UNIVERSITY OF HELSINKI
FACULTY OF SOCIAL SCIENCES

Unity in Diversity: Multilevel Analysis of Religious Tolerance in Indonesia

Ghaida Nasria Azzahra
Master's Programme in Contemporary Societies
Faculty of Social Sciences
University of Helsinki
Master's Thesis
May 2024

Abstract

Faculty: Faculty of Social Sciences

Degree programme: Master's Programme in Contemporary Societies

Study track: Social Data Science

Author: Ghaida Nasria Azzahra

Title: Unity in Diversity: a Multilevel Analysis of Religious Tolerance in Indonesia

Level: Master's thesis

Month and year: May 2024

Number of pages: ii+45+1

Keywords: multilevel, tolerance, diversity, intergroup, religion, Indonesia

Supervisor or supervisors: Kimmo Vehkalahti & Maria Valaste

Where deposited: Helsinki University Library

Additional information: -

Abstract: Tolerance is the pillar that ensures the safety and stability of a nation, especially for a populous and multicultural one such as Indonesia. Societal characteristics, such as diversity, may play a significant role in shaping people's attitudes. Diversity may improve intergroup relations by encouraging positive interaction between members of different groups. On the other hand, it may have a negative influence by aggravating social tension. This thesis aims to examine how intergroup contact and diversity associate with religious tolerance in Indonesia. Due to the hierarchical nature of the problem, multilevel regression analysis was utilized for this study. This thesis also aims to investigate if the application of multilevel regression can provide any additional value to the analysis. Classical single-level linear regression was also employed as a comparison for the multilevel model and to complement the analysis. This thesis utilized a combination of national survey data and official data from government bodies.

The intraclass correlation coefficient (ICC) calculated from the model revealed that around 23% of the variation in tolerance score was due to the grouping structure in the data, which can be considered to be quite high. Religious diversity was found to be statistically significant and was able to explain a substantial amount of the variation in province level. Intergroup interaction was also found to be positively and significantly associated with tolerance score. The multilevel model was also found to correct the underestimation of standard error in single-level regression due to treating a group-level predictor as an individual-level predictor. Thus, the application of multilevel analysis provided additional value to the analysis by revealing province-level inequalities in the data and minimizing the risk of spurious significant results.

Contents

1 Introduction

1.1 Background

Like many other nations born out of colonialism, Indonesia had to carve out its identity based on the borders drawn by its former colonizer (Pisani, 2014). Aceh, the northwestern part of the nation, is around 5,000 kilometers away from Papua in the east, which is just about as far as London is to Tehran (Pisani, 2014). The people of Aceh and Papua are just as different as the people of London and Tehran: they speak different languages, belong to different religions, and observe different traditions. Between these two regions, there are more than 360 ethnic groups who are practicing different major religions as well as a multitude of indigenous beliefs (Kersten, 2017; Pisani, 2014). For a country as plural as Indonesia, tolerance is the glue that secures the stability and unity of the nation.

Tolerance is essential for the wellbeing of a society as it ensures mutual respect and safety. It enables different groups with different ways of life to live side by side in harmony and stimulates trust between them. However, fostering a tolerant attitude within a multicultural society of over a quarter billion people is not a simple task (Marshall, 2018). Intolerance issues are more pronounced for religious differences, which remains to be a "significant marker of difference" in Indonesia while ethnic identity has become less important in recent times (Sidi, 2020, p.11).

Tolerant attitude does not only rely on personal characteristics of individuals, but also on the societal context around them. Different environments might encourage different kinds of social interactions, and thus shaping the attitudes of the people in a way that cannot be explained by their individual characteristics alone (Mulder & Krahn, 2008). For instance, the effect of diversity on intergroup attitude is important to understand. Examining this question may help reveal possible sources of conflicts and ways to mitigate it. This topic has been gaining considerable attention due to the complexity behind it, and multiple studies has been testing its effect on various intergroup outcomes with mixed results (Hewstone & Schmid, 2014; Lundåsen, 2023; Putnam, 2007; Van Assche et al., 2018). Hewstone (2015) suggested that the inconsistent result of previous studies may be due to the differences in history or policy of each country. While there is a growing amount of literature on this topic, the effect of diversity on religious tolerance has not been examined in the Indonesian context.

The effect of diversity is closely related to intergroup contact. Allport's (1979) intergroup contact theory proposed that interacting with a member of a different group can improve intergroup attitudes. Higher diversity may increase the chance for different groups to meet and interact with each other. Pettigrew and Hewstone (2017) warned that failing to account for intergroup contact when examining diversity may lead to a fallacious conclusion. As this would require researchers to consider both individual and group features, they further recommend a multilevel approach.

Therefore, this thesis investigates how intergroup contact and diversity associate with religious tolerance in Indonesia. This study also attempts to answer the previous calls by Pettigrew and Hewstone (2017) and Yusuf et al. (2019) to utilize a multilevel analysis to approach this question. Besides allowing for simultaneous variable analysis in different hierarchical levels, multilevel analysis also has another benefit from a statistical perspective. Multilevel analysis can be a solution for the violation of the assumption of independence that often happen in hierarchical data as observations in the same group tend to be more similar to each other than to the observations in other groups. Uncorrected, this could cause underestimation of standard error and lead to spurious significant results. Additionally, intolerant incidents were not happening uniformly across the country. According to Marshall (2018), around 80 percent of recent intolerant incidents happened in just three provinces: Aceh, West Java, and South Sulawesi. This gives the indication that the level of intolerant attitude varies between regions.

1.2 Research Question

With rising globalism, contact between different groups becomes more frequent and unavoidable. Consequently, understanding the dynamics of tolerance between different fractions is increasingly important to maintain a safe and well-functioning society. Identifying individual and societal level predictors of religious tolerance have important policy implications. Such identification may serve to assist the government in effectively building policies and constructing environments that would foster higher religious tolerance among the people. Using theories of intergroup relations, this thesis aims to examine the determinants of religious tolerance in Indonesia.

Specifically, the research questions for this study are:

1. How does religious diversity and intergroup contact associate with religious tolerance in Indonesia?
2. Does the utilization of multilevel regression in this context provide additional value to the analysis?

1.3 Thesis Structure

Chapter 2 examines relevant social theories on tolerance and intergroup relation as the theoretical framework. Moreover, this chapter also imparts some informational background about religious dynamics in Indonesia to provide the required context. Chapter 3 presents the statistical properties of multilevel linear regression analysis, which will be used as the main analysis tool in this study. This chapter also provides the justification and suitability of the chosen analysis method. Chapter 4 describes the materials used for this study and specifies the variable construction method and analysis procedure. Chapter 5 presents the descriptive findings and interprets the results from the regression models. Finally, Chapter 6 concludes this thesis with a discussion of the findings and an evaluation of the whole research process. This chapter also assesses the limitations of this study and possible opportunities for future research.

2 Tolerance in Society

The characteristics of a society may play a significant role in shaping people's attitudes by "providing the context within which group interaction takes place and attitudes and beliefs are shaped" (Mulder & Krahn, 2008, p.424). Agnew's (1992) general strain theory posited that community characteristics (such as density, inequality, or diversity) may generate a strain that stimulates anti-social tendencies. Empirically, Milligan et al. (2014) confirmed that tolerance level is generally lower in countries with higher inequality and lower economic activity. Similarly, Muzayanah et al. (2020) found that individuals living in higher density areas have lower levels of bridging trust and tolerance, while Mulder and Krahn (2008) reported that residents of larger cities tend to be less supportive of diversities.

To provide the theoretical background for this study, this chapter presents the literature review related to intergroup relations and how societal characteristics, such as diversity, may influence them. The discourse in this chapter is divided into two sections, and it starts with a literature review on how intergroup contact and diversity may play a role in shaping the attitude of an individual. Then, the last section lays out the religious dynamics in Indonesia to delineate the necessary context for the study.

2.1 Intergroup Contact and Diversity

The intergroup contact theory posits that interaction between members of different groups in a conducive environment can effectively reduce prejudice (Allport, 1979). This theory has grown to be one of the most influential theories in social psychology and has been supported by numerous empirical findings (Christ et al., 2014; Mavridis, 2015; Sanjaya, 2022). A meta analysis of 515 studies by Pettigrew and Tropp (2006) revealed that 94% of these studies reported a negative association of intergroup contact and prejudice. Pettigrew and Tropp (2006) suggested that it may be linked to the mere-exposure effect, which postulated that higher exposure to an object can significantly increase liking to that object. The familiarization process caused by intergroup interaction can also act as a force to shift the boundaries between the "us" and "them" dichotomy (Mulder & Krahn, 2008).

Various empirical evidence have demonstrated the versatility of the intergroup contact theory. Although this theory is more commonly utilized in relation to racial and ethnic encounter researches, a meta-analysis by Pettigrew and Tropp (2006) confirmed the wider applicability of this theory, claiming that it can also be extended to various kinds of target groups, such as religious outgroups or homosexuals. For instance, some previous studies have applied this theory to examine interreligious contacts (Kanas et al., 2015; Yusuf et al., 2019). Furthermore, the positive impact of intergroup contact has been shown to be universal across age, gender, or nations (Pettigrew et al., 2011).

Intergroup contact may also have a much wider impact that affects more than the individuals or groups that are directly involved in the interaction. A previous study by Christ et al. (2014) showed that living in an area with a high rate of positive intergroup contact may have a positive change to intergroup attitude, even for people who do not experience intergroup contact themselves. Similarly, Pettigrew et al. (2011) noted how intergroup contact can encourage individuals to be more accepting towards outgroups, including the ones that have never interacted with them. Furthermore, simply being friends with a person who is close to an outgroup member may also help to lessen prejudice, although the effect is not as strong as having direct contact with outgroup members (Pettigrew et al., 2011).

Certainly, not all interactions are positive in nature; negative interactions can happen especially in a high stress or competitive environment (Pettigrew et al., 2011). To create an optimal climate for positive intergroup contact, Allport (1979) postulated that four conditions are needed: equal status between the groups, common goals, intergroup cooperation, and institutional support. However, Pettigrew and Tropp (2006) found that these conditions are not essential; positive results from intergroup contact can still be observed even when the conditions are not fulfilled. Furthermore, they showed that based on various empirical evidence, the impact of intergroup contact still tends to be overwhelmingly positive. They proposed three reasons for this phenomenon: first, positive interactions are much more frequent than negative ones; second, the effect of negative contacts are far smaller in voluntary interactions; and third, people who had a lot of intergroup contact experiences, both positive and negative, still tend be as tolerant as people who only had positive interactions.

While the evidence for positive effect of intergroup contact is quite consistent and strong, the evidence for the effect of diversity is less clear. Previous literature found conflicting results on how diversity affects intergroup relations (Hewstone and Schmid (2014), Putnam (2007), and Van Assche et al. (2018)). Allport (1979) leaned to a more pessimistic view; he postulated that diversity tends to incite clashes of interests, which in turn would intensify conflict and prejudice. He further argued that conflicts are more likely to happen when the proportion of minority groups is higher. This risk is even more significant when the minority group is segregated from the larger population, such as when they are living in a concentrated area. Furthermore, diversity may only succeed in encouraging casual intergroup interactions, which can aggravate prejudice rather than reduce it. Mulder and Krahn (2008) suggested that higher diversity can also increase misinformation or generate more anxiety towards the members of the outgroups. Lastly, Putnam (2007) empirically found that diversity may reduce social trust, increase perception of outgroup threat, and cause withdrawal from social interaction.

On the other hand, Van Assche et al. (2018) suggested that previous studies found mixed results because there are two competing forces that work in opposite directions: positive intergroup interaction and perceived outgroup threat. They found no association between diversity and prejudice in their research, citing that the two competing forces canceled each other out. This dual effect of diversity was also discussed by Pettigrew and Hewstone (2017), who warned that inspecting the effect of diversity without taking into account the role of intergroup contact or segregation may lead to an erroneous conclusion, a mistake they called "the single factor fallacy".

Empirically, Hewstone and Schmid (2014) showed that negative association between diversity and intergroup attitudes found in previous studies were often caused by omitting intergroup contact in the analysis. They explained that increased opportunity for contact does not always result in actual increased intergroup contact, and thus it is crucial to include a measure of actual contact in the analysis. In their study, they found that negative effects of diversity disappeared after accounting for intergroup contact. Along the same lines, Laurence et al. (2019) observed that increasing diversity has a negative impact in segregated communities, but positive impact in integrated communities. Individuals who live in segregated communities have less opportunity to interact with outgroup members, and thus less likely to mediate the feeling of perceived threat (Laurence et al., 2019).

2.2 Religious Dynamics in Indonesia

Historically, trade and immigration has been the main pathway for the spread of various religions in the Indonesian archipelago. According to Juergensmeyer and Roof (2012), Hinduism was recorded to be one of the first foreign religions that influenced the region, estimated to first reach the islands at around 1 CE. Around the 6th century CE, Buddhism arrived from the Indian subcontinent and coexisted with Hinduism (Juergensmeyer & Roof, 2012). Meanwhile, some traces of Islamic influence in the region were found from as early as the seventh century, although Islam had only started to become prominent centuries later. By the 15th century, Islam had spread as far east as the Moluccas (Aritonang & Steenbrink, 2008), and around 20 Muslim kingdoms had been established in Sumatra and Java (Juergensmeyer & Roof, 2012). By the end of the 18th century, Bali has become the last bastion of Hinduism in the archipelago (Aritonang & Steenbrink, 2008) and remained to be so until present times.

Catholicism was brought to the Indonesian archipelago through the Portuguese conquest in the 16th century (Aritonang & Steenbrink, 2008; Goh, 2005). However, according to Goh (2005) the Portuguese were not so keen in pushing their missionary agenda, while Aritonang and Steenbrink (2008) noted that conflict between the Portuguese and the locals often ended up further consolidating Islamic influence, as was the case in Moluccas and Sumatra. Furthermore, when the Dutch came to the islands in the 17th century, they were hostile to the spread of other denominational groups, especially Catholicism (Goh, 2005). Consequently, there was not much evangelistic activity in the archipelago before the 19th century. Nevertheless, Christianity eventually became the second largest religion in modern Indonesia after it gained popularity among many minority ethnic groups such as the Bataks in Sumatra and Minahasans in Sulawesi.

After gaining its independence in 1945, Indonesia was faced with a daunting task of uniting the vast and fragmented archipelago. As Islam was the majority religion among the people, there were many parties who advocated to make Islam as the foundation of the state (Seo, 2012). However, the national leaders realized that proclaiming an Islamic state would cause discontent among the minority groups or even secession, especially in the eastern part of the nation that is Christian majority (Hefner, 2013). As a compromise, the founding fathers kept the religious aspect of the state ideology but removed the reference to a specific religion (Seo, 2012). Hefner (2013) noted that although the constitution of the newly independent nation guaranteed the freedom of religion, there were different interpretations of what was considered as a religion. Indigenous beliefs were not regarded as religions, but instead “cultural belief systems” (Marshall, 2018). In 1965, the government issued a decree that legally recognizes six religions: Islam, Christianity (Protestantism), Catholicism, Hinduism, Buddhism, and Confucianism. Other religions were still allowed to be practiced although they did not receive the same level of support and protection from the state (Fealy & Ricci, 2019; Marshall, 2018).

In the end of 1965, an alleged coup took place which led up to a massacre of around half a million communists and suspected communists (Vickers, 2013). As atheism was associated with communism, the newly installed New Order government then required all citizens to declare their faith as an effort to deter the return of communism (Seo, 2012). To escape prosecution, many adherents of unrecognized religions converted to one the recognized religions, especially Christianity (Seo, 2012). In Java, non-practicing Muslims were disproportionately targeted for the killings, which led to mass conversions to Christianity or Hinduism (Hefner, 2013).

Suryadinata (2005) recorded how the anti-communist purge also engendered a sense of suspicion towards Chinese Indonesians due to their historical connection to the People’s Republic of China. The New Order government started to enforce the assimilation policy towards Chinese Indonesians to diminish their Chinese identity, which eventually led to the ban and derecognition of Confucianism. As a result, adherents of Confucianism had to convert to other religions, mainly Buddhism and Christianity. The ban on Confucianism was later revoked after the New Order government fell, more than three decades after its prohibition (Seo, 2012). By that time, the number of its adherents has dropped very significantly (Suryadinata, 2005).

The fall of the New Order at the end of the 20th century became the catalyst of increasing religious expressions that were previously suppressed (Yusuf et al., 2019). During the New Order era, communal conflicts were relatively contained due to the strict media censorship and military control (Vickers, 2013). Vickers (2013) described how as state control faltered, religious fundamentalism grew stronger and the media were used by all parties to inflame the situation. As is common during economic and political unrest, the years after were marked by a string of violent ethno-religious conflicts, from the Ambon sectarian conflict in 1999-2002 to the Bali bombings in 2002 and 2005 (Vickers, 2013). As economic and political climate grew more stable, extreme violent outbreaks subsided. However, there are still some recent intolerant incidents, one of which was the prosecution of a popular

Christian-Chinese politician by utilizing the blasphemy law. This incident has invited numerous commentaries and studies (Mietzner & Muhtadi, 2018).

In 2017, there was a positive development towards the acceptance of indigenous religions. The Indonesian Constitutional Court deemed that requiring people to choose between the six recognized religions is unconstitutional, and that indigenous beliefs should be recognized as religions (Marshall, 2018). Even so, Marshall (2018) noted that past and present discrimination have led adherents of indigenous religions to report their religions as one of the recognized religions. Consequently, their true number is hard to estimate.

3 Multilevel Regression Analysis

Multilevel regression analysis can be employed to examine the relationship between characteristics at individual level and at societal level (Hox, 2010). Multilevel data have a hierarchical structure, which can take on various forms, such as: pupils nested in schools, patients nested in hospitals, or employees nested in offices. In this study, the data will be structured in the form of individuals within administrative areas (provinces). This choice of nesting was done on the considerations that policies are made on administrative unit levels and for data availability reasons.

Multilevel research questions can also be approached by assigning community-level characteristics to each individual in the sample then running an ordinary (single-level) multiple regression. However, according to Hox (2010) this approach has two shortcomings. First, the statistical tests would treat the community-level values as independent information from the individual samples rather than the groups. This means that the sample size used for the calculation (the number of individual samples) would be much larger than the actual sample size (the number of groups), which can lead to spurious significant results. Second, the results from this approach must be interpreted very carefully as they can easily lead to analytical fallacies. Inferences at community level made from analyzing the data at individual level, or vice versa, can be very misleading. To illustrate this point, Hox (2010) used an example provided by Robinson (1950) found a high correlation level between the percentage of black population and illiteracy level in nine regions (0.95). It may seem reasonable to conclude that black people are more likely to be illiterate, but the correlation at individual level was found to be much lower (0.2), which suggests that individual-level correlation cannot be inferred from the community-level correlation (Robinson, 1950).

This chapter provides the information on the structure and properties of multilevel regression analysis. It starts by introducing classical single-level linear regression as the basis of the substance. Building up from the materials in the first subchapter, the next subchapter covers the specifics of multilevel linear regression models.

Unless stated otherwise, the materials presented in the first subchapter are cited from Applied Regression Analysis by Draper and Smith (1998), while the ones presented in the second subchapter are cited from Multilevel Analysis: Techniques and Applications by Hox (2010). As this thesis compiles information from different sources that utilize different notations, some notations have been altered from their original sources to maintain clarity and consistency.

3.1 Linear Regression

Regression analysis is a tool to estimate the relationships between one or more predictor variables and one continuous response variable (Kleinbaum et al., 1998). When the relationship between these variables is assumed to be linear, a linear regression model can be applied. Mathematically, a linear regression model with K $(k = 1, 2, ..., K)$ predictors and n $(i = 1, 2, ..., n)$ observations can be defined as:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_K X_{Ki} + \epsilon_i \tag{1}$$

where Y_i is the response variable; $\beta_0, \beta_1, ..., \beta_K$ are the parameters that need to be estimated; $X_{1i}, X_{2i}, ..., X_{Ki}$ are the predictor variables; and ϵ_i is the error term. The regression coefficients are commonly estimated by using the least squares approach, although it can also be done through other methods. The least squares method determines the best fitting regression line by minimizing the sum of the squares of the estimated error (residual). The equation for the regression line is:

$$\hat{Y}_i = b_0 + b_1 X_{1i} + b_2 X_{2i} ... + b_K X_{Ki} \tag{2}$$

where $\hat{Y}_i$ is the predicted/fitted value of Y_i and $b_0, b_1, ..., b_K$ are the estimates of $\beta_0, \beta_1, ..., \beta_K$. As per Equation 2, b_k can be interpreted as the difference of $\hat{Y}$ when comparing two subjects with one point of difference in X_k, assuming that all of the other predictors are kept constant (Gelman et al., 2020). Meanwhile, b_0 (intercept) is the value of $\hat{Y}$ when $X_1 = X_2 = ... = X_K = 0$.

3.1.1 Significance Testing

F-test is commonly utilized to evaluate if there is enough evidence to suggest that a regression model has significant predictive power. The null hypothesis is that all of the coefficients in the model, except for the intercept, equal to zero ($\beta_1 = \beta_2 = ... = \beta_K = 0$). In other words, none of the predictors significantly explains the response variable. Meanwhile, t-test is used to evaluate if there is a statistically significant linear relationship between a specific predictor variable and the response variable. The null hypothesis of this test is that the true value of β_k equals to 0.

3.1.2 R-Squared and Adjusted R-Squared

R-squared (coefficient of determination) is a statistical measure that is often used to assess the performance of a regression model. It can be defined as the proportion of total variation in the response variable that can be explained by the regression model. R-squared is calculated by comparing the sum of squares due to the regression, which is the total squared difference between the predicted values to the average of Y, and the sum of squares about the mean value, which is the total squared difference between each observation and the average of Y. Mathematically, it can be defined by the equation:

$$R^2 = \frac{\sum_i (\hat{Y}_i - \overline{Y})^2}{\sum_i (Y_i - \overline{Y})^2} \tag{3}$$

where $\hat{Y}_i$ is the fitted value of Y_i and $\overline{Y}$ is the average value of Y.

R-squared value normally ranges from 0 to 1 although it can also yield negative scores in some cases, for instance if the intercept or the slope are constrained (Chicco et al., 2021). The value of 1 means that all of the variations in the data are explained by the regression model, which can happen when Y_i equals $\hat{Y}_i$. However, in practice it is impossible to attain this value due to the existence of error. Meanwhile, R-squared can be 0 when the regression line is horizontal at $\overline{Y}$ ($\hat{Y} = \overline{Y}$). A value of 0 means that the predictors are not useful in explaining the variation in the data.

The value of R-squared cannot decrease with the addition of a predictor variable in the model, even if the additional predictor does not contribute meaningfully to the model. To account for this phenomenon, adjusted R-squared can be calculated. It is defined as:

$$R_a^2 = 1 - (1 - R^2)\frac{n-1}{n-p} \tag{4}$$

where n is the sample size and p is the total number of parameters in the model. From Equation 4, it can be observed that adjusted R-squared value is penalized by each addition of p.

3.2 Multilevel Linear Regression

Multilevel regression can be applied to analyze hierarchical datasets with one response variable that is measured at the lowest level and multiple predictor variables that can be measured at all existing levels. For example, in the case of a dataset structured as individuals grouped within provinces, the response variable should be measured at individual level, while the predictor variables may be measured at both individual and province level. It is also possible to have more than two levels in the data.

In multilevel regression, both the intercept and slope can be made to vary between groups. To illustrate this point, consider a two-level dataset collected from J $(j = 1, 2, ..., J)$ groups with n_j $(i = 1, 2, ..., n_j)$ individuals in each group. A different regression model with K individual-level predictors can be built for each group. A varying-intercept model, which have the same slope for each group but varying intercept (Gelman & Hill, 2007), can be defined as:

$$Y_{ij} = \beta_{0j} + \beta_1 X_{1ij} + \beta_2 X_{2ij} + ... + \beta_K X_{Kij} + \epsilon_{ij} \tag{5}$$

where Y_{ij} is the response variable for individual i in group j; β_{0j} is the intercept for group j; $\beta_1, \beta_2, ..., \beta_K$ are the parameters that are fixed for all groups; $X_{1ij}, X_{2ij}, ..., X_{Kij}$ are the individual-level predictor variables for each individuals; and ϵ_{ij} is the error term.

The variation of the intercept is explained by introducing Q $(q = 1, 2, ..., Q)$ group-level predictor variables as follows:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_{1j} + \gamma_{02} Z_{2j} + ... + \gamma_{0Q} Z_{Qj} + u_{0j} \tag{6}$$

where β_{0j} is the individual-level intercept for group j; γ_{00} is the group-level intercept; $Z_{1j}, Z_{2j}, ..., Z_{Qj}$ are the values of group level predictors in group j; $\gamma_{01}, \gamma_{02}, ..., \gamma_{0Q}$ are the parameters for the group-level predictors; and u_{0j} is the error term. By combining Equation 5 and Equation 6, it can be seen that if γ_{0q} is positive, the average of Y_{ij} would be higher in groups where Z_{qj} is higher.

Meanwhile, a varying-slope model has a constant intercept but varying slopes for each group (Gelman & Hill, 2007). Contrary to Equation 5, the j subscripts would instead be attached to the β_k parameters to indicate the group they belong to. Thus, the formula would be:

$$Y_{ij} = \beta_0 + \beta_{1j} X_{1ij} + \beta_{2j} X_{2ij} + ... + \beta_{Kj} X_{Kij} + \epsilon_{ij}. \tag{7}$$

Similarly as before, the group-level predictors are used to explain the variation of the individual-level parameters $(\beta_{1j}, \beta_{2j}, ..., \beta_{Kj})$. Thus, a similar model as Equation 6 would be constructed for each of the β_{kj} parameter instead of the intercept, for example an equation for β_{1j} would be:

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_{1j} + \gamma_{12} Z_{2j} + ... + \gamma_{1Q} Z_{Qj} + u_{1j} \tag{8}$$

where β_{1j} is the individual-level β_1 for group j; γ_{10} is the group-level intercept; $Z_{1j}, Z_{2j}, ..., Z_{Qj}$ are the values of group level predictors in group j; $\gamma_{11}, \gamma_{12}, ..., \gamma_{1Q}$ are the parameters for the group-level predictors; and u_{1j} is the error term at group level. Consequently, when γ_{1q} is positive, it can be concluded that β_{1j} would be higher in groups with higher value of Z_{qj}.

It is also possible to have varying-intercept, varying-slope models where both of the intercept and slope are made to vary in each group (Gelman & Hill, 2007). Then, the regression model becomes:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + ... + \beta_{Kj}X_{Kij} + \epsilon_{ij}. \tag{9}$$

In this case, both Equation 6 and Equation 8 are applicable to explain the variation of the parameters in the model.

3.2.1 Intraclass Correlation Coefficient (ICC)

In a hierarchical dataset, observations in the same group tend to be more similar to each other than to observations in other groups, which violates the assumption of independence. Intraclass correlation coefficient (ρ) quantifies the amount of dependence in the data. It can be estimated by building a varying intercept model with no predictors (intercept-only model). When there are no predictors at the lowest level, Equation 5 can be reduced to:

$$Y_{ij} = \beta_{0j} + \epsilon_{ij}. \tag{10}$$

Meanwhile, with no predictors at group level, Equation 6 also reduces to:

$$\beta_{0j} = \gamma_{00} + u_{0j}. \tag{11}$$

Then, Equation 11 can be subsituted into Equation 10 to form:

$$Y_{ij} = \gamma_{00} + u_{0j} + \epsilon_{ij}. \tag{12}$$

Equation 12 does not explain any variance in Y, instead, it only divides the variance into two separate components: u_{0j} which is the variance of the lowest-level error (σ_e^2) and ϵ_{ij} which is the variance of the group-level error (σ_{u0}^2). ICC, which is the proportion of group-level variance to the total variance, can be calculated from these terms. Alternatively, ICC can also be defined as the proportion of variance that is explained by the grouping structure.

Mathematically, the ICC can be expressed as:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2}. \tag{13}$$

3.2.2 R-Squared Measures for Multilevel Models

Compared to the classical single-level regression analysis, calculating the proportion of variance explained by the predictors is more complicated for multilevel regression analysis. This is due to the fact that in multilevel regression, unexplained variance exists at several levels simultaneously. One approach to handle this problem is by examining the changes of the variances by level through a sequence of models that are increasing in complexity. The reduction of variance after the addition of new predictors can be interpreted as the amount of variance at each level that was explained by the new predictors. Hence, a statistical measure analogous to the R-squared for each level can be calculated by measuring the proportion of this reduction to the total error variance.

For example, consider two models constructed from a two-level hierarchical dataset, one is an intercept-only model and the other has some predictors added (comparison model). Then, the proportion of variance explained at the first level would be:

$$R_1^2 = \frac{\sigma_{e|b}^2 - \sigma_{e|m}^2}{\sigma_{e|b}^2} \tag{14}$$

where $\sigma_{e|b}^2$ is the first-level residual variance for the intercept-only model and $\sigma_{e|m}^2$ is the first-level residual variance for the comparison model. This R_1^2 value could be interpreted as the amount of variance explained at the lowest level by the predictors added in the comparison model.

Meanwhile, the proportion of variance explained at second level can be calculated with:

$$R_2^2 = \frac{\sigma_{u0|b}^2 - \sigma_{u0|m}^2}{\sigma_{u0|b}^2} \tag{15}$$

where $\sigma_{u0|b}^2$ is the second-level residual variance for the intercept-only model and $\sigma_{u0|m}^2$ is the second-level residual variance for the comparison model. Similar with the previous one, this R_2^2 value could be interpreted as the amount of variance explained at the second level by the predictors added in the comparison model.

It should be noted that even if the comparison model only contains individual-level predictors, those predictors may also explain the variance at the second level (group level). This can be explained by compositional difference of the groups. This should not be interpreted as a real contextual effect. Meanwhile, adding a group-level predictor does not impact variation at individual level.

The limitation of these R-squared calculations is that in multilevel analysis, it is possible that an addition of a predictor would decrease the variance at one level but increase the variance at the other. Consequently, the R-squared value calculated from Equation 14 or Equation 15 can be negative. A second limitation is that for varying-slope models, the estimated variance can be affected by the scale of the predictors.

4 Data, Methods, and Analysis

This chapter introduces the data source and analysis strategy used in this study. The first section lays out the materials used in this study and the sources. This section also provides a brief overview of the survey data. The next subchapter, Methods and Analysis, is divided into two sections. The first section details the specifics of the variable construction, while the next section specifies how the analysis will be conducted. The last section also presents the justifications for the analysis strategy chosen for the study.

4.1 Data Source

This research utilizes the September 2021 National Socioeconomic Survey (Susenas September 2021, from here on will be referred to as 'Susenas'), a nationally-representative survey held by Statistics Indonesia. Susenas is a large multipurpose household survey routinely held twice a year, in March and September. This study will only use the September 2021 cycle of this survey. The data from this survey is representative at national and provincial level. The target sample in this survey is 75,000 households in all 34 provinces of Indonesia. The response rate for this survey is around 99.6%, thus the data contains observations from 74,701 households.

A specialized sociocultural module, which will be used in this research, is added to the September cycle of this survey in 2021. This module includes questions about religious tolerance, which will be used as the response variable in this research. This module is only added to the survey once every three years, so the September 2021 dataset was the most recent data available at the time of the study. This dataset will provide the individual-level response and predictor variables for this study.

Additionally, provincial-level data collected from Statistics Indonesia and the Indonesian Ministry of Religious Affairs will also be utilized as the province-level predictor variables. To maintain consistency, all of the provincial-level data were chosen from the year 2021. When multiple data points are available within a year, the data point closest to September was chosen.

4.1.1 Overview of the Survey Data

Susenas is a household survey conducted through face-to-face interview with a member of the sampled household. The dataset acquired from this survey can be divided into two types based on the observation level. The first type contains information from all household members. Thus, this dataset contains 272,088 observations which are the total number of household members from the 74,701 sampled households. This dataset contains personal information about each household member, such as age, gender, and education level.

Meanwhile, the second type contains information at household level. Therefore, this dataset only contains 74,701 observations, one from each sampled household. This dataset contains information that relates to the household characteristics, such as household expenditure and type of place of residence. Additionally, this dataset also contains information from the specialized sociocultural module, which includes the questions about religious tolerance and intergroup interaction.

4.2 Methods and Analysis

4.2.1 Variable Construction

The predictor variables for this study are chosen based on the theoretical and empirical evidence discussed in Chapter 2. To account for the community effect, there will be two levels of predictor variables: individual and provincial levels. The response variable and the individual-level predictors will be obtained from Susenas September 2021 data. Meanwhile, the province-level predictors will be collected or constructed based on the data obtained from Statistics Indonesia, the Indonesian Ministry of Home Affairs, and the Indonesian Ministry of Religious Affairs.

Table 1: Variables Used in the Study

Variable	**Source**
Response variable	
Religious tolerance score	Susenas September 2021 (Statistics Indonesia)
Individual-level predictors	
Intergroup contact	Susenas September 2021 (Statistics Indonesia)
Educational attainment	Susenas September 2021 (Statistics Indonesia)
Household expenditure	Susenas September 2021 (Statistics Indonesia)
Place of residence	Susenas September 2021 (Statistics Indonesia)
Age	Susenas September 2021 (Statistics Indonesia)
Gender	Susenas September 2021 (Statistics Indonesia)
Province-level predictors	
Religious diversity	The Indonesian Ministry of Religious Affairs
Gini ratio	Statistics Indonesia
Population density	Statistics Indonesia

A numerical response variable representing an individual's religious tolerance will be constructed from the following questions in the survey data:

Table 2: Response Variable's Source Questions

Number	**Question**
1811B	How would you feel if someone with a different faith from you holds a religious activity around your home?
1812B	How would you feel if one of your household members befriended a person with a different faith from you?
1813B	How would you feel if you are led by someone with a different faith from you?
1815	How would you feel if people who have different faiths from you build a house of worship in your community?

Source: Statistics Indonesia, 2021

For each of these questions, the respondent may answer that they disapprove, slightly disapprove, approve, or strongly approve, coded as numbers from 1 to 4. These questions were only asked to one person in each sampled household. The interviewee can either be the household head or their spouse, however, it cannot exactly be known. Consequently, connecting the response variable to the personal information (age, gender, and educational level) of the interviewee becomes a challenge.

A previous study by Wijaksono (2023) which also utilized Susenas data handled this problem by assuming that the interviewees are all the household heads. However, the spouse is just as likely to be the interviewee, if not more. Thus, to minimize the uncertainty level in the data, this study will utilize a variable that can be used to identify the household member who gave out the most information (the main interviewee) at the start of the interview. As surveys are typically started and completed by the same person, it will be assumed that the interviewees at the start of the survey are also the ones who answered the questions on religious tolerance, unless if they were not the household head or their spouse. In such cases, then the household head would be assumed to have answered the religious tolerance questions.

As was specified in Table 1, this study will utilize six individual-level predictors and three province-level predictors. The individual-level predictors are intergroup contact, educational level, gender, age, household expenditure, and type of place of residence. The main interest of this study among the individual-level predictors is intergroup contact. This variable will be measured as a binary variable with the value of 1 if the respondent has interacted with a person that belongs to a different faith during the last 6 months, and 0 if otherwise. This variable was recorded based on the respondent's own admission. Considering that religious difference is not usually identifiable without obvious markers (such as religious attributes), it is possible that some individuals have experienced intergroup contact unknowingly and thus were recorded as not having any contact. However, it is less likely that intergroup contact without awareness will have much impact on the attitude of a person.

Educational level will be measured as an ordinal variable with four categories: primary or lower, lower secondary school, upper secondary school, and higher education. This variable will be defined as the highest education level completed by the respondent. Thus, an individual who has attended lower secondary school but did not complete it will be recorded as having a primary education. Primary or lower category will act as the reference category.

Household monthly expenditure per capita will be used as a measure of individual economic level. As in many other developing countries, expenditure data in Indonesia is considered to be more reliable than income data (Yusuf et al., 2019). As the distribution of the data is highly skewed, this variable will be transformed into a logarithmic scale. Type of place of residence will be a binary variable with the value of 0 when the respondent resides in a rural area and 1 if the respondent resides in an urban area. The reference category will be urban. Similarly, gender will also be included as a binary categorical variable, where 0 represents male respondents and 1 represents female respondents. The reference category will be male. Meanwhile, the respondent's age in years will be included as a numerical variable. This variable is defined as the respondents' age at their last birthday.

In addition to the individual-level predictors, this study will include three province-level predictors: religious diversity, Gini ratio, and population density. The main interest of this study is the religious diversity variable, which will be measured using the fractionalization index. This index is the most commonly used index to capture ethnic or religious diversity (Montalvo & Reynal-Querol, 2005) and is a variation of the Hirschman-Herfindahl Index. In the context of this study, this index measures the probability that two individuals randomly drawn from a province do not share membership in the same religion. Mathematically, it is defined as (Montalvo & Reynal-Querol, 2005):

$$RD = 1 - \sum_{i=1}^{N} \pi_i^2 \tag{16}$$

where π_i is the proportion of people belonging to religion i. Thus, the index ranges from 0 to 1, with 0 representing the lowest possible level of diversity (everyone in the region has the same religion). This measure will be calculated from a dataset obtained from the Indonesian Ministry of Religious Affairs. The dataset contains the number of adherents of the six recognized religions (Islam, Christianity, Catholicism, Hinduism, Buddhism, and Confucianism) and one "Other" category, by province.

Gini ratio is a measure to represent the level of economic inequality within a society. This index will be used as a measure of economic inequality. The value of 0 represents perfect economic equality, while 1 represents absolute inequality. The data will be obtained from the official publication of Statistics Indonesia. Lastly, population density of each province is calculated from the division of the total number of populations by the area size. This measure will be included as the community characteristics. The data will be obtained from the official publication of Statistics Indonesia.

4.2.2 Analysis Strategy

This study has a relatively small sample size at the second level (34 observations), which can make it harder to obtain accurate estimates. Although the 30/30 rule (a minimum of 30 groups each with 30 observations in it) is often suggested as a general rule, it does not guarantee accuracy since in practice the minimum sample size depends on various factors (Hox & McNeish, 2020).

This study will attempt to minimize this risk by using the Kenward-Roger correction that adjusted the standard error and degrees of freedom when the sample size is small (Hox & McNeish, 2020). Furthermore, this study will use the restricted maximum likelihood (REML) estimation method, which generally produces more accurate estimates compared to the least squares method or full maximum likelihood (FML) method, especially when the group sample size is small (Hox, 2010; Hox & McNeish, 2020).

For group-level variance, restricted maximum likelihood method can provide accurate estimates with as few as 30 groups (Maas and Hox, 2004, as cited in Hox, 2010), or even 6-12 groups (Browne and Draper, 2000, as cited in Hox, 2010). Hox and McNeish (2020) suggested lower minimum group sample sizes (7-25 groups) with REML compared to the FML method (30-75 groups). The drawback of this approach is that with the REML method it is not possible to utilize likelihood ratio tests to compare models (Hox, 2010).

There are two model building strategies for multilevel analysis: top-down or bottom-up (Hox, 2010). The top-down strategy starts from the the most complex model with the maximum number of predictor and random slopes. Then, insignificant effects are gradually removed from the model. However, Hox (2010) suggested the bottom-up modeling approach, which starts with a simple model then gradually adds more complexities. This is because the top-down approach need a longer computation time and can lead to convergence problems.

The bottom-up modeling stages suggested by Hox (2010) are:

1. The first step of this strategy is to construct a varying-intercept empty model, which will be used to calculate the ICC.

2. The second step is to add the individual-level predictors, while keeping the slopes constant an.d the intercept varies. In this step, the contribution of each individual-level predictors can be tested and assessed. Moreover, changes in the variance terms of both individual and province level can be inspected. Adding individual-level predictors may reduce the variance in both individual and group level, when the groups have different compositions (Hox, 2010).

3. The third step is to add the province-level predictors (Equation 6) and examine their significance in the model. The changes of the variance terms in this step may also be analyzed. In contrast with the individual-level predictors, adding group-level predictors should only reduce the group-level variance.

According to the modeling strategy described by Hox (2010), the next step is to assess and construct a varying-slopes model. However, this thesis will only do a varying-intercept model for a couple of considerations. First, the main interest of this study is in analyzing and assessing the fixed coefficients, and a varying-intercept model is already sufficient to achieve the objective. Furthermore, it is recommended to keep the model as simple as possible, especially when a smaller sample size is involved (Hox & McNeish, 2020).

To complement the analysis, classical single-level regression models will be constructed as a comparison to the multilevel models. The single-level regression will be conducted with the least square estimates. For the single-level regression analysis, the modeling process will be done in three stages:

1. The first model will contain all of the individual-level predictors, except for intergroup contact.
2. In the second model, intergroup contact will be added. This was done to examine the impact of intergroup interaction to the estimates of other predictors.
3. In the third model, all of the province-level predictors will be added.

5 Finding and Results

This chapter discusses the findings from the models and their interpretations. The chapter is opened with the exploration of descriptive findings from the data. Then, single-level linear regression models are presented first as the basis and comparison for the more complicated models. Finally, the results from the multilevel linear regression modeling are presented. The modeling process follows the strategy introduced in Chapter 4. The models were run in R with lm function in stats package for the single-level linear regression and lmer function from lmerTest package for the multilevel linear regression.

5.1 Descriptive Statistics

Table 3: Summary Statistics of Categorical Variables

Variables	**n**	**Percentage**	**Tolerance Score***
Intergroup interaction			
No	42402	56.76	2.13 (0.64)
Yes	32299	43.24	2.64 (0.48)
Gender			
Male	25435	34.05	2.41 (0.62)
Female	49266	65.95	2.32 (0.63)
Education			
Primary or lower	35367	47.34	2.26 (0.65)
Lower secondary	13521	18.10	2.35 (0.63)
Upper secondary	18087	24.21	2.44 (0.59)
Higher education	7726	10.34	2.52 (0.58)
Place of residence			
Urban	32255	43.18	2.44 (0.59)
Rural	42446	56.82	2.28 (0.65)
Total	**74701**	**100.00**	**2.35 (0.63)**

* Mean (Std.dev)

Table 3 presents the summary statistics of the categorical predictors. People who have interacted with members of other religions had higher average tolerance scores (2.64) compared to the ones who have not (2.13), and the difference was quite substantial. Around 43.24% of the respondents have interacted with someone from a different religion during the last 6 months before the survey was conducted. There were more women (65.95%) compared to men (34.05%) in the dataset. On average, women had lower tolerance scores (2.32) than men (2.41), although the difference was minor. Average tolerance levels were consistently higher in more highly educated groups. Respondents who lived in urban areas had a slightly higher average tolerance level (2.44) compared to the ones who lived in rural areas (2.28).

Table 4: Summary Statistics of Continuous Variables

Variable	**Mean**	**Std.dev**	**Median**	**Min**	**Max**
Age	46.74	13.68	46	13	98
Expenditure (Original)	4,442,329	4,043,235	3,487,968	209,046	150,110,423
Expenditure (Transformed)	6.55	0.28	6.54	5.32	8.18

The average age of the respondents was 46.7 years old. The youngest respondent was 13 years old, while the oldest was 98 years or older (respondents older than 98 years old were recorded as 98 years old). Meanwhile, the average per capita expenditure was Rp4,442,329 per month. The median was quite far from the average, which means that there were some extreme values at the higher end that pulled the average up. This is in line with the maximum value, which was extremely high compared to the average and median. The skewness was reduced after the log transformation.

Figure 1 presents the average religious tolerance score in each province. Darker color represented a higher level of religious tolerance, which ranged from 1 to 4. Based on this figure, provinces in Kalimantan, East Nusa Tenggara, and Papua had relatively higher tolerance scores, while provinces in Sumatera and Jawa had relatively lower tolerance scores. Nearby provinces tended to have similar average tolerance scores, especially when located on the same island. This is quite expected since the attitude of an individual can be heavily influenced by the attitude of others around them. It also reflects similarities in culture, experience, or values of nearby communities.

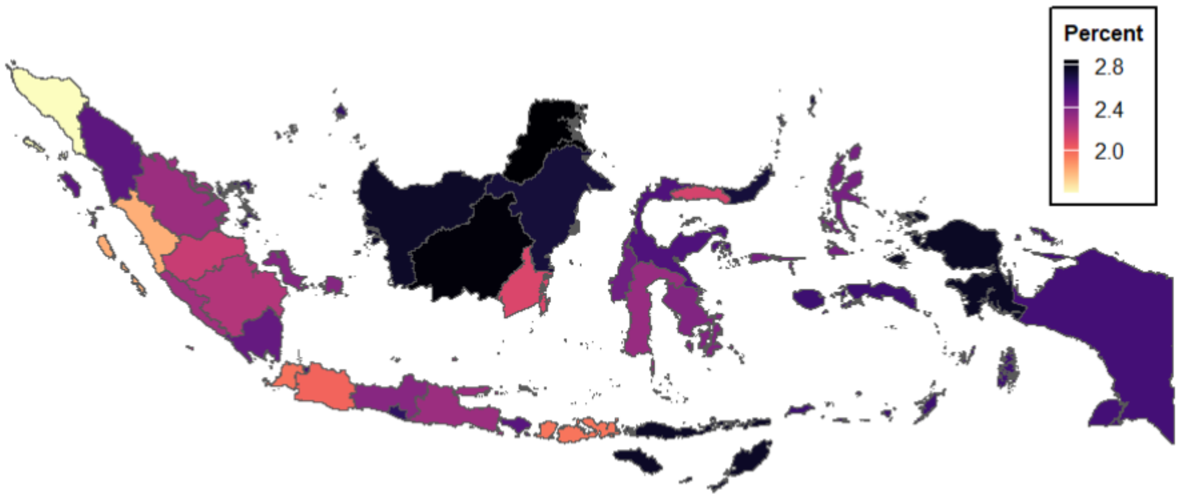

Figure 1: Average Religious Tolerance Score by Province

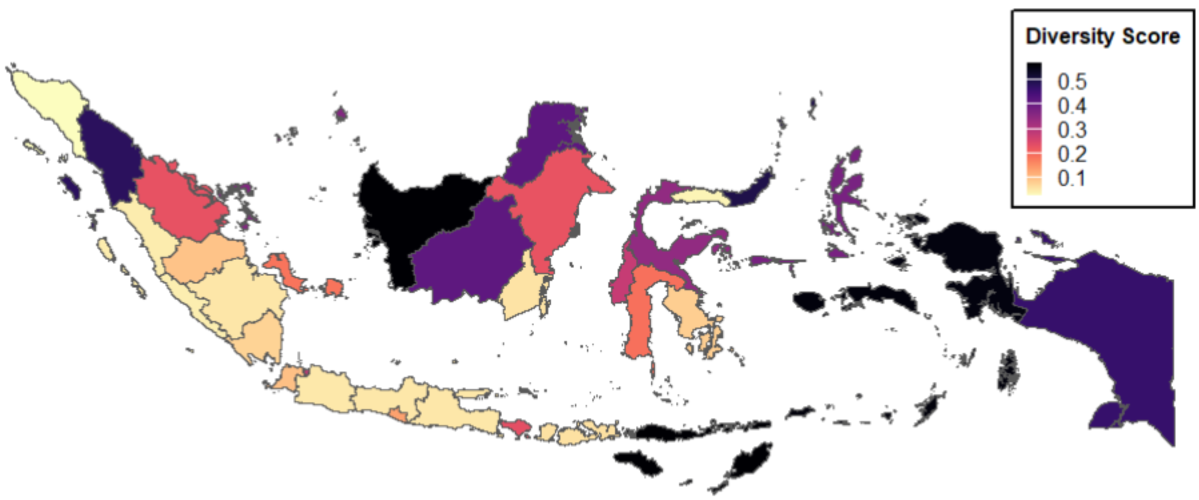

Figure 2: Religious Diversity Score by Province

Meanwhile, Figure 2 shows the religious diversity score by province. Theoretically, the diversity score may range from 0 to 1, but the lowest score in the dataset was 0.03 in Aceh and the highest one was 0.57 in East Nusa Tenggara. Compared to Figure 1, the similarity between neighboring provinces was less pronounced here, with some provinces having much higher, or lower, diversity compared to its neighbors. For instance, North Sumatra was the province with the highest diversity in Sumatra. This is because many of the Bataks, a minority ethnic group who live in that area, are Christian. This was in contrast with Aceh, the province at the northwest corner of Sumatra, which had the lowest diversity index (almost everyone there were Muslims). Nevertheless, by comparing the two figures, it can be observed that provinces with higher diversity also seemed to have higher average religious tolerance scores.

5.2 Single-Level Linear Regression

The modeling in this section was done in three stages. The first model included age, expenditure, gender, education, and type of place of residence. Then, in the second model, intergroup contact was added. Finally, province-level predictor variables (religious diversity, Gini ratio, and population density) were added in the third model. The results are presented in Table 5.

People who have experienced intergroup contact were found to have significantly higher tolerance scores on average (0.48 points in Model 2 and 0.38 points in Model 3), compared to people who have never experienced intergroup contact. The addition of intergroup contact variable increased the amount of variance explained (adjusted R-squared) quite substantially, from 4.16% in Model 1 to 17.57% in Model 2. Furthermore, after the addition of intergroup contact in Model 2, the estimates of other predictors became smaller. There was even a reversal of the sign for the age variable.

Meanwhile, people who live in more diverse provinces were found to have significantly higher tolerance scores. It is important to note that religious diversity was measured in a form of an index that ranged from 0 to 1. The estimate value for religious diversity variable (0.90) can be interpreted as the estimated average score difference between people who live in a place where everyone has the same religion (religious diversity score equals to 0) and people who live in a place where everyone has different religions (religious diversity score equals to 1). In reality, these are quite unusual values; religious diversity index in the dataset only ranges from 0.03 to 0.57. Based on this information, the expected difference between people who live in the least diverse province and the most diverse province is $(0.57 - 0.03)0.90 = 0.49$ points on the tolerance score.

Age was not found to be significant in both Model 1 and 2, but interestingly tolerance was found to be significantly increased with age after the addition of province-level variables in Model 3. Household expenditure was found to be positively and significantly related to tolerance level in all three models, although the effect weakened after the introduction of more variables in Model 2 and 3. Men were found to have significantly higher tolerance scores in all three models compared to women.

Table 5: Single-Level Linear Regression

	Model 1	**Model 2**	**Model 3**
Intercept	1.1441*** (0.0595)	1.7619*** (0.0554)	1.4889*** (0.0599)
Age	-0.0002 (0.0002)	0.0001 (0.0002)	0.0005** (0.0002)
Expenditure	0.1956*** (0.0088)	0.0641*** (0.0082)	0.0439*** (0.0079)
Gender			
Male (ref.)	-	-	-
Female	-0.0889*** (0.0049)	-0.0599*** (0.0045)	-0.0340*** (0.0044)
Education level			
Primary/lower (ref.)	-	-	-
Lower secondary	0.0589*** (0.0065)	0.0391*** (0.0060)	0.0509*** (0.0058)
Upper secondary	0.1124*** (0.0062)	0.0665*** (0.0057)	0.0745*** (0.0055)
Higher education	0.1544*** (0.0084)	0.0968*** (0.0078)	0.1121*** (0.0075)
Place of residence			
Urban (ref.)	-	-	-
Rural	-0.1088*** (0.0048)	-0.0716*** (0.0045)	-0.1175*** (0.0045)
Intergroup contact			
No (ref.)	-	-	-
Yes		0.4816*** (0.0044)	0.3766*** (0.0044)
Religious diversity			0.8964*** (0.0115)
Gini ratio			0.7246*** (0.0613)
Population density			-0.0002 (0.0001)
R^2	0.0417	0.1758	0.2380
Adjusted R^2	0.0416	0.1757	0.2379
F-Statistic	464***	1991***	2121***

Note:

* p<0.05, ** p<0.01, *** p<0.001, standard error in parentheses

Furthermore, people who have completed higher levels of education were found to be significantly more tolerant on average. Compared to the first model, the estimates for each level of education dropped after the addition of intergroup interaction in Model 2 before increasing again in Model 3. A similar pattern can be observed for place of residence variable. People who live in rural areas were found to have significantly lower tolerance scores compared to people in urban areas in all three models, and the estimate is the lowest in Model 2.

Out of the three province-level predictors, population density was the only one found to be not statistically significant. People who live in provinces with higher inequality (higher Gini ratio) were found to significantly have higher tolerance score. Similar with the religious diversity score, it is also important to note that the Gini ratio ranged from 0 to 1. The addition of province-level variables further increased the amount of variance explained from 17.57% in Model 2 to 23.79% in Model 3.

5.3 Multilevel Linear Regression

The multilevel modeling was started by constructing an empty (intercept-only) model in order to estimate the intraclass correlation coefficient. Then, all individual-level independent variables were added in Model 5. Finally, in Model 6 the province-level independent variables were added. The results are presented in Table 6.

In the intercept-only model, the estimated residual variation at individual level was 0.31, while the residual variation at province level was 0.09. By inputting these values into Equation 13, the intraclass correlation coefficient can be calculated as: $0.09/(0.09 + 0.31) = 0.23$. Therefore, around 23% of the variance was attributable to the grouping structure whereas the remaining 77% is attributable to individual variation. This ICC value can be considered to be quite high and may justify the use of multilevel regression.

Between Model 4 and 5, there was a reduction of variability in both individual (0.31 to 0.27) and group level (0.09 to 0.06). Then, adding group-level predictors in Model 6 caused quite a considerable drop in unexplained group-level variability (0.06 to 0.03).

Table 6: Multilevel Linear Regression

	Model 4	**Model 5**	**Model 6**
Intercept	2.4077*** (0.0521)	1.6006*** (0.0679)	1.1766*** (0.2962)
Age		0.0003* (0.0002)	0.0003* (0.0002)
Expenditure		0.1015*** (0.0077)	0.1014*** (0.0077)
Gender			
Male (ref.)	-	-	-
Female		-0.0169*** (0.0042)	-0.0168*** (0.0042)
Education level			
Primary/lower (ref.)	-	-	-
Lower secondary		0.0545*** (0.0055)	0.0545*** (0.0055)
Upper secondary		0.0898*** (0.0053)	0.0899*** (0.0053)
Higher education		0.1266*** (0.0072)	0.1266*** (0.0072)
Place of Residence			
Urban (ref.)	-	-	-
Rural		-0.1183*** (0.0044)	-0.1184*** (0.0044)
Intergroup contact			
No (ref.)	-	-	-
Yes		0.3307*** (0.0042)	0.3305*** (0.0042)
Religious diversity			1.0048*** (0.1699)
Gini ratio			0.5277 (0.8142)
Population density			-0.0003 (0.0012)
Random effects			
Individual	0.3109	0.2710	0.2710
Province	0.0922	0.0653	0.0329
ICC	0.2287	0.1941	0.1083
R_1^2	0.0000	0.1282	0.1282
R_2^2	0.0000	0.2920	0.6429

Note:

* p<0.05, ** p<0.01, *** p<0.001, standard error in parentheses

According to the R_1^2 score, 12.82% of the individual-level variability was explained by the addition of individual-level predictors in Model 5. Interestingly, these individual-level predictors also explained 29.2% of province-level variability, which was quite high. After the addition of group-level predictors, the model explained 64.29% of the variability at province-level.

Most of the findings are in line with the findings from the single-level linear regression models. In Model 5, all of the predictors were found to be significant. Intergroup contact was found to be highly and significantly associated with higher tolerance level. Compared to people who did not experience intergroup contact, people who have the experience on average have tolerance scores that are 0.33 points higher. This estimate is slightly lower than the ones obtained from the single-level models.

The addition of province-level variable did not have a substantial effect on the estimates of individual-level predictors. Out of the three province-level predictors, religious diversity was the only one found to be statistically significant in explaining group-level variation for religious tolerance. The estimate (1.00) was a little bit higher than the one in the single-level linear model. Population density was not found to be statistically significant in both single-level and multilevel models. Contrary to the finding in the single-level model, Gini ratio was not found to be statistically significant. In Model 6, it can be observed that the standard errors for the province-level predictors are substantially larger than the ones in Model 3.

Figure 3 shows the estimated deviation between the overall (mean) intercept to the personal intercept of each province. The value around 0 means that the intercept of that province was about the same as the overall intercept. Values further away to the left (negative values) means that the intercept of that province was lower than the overall intercept, while the ones further away to the right (positive values) are higher. From this figure, it can be seen that the intercept for Aceh province is substantially lower than other provinces. Other than Aceh (-0.46), the differences all ranged from -0.29 in West Sumatra to 0.31 in Lampung.

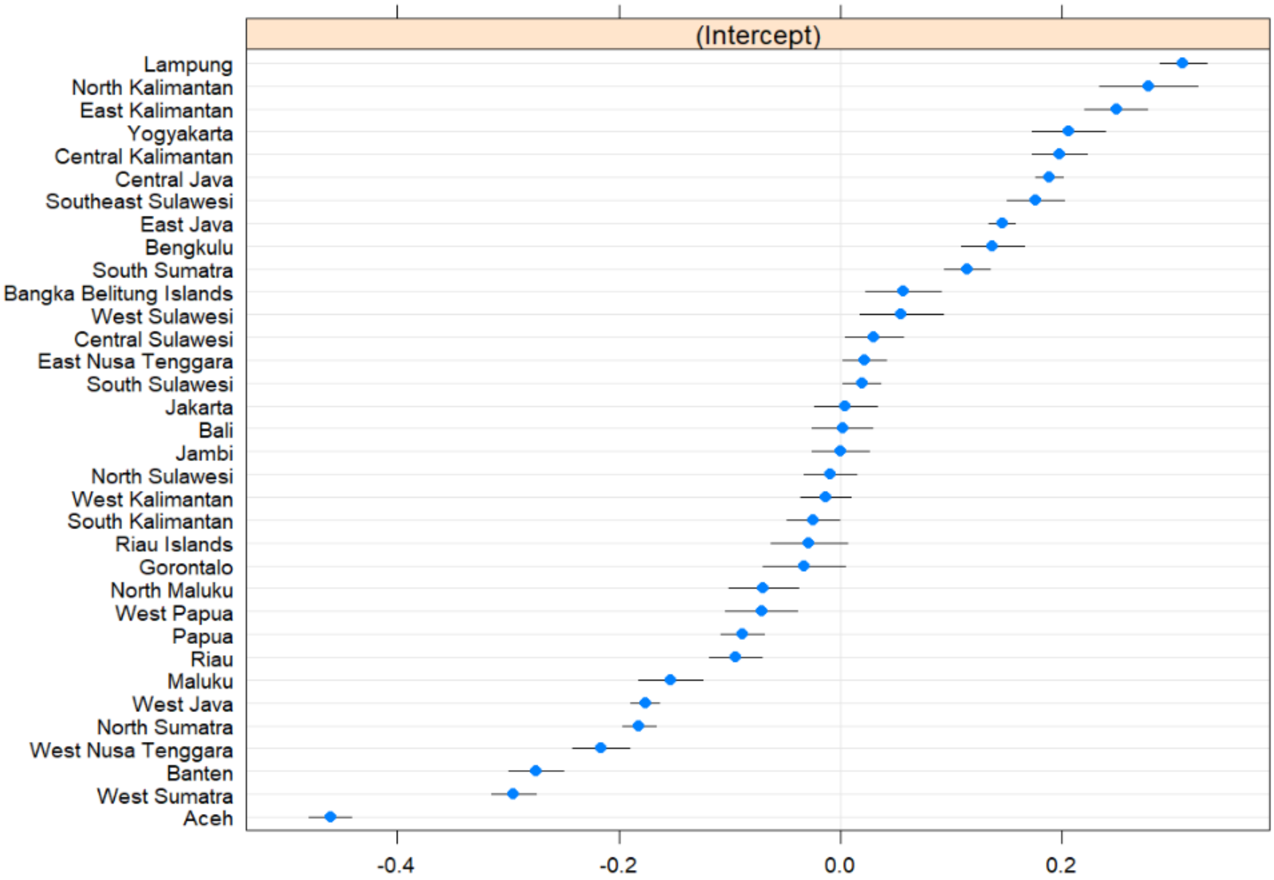

Figure 3: Intercept Deviation by Province

6 Discussion and Conclusions

This chapter is divided into two subchapters. In the first section, further analysis of the results found in Chapter 5 are presented. Furthermore, this section aims to address the implications of these findings and the limitations of the analysis. Then, the last subchapter wraps up the thesis by providing a brief recapitulation of the study and a reflection of the whole process.

6.1 Discussion

In general, the findings in this study are consistent with theories and other findings from previous studies. The findings from the single-level and multilevel regression models are also mostly consistent. However, Gini ratio was found to be significantly and positively related to tolerance level in the single-level regression model but not in the multilevel linear regression model. This may be an example of spurious significant result due to the underestimation of standard error in the single-level linear regression. By comparing the results in Table 5 and Table 6, it can be seen that the standard error for province-level estimates are substantially lower in the single-level model. This phenomenon can arise from the difference of sample sizes used in calculating the estimated standard errors of group-level predictors (Hox, 2010). The true number of sample size for the province-level predictors is 34, but by treating these predictors as individual-level predictors, the sample size was inflated to be the number of individual observations (74,701).

The findings from both single-level and multilevel regression underlined how important accounting for intergroup contact is in analyzing tolerance issues and how excluding it might cause significant omitted variable bias. After the addition of intergroup contact variable in Model 2, the estimates of other predictors became considerably smaller. This indicates that the effect of intergroup interaction was absorbed by other variables in Model 1. By adding this variable in Model 2, the true effect of the other predictors were isolated. For example, it can be observed from Table 5 that the estimate of education level dropped considerably from Model 1 to Model 2, and higher educational level experienced a more substantial decrease. It is possible that people with higher level education tend to have more opportunities to interact with people from other religions. Without the intergroup interaction variable, this effect was calculated as part of the educational level effect.

Another interesting finding related to the impact of intergroup contact variable was the reversal of sign of the age variable. This may be due to an underlying relationship between intergroup contact and age, although the models showed no significant multicollinearity according to the generalized variance inflation factor (GVIF) statistics (more detailed information available in Appendix A). Johnston et al. (2018) demonstrated how such a phenomenon can still be observed despite low VIF values. When intergroup interaction was ignored, the model shows a slight, non significant tendency that younger people have a higher level of tolerance. However, it is likely that the underlying cause of this relationship is intergroup interaction is a more common experience among younger people, and people who have experienced intergroup interaction tend to have higher tolerance scores. so that when it has been considered, the relationship between age and tolerance level changed. In other words, when people who have similar intergroup interaction experience are compared, older people tend to have higher tolerance scores. The effect became stronger after province-level variables had been considered.

The ICC value from the multilevel model suggested that there was a quite sizable variation in tolerance level between provinces. Religious diversity was found to explain a substantial proportion of this variation. In both single-level and multilevel models, provinces with higher diversity were significantly found to have higher average tolerance scores. This finding may be proof that at least in the context of this study, the positive effect of intergroup contact still has a stronger impact compared to the outgroup threat that is induced by higher diversity. This may be because the regions with higher levels of religious diversity had increased their diversity gradually. Negative effects for diversity are usually found in places where the number of outgroup members increased rapidly. In such case, the negative effect from the heightened sense of outgroup threat tend to be more prominent (Pettigrew & Hewstone, 2017). The significance of both diversity and intergroup contact in the model suggested that diversity itself has a positive effect independent of individual's experience of intergroup interaction. This may be explained by the indirect effect of intergroup contact found in previous studies (Christ et al., 2014; Pettigrew et al., 2011) that the impact of intergroup contact may affect even individuals that are not directly involved in the contact themselves. Additionally, people who live in a more homogeneous society may also feel more comfortable voicing intolerant views as they are less likely to be confronted by outgroup members.

The addition of individual-level predictors in Model 5 caused the reduction of variance terms in both individual and province level. This phenomenon is generally caused by unequal composition between groups instead of contextual effect (Hox, 2010). Since the composition of the individuals in each province is different, the addition of individual-level predictors will also explain some of the variation in province level. For example, some provinces may have a higher proportion of highly educated individuals than others. The reduction was also quite substantial relative to the total variance, which indicates that the compositional difference between provinces was quite significant. By contrast, the addition of province-level variables only reduced the variance at province level because they cannot predict individual-level variation (Hox, 2010).

Besides the aforementioned main findings, there are also some additional findings in this study. Age and household expenditure were found to be positively related to tolerance level. Men were found to have higher tolerance on average compared to women. People who live in urban areas tend to be more tolerant on average compared to people in rural areas, which may be explained by people in urban areas having higher exposure, either direct or indirect, to outgroup members. Regarding education level, a previous study on religious intolerance in Indonesia discovered an interesting finding that in recent times, highly educated people were more intolerant than less educated ones (Mietzner & Muhtadi, 2019), however, this study found that people with higher education level are consistently more tolerant on average. It should be noted that although the questions used to measure religious tolerance were quite similar with this study, Mietzner's (2019) study focused on one religious group, the questions were more detailed, and the data was collected a few years before this study.

One significant limitation of this study is the inability to account for differences in societal norms and values. Societal values may significantly shape an individual's attitude towards outgroup members. Intolerant attitudes can also be shaped due to the need to conform to the common norms in the society (Allport, 1979). For example, societies that place higher value on conformity and tradition may discourage assimilation with foreign influences. Thus, individuals who are living in such an environment may tend to be less inclined to seek intergroup relations or even avoid it due to the fear of being antagonized by the people around them. As Indonesia is an extremely plural society, identifying these different societal values and analyzing how they impact intergroup relations and attitudes may be a potential area for future research.

Another important limitation of this study is the crude measurements, especially for religious tolerance and intergroup contact. The intergroup contact variable in this study lacks the detail about the kind and quality of the interaction, which may reveal important information that could add value to the analysis. Allport (1979) theorized that the impact of intergroup contact would largely depend on the nature of the contact itself. Negative or unequal interactions where one party has more power over the other might strengthen prejudice instead of reducing it (Allport, 1979). Considering that religious differences tend to be less visible and that the variable was recorded based on the respondent's own admission, it is likely that the variable is inclined to record interactions between individuals that are already familiar enough with each other to know about their religious differences. Consequently, these interactions are more likely to be positive. This factor, along with the previous finding that positive encounters are far more frequent than negative ones in general (Pettigrew et al., 2011), may contribute to the strong positive impact of intergroup contact found in this study.

Meanwhile, religious tolerance was only calculated from four different aspects that were measured with a four-point scale. Nevertheless, it includes various different aspects of intolerance: the question on being led by an outgroup member portrays intolerance in political aspect, the question on being friends with an outgroup member portrays intolerance in personal aspect, while the questions on outgroups member holding events or building a place of worship nearby portrays intolerance in social and religious aspect.

Additionally, there are two competing views regarding how Likert-scale data should be analyzed. Some scholars such as Jamieson (2004) view Likert-scale as an ordinal data, thus parametric analysis or even arithmetic manipulations such as summation or average cannot be used. However, others (Carifio & Perla, 2008; Norman, 2010; Willits et al., 2016) have argued that composite scores of Likert items can be treated as interval data, and thus parametric methods can be utilized to analyze it. According to Carifio and Perla (2008), treating Likert data strictly as ordinal may prevent researchers from benefitting from more powerful and nuanced analysis methods, such as multiple regression.

Aside from the crude measurement issues, there was some uncertainty in the data due to the lack of identifier variables. The method chosen to handle this issue and its justification was described in Section 4.2.1. Furthermore, it was also not possible to do analysis on regency/city level instead of province level with the survey data, which would resolve the smaller sample size problem and give the analysis more granularity. These aforementioned issues are common shortcomings of working with a preexisting secondary dataset. Since the dataset was originally produced for another purpose, it may not be perfectly fitted to the need of the study. Even so, utilizing a secondary dataset is still more advantageous in general since it is cost and time effective, it enables analysis on a larger scale, and it ensures the quality of the dataset.

A possible area for future research is to use longitudinal analysis to reveal information that cannot be obtained through cross-sectional research designs. Changes in social and population structures may have influenced attitudes in a way that is obscured in cross-sectional analysis. For instance, Allport (1979) theorized that rapid increase of minority group proportion in a society can intensify prejudice. Therefore, higher levels of prejudice may be observed in areas that had just experienced an influx of minority groups in recent times, regardless of the current state of diversity in the area. Rapid changes may heighten the sense of outgroup threat in a society while the positive impact from intergroup contact is still minimal (Pettigrew & Hewstone, 2017).

Lastly, future research may be able to examine in various different contexts if the effect of intergroup contact would differ depending on the status of the individual. In their meta analysis, Pettigrew and Tropp (2006) found that the majority group was affected more strongly by intergroup contact compared to the minority group, although the difference is small. Barlow et al. (2013) posited the “wallpaper effect” to explain this phenomenon: for a member of a minority group, intergroup contact is the norm, and consequently it has less transformative effect on their attitude. However, Schmid et al. (2017) failed to replicate this effect in a different context. Additionally, they also found that intergroup contact can sometimes have a stronger effect for the minority groups compared to the majority group. Replication in various different contexts is still needed to test this hypothesis (Pettigrew & Hewstone, 2017).

6.2 Conclusions

For a populous and highly multicultural country such as Indonesia, maintaining intergroup tolerance is compulsory to keep a functioning and safe society. This thesis aims to identify how intergroup interaction and diversity may affect tolerance level in society. To inspect this question, multilevel linear regression analysis was utilized. From a statistical standpoint, this analysis can be used as a solution to the violation of independent assumption in hierarchical datasets. From a theoretical standpoint, the attitude of an individual is heavily shaped by the characteristics of their community, so an analysis tool that can be used to inspect both individual- and community-level variables is needed to examine this issue in a more comprehensive way.

The analysis was done by utilizing a national survey data complemented with province-level information compiled from various sources. First, descriptive findings were explored through tables and figures. To complement the analysis, single-level linear models were first constructed before advancing on the multilevel linear regression models. The models were run in R with stats package for the single-level linear regression and lmerTest package for the multilevel linear regression.

Intergroup interaction was strongly and significantly found to be associated with higher religious tolerance level. According to the final model, individuals who have interacted with members of other religions tend to have tolerance scores that are about 0.33 points higher on average compared to individuals who have not experienced intergroup interaction recently. The impact of this variable was so strong that there is evidence that omission of this variable may lead to misleading results.

From the multilevel analysis result, it was revealed that around 23% of the variation in tolerance level was due to the grouping structure instead of individual differences. This indicate that the inequality of tolerant attitudes between provinces is quite high. Religious diversity was found to explain a substantial portion of this variation. The individual-level predictors were able to explain around 29.2% of the variation in the province-level, and after the addition of province-level predictor to the model, the amount of variance explained in the province-level became 64.29%.

Religious diversity was found to be significantly and positively associated with religious tolerance level. This indicates that in the context of this study, the positive effect that is caused by higher level of exposure to outgroup members is still higher than the negative effect from outgroup threat. In the multilevel model, religious diversity was the only province-level predictor that was found to be significant.

The Gini ratio was found to be significant in the single-level regression model, however, it was not found to be significant in the multilevel model. This may be caused by the underestimation of standard error from treating the variable as an individual-level predictor, which was corrected in the multilevel analysis. Thus, the utilization of multilevel analysis provided an added value to the study by revealing the disparities between provinces and which group-level predictors were able to explain this difference.

This study has some limitations, which includes the omission of cultural characteristics, crude measurements of the response variables, smaller sample size on group-level, and some uncertainties in the dataset. Some precautions have been implemented to minimize the impact, however, it is still important to acknowledge these limitations. On the other hand, these limitations may also become a potential for future research. For instance, future research may focus on identifying cultural differences in the country and analyzing their impact on religious tolerance. Another possible area of future research is investigating how the effect of intergroup contact may differ for majority and minority group members. Future research can also use longitudinal analysis to examine the question and uncover information that cannot be obtained from cross-sectional data.

References

Agnew, R. (1992). Foundation for a General Strain Theory of Crime and Deliquency. *Criminology*, *30*(1), 47–88. https://doi.org/10.1111/j.1745-9125.1992.tb01093.x

Allport, G. W. (1979). *The Nature of Prejudice* (Unabridged, 25th Anniversary Ed). Addison-Wesley Pub. Co.

Aritonang, J. S., & Steenbrink, K. A. (Eds.). (2008). *A History of Christianity in Indonesia* [OCLC: ocn228370027]. Brill.

Barlow, F. K., Hornsey, M. J., Thai, M., Sengupta, N. K., & Sibley, C. G. (2013). The Wallpaper Effect: The Contact Hypothesis Fails for Minority Group Members Who Live in Areas with a High Proportion of Majority Group Members (M. Costantini, Ed.). *PLoS ONE*, *8*(12), e82228. https://doi.org/10.1371/journal.pone.0082228

Carifio, J., & Perla, R. (2008). Resolving the 50-Year Debate Around Using and Misusing Likert Scales. *Medical Education*, *42*(12), 1150–1152. https://doi.org/10.1111/j.1365-2923.2008.03172.x

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The Coefficient of Determination R-Squared is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *PeerJ Computer Science*, *7*, e623. https://doi.org/10.7717/peerj-cs.623

Christ, O., Schmid, K., Lolliot, S., Swart, H., Stolle, D., Tausch, N., Al Ramiah, A., Wagner, U., Vertovec, S., & Hewstone, M. (2014). Contextual Effect of Positive Intergroup Contact on Outgroup Prejudice. *Proceedings of the National Academy of Sciences*, *111*(11), 3996–4000. https://doi.org/10.1073/pnas.1320901111

Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis* (3. ed). Wiley.

Fealy, G., & Ricci, R. (2019). *Contentious Belonging: The Place of Minorities in Indonesia*. ISEAS Publishing.

Gelman, A., & Hill, J. (2007). *Data Analysis using Regression and Multilevel/Hierarchical Models* [OCLC: ocm67375137]. Cambridge University Press.

Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Stories* [OCLC: 1379291547]. Cambridge University Press.

Goh, R. B. H. (2005). *Christianity in Southeast Asia*. Institute of Southeast Asian Studies. Maps on lining papers.

Hefner, R. W. (2013). The Study of Religious Freedom in Indonesia. *The Review of Faith & International Affairs*, *11*(2), 18–27. https://doi.org/10.1080/15570274.2013.808038

Hewstone, M. (2015). Consequences of Diversity for Social Cohesion and Prejudice: The Missing Dimension of Intergroup Contact. *Journal of Social Issues*, *71*(2), 417–438. https://doi.org/10.1111/josi.12120

Hewstone, M., & Schmid, K. (2014). Neighbourhood Ethnic Diversity and Orientations Toward Muslims in Britain: The Role of Intergroup Contact. *The Political Quarterly*, *85*(3), 320–325. https://doi.org/10.1111/1467-923X.12098

Hox, J. (2010). *Multilevel Analysis: Techniques and Applications, Second Edition*. Routledge.

Hox, J., & McNeish, D. (2020, February). Small Samples in Multilevel Modeling. In *Small Sample Size Solutions* (1st ed., pp. 215–225). Routledge. https://doi.org/10.4324/9780429273872-18

Jamieson, S. (2004). Likert Scales: How to (ab)use Them. *Medical Education*, *38*(12), 1217–1218. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Johnston, R., Jones, K., & Manley, D. (2018). Confounding and Collinearity in Regression Analysis: A Cautionary Tale and an Alternative Procedure, Illustrated by Studies of British Voting Behaviour. *Quality & Quantity*, *52*(4), 1957–1976. https://doi.org/10.1007/s11135-017-0584-6

Juergensmeyer, M., & Roof, W. C. (Eds.). (2012). *Encyclopedia of Global Religion*. SAGE Publications.

Kanas, A., Scheepers, P., & Sterkens, C. (2015). Interreligious Contact, Perceived Group Threat, and Perceived Discrimination: Predicting Negative Attitudes among Religious Minorities and Majorities in Indonesia. *Social Psychology Quarterly*, *78*(2), 102–126. https://doi.org/10.1177/0190272514564790

Kersten, C. (2017). *A History of Islam in Indonesia: Unity in Diversity*. Edinburgh University Press.

Kleinbaum, D. G., Kupper, L. L., & Muller, K. E. (1998). *Applied Regression Analysis and Other Multivariable Methods* [OCLC: 777774797]. Duxbury Press.

Laurence, J., Schmid, K., Rae, J. R., & Hewstone, M. (2019). Prejudice, Contact, and Threat at the Diversity-Segregation Nexus: A Cross-Sectional and Longitudinal Analysis of How Ethnic Out-Group Size and Segregation Interrelate for Inter-Group Relations. *Social Forces*, *97*(3), 1029–1066. https://doi.org/10.1093/sf/soy079

Lundåsen, S. W. (2023). Intergroup Contacts, Neighborhood Diversity, and Community Trust: The Asymmetrical Impact of Negative and Positive Experiences. *Journal of International Migration and Integration*, *24*(1), 163–188. https://doi.org/10.1007/s12134-021-00932-z

Marshall, P. (2018). The Ambiguities of Religious Freedom in Indonesia. *The Review of Faith & International Affairs*, *16*(1), 85–96. https://doi.org/10.1080/15570274.2018.1433588

Mavridis, D. (2015). Ethnic Diversity and Social Capital in Indonesia. *World Development*, *67*, 376–395. https://doi.org/10.1016/j.worlddev.2014.10.028

Mietzner, M., & Muhtadi, B. (2018). Explaining the 2016 Islamist Mobilisation in Indonesia: Religious Intolerance, Militant Groups and the Politics of Accommodation. *Asian Studies Review*, *42*(3), 479–497. https://doi.org/10.1080/10357823.2018.1473335

Mietzner, M., & Muhtadi, B. (2019). The Mobilisation of Intolerance and Its Trajectories: Indonesian Muslims' Views of Religious Minorities and Ethnic Chinese. In G. Fealy & R. Ricci (Eds.), *Contentious Belonging: The Place of Minorities in Indonesia* (pp. 155–174). ISEAS Publishing.

Milligan, S., Andersen, R., & Brym, R. (2014). Assessing Variation in Tolerance in 23 Muslim-Majority and Western Countries: Tolerance in Muslim-Majority and

Western Countries. *Canadian Review of Sociology*, *51*(3), 239–261. https://doi.org/10.1111/cars.12046

Montalvo, J. G., & Reynal-Querol, M. (2005). Ethnic Diversity and Economic Development. *Journal of Development Economics*, *76*(2), 293–323. https://doi.org/10.1016/j.jdeveco.2004.01.002

Mulder, M., & Krahn, H. (2008). Individual- and Community-level Determinants of Support for Immigration and Cultural Diversity in Canada. *Canadian Review of Sociology*, *42*(4), 421–444. https://doi.org/10.1111/j.1755-618X.2005.tb00848.x

Muzayanah, I. F. U., Nazara, S., Mahi, B. R., & Hartono, D. (2020). Is There Social Capital in Cities? The Association of Urban Form and Social Capital Formation in the Metropolitan Cities of Indonesia. *International Journal of Urban Sciences*, *24*(4), 532–556. https://doi.org/10.1080/12265934.2020.1730934

Norman, G. (2010). Likert Scales, Levels of Measurement and the "Laws" of Statistics. *Advances in Health Sciences Education*, *15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Pettigrew, T. F., & Hewstone, M. (2017). The Single Factor Fallacy: Implications of Missing Critical Variables from an Analysis of Intergroup Contact Theory. *Social Issues and Policy Review*, *11*(1), 8–37. https://doi.org/10.1111/sipr.12026

Pettigrew, T. F., & Tropp, L. R. (2006). A Meta-Analytic Test of Intergroup Contact Theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Pettigrew, T. F., Tropp, L. R., Wagner, U., & Christ, O. (2011). Recent Advances in Intergroup Contact Theory. *International Journal of Intercultural Relations*, *35*(3), 271–280. https://doi.org/10.1016/j.ijintrel.2011.03.001

Pisani, E. (2014). *Indonesia Etc: Exploring the Improbable Nation*. Granta.

Putnam, R. D. (2007). *E Pluribus Unum* : Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies*, *30*(2), 137–174. https://doi.org/10.1111/j.1467-9477.2007.00176.x

Robinson, W. S. (1950). Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, *15*(3), 351. https://doi.org/10.2307/2087176

Sanjaya, M. R. (2022). New Evidence on Ethnic Diversity and Social Capital in Indonesia. *Jurnal Ekonomi & Studi Pembangunan*, *23*(1), 35–50. https://doi.org/10.18196/jesp.v23i1.13863

Schmid, K., Wölfer, R., Swart, H., Christ, O., Al Ramiah, A., Vertovec, S., & Hewstone, M. (2017). The "Wallpaper Effect" Revisited: Divergent Findings on the Effects of Intergroup Contact on Attitudes in Diverse Versus Nondiverse Contexts. *Personality and Social Psychology Bulletin*, *43*(9), 1268–1283. https://doi.org/10.1177/0146167217711929

Seo, M. (2012). Defining 'Religious' in Indonesia: Toward Neither an Islamic Nor a Secular State. *Citizenship Studies*, *16*(8), 1045–1058. https://doi.org/10.1080/13621025.2012.735028

Sidi, B. A. (2020). *Unity and Diversity: National Identity and Multiculturalism in Indonesia* [Doctoral dissertation]. University of Otago.

Suryadinata, L. (2005). Buddhism and Confucianism in Contemporary Indonesia: Recent Developments. In T. Lindsey, H. Pausacker, & C. A. Coppel (Eds.), *Chinese Indonesians: Remembering, Distorting, Forgetting* (pp. 77–94). Institute of Southeast Asian Studies.

Van Assche, J., Asbrock, F., Dhont, K., & Roets, A. (2018). The Diversity Challenge for High and Low Authoritarians: Multilevel and Longitudinal Effects Through Intergroup Contact and Threat. *Personality and Social Psychology Bulletin*, *44*(8), 1163–1179. https://doi.org/10.1177/0146167218764653

Vickers, A. (2013). *A History of Modern Indonesia: Second Edition* (Second Edition). Cambridge University Press.

Wijaksono, A. (2023). Ethnic and Religious Tolerance in Indonesia. *Journal of Developing Economies*, *8*(2), 316–325. https://doi.org/10.20473/jde.v8i2.46417

Willits, F. K., Theodori, G. L., & Luloff, A. E. (2016). Another Look at Likert Scales. *Journal of Rural Social Sciences*, *31*(3), 126–139. https://egrove.olemiss.edu/jrss/vol31/iss3/6

Yusuf, A. A., Shidiq, A. R., & Hariyadi, H. (2019). On Socio-Economic Predictors of Religious Intolerance: Evidence from a Large-Scale Longitudinal Survey in the Largest Muslim Democracy. *Religions*, *11*(1), 21. https://doi.org/10.3390/rel11010021

Appendix A: Multicollinearity Check

Table A1: Multicollinearity Check

	GVIF	df	aGSIF
Model 1			
Age	1.16	1	1.08
Expenditure	1.22	1	1.10
Gender	1.04	1	1.02
Education	1.33	3	1.05
Place of residence	1.13	1	1.06
Model 2			
Age	1.16	1	1.08
Expenditure	1.25	1	1.12
Gender	1.04	1	1.02
Education	1.34	3	1.05
Place of residence	1.14	1	1.07
Intergroup contact	1.07	1	1.03
Model 3			
Age	1.16	1	1.08
Expenditure	1.26	1	1.12
Gender	1.05	1	1.03
Education	1.35	3	1.05
Place of residence	1.21	1	1.10
Intergroup contact	1.18	1	1.09
Religious diversity	1.21	1	1.10
Gini ratio	1.24	1	1.11
Population density	1.18	1	1.09