical results, whether examining p-values in frequentist approaches or credible intervals in Bayesian methods. This choice can lead to different substantive conclusions about the nature of the relationships in question.

In MLMs, each group-level effect is estimated independently, allowing for direct comparisons between any groups without the need for a designated reference category. Interactions estimate the effect dependently on a specified reference group. The MLM approach eliminates the potentially problematic practice of privileging one group as the baseline, which is inherent in interaction models where effects are estimated relative to a reference group. MLMs provide a more equitable representation of group differences, as no one group is considered the normative baseline. This is valuable in research contexts where designating a 'standard' group, particularly White or men identifiers, could reinforce unwarranted hierarchies. Lastly, a better intersectional model needs to not only be empirically sound, but theoretically sound. Intersectionality is not just about the identity intersections, but capturing the structure of power dynamics that create these shared experiences. By capturing the higher-order structural effects, the MLM more closely resembles what is theoretically described in the literature. A researcher could specify a Bayesian interaction term model and come away with reasonable findings, but if their priority is to both theoretically and empirically understand intersectionality, they would be better suited with a MLM.

The Ideology Partisanship Paradox - ANES Example

The synthetic simulation made it evident that theoretically and empirically, a BMLM is a logical choice for intersectional modeling. To support this, I show how the BMLM performs in a practical example using the 2020 ANES. This exercise demonstrates what is at risk if a researcher chooses an intersectional modeling approach ill-suited to the data context, and also offers novel substantive findings to the partisanship literature. Partisanship research is vital to the field as this political moment is both highly polarized and calcified around identity-based lines (Sides et al. 2023). Intersectional lenses on longstanding questions of American politics stand to further illuminate the heavily entrenched dividing lines of politics.

To motivate the data context, I feature Table 3, a weighted cross-tabulation for race, gender, and class categories in the ANES. In this case, we can see how small subsamples are particularly for racial and ethnic groups in the highest and lowest class categories. While a BMLM does not remedy all data scarcity issues, by combining partial pooling and Bayesian estimation techniques, researchers can use the data more efficiently. Using the minimum of 20 groups and 10 individuals per group parameters, Black and Hispanics in the highest class category will likely be unstable regardless of modeling choice. However, using the BMLM helps in particular with lower class categories of minorities which have small, but workable sample sizes.

2020 American National Election Survey - Weighted				
	Lower Class	**Working Class**	**Middle Class**	**Upper Class**
Asian Woman	3	14	52	8
Asian Man	1	30	67	11
Black Woman	43	112	92	11
Black Man	21	78	98	9
Hispanic Woman	22	142	95	7
Hispanic Man	10	125	122	6
Multiple Race Woman	18	29	64	1
Multiple Race Man	5	32	34	5
Native American Woman	5	14	14	2
Native American Man	3	41	43	10
White Woman	79	504	913	78
White Man	61	502	952	93

Table 3: 2020 American National Election Study Subgroup Sizes

An area ripe for academic interrogation is the relationship between partisanship and ideology. Scholarship has shown that the finding that ideology translates to a person's partisan identification (PID) is driven largely by White Americans. Some scholarship leverages that group identity and the ties around race play a more central role in determining PID than ideology (Abromowitz 2010; Hajnal and Lee 2011; White and Laird 2020). There are a few theories as to why this is the case. Some find that ideology has different meanings and underpinnings for Black Americans, and many will identify as ideologically conservative while voting for Democrats (Jefferson 2020). Others, such as White and Laird (2020), demonstrate that this phenomenon is propelled by Black Americans through racialized social constraint. This suggests that, despite ideological differences, the shaping influence of race and oppression necessitates coalitions around the Democratic party. Their work also shows evidence that this divide could also be present for Latino(a)s (p. 12). There is not yet work that studies how this paradox varies by other group identities like gender and class which may also constrain and mediate ideology to predict PID. To study this phenomenon observationally, one would need to specify a four-way interaction between race, gender, class, and ideology. In this piece, I show that the ideology/partisanship paradox is indeed present for certain subsets of the Latino(a) community. However, these findings would be obscured if one were to use an interaction term because of the small sample sizes across race, gender, class, and their interaction with ideology.

White and Laird (2020) introduce the concept of racial social constraint, which helps explain why similar ideological and partisan disconnects may exist for other racial groups. This framework is particularly useful for understanding why some subgroups of Latino(a) Americans experience the ideology/partisanship paradox. Theoretically, we would expect certain subsets of the Latino(a) community to share political attitudes with Black Americans. Recent research highlights coalition links between these groups, suggesting that people of color (POC) may take cues from Black Americans' experiences of racialized social constraint, given their shared histories of oppression and racialization in the U.S. (Gershon et al. 2019; Jones and Brewer 2020; Pérez 2021).

We think this will be pronounced for certain parts of the Latino(a) community. In particular, Latinas who are native-born and/or Afro-Latinas, will have similar attitudes to Black American women because of shared experiences as WOC (Brown 2014; Mallard 2022; Matos et al. 2023). This concept dates back to some of the earliest iterations of Black feminism and the Collective (1974) where activists connected the struggles and goals of women in minority communities. Since then, research has shown that there is a coalition between these groups that is a meaningful political marker for attitudes and participation (Matos et al. 2023). Research on the political preferences of Black women and Latinas also shows that they vote similarly for Democratic candidates (Junn and Masuoka 2020). Further, Latinas are more liberal than the men in their community (Bejarano 2014).

Afro-Latinidad taps into both Latino(a) experience and blackness due to their ties with both the ethnic experience and racial experience of Latino(a) and Black communities, so there are subsets of Latinas who share direct ties to other WOC. The culmination of these factors could create a climate where the shared racial constraints within the Black community translate directly to segments of the Latina community who share similar struggles and political goals.[25] Ultimately, there are many reasons to believe the ideology/partisanship paradox will exist for Latinas because of shared intersectional political experiences and consistent patterns of common political behavior between Black women and Latinas. It also motivates broader consideration of intersectionality and the inclusion of class, as Black and Latina feminism across the spectrum of interpretive, activist, and empirical work shows that these identities are often interwoven. This is the motivation for applying an intersectional lens to the ideology/partisanship paradox, and why one might expect shared patterns between Latinas and Black women in this case.

It is also important to consider these phenomena might not exist uniformly for Latino(a)s because it is an incredibly diverse pan-ethnic group (Licea 2020). Certain segments like Latinas are steady supporters of the Democratic party and are more liberal, and some national origin groups, like Cubans, are traditionally more conservative and Republican-leaning (Bejarano 2014; Junn and Masuoka 2020). Therefore, an intersectional lens that accounts for the diversity within racial groups is the ideal way to interrogate across and within group variation. The BMLM however does not remedy all data scarcity issues in the ANES, and it is not feasible to disaggregate via national origin and other markers of diversity in the Latino(a) community. This work is meant to be an early step in interrogating this paradox for this community, and future work from intersectional scholars can disaggregate further.

As previously mentioned, to test for the paradox with an intersectional lens, one has to specify a four-way interaction with race, gender, class, and ideology to predict PID. Ideology is measured using a zero-centered seven-point scale with -3 being the most liberal, and 3 being the most conservative. Party identification (PID) is predicted on the same scale, with a zero-centered variable with strong Democrat at -3 and strong Republican at 3. Class is measured as subjective social status (lowest class, working class, middle class, and highest class), and race and gender are measured as standard factor variables.[26] Racial groups included in the full model are Black, Latino(a), Asian American, Multiracial, and White.[27] Results and the discussion are centered around two sub-groups of interest in the ANES for intersectional purposes, Black and Latino(a) of different class groups. This choice was made to center WOC in the intersectional analyses as argued by Nash in 2018, but to also investigate if there is heterogeneity among Black Americans and heterogeneity for other POC.[28]

In Figure 4, I show the results for a subset of the four-way interaction coefficients for WOC fit with an ordinary least squares (OLS) model. The reference level is White men in the lowest class category.

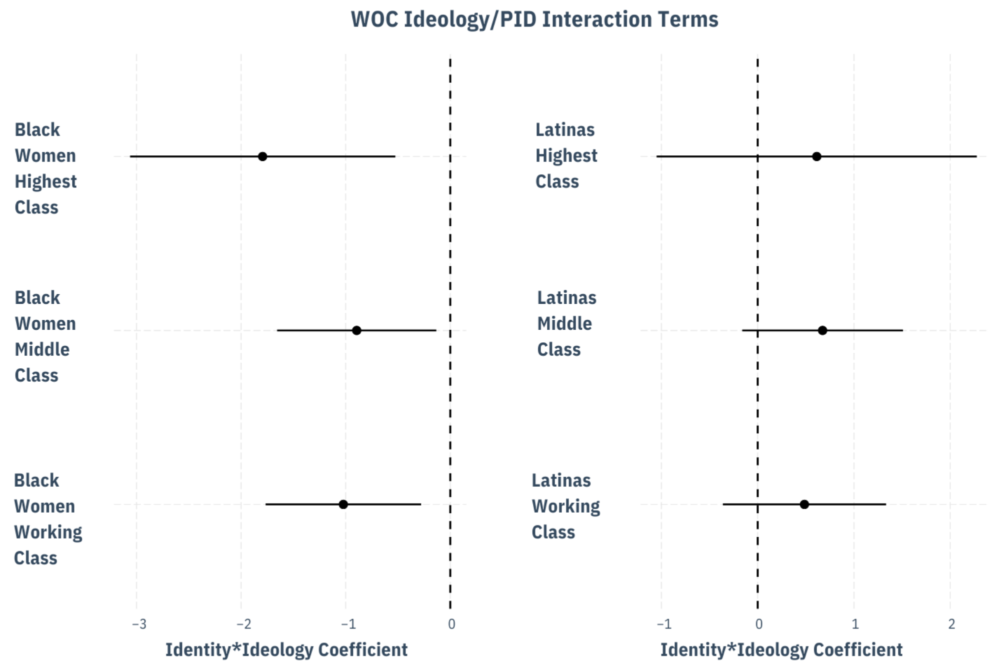

Figure 4: Four Way Interaction in the ANES Predicting PID
*The dependent variable is the seven-point partisanship scale, and the coefficients are the interaction estimate between race*gender*class*ideology.*

While we can uncover the ideology/partisan paradox for Black women (negative coefficients for the four-way interaction), it is not possible to determine the same for Latinas in this case. Each coefficient is insignificant, which we know from the synthetic example to potentially be due to small subsample sizes rather than a lack of pattern in the data. To follow the approach of the synthetic simulation, a BMLM was fit with an intersectional grouping variable, and ideology was allowed to vary by group. This allows us to determine the group-level impact of an identity and ideology has on party ID. Similar to the four-way interaction, it allows us to understand the slope of how ideology relates to partisanship when identity mediates.

An important part of the Bayesian workflow is incorporating prior knowledge into the estimation process. In this case, priors were specified by modeling the same parameters in the 2016 ANES. The results from the identity, ideology, and PID estimation were then used as the baseline priors for the model using the 2020 ANES. From there, I used my understanding of aforementioned intersectionality literature and literature of POC/WOC coalitions to support the prior that for the grouping variables (all communities of color other than White), there likely is substantial variation on how ideology impacts partisanship. I introduced stronger priors for intersectional groups by increasing the size of the group-level standard deviation for ideology. This increases the amount by which ideology can vary by group.[29]

This is an accessible approach to prior specification in a multilevel context. However, it's worthwhile to note that the benefit of an interaction term approach in the *brms* package it is easier to add specific group-level priors for the coefficients of each intersectional sub-group—for example specific priors for Black women in a certain class. To do this in a Bayesian multilevel context would require programming directly in Stan and not using *brms*, which decreases the accessibility of the method. There is a tradeoff then between choosing a more complicated coding approach with more specific group-level priors, or a simpler one with less specificity. As previously mentioned, choosing to move into the Bayesian interaction framework loses the theoretical benefit of multilevel structure and the benefits of partial pooling. Therefore it is largely up to the researcher to pick their priority; however, this piece recommends starting with the simpler coding strategy with broader informative priors as these models are still able to teach us a lot about intersectional political life. It is always important to make sure that in any Bayesian context, the priors are not wholly driving results and should operate in a balanced way. In this case, checks were done to ensure the results from the informative priors were not so starkly different from weakly informative defaults in *brms* or the baseline 2016 ANES priors that it is unreasonable. Figure 5 demonstrates the group-level coefficients for ideology fit with a

BMLM to predict PID.

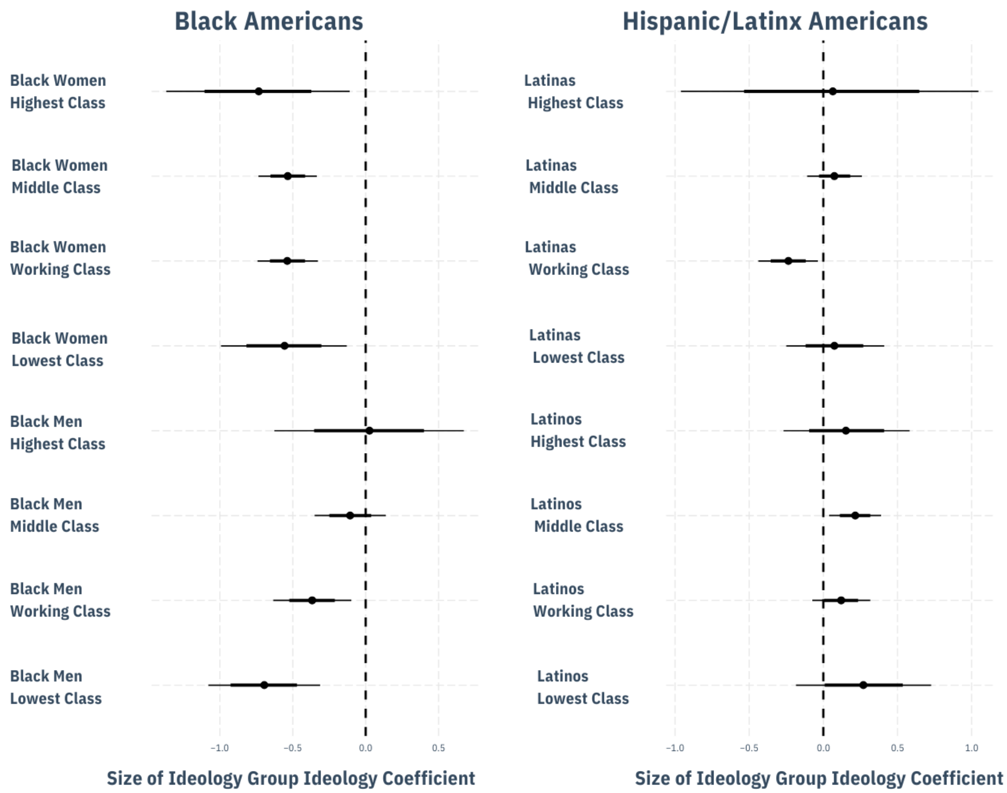

Figure 5: BMLM in the ANES Predicting PID
The dependent variable is the seven-point partisanship scale, and these coefficients should be assessed as the group level deviation from the individual level fixed effect for ideology.

Substantively, a negative coefficient means that an increase in ideology (becoming less liberal or more conservative) does not mean a corresponding increase in the intensity of PID. The most important finding is that Latinas in the working class—who in the interaction framework (Figure 4) are not significantly different from White men in the lowest class category—i.e. also demonstrate the disconnect between ideology and PID. Substantively, a researcher fitting an interaction model on its own would come away thinking that Latinas are not different from White Americans in that class group. However, as demonstrated in Figure 5 there is indeed a negative coefficient in Latinas in the working class for ideology similar to what happens for Black Americans (shown by negative group level effect). The risk of an interaction approach is that it misses a finding for a key group of interest in American politics, especially in the context of intersectional research. It is also important to point out that even though there are insignificant findings for many of the Latino(a) subgroups, it does not discredit that these groupings are important intersectional identities. Rather, we know from a long academic lineage of intersectionality that they are impactful, but their influence in a given study can be more or less pronounced based on the sample size or outcome in question. This is a direct departure from the interpretations of Block et al. (2023), which evaluate intersectionality's existence in each regression/study context. In that case, null results for all identities would suggest that intersectionality is not present. Recognizing intersectionality requires moving beyond null results in individual studies and acknowledging the broader theoretical and empirical foundations that establish its significance. I argue that fully null results for identities mean that under the specific conditions for that outcome variable and that specific sample, the measurable impact of intersectionality is not detected. It does not lay any claim overall on intersectionality's existence.[30]

For Black Americans, the substantive takeaways are similar in both interaction and BMLM frameworks, however, we can see the gains in precision (smaller overall error) in the BMLM context. An intersectional lens also adds to the substantive literature on partisanship and ideology showing evidence of only certain class groups being associated with the paradox. For Black women and men in the highest class and Black men in the middle class, estimates are not significantly different from zero, showing that this paradox may be driven by the two lower class categories within the racial group. Further, there seems to be a divide between Black men and women in the middle class, with Black women's ideology coefficient higher than for Black men, meaning the partisanship and ideology disconnect is larger. To solidify this finding, additional investigation is needed in a more data-rich environment like in Barreto et al. (2018).

To directly compare the slopes, working-class women across different racial groups in the BMLM context are featured in Figure 6. Negative coefficients from the previous plots reflect a flatter slope and positive coefficients reflect a steeper slope. This figure shows the effect of increasing ideology, and how for some groups an increase in ideology does not mean an identical increase in PID. In this case, comparing women in the working class across these racial groups shows that for White women their ideology accurately reflects their PID, and that relationship weakens for other groups. The paradox for Latinas is clear here as their slope is significantly flatter than White women in this class group, which shows that their ideology isn't as cleanly linked to their PID. Comporting with the extant literature, Black women in this case show the flattest slope which substantively means that increases in the strength of conservatism do not perfectly translate to increasing strength of Republican identity.

Figure 7 comprehensively shows predictions for partisanship and how they vary by ideology across White, Black, Latina, and Latino Americans in the working class. Key findings here show that beyond -2 on the ideology scale (liberal) we see that the slope for Latinas in the working class is significantly flatter than that of Latinos. This means that increases in ideology for Latinas lead to smaller increases in PID than for Latinos in the same class group. Latinos are closer in slope to White men and women than they are to their racial group counterparts. While their group-level credible interval was not significantly different from zero, it is possible that under further investigation, this finding would become clearer in more data-rich circumstances. The paradox for Black men and women in this case is stark, showing

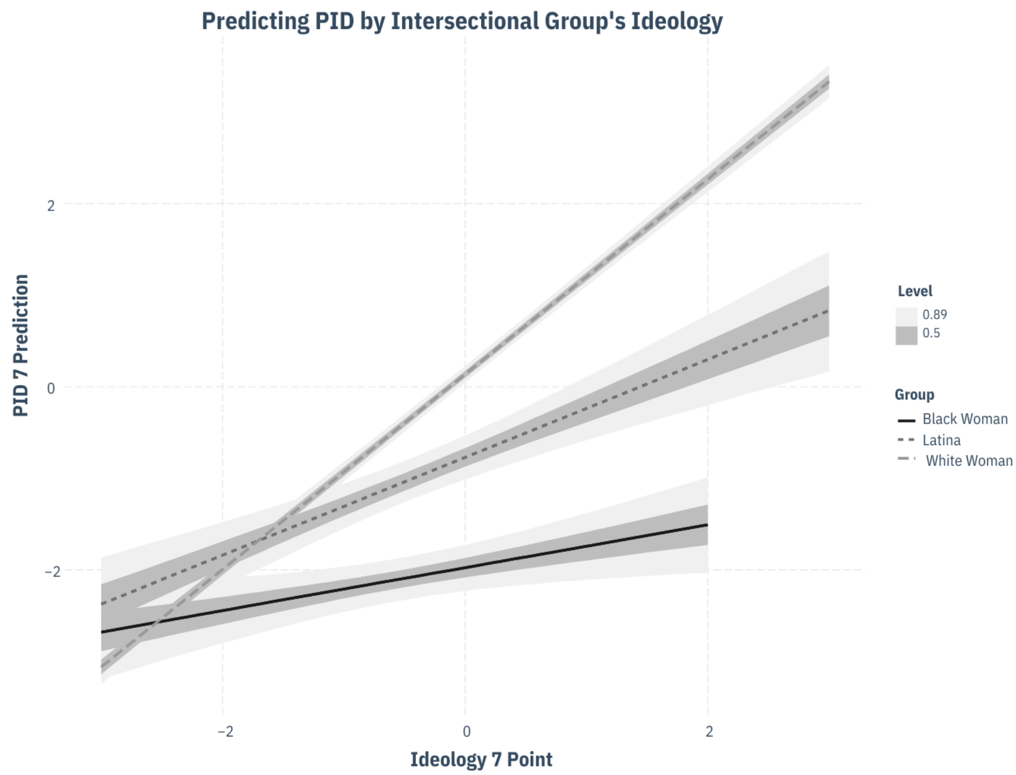

Figure 6: Working Class Women Comparison Predicting PID
The y-axis is the seven-point PID scale, and the x-axis is the five-point ideology scale. The lines are the predicted PID by race, gender, and class group from the BMLM looking at just women.

that becoming less liberal does not mean becoming a stronger Republican. Figure 7 also shows some tentative evidence that as Black women become less liberal, their PID score goes up at a slower rate than it does for Black men in this class group.

Lastly, this BMLM featured a VPC of 31% which shows that 31% of the total variation explained by this model can be attributed to our grouping variable. This demonstrates the importance of identity in understanding partisanship, as well as how ideology translates to partisanship. This work provides evidence that the racialized social constraint translates to other subsets of racial groups potentially through

Figure 7: Working Class Racial Comparison Predicting PID
The y-axis is the seven-point partisanship scale, and the x-axis is the five-point ideology scale. The lines are the predicted PID by race, gender, and class group from the BMLM, looking at all groups.

the coalitions between WOC. Furthermore, cross-validation (CV) approaches like leave-one-out CV (LOOCV) were used to compare the two best applications of both approaches, a Bayesian interaction term, and the BMLM. In the synthetic simulation, it was clear that Bayesian methods outperformed frequentist approaches, so the two best methods (the Bayesian interaction and the BMLM) were chosen for comparison, and frequentist models were omitted assuming they would not predict as well in the real world. LOOCV holds out one data point and uses it to train predictions for the rest of the data. It does this for as many observations as are done in the data. Performance is assessed by how well the model predicts the "new" data. It is computationally expensive in larger datasets but more efficient and accurate in our small sample-size environment. In addition, LOOCV is built in to be compatible with *brms*.

The LOOCV comparison reveals that the BMLM performs better than the Bayesian interaction model with the difference in expected log predicted density (ELPD) for the interaction model at -95.2 with a standard error of 17.9.[31] However, in both models, it is important to note that there were less than three observations in each case with a Pareto K diagnostic of over 0.7. This denotes that in both cases there are observations that could be skewing these results.[32] Further, the Bayesian R^2 for the BMLM was 64%, and the Bayesian interaction was 62% with no significant overlap in the estimated error, showing that the BMLM better captured the variation in this ANES example. Therefore, I pose the results of the CV as only tentative empirical evidence for BMLM performance, but combined with the theoretical benefits, it is still the ideal choice.

Overall, this section shows that many canonical findings in political science are reshaped by considering not just race, but the intersectional lens of race, gender, and class. These social identities often play mediating roles in political processes in a way that changes substantive findings. In addition, this ANES application highlights the dangers of a higher-order interaction approach in small sample-size environments. Researchers could find important intersectional patterns obscured in a challenging data environment. A BMLM is recommended as a best practice to fix this issue.

Conclusion

There is much work to be done to use and operationalize intersectionality methodologically. As Simien astutely observed in 2007, "Political scientists must construct new theories and methodological approaches that address the complex processes through which social categories shape and, in effect, determine political outcomes." This is a necessary and largely underserved part of the discipline. Research like this is particularly important for scholars of American politics who are interrogating U.S. democracy while focusing on those for whom the democratic promise is unrealized.

This method is most salient for scholars incorporating intersectionality into their work quantitatively with often unsatisfactory results. These scholars are doing work vital to political science at a sociopolitical moment of heightened unrest in the United States, so they deserve methods tantamount to their task. This model is not meant to be a final solution or silver bullet (as Hancock (2007b) would say), but a step towards better methods for an important charge of the discipline.

In this piece, I have argued for BMLMs to account for the heterogeneities produced by race, gender, and class intersections. The BMLM was able to show the intersectionality effects more precisely than current methods. To be sure, for both cases of intersectionality and multidimensionality, if the dataset has large sample sizes for each subgroup the gained precision from the BMLM is no better than the interaction term in frequentist regression (Gelman and Hill 2006). However, intersectional quantitative scholars consistently face data limitations, and large datasets are rare. In addition, there will always be a need in American politics to explore historical data like in the ANES, which will always have these issues. There may also be contexts in which a Bayesian interaction term model is suited for a researcher who wants predictive accuracy gains and more granular group priors than available in *brms*. However, the BMLM will still estimate helpful terms such as the VPC/ICC, and maintain the theoretical and empirical benefits of partial pooling and MLM hierarchical structure.

Applying this model to the ideology/partisanship paradox illuminated the experience of working-class Latinas that would be obscured if an interaction term approach were used. This shows that there are real substantive implications to political science knowledge production if the correct modeling choice is not used. In sum, the goal of this article is to show this tool has utility in the study of political behavior and attitudes in the U.S., as intersectionality and multidimensional identities undergird the lived experience that shapes these areas of study. It is a methodologically rigorous, and theoretically sound modeling tactic that builds intersectional and multidimensional context into research practices.

Notes

1. Political science, especially comparative politics researchers, readily use MLMs to account for the effects of geographical units like countries.

2. I used BMLM to refer to Bayesian multilevel models and MLM to refer to multilevel models in either frequentist or Bayesian settings.

3. Data for the main analysis is available on the ANES website. Synthetic data is available through replication files in the supplemental materials.

4. Scholars and activists include but are not limited to: Gloria Anzaldua, Combahee River Collective, Patricia Hill Collins, Anna Julia Cooper, Ida B. Wells, and Maria Stewart.

5. Jordan-Zachery (2007) cautions treating intersectionality as a method to pull out of a researcher's bag of tricks because it's not just a method but a lived experience. This work encourages scholars to meaningfully engage with the intersectionality literature to understand these nuances and employ the method carefully.

6. There are three main ways to "pool," or combine the data: complete, no pooling, and partial pooling. Partial pooling is the happy medium. "Multilevel modeling partially pools the group-level parameters at their mean to split the difference between the two extremes of no pooling and complete pooling (Gelman and Hill 2006)."

7. They will not be considered further as an intersectional method.

8. Parsimony means simplicity and means the models are simple to interpret. It is a desired aspect of a statistical model.

9. I do not detail this method further because of these limitations. However, if a researcher is in a large sample size environment and does not need to compare groups directly, it can be a reasonable choice to study identity.

10. The words effect and estimate will be used interchangeably to talk about MLM coefficients as it is standard to refer to coefficients as random or mixed effect. Neither is being used in a causal sense as this is all observational data.

11. The grouping variable is a factor which, for example, has separate levels for Black women, Black men, Latinas, and Latinos, Asian men, Asian women, and White Women and White men in four different class categories (lowest, working, middle, highest). Multidimensional groups are specified in *one* grouping variable to maintain their interwoven effects.

12. For an introduction to Bayesian methods see Gelman et al. (2013), Clark (2018) and Johnson et al. (2022).

13. LKJ priors of 1 are the default and are uniform correlation priors.

14. See MCMC Pack, RCPP, Bayes M.

15. Also see rstanarm.

16. In Evans et al. (2018)'s body of work, they use a similar concept, the proportional change in variance (PCV), to study group-level variation and its impact.

17. I recommend using the *performance* R package to compute VPC/ICC for *brms* models.

18. If additional formal geographic elements (such as states) are relevant, they should be treated as another grouping variable, consistent with the traditional use of MLMs in Political Science.

19. Focusing in on these two groups was done to ease interpretation

20. These effects were built to be relatively large and exaggerated to show apparent differences in this synthetic context. This does not mean that, substantively, there will be an FT where Black and White women are diametrically opposite. This phenomenon still occurs when the effects are not as exaggerated.

21. *brms* was designed to have weakly default informative priors which sway the results as little as possible and account for the bias that can occur with naive uninformative priors as previously cited (Bürkner 2017).

22. This is also supported by a Monte Carlo simulation detailed in Appendix part A. See Appendix A for information about average standard error, rates of "insignificance", and variation of coefficients in both models.

23. These values are the random intercept effects because no random coefficients were specified in this model. These values can be interpreted as the estimated group-level impact on the DV.

24. Out of sample performance or OOS performance allows a researcher to see which model would best predict new data.

25. In an ideal world I would be able to disaggregate to assess the influence of Afro-Latinas specifically, but this dataset does not allow it. In further work, assessing the role of which subsets of Latinas are most likely to show this phenomenon is vital. Afro-Latinas are included in this piece to recognize the multitude of theoretical reasons why Latinas might behave like Black women in the literature.

26. While there is noted bias of measuring class in this way as many overstate their presence in the middle class, this measurement is simple enough to reduce class down to four different categories.

27. I fit a model to include all racial groups, but the results featured focus on Black and Latino(a) people because there is the most reason to believe there would be shared patterns.

28. Ideally, Asian American and multiracial women will be included in future work on this issue, but it is beyond the scope of this methodology-oriented piece to fully disaggregate every intersectional angle.

29. The group-level parameters come from a multivariate normal distribution with a mean of zero, and these parameters can be thought of as adjustments to the population-level ideology coefficient.

30. In some outcomes, we might expect that Latino(a)s in different racial and class categories operate similarly, in the same way, we might expect Black individuals across these groups to operate similarly. If the BMLM suggests that race, gender, or class alone accounts for most of the variation (i.e., null results for multiple intersectional identities), then a useful approach is to treat the BMLM as a robustness check. Researchers could then pool gender or class categories together in the main analysis to focus on race.

31. ELPD is the conventional metric used to assess the predictive accuracy of a model in LOOCV. Larger ELPD corresponds to more accurate models, so the best model in the comparison will have an ELPD difference of zero, and the worst model will be the deviation from that value.

32. This comes from small group sub-sample sizes.

References

Abromowitz, A. (2010, June). *New Directions in American Political Parties*. Routledge. Google-Books-ID: ZNuOAgAAQBAJ.

Ackerly, B. and J. True (2010, September). Back to the future: Feminist theory, activism, and doing feminist research in an age of globalization. *Women's Studies International Forum 33*(5), 464–472.

Alexander-Floyd, N. G. (2012). Disappearing Acts: Reclaiming Intersectionality in the Social Sciences in a Post-Black Feminist Era. *Feminist Formations 24*(1), 1–25.

Barreto, M. A., L. Frasure-Yokley, E. D. Vargas, and J. Wong (2018). Best practices in collecting online data with Asian, Black, Latino, and White respondents: evidence from the 2016 Collaborative Multiracial Post-election Survey. *Politics, Groups, and Identities 5503*, 1–10. Publisher: Taylor & Francis.

Bejarano, C. E. (2014, March). Latino Gender and Generation Gaps in Political Ideology. *Politics & Gender 10*(01), 62–88.

Bell, A., D. Holman, and K. Jones (2019, April). Using Shrinkage in Multilevel Models to Understand Intersectionality. *Methodology 15*(2), 88–96. Publisher: Hogrefe Publishing.

Berelson, B., P. Lazarsfeld, and W. McPhee (1954). *Voting: A study of opinion formation in a presidential campaign.* University of Chicago Press.

Bittner, A. and E. Goodyear-Grant (2017). Sex isn't Gender: Reforming Concepts and Measurements in the Study of Public Opinion. *Political Behavior 39*(4), 1019–1041. Publisher: Springer US.

Block, R., M. Golder, and S. N. Golder (2023, January). Evaluating Claims of Intersectionality. *The Journal of Politics*, 000–000. Publisher: The University of Chicago Press.

Brambor, T., W. R. Clark, and M. Golder (2006). Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis 14*(1), 63–82. Publisher: [Oxford University Press, Society for Political Methodology].

Brown, N. E. (2012). Negotiating the Insider/Outsider Status: Black Feminist Ethnography and Legislative Studies. *Journal of Feminist Scholarship 3*, 17.

Brown, N. E. (2014, October). Political Participation of Women of Color: An Intersectional Analysis. *Journal of Women, Politics & Policy 35*(4), 315–348.

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software 80*(Plummer 2013).

Cho, S., K. W. Crenshaw, and L. McCall (2013, June). Toward a Field of Intersectionality Studies: Theory, Applications, and Praxis. *Signs: Journal of Women in Culture and Society 38*(4), 785–810.

Clark, M. (2018). Bayesian Basics: A conceptual introduction with application in R and Stan. *E-Pub.*, 57.

Collective, C. R. (1974). The Combahee River Collective Statement.

Converse, P. E., A. Campbell, W. E. Miller, and D. E. Stokes (1961). Stability and Change in 1960: A Reinstating Election. *The American Political Science Review 55*(2).

Crenshaw, K. W. (1989). Demarginalizing the Intersection of Race and Sex : A Black Feminist Critique of Antidiscrimination Doctrine. *Chicago Legal*, 139.

Crenshaw, K. W. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *STANFORD LAW REVIEW 43*, 60.

Davis, K. (2008). Intersectionality as buzzword. *Feminist Theory 9*, 67–85.

Dawson, M. (1994). *Behind the Mule: Race and Class in African American Politics.* Princeton University Press.

Evans, C. R., L. N. Borrell, A. Bell, D. Holman, S. V. Subramanian, and G. Leckie (2024, June). Clarifications on the intersectional MAIHDA approach: A conceptual guide and response to Wilkes and Karimi (2024). *Social Science & Medicine 350*, 116898.

Evans, C. R., G. Leckie, and J. Merlo (2020, January). Multilevel versus single-level regression for the analysis of multilevel information: The case of quantitative intersectional analysis. *Social Science & Medicine 245*, 112499.

Evans, C. R., D. R. Williams, J.-p. Onnela, and S. V. Subramanian (2018). A multilevel approach to modeling health inequalities at the intersection of multiple social identities. *Social Science & Medicine 203*(November 2017), 64–73. Publisher: Elsevier.

Forman-Rabinovici, A. and H. Mandel (2023, June). The Prevalence and Implications of Gender Blindness in Quantitative Political Science Research. *Politics & Gender 19*(2), 482–506.

Frasure-Yokley, L. (2018). Choosing the Velvet Glove : Women Voters , Ambivalent Sexism , and Vote Choice in 2016. *Journal of Race, Ethnicity, and Politics*, 1–23.

García Bedolla, L. (2007). Intersections of Inequality: Understanding Marginalization and Privilege in the Post-Civil Rights Era. *Politics & Gender 3*(02), 232–248.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013). *Bayesian Data Analysis*. CRC press.

Gelman, A. and J. Hill (2006). Data Analysis Using Regression and Multilevel Modeling. *Cambridge Univeristy Press*.

Gershon, S. A., C. Montoya, C. Bejarano, N. Brown, S. Allen, C. Montoya, C. Bejarano, and N. Brown (2019). Intersectional linked fate and political representation. *Politics, Groups, and Identities 7*(3), 642–653. Publisher: Taylor & Francis.

Hajnal, Z. L. and T. Lee (2011, February). Why Americans Don't Join the Party: Race, Immigration, and the Failure (of Political Parties) to Engage the Electorate. In *Why Americans Don't Join the Party*. Princeton University Press.

Hancock, A.-M. (2007a). Intersectionality as a Normative and Empirical Paradigm. *Politics & Gender 3*(2), p.248.

Hancock, A.-M. (2007b). When multiplication doesn't equal quick addition: Examining intersectionality as a research paradigm. *Perspectives on Politics 5*(1), 63–79.

Hancock, A.-M. (2019). Empirical Intersectionality : A Tale of Two Approaches A Tale of Two Approaches. In *The Palgrave Handbook of Intersectionality in Public Policy*, pp. pp. 95–132. Palgrave Macmillan.

Hawkesworth, M. (2003, November). Congressional Enactments of Race–Gender: Toward a Theory of Raced–Gendered Institutions. *American Political Science Review 97*(4), 529–550.

Hesse-Biber, S. and P. Leavy (2007). *Feminist Research Practice*. 2455 Teller

Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc.

Jardina, A. (2019, February). *White Identity Politics* (1st edition edition ed.). Cambridge, United Kingdom: Cambridge University Press.

Jefferson, H. (2020, July). The Curious Case of Black Conservatives: Construct Validity and the 7-point Liberal-Conservative Scale.

Johnson, A., M. Ott, and M. Dogucu (2022). *Bayes Rules! An Introduction to Applied Bayesian Modeling*. Routledge.

Johnston, R., K. Jones, and D. Manley (2018, February). Age, sex, qualifications and voting at recent English general elections: an alternative exploratory approach. *Electoral Studies 51*, 24–37.

Jones, K., R. Johnston, and D. Manley (2016, January). Uncovering interactions in multivariate contingency tables: a multi-level modelling exploratory approach. *Methodological Innovations 9*, 2059799116672874. Publisher: SAGE Publications Ltd.

Jones, P. E. and P. R. Brewer (2020). Elite cues and public polarization on transgender rights. *Politics, Groups, and Identities 8*(1), 71–85. Publisher: Taylor & Francis.

Jordan-Zachery, J. S. (2007). Am I a Black Woman or a Woman Who Is Black ? A Few Thoughts on the Meaning of Intersectionality. *Politics & Gender 3*(2), 254–263.

Junn, J. (2007). Square Pegs and Round Holes: Challenges of Fitting Individual-

Level Analysis to a Theory of Politicized Context of Gender. *Politics and Gender 3*(1), 124–134.

Junn, J. (2017, April). The Trump majority: white womanhood and the making of female voters in the U.S. *Politics, Groups, and Identities 5*(2), 343–352.

Junn, J. and N. Masuoka (2008). Asian American identity: Shared racial status and political context. *Perspectives on Politics 6*(4), 729–740.

Junn, J. and N. Masuoka (2020, December). The Gender Gap Is a Race Gap: Women Voters in US Presidential Elections. *Perspectives on Politics 18*(4), 1135–1145.

Kinder, D. R., L. M. Sanders, and L. M. Sanders (1996, July). *Divided by Color: Racial Politics and Democratic Ideals.* University of Chicago Press.

Kruschke, J. K., H. Aguinis, and H. Joo (2012). The Time Has Come: Bayesian Methods for Data Analysis in the Organizational Sciences.

Licea, R. (2020). Is "Latino" useful? Diversity, commonality, and politics. *Sociology Compass 14*(12), e12836. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/soc4.12836.

Lüdecke, D. (2024). Intraclass Correlation Coefficient (ICC) — icc.

Maas, C. J. M. and J. J. Hox (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology 1.*

Mallard, I. G. (2022, August). *The Politics of Being Afro-Latino/Latina: Ethnicity, Colorism, and Political Representation in Washington, D.C.* Rowman & Littlefield. Google-Books-ID: qeh9EAAAQBAJ.

Mason, L. (2018, April). *Uncivil Agreement: How Politics Became Our Identity.* University of Chicago Press.

Matos, Y., S. Greene, and K. Sanbonmatsu (2023, May). The politics of "women of color": a group identity worth investigating. *Politics, Groups, and Identities 11*(3), 549–570.

McCall, L. (2005, March). The Complexity of Intersectionality. *Signs: Journal of Women in Culture and Society 30*(3), 1771–1800.

McNeish, D. (2016, September). On Using Bayesian Methods to Address Small Sample Problems. *Structural Equation Modeling: A Multidisciplinary Journal 23*(5), 750–773.

McNeish, D. and L. Stapleton (2016, June). The Effect of Small Sample Size on Two Level Model Estimates: A Review and Illustration. *Educational Psychology Review 28*.

Merlo, J. (2018, April). Multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA) within an intersectional framework. *Social Science & Medicine 203*, 74–80.

Montoya, C. (2020, July). Intersectionality and Voting Rights. *Political Science & Politics 53*(3), 484 – 489.

Nash, J. (2018). *Black Feminism Reimagined: After Intersectionality.* Duke University Press.

Ojeda, C. and C. M. Slaughter (2019). Intersectionality, depression, and voter turnout. *Journal of Health Politics, Policy and Law 44*(3), 480–504.

Paccagnella, O. (2011, January). Sample Size and Accuracy of Estimates in Multilevel Models: New Simulation Results. *Methodology 7*(3), 111–120.

Park, J. and H.-T. Yu (2018, October). Recommendations on the Sample Sizes for Multilevel Latent Class Models. *Educational and Psychological Measurement 78*(5), 737–761.

Peugh, J. L. (2010, February). A practical guide to multilevel modeling. *Journal of School Psychology 48*(1), 85–112.

Pérez, E. O. (2021, August). *Diversity's Child: People of Color and the Politics of Identity*. University of Chicago Press. Google-Books-ID: EEU6EAAAQBAJ.

Ramazanoglu, C. and J. Holland (2007). Feminist Methodology.

Rupp, A. A., D. K. Dey, and B. D. Zumbo (2004). To Bayes or Not to Bayes, From Whether to When: Applications of Bayesian Methodology to Modeling. *Structural Equation Modeling 11*(3), 424–451. Place: US Publisher: Lawrence Erlbaum.

Sides, J., C. Tausanovitch, and L. Vavreck (2023, September). *The Bitter End: The 2020 Presidential Campaign and the Challenge to American Democracy*. Princeton University Press. Google-Books-ID: J2m1EAAAQBAJ.

Sides, J., M. Tesler, and L. Vavreck (2018, October). *Identity Crisis: The 2016 Presidential Campaign and the Battle for the Meaning of America*. Princeton University Press.

Simien, E. M. (2007). Doing Intersectionality Research: From Conceptual Issues to Practical Examples. *Politics and Gender 3*(2), 264–271.

Smid, S. C., D. McNeish, M. Miočević, and R. Van De Schoot (2020, January). Bayesian Versus Frequentist Estimation for Structural Equation Models in Small

Sample Contexts: A Systematic Review. *Structural Equation Modeling: A Multidisciplinary Journal 27*(1), 131–161.

Smooth, W. (2011). Standing for women? which women? the substantive representation of women's interests and the research imperative of intersectionality. *Politics and Gender 7*(3), 436–441.

Spry, A. D. (2018). Identity in American Politics: A Multidimensional Approach to Study and Measurement. *Columbia Academic Commons*, 268.

Stauffer, K. E. and D. Z. O'Brien (2018, June). Quantitative Methods and Feminist Political Science. In *Oxford Research Encyclopedia of Politics*. Oxford University Press.

Stegmueller, D. (2013). How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches. *American Journal of Political Science 57*(3), 748–761. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12001.

Tesler, M. (2016, April). Post-Racial or Most-Racial?: Race and Politics in the Obama Era. In *Post-Racial or Most-Racial?* University of Chicago Press.

Wasserstein, R. L. and N. A. Lazar (2016, April). The ASA Statement on p -Values: Context, Process, and Purpose. *The American Statistician 70*(2), 129–133.

Wasserstein, R. L., A. L. Schirm, and N. A. Lazar (2019, March). Moving to a World Beyond "p \textless 0.05". *The American Statistician 73*(sup1), 1–19.

Weldon, S. L. (2006, June). The Structure of Intersectionality: A Comparative Politics of Gender. *Politics & Gender 2*(02).

White, I. K. and C. N. Laird (2020, February). Steadfast Democrats: How Social Forces Shape Black Political Behavior. In *Steadfast Democrats*. Princeton University Press.

Wilkes, R. and A. Karimi (2023, December). What does the MAIHDA method explain? *Social Science & Medicine*, 116495.

Zyphur, M. J. and F. L. Oswald (2015, February). Bayesian Estimation and Inference: A User's Guide. *Journal of Management 41*(2), 390–420. Publisher: SAGE Publications Inc.

A Appendix

A.1 Monte Carlo Simluation

The following plots from this section are from a Monte Carlo simulation that compared data generated and model performance in the main text over 100 iterations. I randomly generated the synthetic data 100 times with 3,000 individuals in the sample, representatively split up among identities. I then fit a frequentist interaction term model and the *brms* BMLM equivalent. The following plots compare the performance of the two models over 100 simulated runs.

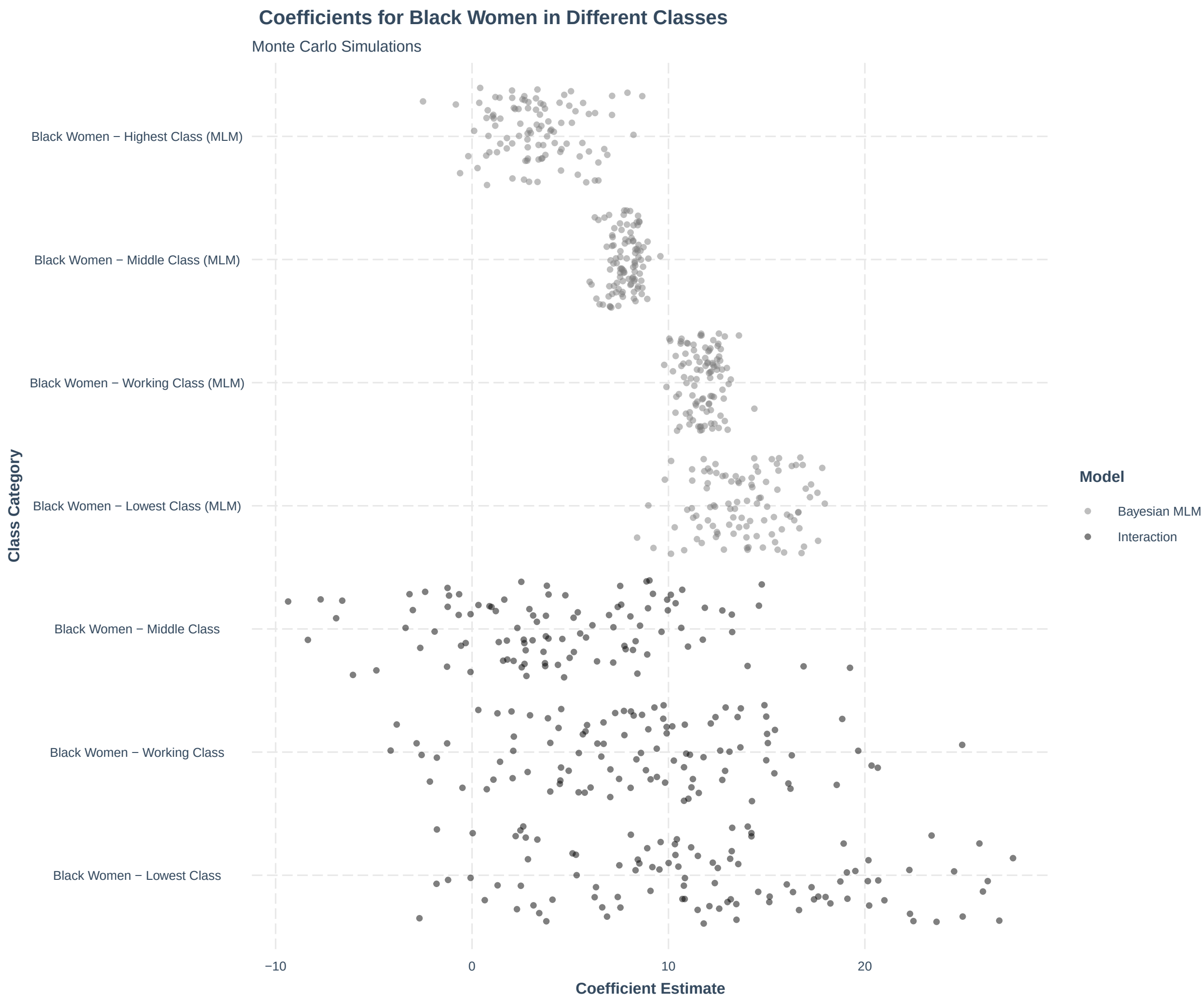

Figure 8: This figure shows the raw coefficient variation between the BMLM and the interaction.

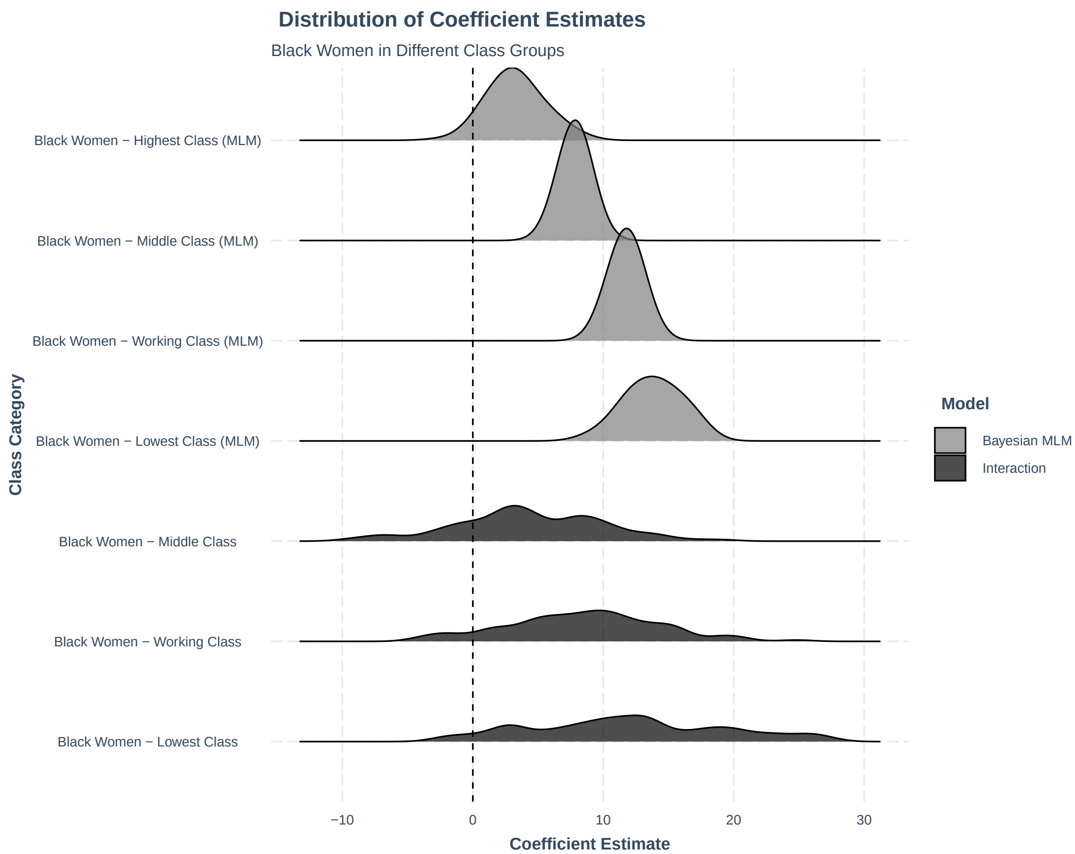

Figure 9: This figure shows the distribution of the raw coefficient variation between the BMLM and the interaction.

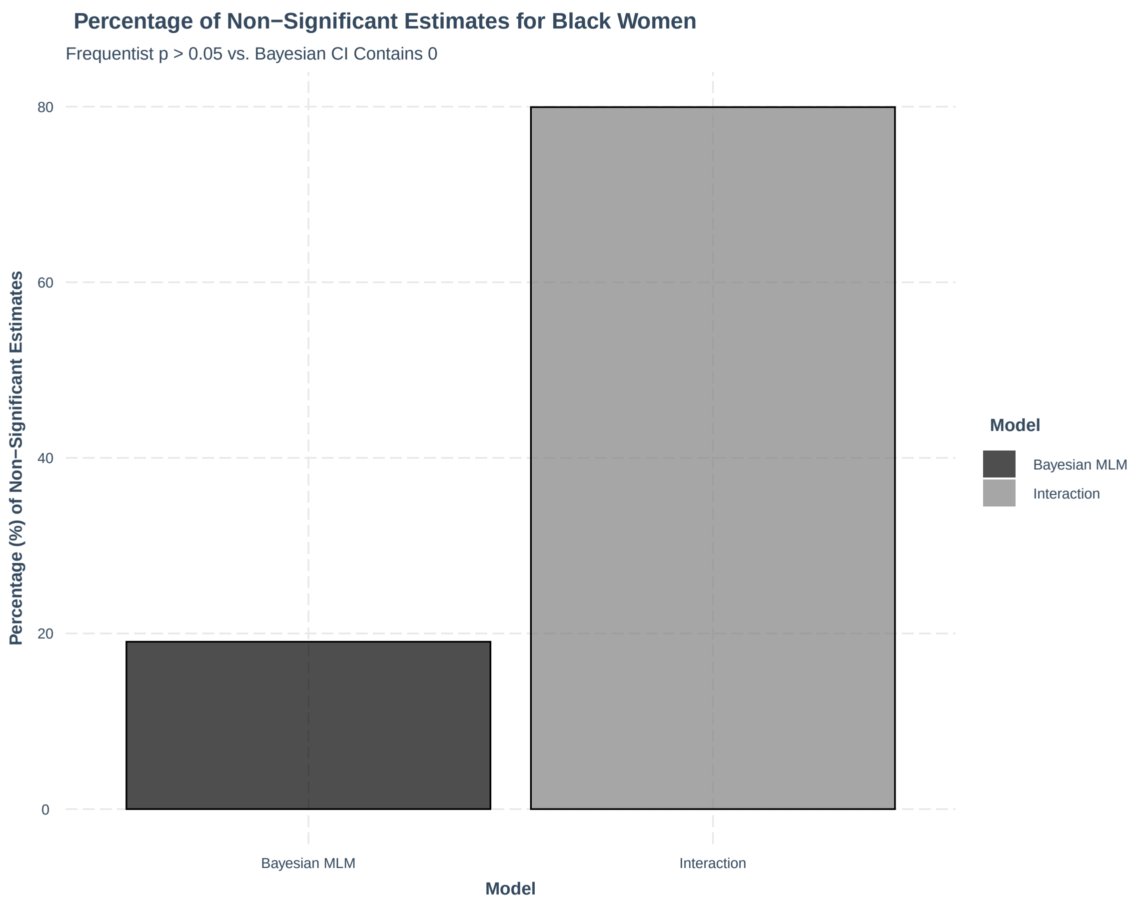

Figure 10: This figure shows the percentage of insignificant p-values in the Bayesian context and the interaction term context for Black women in the different categories. P-values are not traditionally used in Bayesian stats, so I calculated the simplest equivalent which is whether or not the credible interval contained 0 or not.

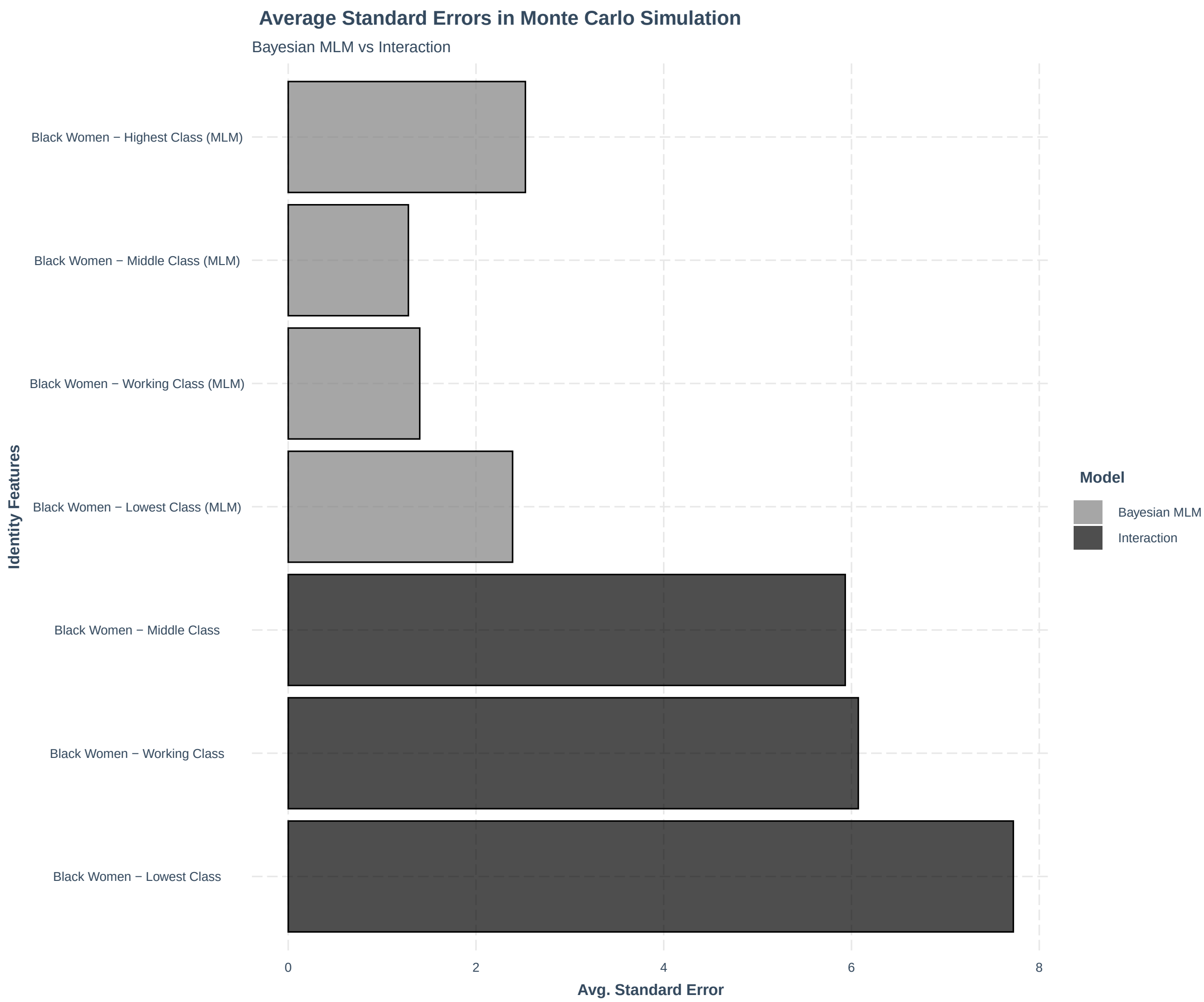

Figure 11: This figure shows the average standard error amount in the Bayesian context and the interaction term context for Black women in the different categories. I calculated the closest equivalent to standard error in the Bayesian context. Since Bayesian models do not produce standard errors in the same way as frequentist models, I estimated the standard error by taking the difference between the 97.5th and 2.5th percentiles of the posterior distribution (which represents the 95% credible interval) and dividing it by (2×1.96). This formula assumes a roughly normal posterior distribution, where 1.96 corresponds to the standard deviation multiplier for a normal distribution's 95% confidence interval.

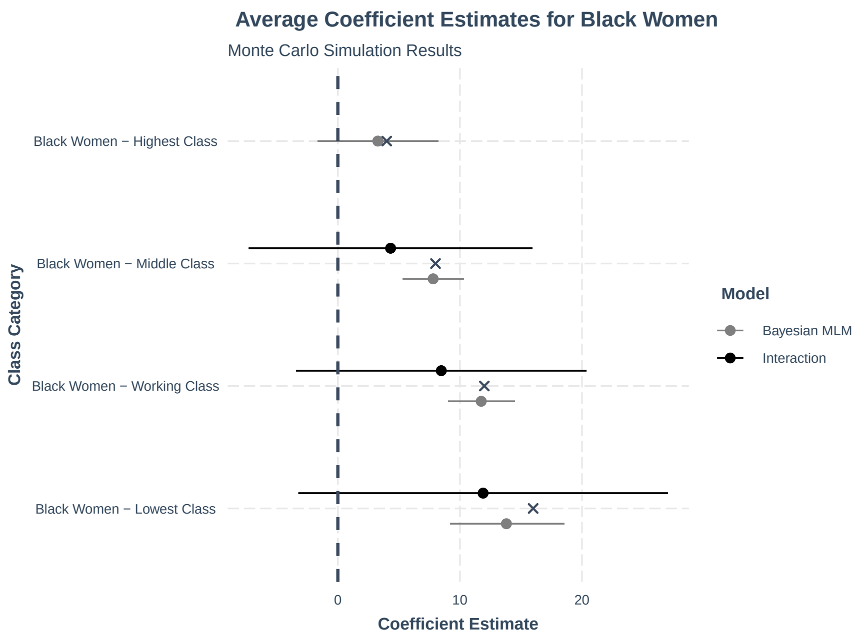

Figure 12: This figure shows the dot and whisker of the average coefficient and standard error estimate for both models compared to ground truth (X). We see that both models are fairly accurate, but the BMLM is again more precise.

A.2 RMSE and Bayes RMSE for Synthetic

Method	RMSE
Indicator Frequentist	7.03
Interaction Frequentist	5.79
Multilevel Frequentist	5.79
Bayesian Multilevel Model	4.71

A.3 LOOCV for ANES

Method	ELPD Difference	Standard Error
BMLM	0.0	0.0
Bayes Interaction	-95.2	17.9

A.4 Priors from 2016 ANES

```
library(brms)
prior_mod <- brm(pid_zero ~ ideo_zero + (1 + ideo_zero | intersectional),
                        data = priors_2016,
                        control = list(adapt_delta = 0.95,
                                       max_treedepth = 10),
                        warmup = 2000,
                        iter = 5000,
                        # prior = priors_ideo,
                        seed = 1234,
                        chains = 4,
                        cores = 4)
#Take results from summary to build priors

summary(prior_mod)
```

Priors from 2016 ANES Part 2

Adust the standard deviation (sd) parameter for the coefficient ideology, and add LKJ prior. If someone wants to run the defaults, skip this step and run the model on the next page. I allowed the sd parameter by slightly more than a whole unit of the PID scale the capture the paradox for communities of color.

```
library(brms)
priors_ideo <- c(
  prior (normal(0.77, 0.07),  b,  coef = ideology_zero),
  prior(normal(-0.53 , 0.16), Intercept),
  prior(lkj_corr_cholesky(.5), grou = intersectional, class = cor),
  prior(normal(1.2, 0.07), sd, coef = ideology_zero, intersectional),
  prior(normal(0.82, 0.15), sd, coef = Intercept, intersectional),
  prior(normal(1.48, 0.02), sigma))
```

The ANES Model

The structure of the model remains the same for both the synthetic and ANES BMLM, the only difference is there is a random slope for ideology denoted by 1+ ideology_zero where the synthetic only has 1.

```
MM_stan_pid_mod <- brm(pid7_zero ~ ideology_zero +
                                    (1 + ideology_zero | intersectional),
                      data = anes_2020_x,
                      control = list(adapt_delta = 0.92,
                                     max_treedepth = 10),
                      warmup = 2000,
                      iter = 5000,
                      prior = priors_ideo,
                      seed = 1234,
                      chains = 4,
                      cores = 4)
# Fit Bayes R^2 Metric

bayes_R2(MM_stan_pid_mod)
fit1 <- loo(MM_stan_pid_mod)
#use the loo_compare() function to compare the performance of
#the analogous interaction model using output from fit1
```