

University of Denver

Digital Commons @ DU

Electronic Theses and Dissertations

Graduate Studies

1-1-2017

An Evaluation of Critical Realignment Theory: Comparing Bayesian and Frequentist Approaches

Tara A. Rhodes
University of Denver

Follow this and additional works at: <https://digitalcommons.du.edu/etd>



Part of the [Political Science Commons](#), and the [Statistics and Probability Commons](#)

Recommended Citation

Rhodes, Tara A., "An Evaluation of Critical Realignment Theory: Comparing Bayesian and Frequentist Approaches" (2017). *Electronic Theses and Dissertations*. 1259.
<https://digitalcommons.du.edu/etd/1259>

This Dissertation is brought to you for free and open access by the Graduate Studies at Digital Commons @ DU. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ DU. For more information, please contact jennifer.cox@du.edu, dig-commons@du.edu.

An Evaluation of Critical Realignment Theory:
Comparing Bayesian and Frequentist Approaches

A Dissertation

Presented to

the Faculty of the Morgridge College of Education

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

Tara A. Rhodes

June 2017

Advisor: Antonio Olmos, Ph.D.

©Copyright by Tara A. Rhodes 2017

All Rights Reserved

Author: Tara A. Rhodes
Title: An Evaluation of Critical Realignment Theory: Comparing Bayesian and Frequentist Approaches
Advisor: Antonio Olmos, Ph.D.
Degree Date: June 2017

ABSTRACT

Prior to this study, critical realignment theory, which presupposes eras of substantial and sustained swings in American political party dominance, had only been evaluated using the classical, frequentist approach to modeling. However, potential for more information concerning these electoral phenomena exists given a shift in the design and approach to realigning elections. This study sought to explore those options through one particular alternative to the classical approach to statistics—in this particular case, the Bayesian approach to statistics.

Bayesian methods differ from the frequentist approach in three main ways: the treatment of probability, the treatment of parameters, and the treatment of prior information. This study sought to understand the effect of these differences as it applied to critical realignment theory: namely, what contribution is made in understanding the occurrence of these eras from each statistical approach? Does the Bayesian approach provide any improvements over the classical approach in terms of understanding critical realignment theory? This first set of research questions was asked from a political viewpoint, but a second set of research questions was also posed from a methodological viewpoint: What methods exist to formally compare these two statistical approaches, and what is the relative strength of each method? Using the most efficient method of comparison, is any

further information gained concerning critical realignment theory, and is any further information gained concerning each statistical approach?

Using multiple linear regression, results were similar across approaches. For the presidential data, critical elections were found in 1860 and 1932. This was replicated in the congressional models, with one additional realigning election found in 1996. As for additional information gained, Bayesian methods aided in understanding in some ways, but the classical approach also retained some benefit. Furthermore, these two statistical approaches were formally compared to one another, highlighting the comparison between credible intervals and confidence intervals. While these intervals are traditionally considered counterparts, this is not a direct comparison. These intervals represent different concepts, relating to underlying differences in the statistical approach. This, however, reiterates the strong role of correct interpretation as it pertains to results.

Keywords: critical realignment theory, critical elections, Bayesian methods

TABLE OF CONTENTS

Chapter One: Introduction	1
Background Information	2
Statement of the Problem	6
Purpose of Study	7
Scope and Limitations of Study	11
Definition of Terms	12
Organization of Study	13
Critical Realignment Theory	14
Bayesian Approach to Statistics	29
Misconceptions about Indicators of Statistical Significance	43
Chapter Two: Method	47
Data	47
Analysis Plan	51
Chapter Three: Results	61
Power Analysis	61
Classical Linear Regression Analysis	62
Bayesian Linear Regression Analysis	70
Formal Comparison	89
Chapter Four: Discussion	101
Summary of Results	101
Importance of Findings	104
Implications for Theory and Practice	108
Study Limitations	110
Avenues for Future Research	111
Chapter Five: Conclusion	113
Bibliography	124
Appendices	135
Appendix A	135
Appendix B	142
Appendix C	227

LIST OF TABLES

Table 1. Descriptive statistics for the presidential classical regression models	65
Table 2. Regression coefficients for the presidential classical regression models	65
Table 3. Descriptive statistics for the congressional classical regression models	69
Table 4. Regression coefficients for the congressional classical regression models	69
Table 5. Means and variances for the informative prior distributions for the presidential Bayesian models.....	71
Table 6. Means and variances for the informative prior distributions for the congressional Bayesian models	71
Table 7. Convergence diagnostics for the presidential Bayesian models using a non-informative prior distribution	73
Table 8. Convergence diagnostics for the presidential Bayesian models using an informative prior distribution.....	74
Table 9. Prior and posterior distribution information for the non-informative Bayesian presidential models, comparing the first and second eras.....	77
Table 10. Prior and posterior distribution information for the informative Bayesian presidential models, comparing the first and second eras.....	78
Table 11. Prior and posterior distribution information for the non-informative Bayesian presidential models, comparing the second and third eras.....	78
Table 12. Prior and posterior distribution information for the informative Bayesian presidential models, comparing the second and third eras.....	78
Table 13. Prior and posterior distribution information for the non-informative Bayesian presidential models, comparing the third and fourth eras	79
Table 14. Prior and posterior distribution information for the informative Bayesian presidential models, comparing the third and fourth eras	79
Table 15. Prior and posterior distribution information for the non-informative Bayesian presidential models, comparing the fourth and fifth eras.....	80
Table 16. Prior and posterior distribution information for the informative Bayesian presidential models, comparing the fourth and fifth eras.....	80
Table 17. Convergence diagnostics for the congressional Bayesian models using a non-informative prior distribution	81
Table 18. Convergence diagnostics for the congressional Bayesian models using an informative prior distribution.....	82
Table 19. Prior and posterior distribution information for the non-informative Bayesian congressional models, comparing the first and second eras.....	85
Table 20. Prior and posterior distribution information for the informative Bayesian congressional models, comparing the first and second eras.....	85
Table 21. Prior and posterior distribution information for the non-informative Bayesian congressional models, comparing the second and third eras	86
Table 22. Prior and posterior distribution information for the informative Bayesian congressional models, comparing the second and third eras	86
Table 23. Prior and posterior distribution information for the non-informative Bayesian congressional models, comparing the third and fourth eras.....	87

Table 24. Prior and posterior distribution information for the informative Bayesian congressional models, comparing the third and fourth eras	87
Table 25. Prior and posterior distribution information for the non-informative Bayesian congressional models, comparing the fourth and fifth eras	88
Table 26. Prior and posterior distribution information for the informative Bayesian congressional models, comparing the fourth and fifth eras	88
Table 27. Ranking of comparison methods for comparisons between the classical and Bayesian statistical approaches	95
Table 28. Differences in point estimates and frequentist confidence intervals and Bayesian credible intervals for the presidential models	98
Table 29. Differences in point estimates and frequentist confidence intervals and Bayesian credible intervals for the congressional models	98

CHAPTER ONE: INTRODUCTION

On May 4, 2016, the New York Daily News' front page depicted a red, white, and blue elephant in a casket. The caption read: "Dearly beloved, we are gathered here today to mourn the GOP" (New York Daily News, 2016). This political cartoon referenced the previous day, where Republican candidate Donald Trump had won the Indiana primary. This win all but guaranteed Trump the party nomination, which led to a strong reaction by some Republican loyalists. Writing for the Atlantic, Ball commented that not only were conservatives lining up to hand in their Republican registrations, but with this nomination, the "old party establishment went into exile, perhaps never to return" (Ball, 2016). This scene clearly illustrates a shift of some kind in party systems. While future electorates and researchers will decide the outcome of the 2016 general election, historians and academics that study critical realignment theory may not be surprised by this turn of events. Critical realignment theory presupposes these kinds of shifts in party systems typically once a generation, practically occurring through the methods described above. The mobilization, conversion, or as illustrated here, the demobilization of partisan voters occur, initiating the change in party dominance.

Despite critiques of the theory, the timeliness of these events illustrates the importance of continued study of these types of elections in the field. Empirical analyses of the theory to date have only taken a classical statistics approach,

evaluating national election returns, state and county election returns, and employing a variety of statistical modeling techniques. Prior to this study, the approach of Bayesian modeling had yet to be applied to this topic. Consequently, this study sought to fill that gap by applying Bayesian modeling to critical realignment theory, with the focus of trying to understand the unique contribution of Bayesian techniques to the question of critical elections.

Background Information

Before addressing this question, information concerning each concept is first presented. The next section highlights the essential aspects of critical realignment theory and Bayesian methods. After providing a brief overview of critical realignment theory, particular attention is paid to mechanisms that cause these types of elections as well as critiques of the genre. In regard to Bayesian modeling, a brief introduction to the approach is presented, followed by a discussion of the main differences between the classical approach to statistics and the Bayesian approach to statistics. Lastly, a brief overview of the controversy concerning measures of statistical significance from the classical perspective is also introduced, illustrating the necessity for Bayesian modeling.

At its core, critical realignment theory presupposes different realignment eras or party systems within American electoral history. These realignment eras or party systems are demarcated and differentiated by the occurrence of a critical election. For many, critical realignment theory understands electoral history to be cyclical, with each cycle beginning with a critical election. Three main realigning elections, also referred to as the canon elections, have been hypothesized: 1860,

1896, and 1932 (Norpoth & Rusk, 2007). Other elections, such as 1964 and 2008, have been under consideration, but these have not been added to the main body of literature concerning realigning, or critical, elections. A consensus among researchers pertaining to the status of these hypothesized critical elections has not been reached due to results differing based on methods employed, data utilized, and perspectives taken.

Potential causal mechanisms for this group of elections fall into three categories: conversion of voters, referred to as the conversion thesis; demobilization of party supporters and party-affiliated voters, referred to as the demobilization thesis; and the mobilization of inactive or other new voters, referred to as the mobilization thesis. Conversion of voters relates to party identification, where, for a variety of reasons, individuals change their party attachment from one party to the opposing party. The mobilization thesis relates to the addition of new voters to the electorate, affecting the electoral makeup, and the demobilization thesis relates to the alienation of previous party supporters, as mentioned above (Darmofal & Nardulli, 2010).

Previous information presented illustrates support for the theory; however, not all researchers see value in critical realignment theory. Criticism of the genre can be grouped into three main points: first, the empirical validity of the theory; second, the addition of the genre to the body of political science literature; and third, the relevancy of the genre to the present day. Regarding empirical validity, critics of the genre find it difficult to replicate studies completed in the field, citing data availability and integrity concerns. Concerning the addition of the

genre to the field of political science, some critics find critical realignment theory to be limiting and narrow, forcing researchers and students of political science to unnecessarily see cyclical patterns in American electoral history. Moreover, they argue that this perspective then limits further exploration into other aspects of American elections. Lastly, critics argue that according to the traditional understanding of critical realignment theory, a critical election has not occurred since 1932, which begs the question of present-day relevance (Mayhew, 2002). For these reasons, not all researchers fully support the advancement of critical realignment theory.

Turning away from critical realignment theory, the next section of this discussion provides a brief introduction to Bayesian modeling. Beginning with its namesake, Thomas Bayes, an English minister in the early 1700s, understood rationality as a probabilistic matter: one's understanding approximates truth as more evidence is gathered. This understanding was greatly influenced by Isaac Newton's work, which suggested that nature, after much observation, follows regular and predictable patterns (Silver, 2012). Richard Price, a friend and colleague of Bayes, published this concept of probability posthumously, although the concept lay dormant until rediscovered by French mathematician Pierre Simon Laplace (McGrayne, 2012).

With this understanding, Bayes' theorem is based on conditional probability: where the probability of one event is predicated on the occurrence of another event. Within this framework, classical statistics seeks to answer the question of the probability of a set of outcomes given a specified hypothesis,

whereas Bayesian modeling instead assumes the truth of the data and seeks to answer the question of the probability of the hypothesis given these outcomes. While more information pertaining to this discussion and a formal presentation of Bayes' theorem is provided in the literature review, this is the main point of difference between the classical approach to statistics and the Bayesian approach to statistics. In the classical approach, probability is understood as a long-run sampling frequency of a certain event occurring, assuming constant conditions across samples. In the Bayesian approach, probability is understood more subjectively as a degree of uncertainty (van de Schoot et al., 2013).

The second main difference between the classical approach to statistics and the Bayesian approach to statistics concerns the treatment of prior information. In Bayesian modeling, prior information is included in the analysis, as the target parameter, or underlying parameter of interest, is assumed to be random. This is differentiated from the classical approach to statistics, where the underlying parameter of interest is assumed to be fixed and simply needs to be uncovered by repeated sampling. Thus, there is no need for the inclusion of prior information—the method will result in the 'true' parameter (Stokes, Chen, & Gunes, 2014). The mechanics of inclusion of prior information as well as the impact on results will be addressed in the literature review.

Stemming from the conceptual differences discussed above, a third difference between the classical approach to statistics and the Bayesian approach to statistics is the emphasis on p -values. Much controversy exists around p -values, but traditionally p -values are used as a measure to indicate statistical significance.

Issues regarding the use of this indicator arise as p -values are largely misunderstood from a definitional standpoint and are associated with a significance level, which is arbitrarily chosen and greatly influences researchers with regard to publication (Kirk, 1996). In Bayesian modeling, p -values lose significance as an indicator as Bayesian analyses approach the null hypothesis from a different perspective. Given the assumption that the parameter of interest is random, Bayesian analyses result in and utilize a posterior distribution. It is the location and the variance of this distribution that aids the researcher (van de Schoot et al., 2013).

Statement of the Problem

As discussed previously, one main difference between the Bayesian approach to statistics and the classical approach to statistics is the treatment of probability. As was mentioned, probability is treated as a long-term frequency of a particular event occurring in the classical approach to statistics, but is viewed as the degree of uncertainty concerning the occurrence of a particular event from the Bayesian viewpoint. To illustrate this difference, consider a coin toss. The classical perspective takes a very clinical approach to the coin toss: all conditions must be the same across every toss. Each coin toss is considered a random replicate of all other coin tosses. However, maintaining precisely the same conditions for each coin toss is an extremely difficult task, even in a completely controlled environment. More importantly, however, this requirement of precisely the same conditions rarely occurs in social science situations. Elections illustrate this point, as changes in voters, salient issues, and candidates occur at each

election. Consequently, the treatment of probability from the classical viewpoint does not adequately capture the social situation of and social dynamics inherent within elections.

The inadequacy concerning the treatment and application of probability from the classical perspective is only compounded by the controversy occurring over the use and application of p -values as they relate to study results. The lack of definitional understanding, the use of an arbitrary significance level, and the dichotomous nature in the application of significance of the p -value severely limits the practical significance and interpretation of results. This limited contribution of the p -value in a practical sense compounded with the conceptual differences in probability between the classical approach and the Bayesian approach to statistics is what called for a reevaluation of critical realignment theory from a new perspective.

Purpose of Study

Thus, the purpose of this study was three-fold: first, to evaluate critical realignment theory from a new perspective; second, to expand the application of Bayesian modeling to a new field; and third, to formalize an empirical method of comparison between classical statistics and Bayesian statistics. This resulted in two sets of research questions. The first set of research questions was concerned with the qualitative contribution of this study to the field of political science.

These questions are given below:

1. Viewed from a national perspective, what contribution does the classical approach to statistics make in terms of increasing understanding regarding the occurrence of critical elections?
2. Viewed from a national perspective, what contribution does the Bayesian approach to statistics make in terms of increasing understanding regarding the occurrence of critical elections?
3. With regards to the identification of critical elections, does Bayesian modeling provide improvements, and, if so, what improvements over the classical approach?

While answering these research questions entailed empirical analysis, the focus of these questions was substantive in nature. The reason for addressing these questions from a qualitative viewpoint was because each approach conceptualizes probability, parameters, and prior information differently. The real intent of these questions was to understand how these different conceptualizations affect the practical result provided at the end of the analysis. In other words, the goal of these questions was to highlight qualitatively the difference in understanding gained surrounding critical realignment theory from the two different approaches. The answers to these particular research questions detail the contribution to the field of political science.

The second set of research questions was concerned with the quantitative contribution of this study to the field of research methods and statistics. In addition to applying Bayesian methods to a new field of study, this study also sought to formalize an empirical method of comparison between the classical approach to statistics and the Bayesian approach to statistics. This was done by first evaluating current methods of comparison for their relative strength, and then applying the most efficient method of comparison to the topic of critical realignment theory to see what additional information can be learned about the theory, but also about the two approaches. The most efficient method of comparison was defined as the method with the highest relative strength. This set of research questions is given below:

1. Given different methods of comparison between the classical approach to statistics and the Bayesian approach to statistics, what are the relative strengths of each method?
2. Using the most efficient method of comparison, is any further information gained in applying this method to critical realignment theory, and if so, what is that contribution?
3. By applying the most efficient method of comparison to the example of critical realignment theory, is any further information gained regarding the two approaches, and if so, what is that contribution?

The focus of this set of questions was quantitative in nature and critical realignment theory is used as an example to illuminate differences in the two approaches. The goal of these questions was to provide a formal, empirical method of comparison between the two approaches, and through the process, highlight the relative strengths and weaknesses of each method. If, through a comparative lens, the implications for each approach can be realized, researchers would be better informed as to when to apply each approach appropriately. This details the contribution to the field of research methods and statistics.

To answer the first set of questions, the occurrence of critical elections at the traditionally understood time intervals (1860, 1896, and 1932) with the addition of the 1964 election were first evaluated through multiple regression as applied from the classical approach. This was then replicated by applying multiple regression, but from the Bayesian perspective. Given the use of prior information in Bayesian modeling, sets of regressions both including and excluding prior information were run to assess the impact of this added knowledge. This process was applied to both presidential and congressional elections, accounting for the differences in realignment eras for Congress. Results from these models addressed the first set of research questions.

To answer the second set of research questions, different methods of comparison were first evaluated conceptually for their relative strength. Then, the most efficient method was applied to critical realignment theory to see if any additional information can be gained regarding the theory. Using critical

realignment theory as an example, the most efficient method was again applied, but this time the focus was on information gained regarding the two approaches.

Scope and Limitations of Study

The scope of this study was limited to an evaluation of the traditional framework of critical realignment theory. This means that this study took a national, structural approach and was limited to an evaluation of the canon elections with the addition of the 1964 election. Given the national scope, this study only utilized national indicators: presidential election returns and share of U.S. House seats, spanning the period from 1828 to 2008.

Limitations to this study also existed. The most impactful limitation was the conceptualization and operationalization of critical realignment theory from such a viewpoint as the structural one depicted here. As will be demonstrated in the literature review, the viewpoint of the researcher can affect the interpretation of the results. Tangentially, a second limitation was the conceptualization and rationale behind specifying the fourth critical election for congressional analysis at a different timepoint than the presidential analysis. This is based on the developments within the field and is supported by historical data; however, this conceptualization can affect study results. Thirdly, the use of national indicators, as opposed to sub-national indicators, can also bias results. Fourthly, utilizing U.S. election data, this study was reliant upon the accuracy of data gathered by published governmental data sources, such as the Office of the Clerk within the U.S. House of Representatives and the Guide to U.S. elections, published by the CQ Press.

Definition of Terms

Within academic circles pertaining to critical realignment theory, different researchers use different vocabulary to describe different electoral phenomena. One example of this is the use of the term realignment, compared to realigning era, critical election, or even epoch. For clarity within this study, those terms are defined within this section. As demonstrated in Figure 1 below, a critical election is the specific election, either general or congressional, at which the change in party dominance is first seen. A realigning era is the political context or atmosphere in which the critical election occurs; thus it spans time both before and after the critical election. A realignment, then, is the span of time in which the change in party dominance is sustained. This encompasses the critical election, as that is when the change in party dominance is first realized. For the purposes of this study, realignment, era, and epoch are used interchangeably.

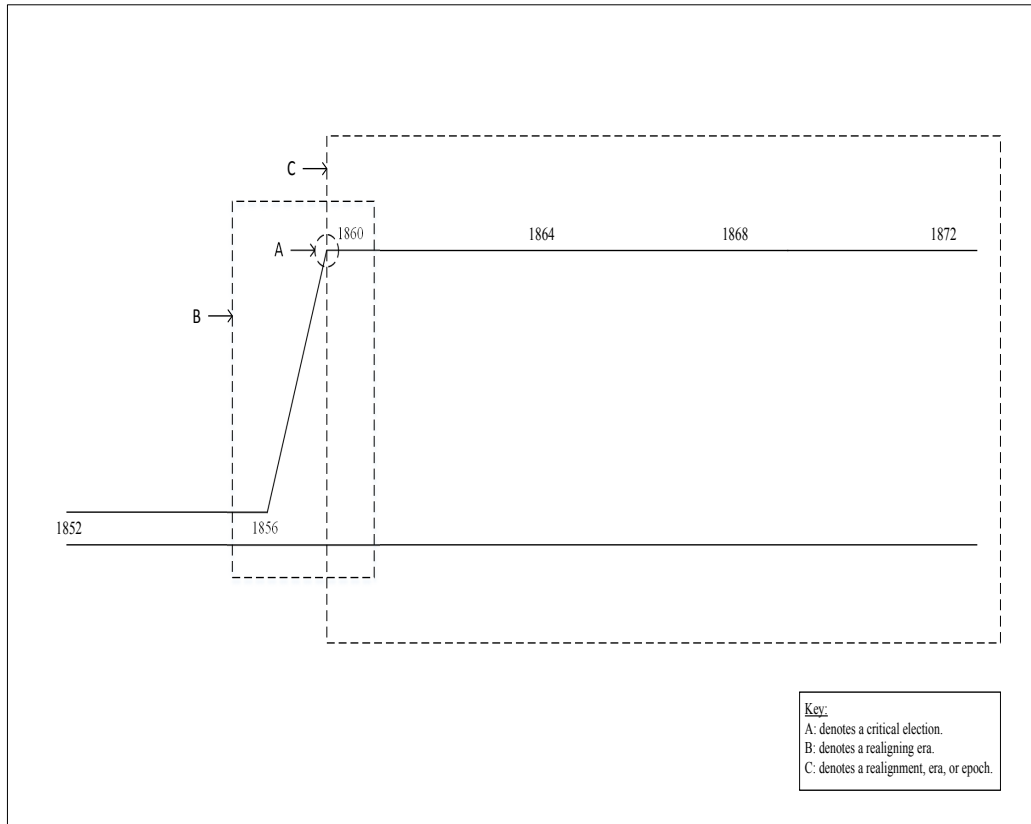


Figure 1. Illustration of terms pertaining to critical realignment theory.

Although this may be self-evident, one other important differentiation pertains to the vocabulary surrounding the different approaches to statistics utilized in this study. References to the Bayesian approach to statistics will be referred to as such; however, the classical approach to statistics is also referred to as the frequentist approach. Presumably, this is due to the understanding of probability within this approach.

Organization of Study

As discussed above, Bayesian modeling had yet to be applied to the question of critical realignment theory. This study sought to fill that gap in knowledge, and by doing so, better reflect the social situations of and inherent

dynamics within elections. To reach this goal, two sets of models were run: one set of models from the classical approach to statistics, and a second set of models from the Bayesian perspective, across which comparisons were made. The importance of this study stems from exploration of the application of Bayesian modeling to the question of critical elections, but also by bringing Bayesian techniques into a new genre within political science. Limitations to this study exist, such as the accuracy of data and the operationalization of critical realignment theory. The next section of this study provides a review of the literature, including overview of the origins and current work in the arena of critical realignment theory. It also includes a description of Bayesian modeling and discusses differing perspectives on indicators of statistical significance in the classical approach. The literature review is followed by a methods section, which describes the data utilized in and analysis plan for this study. Results are then presented, followed by a discussion of the practical significance of such results and of the study as a whole.

Critical Realignment Theory

Critical realignment theory was first introduced to the arena of political science in the late 1950s, enjoying the height of its study through the 1970s. Many researchers today reference V.O. Key as providing the basis of the theory, with Schattschneider (1960), Burnham (1970), Sundquist (1973), and Kleppner (1987) as main contributors (Brunell, Grofman, & Merrill III, 2012; Bullock, Hoffman, & Gaddie, 2006; Campbell, 2006; Stonecash & Silina, 2005). Beginning with a theoretical summary, this section of the literature review starts with a discussion

of the theory from the perspective of these five writers before evaluating practical and current work being completed in this field. This section of the literature review closes by discussing those that disagree with the theory and their rationale for doing so.

The founding theorists conceptualized critical, or realigning, elections from different perspectives, resulting in similar but yet differentiated definitions. Starting from a framework of elections more broadly, Key (1955) perceived critical elections as stemming from a hypothetical typology of elections; thus, these elections were simply one of many types. He did not necessarily advocate for the full development of a typology, but his definition represents this framework. He defined elections as acts of “collective decision,” occurring in a timeline of previous and subsequent behavior. Realigning elections, then, are also acts of collective decision, but where the outcome of the election results in an alteration of party cleavages. Key went further, and stated that the true differentiating feature of realignments is that the sharp change in party lines persists for multiple succeeding elections. This second statement of sustaining new party cleavages within the electorate is seen as necessary, from a definitional standpoint, for all of the other major contributors (Burnham, 1970; Kleppner, 1987; Schattschneider, 1960; Sundquist, 1973).

Kleppner (1987) continued to advocate for Key’s broad framework of a typology of elections, but also extended this perspective and began to link this definition to underlying causes. Following Key, Kleppner perceived critical realignments as partitioning electoral history into times of relative stability, but

also stated that critical realignments should be understood as aggregate-level phenomena that are shaped by any one, or a combination, of several possible patterns of individual behavior. Burnham moved this perspective one step further, shifting the focus from a broader framework of elections to a framework of collective social action. The shift in focus is evident in how Burnham defined critical realignments: “eras ... marked by short, sharp reorganizations of the mass coalitional bases of the major parties which occur at periodic intervals on the national level” (Burnham, 1970, p. 10). He moved from Key and Kleppner’s national viewpoint to a grassroots, coalitional viewpoint, emphasizing the role of the individual in his or her party base.

As previously stated, Burnham retained the aggregated aspect present in Key and Kleppner’s perspectives, but shifted from viewing critical realignments in the macro context of electoral history to viewing critical realignments as movements of the social base of the parties. This shift is evident in Sundquist’s writings, where he also emphasized a grassroots and more humanistic approach to critical realignments. Sundquist (1973) defined critical realignments as an organic change in the party system, where the political norm shifts. Typically, this results in the relocation of the ‘line’ or cleavage between the two party bases, but Sundquist was careful to note that significant shifts in relative party strength can also occur even if the line were to remain fixed. Schattschneider (1960) presented a similar argument, seeing critical elections as changes in political cleavages. He referred to these changes as sectional alliances, but argued that sectionalism actually depresses party organization. This is because sectional alliances can cross

parties and draw new cleavages across the electorate. However, as is evident from this discussion, Schattschneider, Sundquist, Burnham, and to some degree Kleppner, focused on the coalitional, individual, and social aspect of critical realignments, which is differentiated from Key's macro viewpoint of critical elections within the electoral history.

As discussed, slight differences exist in the perspectives of the main contributing writers to the theory of critical realignment. However, these main contributors tended to describe the characteristics of critical elections in the same way. Critical elections are characterized by deeply concerned and highly polarized voters (Key, 1955), where the ideological distance between parties increases (Burnham, 1970). Voter turnout increases and these elections redivide the electorate along new sets of cleavages at the national level (Key, 1955; Schattschneider, 1960), resulting in significant transformations of policy as voting patterns change (Burnham, 1970; Kleppner, 1987). However, this dynamic is also contingent on the size of the group or party, as well as the behavior of other groups or parties (Kleppner, 1987). Furthermore, these changes persist, and are not simply an interruption from the current political system or norm (Burnham, 1970; Key, 1955; Kleppner, 1987; Schattschneider, 1960; Sundquist, 1973).

Driving Factors behind Realignments. From this introduction, it is apparent that critical elections are worthy of study. As Darmofal and Nardulli (2010) state:

The reason for this interest is clear: in contrast to normal voting eras, during critical realignments citizens reject their habitual voting behaviors to hold political elites accountable and forge non-incremental change in

policy outputs. As a consequence, a central concern regarding critical realignments is the identification of the particular changes in voting behavior through which citizens are exerting elite accountability. (p. 256)

One reason for critical realignments is highlighted in this statement—that is, to hold political elites accountable. However, many other ideas have been formed about how realigning elections occur. In this next section, the discussion turns to these ideas: otherwise referred to as the conversion thesis, mobilization thesis, and demobilization thesis.

Conversion Thesis. At its core, the conversion thesis relates to party identification. Committed individuals, for a variety of reasons, change their party attachment from one party to the opposing party (Darmofal & Nardulli, 2010). Such conversion will result in a critical election if it occurs for a large enough number of the electorate (Burnham, 1970; Darmofal & Nardulli, 2010; Kleppner, 1987; Sundquist, 1973; Zingher, 2014). This occurs for the following three reasons: strength of local and state parties, group membership, and the rise of divisive issues. Firstly, the strength of the local and state parties can incite change in party identification (Darmofal & Nardulli, 2010). This is due to the level of activism present within the party at the local level. Secondly, membership in certain social groups can influence individual partisanship. This creates a restructuring of party coalitions whenever the voting behavior of these groups change, and as the ratio of these groups within the electorate change (Zingher, 2014). Thirdly, the rise of divisive issues can also cause changes in party identification. As these polarizing issues enter the arena of political discourse,

tensions within the political system can arise, causing party leaders to become more dogmatic and rigid in party norms, party platforms, and party processes. The rigidity of the established party leadership causes these concepts of party norms, platforms, and processes to become more polarizing instead of integrative. This creates sometimes emotional, but almost always disaffected, voters, which can lead to individual conversion and a change in party identification (Burnham, 1970). These ideas, related to the change in individual partisanship, are generally what constitute the conversion thesis.

Mobilization Thesis. While the conversion thesis focuses on a change in party identification, one could conceptualize the mobilization thesis as individuals gaining a sense of party attachment. The mobilization thesis revolves around the idea of inciting previous non-voters to vote (Darmofal & Nardulli, 2010). The incitement of these previous non-voters is a reflection of the political climate at the time. The high intensity and frequent political stimuli within the current political climate affects these new voters differently than more experienced voters (Andersen, 1979; Beck, 1982; Wanat & Burke, 1982). The voting behavior of this group of voters changes the fabric of the electorate that ultimately can cause a critical election (Sundquist, 1973). These voters generally come from three main populations: the local immigrant population, coming-of-age voters, and inactive voters (Kleppner, 1987; Zingher, 2014).

Demobilization Thesis. Differentiated from the mobilization thesis, the demobilization thesis focuses on the alienation of active, partisan voters. This alienation can occur through a couple of different avenues: firstly, intentionally by

the party system through new regulations on voter registration, such as the increased residency requirements in the late 1920s; or secondly, unintentionally through individual or group disillusionment with the party platform (Kleppner, 1987). Either method results in the same conclusion: previous voters differentially abstain during an election cycle, which offsets the balance of the parties, shifts party cleavages, and results in a critical election. Demobilization, or differential abstention, explains more electoral change prior to the 1950s and 1960s, whereas conversion appears to be the main contributor of electoral change post the Civil Rights era (Shively, 1992).

The conversion thesis, mobilization thesis, and demobilization thesis provide some insight into how critical elections occur. However, changes may not all occur within the same election cycle. The complexity of the American electoral system is too great to assume that the effect of grassroots movements or top-down approaches will be felt immediately within the electorate. Key (1955) realized this, denoting a difference between critical realignments and secular realignments, also referred to as the lingering ‘drift’ toward a different party identification. These two types of realignments are not necessarily distinct; one can think of a secular realignment as the “aftershock” of a critical election (Sundquist, 1973). Furthermore, regions may shift at different points, and different indicators may be affected to varying degrees (Bullock, Hoffman, & Gaddie, 2006). However, given that the basis of critical realignments involves the transformation of the political norm or system, they should be thought of as national, macro events, and thus analyzed as such (Kleppner, 1987).

While these might be macro events, very real, sociological aspects accompany the occurrence of such events. As political cleavages shift or change, new groups of disaffected voters emerge. These are individuals who are dissatisfied with the current political climate, and sometimes so upset that they seek political avenues, such as activism, to have their voice heard (Schofield, Miller, & Martin, 2003). A lack of trust oftentimes undergirds the level of dissatisfaction, aimed at the governing party or the leading candidates in an election cycle (Brooks, 2016). This further cements the political divide, creating a highly polarized, divisive, and hot climate in which social trust must be rebuilt in order to overcome gridlock.

Hypothesized Electoral Eras. Given this theoretical basis of critical realignment theory, the discussion now turns toward previous work completed on the topic. For most researchers, the main question is still whether a realigning election has occurred at specific timepoints, although the traditional “canon” elections are 1860, 1896, and 1932 (Norpoth & Rusk, 2007). The election of 1964 has since been under evaluation as to whether it can be deemed part of the canon, and some researchers do consider it as such. However, this section will evaluate each proposed election in turn, paying particular attention to measurement, method, and results after providing some background information on the political climate at the time.

Considered the first election of the canon (Norpoth & Rusk, 2007), the election of 1860 is considered a realignment for two main reasons: firstly, the electoral returns demonstrated a newly created division among the electorate

(Schofield, Miller, & Martin, 2003); and secondly, the outbreak of the Civil War provided evidence of political shifts internally (Hawley & Sagarzazu, 2012). Regarding the newly created division among the electorate, prior to 1852 Democratic and Whig vote shares were comparable. Neither party dramatically outperformed the other party with reference to general elections. However, in the election of 1860, this dynamic changed. The Whig party candidate, Bell, only won three states, and the two Democratic party candidates took ten states in the South. Of course, the Republican contender, Abraham Lincoln, won a majority of the popular vote in 15 northern and western states, winning the presidency. However, this election illuminated the split between Whig and Democratic party vote shares, suggesting a realignment of electoral support (Schofield, Miller, & Martin, 2003). Current work on this topic, however, challenges this historical account. Using county-level data and analyzing vote transfers through ecological inference models, evidence for a realignment in 1860 is not found (Hawley & Sagarzazu, 2012). Taking a national perspective, some evidence is found in House seats; however, the durability of the realignment is questioned if one accounts for a potential realignment in 1874, if one excludes the South due to the advent of Civil War, and if one considers a realignment to be a shift in party dominance (Norpoth & Rusk, 2007).

The second election of the canon (Norpoth & Rusk, 2007) continued to favor the Republicans. The economic panic during 1893 under Democratic control greatly aided the Republican party, allowing Republicans to propagate prosperity and place blame for unemployment on the Democrats. Such

propagation was highly effective, reducing the linkages between class affiliation (i.e., working class) and party affiliation. Consequently, net movement across at least the New England states was toward Republicans, and only wavered in degree (Key, 1955). Simultaneously, the Populist movement swept the South and threatened to overthrow the current political order. Conservatives within both parties were so concerned and reacted so strongly to this movement that the legacy and southern ties of the Reconstruction arose again, creating a noncompetitive, one-party, sectional 'Solid South.' A similar dynamic happened in the North among the conservative business community. Concerned with the nomination of William Jennings Bryan on a Populist platform, the northern business community sought to work against him (Schattschneider, 1956). In Schattschneider's words, "the resulting alignment was one of the most sharply sectional political divisions in American history" (1956, p. 201). Empirical work on this election, however, provides mixed results. Burnham (1970) utilized regression residuals, systematically moving through comparison pairs of elections within ten-year spans. He compared the average mean difference in residuals over the ten-year spans, and his analysis resulted in the identification of a realignment between 1893-95 and 1927-31. However, replicating his analysis, Stonecash and Silina (2005) disagree with Burnham. They argue that the change was not abrupt, leading these authors to advocate for more evaluation of gradual change when considering realigning eras. Campbell (2006) utilized similar data sources as Burnham, although performed a series of multiple regressions instead of a

residual analysis. He concluded that the election did result in a change in party dominance.

After a period of Republican hegemony in politics, the Democrats made gains in New England states with candidate Alfred E. Smith in the 1928 general election. This was largely due to the mobilization of the local immigrant population, namely low-income, urban Catholic voters. In 1932, Roosevelt retained these gains and sustained the realignment. Key (1955) evaluated the possibility of a realignment circa 1932 by comparing two cities in New England. He demonstrates that while this trend could have started in 1920, evidence for the realignment is most convincing in 1928. In 1920, the difference in Democratic percentage of presidential vote between these two cities was approximately 5 percentage points. However, by 1924, the difference grew to approximately 26 percentage points, and by 1928, the difference in Democratic percentage of the presidential vote measured at 42 percentage points. This illustrates that the campaign of Alfred Smith in the 1928 election created a new cleavage across the electorate, culminating in a critical election (Key, 1955). Looking qualitatively, the election of 1932 also signaled a shift in the conceptualization of party systems. Within the context of the Great Depression, voters used the only political instrument available to them--the Democratic party--to overthrow or cast out the Republican party. This action was not taken because the Democratic party was so well-prepared for the challenge of the Great Depression, but because the electorate was choosing to hold the Republican party responsible, bringing about the advent of the responsible party system (Schattschneider, 1960). This change

was so great that Schattschneider (1960) referred to it as the “revolution of 1932” (p. 206). More recent work on this election supports the conclusion of a critical realignment. Variations in method, such as multiple regression, geographically weighted regressions, ecological inference models, or simply bivariate tests between pairs of elections, do not change this result. Similarly, utilizing national, subnational, or county-level data on presidential vote returns or vote transfers also does not change this conclusion (Brunell, Grofman, & Merrill III, 2012; Campbell, 2006; Darmofal, 2008; Hawley & Sagarzazu, 2012; Kantor, Fishback, & Wallis, 2013).

While the election of 1964 is not a part of the canon of critical realignments as traditionally understood, some have argued for its demarcation as such given the regional importance of the election and its effect on southern white voters (Black & Black, 1992; Carmines, Huckfeldt, & McCurley, 1995; Carmines & Stimson, 1989). At that time, the Republican party was not the favored party of most white southerners given the legacy of the Reconstruction. For the same reason, the Republican party had also attracted many black voters. However this alignment began to change in the 1960s. The Democratic party was becoming more liberal on racial issues as the Republican party was becoming more conservative. The outcome of these dynamics resulted in the Republican nomination of Barry Goldwater, and effectively instituting change in the positions of the two parties on race (Buchanan, 2002; Shelley, Zerr, & Proffer, 2007). One main requirement of a critical election is a change in party loyalties; racial issues of the time provided that impetus (Carmines & Stimson, 1989).

As one can see, the Civil Rights movement carried immense impact on political attitudes, and most researchers agree that a realignment occurred (Buchanan, 2002; Burnham, 1970; Carmines & Stimson, 1989; Feinstein & Schickler, 2008; Schofield, Miller, & Martin, 2003). However, explanations behind the occurrence of the realignment differ among researchers. Some take a more traditional view, arguing that U.S. politics necessitates two dimensions of policy. This means that whatever position presidential candidates adopt, there is always a group of disaffected voters. These voters may be mobilized by third parties, or absorbed into other dominant parties. Realignments are the result of these policy compromises, changes, or stances (Schofield, Miller, & Martin, 2003). Others draw a more complex view, stating that political transformations emerge from the intersection of multiple policy trajectories. For example, the party system was reshaped in the 1930s as the Democrats embraced New Deal liberalism, which then intersected with a second trajectory of civil rights as grassroots activists pushed this issue onto the national scene (Feinstein & Schickler, 2008). A third explanation revolves around issue evolution. Issue evolution is a process by which party coalitions can change, as voting defections among partisans occur and links between citizen and party are broken. These are issues that arise from the old party system and introduce tension into a newly forming party system. These issues capture the public's attention for a longer period of time, and tend to be salient in the minds of voters (Carmines & Stimson, 1989). As is evident, all of these explanations could explain the election of 1964.

Up until this point in this discussion, the trends discussed are fairly consistent across presidential vote returns and congressional distributions of House seats. However, that dynamic changes around the 1960s. While the Republicans made inroads in presidential voting in the 1950s and 1960s as discussed above, this change was not immediately reflected in the distribution of House seats. Between 1954 and 1980, Democrats maintained the majority by 16 percentage points, on average. However, this gap dwindled to 2 percentage points by 1984, and by 1994, Republicans regained majority status in the House for the first time in forty years (Campbell, 2006). Some perceived this delayed Republican victory as a reflection of a long-term shift in party loyalties within the electorate (Abramowitz & Saunders, 1998). Regardless, this is further evidence of how different indicators can yield different results.

Critiques of the Genre. In spite of the discussion above, not all researchers agree with Key's seminal proposal of critical elections. David Mayhew (2002) is a strong critic not only of the empirical work completed, but of the entire genre. His critique is based on three main points: firstly, the validity of the theory; secondly, the "illuminative power of the genre" (p. 35); and thirdly, the lack of relevancy of the theory to the present day. Concerning validity, some researchers, including Mayhew, have found it difficult to replicate and carry out previous work done on critical realignments. While some data are available and replication can be attempted, this is not always feasible (Lichtman, 1976; Mayhew, 2002). Concerning Mayhew's second point, he argues that it has always been obvious that certain elections are more important than others. Consequently,

he is interested in the additive effect of this genre to arena of political science and its contribution to the study of elections. Other researchers agree, stating that periodization of American electoral history is helpful, although they find flaws in the realignment framework. While not the initial intent of Key (1955), many researchers see realignment theory as purely dichotomous: either an election is realigning, or it is not. This dichotomy creates dissonance when attempting to classify elections (Carmines & Stimson, 1989). Another flaw is the sole focus on realignments. Realignments are time-bound and geographically and chronologically constrained. Thus, while rhythms may exist within American electoral history, realignment theory has not given enough consideration to the constraints placed on elections (Shafer, 1991; Silbey, 1991). A third flaw is a gross oversimplification of party change, which has resulted in a constricted view of American political history and a demotivation among political researchers to more fully understand a potentially more intricate pattern of stability and change (Lichtman, 1976).

Mayhew's third point of relevancy also deserves some discussion. At the time of writing his critique, a critical election had not been identified in the last 60 years. He argues that for a cyclical theory, this presents a problem. He hypothesizes that with more advanced survey techniques, parties are able to better understand their party base and supporters and pinpoint the median voter. This has reduced the amount of polarization in general elections, and thus realignments (Mayhew, 2002). However, Sundquist (1973), one of the main contributors, argues that the failure of a realignment in the 1960s was due to the lack of a

triggering event. Regardless, the relevancy of critical realignment theory is questioned as realignments have not occurred as predicted, or at all (Gans, 1985; Mayhew, 2002; Silbey, 1991).

From a review of the literature, more insight is gained regarding the background of the 1860, 1896, 1932, and 1964 general elections. Controversy exists regarding the denotation of these elections as realigning, and current work in this field brings no resolution. Reasons for realignment are also discussed, as well as critiques to the genre and differences between presidential and congressional elections. However, in reviewing methods utilized in evaluating these elections, Bayesian techniques have yet to be applied. The next section provides an introduction to Bayesian modeling.

Bayesian Approach to Statistics

The next section provides an introduction to Bayesian modeling. It begins with some background regarding the origins of the theory, moves through a formal presentation of Bayes' theorem, and discusses components of Bayesian inference and model fit procedures before ending with a comparison to a classical statistics approach.

Origins of the Bayesian Approach. Thomas Bayes was an English minister in the late 1700s. He grew up in Hertfordshire, a southeastern county in England, but gained his education from the University of Edinburgh (Silver, 2012). Despite few publications, Bayes was elected as a Fellow of the Royal Society on November 4, 1742 (The Royal Society, 2017), and likely served as a mediator of intellectual debates. Although published posthumously by friend and

colleague Richard Price in 1763, one of Bayes' more famous works, "An Essay toward Solving a Problem in the Doctrine of Chances," focused on the formation of probabilistic beliefs as new data are encountered.

Bayes was greatly influenced by and a strong advocate of Isaac Newton's work, which suggested that nature follows predictable patterns. Thus, the argument made by Bayes and Price is not that the world is naturally probabilistic, but that one's knowledge is gained through probabilistic means. The example provided by Bayes and Price concerns a caveman: the caveman emerges from a cave, and sees the sun for the first time. He is unsure whether this is a typical occurrence, but as the sun rises each sequential morning, he gains confidence that this is a permanent fixture of nature. In the Bayesian viewpoint, then, learning is done through approximation: as more evidence is gathered, it more clearly reflects truth (Silver, 2012).

Although published by his friend in 1763, Bayes' thoughts on the topic lay dormant for about a decade. Working a little after Bayes, Pierre-Simon Laplace, a French mathematician, independently rediscovered Bayes' mechanism and published his work in 1774. However, due to Price and a visit to Paris, Laplace eventually learned of Bayes' earlier work and credited him with the idea. Regardless, Laplace contributed substantially to the promulgation of Bayes' theorem, as he derived the formal statement of the theory (McGrayne, 2012).

Advancement of the theory persisted; however, concerns about the subjective nature of the prior probabilities led to debate regarding the entire approach. These conversations surrounding Bayesian methods continued, and

given the rise of the ‘classical’ approach promulgated by R.A. Fisher and Karl Popper in the 1920s, Bayesian methods were scarcely taught in universities and even then, it was more so to dismiss the approach. In recent times, however, there has been some return of the method to the university level (Howson & Urbach, 2006).

Mechanics of Bayesian Inference. Using Bayes’ ideas and Laplace’s formalization of the theory, this discussion turns toward an applied discussion of Bayes’ theory. At its core, Bayes’ theorem is a statement of conditional probability. Conditional probability is an expressed degree of uncertainty based on some prior knowledge (Downey, 2012). For example, suppose one is interested to know the probability of an incumbent maintaining his House seat. Using the results of the 2012 House races, an incumbent had a 90% chance of winning his race (Giroux, 2012). However, suppose boundary lines for House districts moved. Now, the probability cannot be appropriately estimated at 90% as districts have changed. Here, the question changes from the percent chance of an incumbent maintaining his seat to the percent chance of an incumbent maintaining his seat, given that district boundaries changed. This second probability is conditional, as it accounts for other factors that are unique to this race. The notation for this is particular probability is $p(A|B)$, which is read as “the probability of A given that B is true” (Downey, 2012, p. 2), or in the context of the example, as the probability of the incumbent maintaining his House seat given that district boundaries changed.

With the concept of conditional probability now presented, Bayes' theorem takes this form (Downey, 2012; Stokes, Chen, & Gunes, 2014):

$$p(\theta|y) = \frac{(p(\theta)p(y|\theta))}{p(y)} \quad (1)$$

and where $p(y)$ is understood as:

$$\int p(\theta)p(y|\theta)d(\theta) \quad (2)$$

and where θ is understood as the unknown parameter of interest, and y represents the observed data. This means that $p(\theta)$ represents the prior distribution, $p(y)$ represents the probability of the observed data, and $p(\theta|y)$ represents the probability of the unknown parameter conditional on the observed data. These topics are addressed more fully in the next paragraph. This theorem is what forms the basis of Bayesian inference. It uses probabilities that are conditional on data to express beliefs about unknown quantities (Downey, 2012; StataCorp, 2015; Stokes, Chen, & Gunes, 2014). The conditional nature of Bayes' theorem means that Bayesian inference has the ability to update beliefs about model parameters by accounting for additional data (Downey, 2012; StataCorp, 2015; Stokes, Chen, & Gunes, 2014; van de Schoot et al., 2013), as model parameters are assumed to be random. It is because of this assumption of the randomness of model parameters that prior knowledge can be incorporated (StataCorp, 2015). If it were assumed that model parameters were fixed—as in the frequentist approach—the addition of prior knowledge would not carry an effect on the parameters or on the analysis. This is a main point of difference between the classical and Bayesian approaches and relates back to the Bayesian understanding of probability. Instead

of understanding probability as a long-run frequency, probability is understood as an expressed degree of uncertainty. Treating probability this way implies that parameters are random, which brings about the necessity or the opportunity for the use of prior distributions. It is the fundamental difference in the understanding of probability that allows for these effects to be seen.

Thus, Bayesian inference has three main components: the prior distribution, the evidence at hand (also referred to as the likelihood), and the posterior distribution. The prior distribution is combined with the evidence at hand to create the posterior distribution (StataCorp, 2015; Stokes, Chen, & Gunes, 2014). The evidence at hand represents the data collected or gathered for the current analysis, whereas the prior distribution is a reflection of prior knowledge about the topic. More specifically, since a prior distribution must be chosen for each model parameter, the variance of the prior distribution reflects the level of uncertainty regarding the population value of that parameter. The larger the variance of the prior distribution, the higher the level of uncertainty (van de Schoot et al, 2013). Please refer to Figure 2 to see this concept displayed visually. The selections of distributions in Figure 2 illustrate various levels of prior knowledge concerning the average math ability for a group of students.

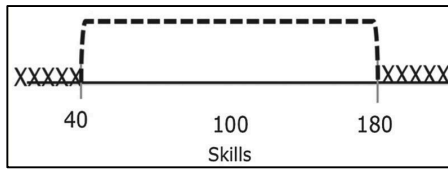


Figure 2a

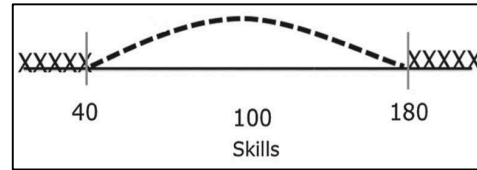


Figure 2b

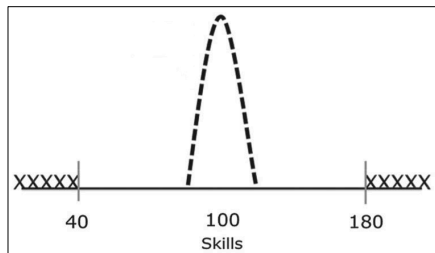


Figure 2c



Figure 2d

Figure 2. Prior distributions illustrating varying levels of uncertainty. Adapted from *A Gentle Introduction to Bayesian Analysis: Applications to Developmental Research*, p. 5, by van de Schoot et al., 2013, Society for Research in Child Development.

Figure 2 displays four prior distributions concerning the average math ability for a group of students. Assuming that this is assessed via a skills test where the range of possible scores is 40 through 180, Figure 2a illustrates a non-informative prior. Each value between 40 and 180 is equally likely to be the mean of the group of students. This represents an assumption that nothing is known about the mean math ability for the given population prior to the start of the study. This is in contrast to informative priors, which are displayed in Figures 2b-2d.

Figure 2b represents the expectation that the mean is 100, as opposed to a very low score or a very high score, but substantial uncertainty exists concerning this expectation because the scores vary from very low to very high. The variance of this distribution is greater than that of Figure 2c. Figure 2c also expects a mean math ability of 100, but with less uncertainty. Figure 2d displays an assumption of higher uncertainty concerning the mean math ability, but expects a lower mean score for the population. As is evidenced by this figure, prior distributions can be more or less informative. However, it is important to realize that while a non-informative prior can appear as more objective, it does not represent complete ignorance about the parameter in question. There is a degree of subjectivity associated with the choosing of any prior distribution (Stokes, Chen, & Gunes, 2014). A second, important classification of prior distributions concerns the degree of conjugacy between the prior and posterior distributions. The prior distribution is considered to be conjugate if it and the resulting posterior distribution are found in the same family of distributions. Conjugate priors are used more frequently for mainly two reasons: firstly, their use simplifies computations; and secondly, the resulting posterior distribution becomes interpretable as additional data and thus can be used to update the analysis as the next prior distribution (Gelman, Carlin, Stern, & Rubin, 1995). However, conjugate priors may not necessarily represent the model parameters realistically and due to the limited number of conjugate priors, the overuse of these distributions limits the flexibility found in Bayesian modeling (StataCorp, 2015).

As previously stated, the likelihood is combined with prior information to create the posterior distribution. Less uncertainty should exist in the posterior distribution, given the inclusion of both prior information and data at hand (Gelman, Carlin, Stern, & Rubin, 1995; van de Schoot et al., 2013). Gelman references this as a “compromise,” stating that the posterior distribution is centered at a point of compromise between the prior distribution and the data, and this point of compromise is increasingly controlled by the data as the sample size increases (Gelman, Carlin, Stern, & Rubin, 1995). See Figure 3 for a visual display of this concept. However, given that the prior distribution and the likelihood are combined mathematically through integrals (Holmes, n.d.), the posterior distribution is obtained via simulation using Markov chain Monte Carlo (MCMC) methods (Stokes, Chen, & Gunes, 2014; van de Schoot et al., 2013). These methods generate a series of samples from the target distribution and compute the posterior estimates of interest using Monte Carlo Markov chains—a numerical integration method that finds the expectation of an integral (Stokes, Chen, & Gunes, 2014). It is from this simulated distribution that point estimates are derived. The posterior distribution reflects all current information known concerning the parameter: the location of the distribution is summarized by the mean, median, and mode, and the variation is summarized by the standard deviation and interquartile range. The mean represents the posterior expectation of the parameter and the mode may be interpreted as the single “most likely” value of the parameter, given the data. These statistics are reported as results of the analysis, in addition to a report of posterior uncertainty, or variance (Gelman,

Carlin, Stern, & Rubin, 1995). Oftentimes, the credibility interval, also referred to as the posterior probability interval (PPI), is reported, which is the counterpart of the frequentist confidence interval (StataCorp, 2015; van de Schoot et al., 2013).

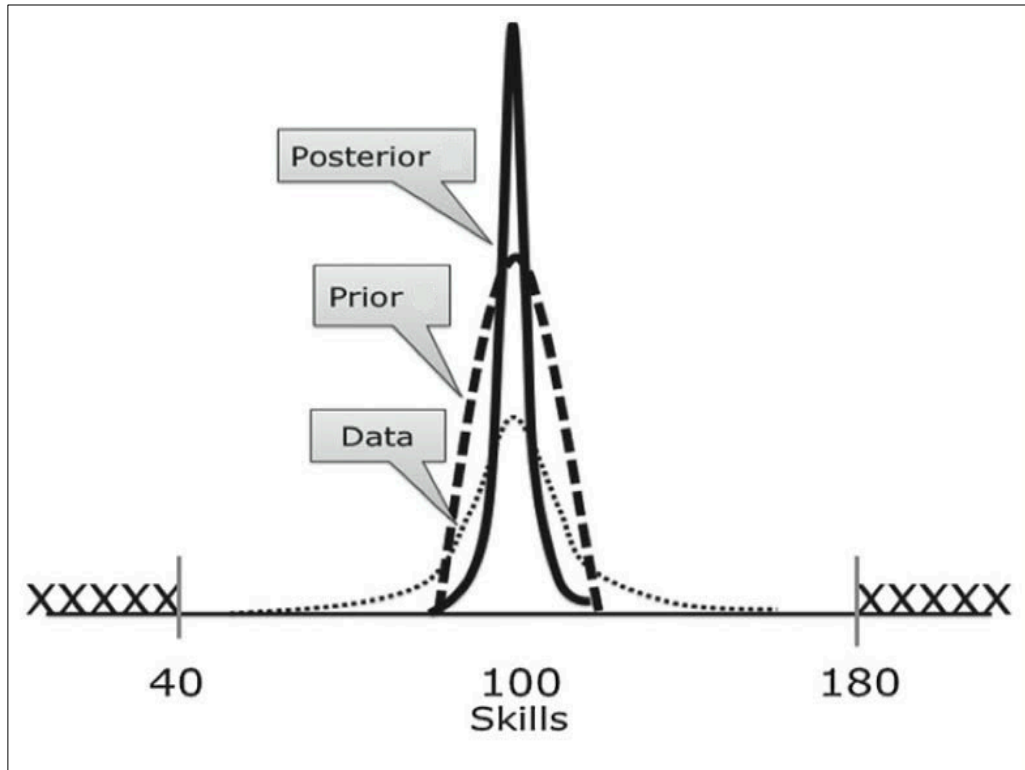


Figure 3. Visual display of the combination of the prior distribution and the likelihood to create the posterior distribution. Adapted from *A Gentle Introduction to Bayesian Analysis: Applications to Developmental Research*, p. 8, by van de Schoot et al., 2013, Society for Research in Child Development.

When comparing two Bayesian models, model fit is also reported. It is traditionally assessed through the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and deviance information criterion (DIC) (StataCorp, 2015; Stokes, Chen, & Gunes, 2014). These fit indices do not provide overall model fit, however, but instead are used as comparative statistics across models. The BIC is more conservative than the AIC, but all three are appropriate for non-informative priors in Bayesian modeling. A difference is that the DIC is designed specifically for Bayesian estimation that involves MCMC sampling. An additional type of fit index, the Bayes Factor (BF), has also been developed and represents the ratio of the marginal likelihoods of two competing models. A Bayes Factor essentially computes the relative probabilities of how well each model fits the data compared to the base model (StataCorp, 2015). While this is helpful when comparing two models directly, it is computationally very difficult—so much so that it led to the development of the other aforementioned indices (Berg, Meyer, & Yu, 2012).

Differences between the Bayesian and Frequentist Approach to Statistics. Given this presentation of Bayesian methods, it is evident that many differences exist between the Bayesian approach and the frequentist approach. These differences are summarized primarily in the expression of probability, the treatment of parameters, and the reporting of statistics. Fundamentally and at its core, the Bayesian approach to statistics views probability as “the subjective experience of uncertainty” (van de Schoot et al., 2013). This is in contrast to the frequentist approach, which perceives probability as the frequency of a particular

event, given repeated sampling. This difference in understanding regarding the nature of event probability provides the basis for the difference in understanding regarding parameters between the two approaches. For the frequentist approach, parameters are understood as unknown but fixed and constant quantities across samples, which are reflected by a fixed parameter associated with some level of error due to sampling (StataCorp, 2015; Stokes, Chen, & Gunes, 2014). That is why repeated random sampling is so important to the frequentist methodologist: the frequentist analysis answers questions based on the distribution of statistics from repeated hypothetical samples, which are generated by the same process. However, a Bayesian analysis seeks to answer questions also based on the distribution of parameters, but conditional on the observed sample. The Bayesian approach assumes that the observed data are fixed; the model parameters are allowed to vary and are treated as random (StataCorp, 2015).

This difference in probability and thus parameters leads to a difference in reporting statistics and subsequently, results. As noted above and unique to Bayesian methods, common summaries of location for the posterior distribution are the mean, median, and mode, and common summaries of variance are the standard deviation and interquartile range (Gelman, Carlin, Stern, & Rubin, 1995). The frequentist approach relies on confidence intervals and hypothesis testing of the model and of parameters. Given that comparisons are oftentimes made between resulting models of the frequentist and Bayesian approaches, the Bayesian approach has incorporated its version of these statistics—also using hypothesis testing and developing posterior probability intervals, the latter of

which are the counterparts to confidence intervals. The interpretation of the frequentist 95% confidence interval is that with repeated sampling and computations of the confidence interval each time, 95% of these intervals will contain the true value of the parameter. Thus, for any given single confidence interval, the probability that the true parameter is in that interval is either zero or one. However, the Bayesian credible interval, or posterior probability interval, provides a range for a parameter such that the probability that the parameter lies in that range is 95%, and not zero or one (StataCorp, 2015; van de Schoot et al., 2013). Regarding hypothesis testing, the frequentist approach answers the question of how likely are the observed data, given that the null hypothesis is true, whereas the Bayesian approach answers the question of how likely is the null hypothesis, given the observed data (StataCorp, 2015).

Methods of Comparison between the Frequentist and Bayesian

Approach to Statistics. Continuing in this vein of comparisons between a classical approach and the Bayesian approach to statistics, very little work has been done in terms of formalizing a method of comparison between the two approaches. Most work on this point takes a very narrow approach; researchers are interested to know about the comparative utility of each approach but limited to their specific case. Consequently, many studies have been done that compare the classical approach to the Bayesian approach, but the method of comparison varies widely based on research design and variable construction. In spite of the variation, attempts at a comparison method can be grouped into three main categories: a

comparison to a known underlying estimate; a comparison in terms of bias; and a comparison of frequentist confidence intervals to Bayesian credible intervals.

Much work on this topic uses simulations, and studies that make comparisons to a known underlying estimate fall into this category. These studies generally first generate a known distribution, reliability estimate, or point estimate, and then evaluate the closeness of results of the different modeling approaches to this known estimate (Betti, Cazzaniga, & Tornatore, 2011; Guikema, 2005). Simulation work is also involved in the second comparison method. In this comparison method, studies use simulations and generate a known distribution or estimate. However, instead of providing a simple, direct comparison to the known estimate, studies that utilize this comparison method calculate the amount of bias in the models. This generally requires a comparison to the known estimate; however, it takes the analysis one step further by evaluating parameter bias, looking at statistical power, or assessing credible intervals (Bennett, Crowe, Price, Stamey, & Seaman, Jr., 2013; Price, 2012). The last method of comparison is more straightforward, as it compares frequentist confidence intervals to Bayesian credible intervals (Liu, Yang, Qiang, Xiao, & Shi, 2012; Stegmüller, 2013). While simulations have been used, they are not required for this method as empirical data can provide the necessary points for comparison. Out of the three methods, this is most likely most direct in situations involving empirical, or not simulated, data.

While methods of comparison fall generally into these three categories, some outliers remain. Based on the variable construction, one study utilized kappa

scores to compare classifications of hospitals. The kappa scores were used to show level of agreement between the classical approach and the Bayesian approach in terms of hospital classification (Austin, Naylor, & Tu, 2000). Another study made no formal comparison and instead, simply compared results qualitatively (Coory, Wills, & Barnett, 2009). This illustrates the wide breadth of methods available to researchers wishing to compare results of the Bayesian and classical approaches. While this may be good for the researcher in terms of flexibility and applicability of method, it also is one point within the field where standardization could occur.

In summary, there are both advantages and disadvantages to the Bayesian approach to statistics. The inclusion of prior information provides not only more balanced results (StataCorp, 2015), but also requires reflection on work already completed in the field (van de Schoot et al., 2013). The Bayesian approach is also seen as more comprehensive and exact as it utilizes the entire posterior distribution of model parameters (StataCorp, 2015), resulting in a more direct expression of uncertainty (van de Schoot et al., 2013). Results are also more intuitive and straightforward in interpretation (StataCorp, 2015; Stokes, Chen, & Rubin, 1995; van de Schoot et al., 2013). However, specifying prior information, given its subjective nature, is seen as controversial by some and increases the complexity of both computations and the model. Essentially, if one is interested in repeated-sampling inference regarding parameters, then the frequentist approach is the more appropriate method. However, if one is interested in the probability that the parameter of interest belongs to some pre-specified interval, then the

Bayesian approach is the more appropriate method (StataCorp, 2015). Due to the nature of elections and the treatment of probability as uncertainty and parameters as random, this understanding illustrates why Bayesian methods are more appropriate for the subject at hand than the frequentist approach.

Misconceptions about Indicators of Statistical Significance

In March 2016, the American Statistical Association (ASA) issued a statement concerning statistical significance and outlined a series of principles to improve the quality surrounding the conduct and interpretation of statistics, particularly highlighting the p -value. Many of these statements codified previous controversial, or at least non-traditional, thoughts concerning the use of p -values in research. This next section addresses those discussions concerning the controversy surrounding and misconception of the p -value.

One point of misconception concerning the p -value is its definition. A correct definition of the p -value is as follows: The p -value represents the probability of observing data as extreme or more extreme than the data collected, under an assumption of no effect or that the null hypothesis is true (Goodman, 1999; Wilkinson, 2014). Oftentimes the researcher misconstrues this definition, potentially due to a misordering of the conditional probability. The p -value provides the probability of $p(D|H_0)$, or the probability of these data, given the null hypothesis, instead of $p(H_0|D)$, or the probability of this hypothesis, given the observed data. The latter represents what most researchers may wish to conclude, although that would be incorrect (Falk & Greenbaum, 1995; Gill, 1999; Gliner, Leech, & Morgan, 2002; Gross, 2015; Kirk, 1996). The p -value does not provide

the probability of either the null or alternative hypothesis being true (Minium, King, & Bear, 1993). This also coincides with the second principle articulated by the ASA: “*P*-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone” (American Statistical Association, 2016).

One point of controversy surrounding the *p*-value is its usefulness as it relates to the null hypothesis. The null hypothesis can be rejected unless the effect is exactly zero with a large enough sample or enough statistical power (Gill, 1999; Gliner, Leech, & Morgan, 2002; Gross, 2015; Kirk, 1996; Meehl, 1978). This detracts from its usefulness, but also highlights that a ‘statistically significant’ result does not necessarily indicate a meaningful result. The strong emphasis on *p*-values has deterred further investigation into more meaningful evaluations of measurement and has blurred the distinction between *p*-values and effect sizes (Gliner, Leech, & Morgan, 2002; Gross, 2015; Rothman, 2014). The ASA summarized it this way, stating that “statistical significance ... does not measure the size of an effect or the importance of a result” (American Statistical Association, 2016).

A second point of controversy relates to the arbitrary nature of the significance level. By setting a fixed level of significance, the researcher turns a distribution of probability or uncertainty into a dichotomous decision of either rejecting or failing to reject the null hypothesis (Kirk, 1996). Creating this dichotomy and this distinction presumes no measurement error, an assumption that very few social scientists would be willing to defend (Gill, 1999). Treating

the significance level in this way can also influence conclusions. For example, presume two researchers obtain identical treatment effects from a set of data; however, one measures at a significance level of .05, whereas the other measures at a significance level of 0.01. Different conclusions are drawn concerning the treatment effect, given the significance level (Kirk, 1996). Furthermore, a small change in a group mean or a regression coefficient can cause a different conclusion to be drawn, which is why some advocate measuring the statistical significance of the difference of two results instead of the difference in significance levels between two results (Gelman & Stern, 2006). The ASA highlighted this discrepancy as well, stating that “scientific conclusions and business or policy decisions should not be based only on whether a p -value passes a specific threshold” (American Statistical Association, 2016).

Lastly, the implications of these misconceptions surrounding the definition and use of p -values are great. Using a misconstrued statistic creates a poor classification system for results, causing researchers to preoccupy themselves with ‘statistical significance’ and ending ultimately with a publication bias (Gross, 2015; Rothman, 2014). This focus on significance has also led to a dichotomous view of relationships that are better handled in quantitative or probabilistic terms (Rothman, 2014). Furthermore, it undermines and inhibits progress in the field. Researchers lose incentive to specify more precise hypotheses and explore competing hypotheses. Likewise, even a fair treatment of the p -value results in an understanding just as narrow, as it provides no further

information than what is already known about the state of the world (Gigerenzer, 1998; Gill, 1999).

From this review, the need to bring a new perspective to critical realignment theory becomes clear. As demonstrated above, current work has only evaluated critical realignment theory through the classical approach to statistics. This is concerning given the misconceptions surrounding p -values and the treatment of probability. However, the Bayesian approach to statistics corrects for both of these concerns, treating probability as the degree of uncertainty and the parameters as random. The next section of this discussion turns to the application of Bayesian modeling to the question of critical realignment theory.

CHAPTER TWO: METHOD

To assess if Bayesian methods improve the usefulness of critical realignment theory in comparison to the frequentist approach, two sets of models were estimated: one pertaining to the presidential popular vote, and one pertaining to U.S. House seats. Each set of models was analyzed using both ordinary least squares (OLS) regression and Bayesian linear regression. The results of these two statistical approaches were then compared. Analyses relating to the classical approach utilized SPSS version 24 and analyses relating to the Bayesian approach utilized SAS 9.3.

Data

Presidential Popular Vote Returns. The data for the first set of models were taken from a compilation of presidential popular vote returns denoted in Congressional Quarterly's Guide to U.S. Elections. For each general election between 1828 and 2008, the total number of votes, the number of votes for the Democratic candidate, the number of votes for the Republican candidate, and the number of votes for both a third party and a fourth party candidate, respectively, were collected. Limiting this analysis to the two major political parties, votes for the third and fourth parties were discarded. This resulted in a sample size of 45 elections.

The dependent variable for this analysis was the Democratic two-party percentage of the presidential vote from 1828 to 2008. While the raw data collected from the Congressional Quarterly included a variable containing the percentage of Democratic vote, this percentage accounted for the third and fourth party voting that occurred. Consequently, the Democratic percentage for the dependent variable was recalculated to appropriately capture only the two-party percentage. There was no concern for floor or ceiling effects on the dependent variable in the presidential analysis or in the congressional analysis as the dependent variable is measured on a scale from 0 to 1, or as a percentage between the values of 0 and 100.

The other variables in the model included a set of five dummy variables and one constant. The set of five dummy variables were coded to identify the five hypothesized epochs or eras in American electoral history. For example, the first epoch is hypothesized from 1828 to 1856. Thus, all general elections within and including the endpoints of that range would be coded as one for the 1828 era variable, whereas all other general elections are coded as zero. This was done for all five epochs, resulting in the five dummy-coded era variables. However, introducing all five of those variables into the model simultaneously would result in perfect multicollinearity. Consequently, the dummy variable for the era to which comparisons are being made was excluded from the model. In its place, however, was a constant, which corresponds to the mean of the dependent variable for that era. It is this constant that allowed for a comparison to the excluded era.

Share of U.S. House Seats. Data for this second set of models were taken from the Office of the Clerk within the U.S. House of Representatives and the Historical Statistics of the United States. For each House election between 1828 and 2008, the number of U.S. House seats held by Republicans and Democrats was gathered. Limiting the analysis to the two major parties, the sum of total seats was calculated. Then, using this calculated sum, the percentage of U.S. House seats held by the Democrats was calculated. This two-party percentage of Democratically-held U.S. House seats was the dependent variable for this set of models. The sample size for this analysis was 90 congressional elections.

Similar to the presidential popular vote models, the other variables in this series of models included a set of four dummy variables and one constant. However, one additional variable was added to the model: a measure of the effect of a general election running simultaneously with the House election. Campbell referred to this as the “on-year presidential surge and the midterm decline” (Campbell, 2006). The variable represented the difference between the Democratic presidential candidate’s vote percentage and 50 percent. It took a positive value in the “on year”, or the general election year, and took a negative value in the midterm cycles. This was to control for the surge and decline effects; between the on-year and midterm cycle, the midterm decline cancels the on-year surge.

While these two sets of analyses are similar, one important difference exists between them. As discussed previously in the literature review and shown visually in Figure 4, the timing of the epochs differ slightly between the

congressional analysis and the presidential analysis. The reasons for this have already been outlined in a previous chapter; mentioning this here is for the coding of the data. For example, for the fourth presidential epoch spanning general elections from 1932 to 1960, all general elections within this timeframe and inclusive of the endpoints, were coded as one. All other general elections, from 1828 to 2008, were coded as zero. For the analysis of the fourth congressional epoch, all midterm and general elections between 1932 and 1994 inclusive of the endpoints, were coded as one. All other general and midterm elections between 1828 and 2008 were coded as zero.

One other unique feature of this type of data relates to the formation of the present-day political parties in the United States. In recent history, mainly two parties, the Republicans and Democrats, have dominated American politics. However, while the basic ideals of these parties have not changed over the years, the names of these parties have. Other labels that have been used with regards to American political parties and affiliation include Whigs, Democratic Republicans, National Republicans, among others. However, these names changes are only relevant for one election within this analysis. For all but the 1912 election, the largest contenders, in terms of percentage of the total vote, were the Democrats and Republicans. However, four candidates ran in the 1912 general election: Woodrow Wilson, representing the Democratic party; Theodore Roosevelt, representing the Progressive party; William H. Taft representing the Republican party; and Eugene V. Debs representing the Socialist party. In this particular election, the largest contenders, in terms of percentage of the total vote, were

Woodrow Wilson and Theodore Roosevelt (CQ Press, 2010). Thus, to remain consistent with using the largest two contenders percentage-wise, this election only compared the Democratic percentage vote to the Progressive percentage vote. Theoretically, this is appropriate as the Progressive party was understood as a more active branch of the Republican party (Milkis, 2012).

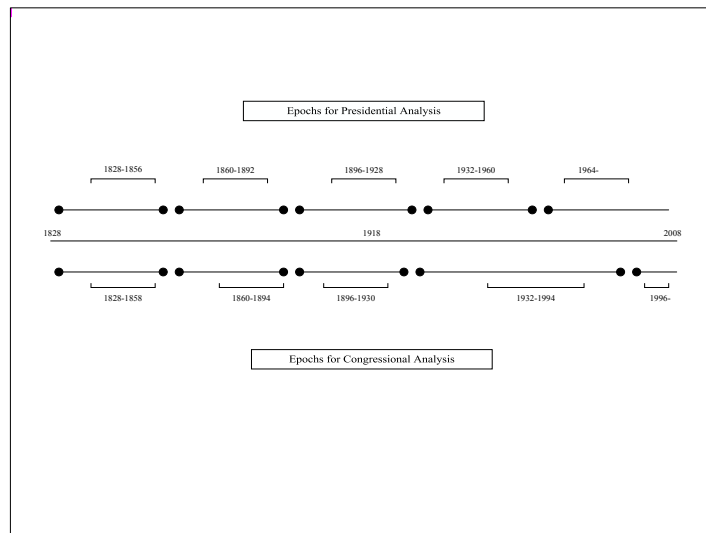


Figure 4. Diagram illustrating the five hypothesized epochs of American electoral history, specifying elections for the congressional and presidential analyses, respectively.

Analysis Plan

The analysis plan for the first set of research questions posed in this study was comprised of three sets of multiple regressions for the analysis of the presidential popular vote and three sets of regressions for the analysis of the U.S. House seats. The first set of regressions within each analysis (i.e., either presidential or congressional) sought to replicate and extend the results found in Campbell's 2006 study to illustrate the contribution of classical multiple

regression to the question of critical elections. The second set of regressions examined the effect on and influence of Bayesian methods with regard to this research question, setting up the model similarly to the first set but specifying a non-informative prior distribution. The third set of regressions mimicked the second set of regressions, but included an informative prior distribution. Each set of regressions had four unique regressions within it; this was to individually test the five different epochs associated with critical elections and realignment theory.

To answer the second set of research questions raised in this study, the relative strength of the comparison methods discussed in the literature review was first assessed. From here, the most efficient method was identified and then applied to the current comparison being made between the frequentist and Bayesian approaches on the topic of critical realignment theory.

Frequentist Approach. To answer the first research question, which examines critical realignment theory from the classical statistical approach, a set of multiple linear regressions was conducted. The full set of regressions evaluates critical realignment theory as a whole, and each single regression compared one baseline era to its corresponding comparison era. Given that there are five hypothesized eras, this resulted in four regressions to test the entire theory.

This procedure had multiple steps. However, before moving forward with the construction of variables and analysis plan, a power analysis was first conducted to ensure enough statistical power existed to carry out the analysis. This is because an underpowered study will cause a true difference in outcomes to go undetected, and an overpowered study will find a meaningless effect. An

acceptable range of statistical power is from 0.80 to 0.90 (Adams-Huet & Ahn, 2009).

Second, after meeting the requirement of statistical power, variables were manipulated as described above, resulting in five predictors in each regression. Those five predictors include: a constant held to the mean of the dependent variable for the baseline era, and four dummy-coded variables representing the remaining four hypothesized eras. Descriptive statistics were also run at this time.

The third step of the process involved checking the assumptions of linear regression. The assumptions of linear regression include independence of observations, independence of errors, normality, linearity, homoscedasticity of residuals, and absence of multicollinearity. Independence of observations implies that each observation is stand-alone; one observation does not affect another observation. The independence of observations is generally assessed through an evaluation of data collection methods. Independence of errors implies non-correlated errors, and the Durbin-Watson statistic is used to assess this assumption. Normality is assessed for each variable in the model, and this is done through kurtosis and skewness statistics. The assumption of linearity speaks to the type of relationship between each independent and the dependent variable, and is typically assessed through an evaluation of observed versus predicted values using scatterplots. Homoscedasticity means constant error variance, and is assessed by evaluating scatterplots of residuals against predicted values. Lastly, the absence of multicollinearity means that independent variables are not highly

correlated with one another, and this is typically assessed through a variance inflation factor.

After assessing these assumptions, the next step in the process was to run the models. All variables were entered simultaneously into the model, and for this analysis, the model took the following form:

$$\hat{y} = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 \quad (3)$$

where \hat{y} represents the Democratic two-party percentage vote for the given election, β_0 represents the mean of the dependent variable for the baseline era, β_1 represents the regression coefficient for the comparison era, x_1 represents the coding of the comparison era, β_2 represents the regression coefficient for the first control era, x_2 represents the coding of the first control era, β_3 represents the regression coefficient for the second control era, x_3 represents the coding of the second control era, β_4 represents the regression coefficient for the third control era, and x_4 represents the coding of the third control era. As stated previously, four regressions were run in order to test each baseline era against its corresponding comparison era.

After running the models, the fifth step in the process was to review the results and then assess model fit. The F -test was used to assess whether the set of independent variables collectively predicted the dependent variable; however, the focus of this study was on the results of the t -tests for the regression coefficients in addition to the value of coefficients to determine the presence and impact of a

realignment. Given the constant and the dummy coding of the eras, one era was tested against the constant at a time. Thus, the resulting coefficient represented the difference in percentage points of the mean of the dependent variable for the baseline era and the comparison era. The t -test for this coefficient then indicated whether the difference is statistically significant. The R^2 value was also reported to assess model fit and evaluate how much variance is accounted for in the model.

Bayesian Approach. After completing this first set of regressions, the research question was then examined from the perspective of Bayesian methods. The models carried the same specification as in the classical ordinary least squares regression, given that a point of interest was to compare the contribution of each method on the topic of critical realignments.

Several steps also existed in the Bayesian analysis. The first step was to choose a probability model for the data. This is similar to choosing a data model in the classical approach. It involves selecting a probability distribution for the data if the parameters of interest were known. If the assumption is made that observations are independent and covariates will be included in the model, then a probability function of the form $p(y_i | x_i, \theta)$ would be used, where y_i are the data values to be observed, x_i is the covariate information, and θ is the vector of unknown parameters. This can take the form of a Bernoulli distribution or a normal distribution, for example.

After selecting the data model, the second step in Bayesian analysis was to select a prior distribution. This distribution represented current knowledge regarding the unknown parameters prior to data being observed. There are two

main types of prior distributions: non-informative prior distributions, also referred to as reference distributions, and informative prior distributions. The non-informative prior distribution is more objective and assigns equal probability to all values of the parameter. This assumes that no prior knowledge exists regarding the parameter. This is in contrast to an informative prior distribution, which assumes some level of knowledge exists about the parameter. This knowledge can be gained through substantive information known by the researcher performing the analysis, through eliciting expert opinion, or through meta-analysis. This step was completed for each unknown parameter. For this particular study, expert opinion was gathered from a researcher in the field who specializes in critical realignment theory. Information from this source was translated into a distribution and used as the prior for the respective parameter.

After selecting the prior distributions, the next step was to observe or collect the data. These data were used to create the likelihood function, or more simply, the likelihood. This likelihood is a joint probability function, and treats the data as fixed quantities. The likelihood is given by:

$$L(\theta|y) = p(y_1, \dots, y_n|\theta) = \prod_{i=1}^n p(y_i|\theta) \quad (4)$$

where θ represents the unknown parameter, and y represents the observed data. Applying this likelihood to the example at hand, one might understand it in the following way: the likelihood of a certain mean difference in the Democratic two-party percentage between two realignment eras is conditional on the observed

election returns, which is equal to the probability of such election returns, given a certain mean difference between realignment eras. This is also equal to the product of the probability of the observed election returns conditional on a certain mean difference for all observed data points. This likelihood function assumes that the data values $y = (y_1, \dots, y_n)$ were obtained independently.

The fourth step in the process was to create the posterior distribution for each unknown parameter by combining the prior distribution with the likelihood, or the data at hand. To do this, Bayes' theorem was applied:

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{\int p(\theta)p(y|\theta) d\theta} = \frac{p(\theta)L(\theta|y)}{p(y)} \propto p(\theta)L(\theta|y) \quad (5)$$

where “ \propto ” means “is proportional to”, θ represents the unknown parameter, and y represents the observed data. This formulation of Bayes' theorem allows the reader to see where the prior information is combined with the data at hand. At the right side of the equation, the likelihood function, $L(\theta|y)$, is multiplied by $p(\theta)$, which represents the prior distribution (Glickman & van Dyk, 2007). This means that the posterior distribution is proportional to the product of the prior distribution and the likelihood function. Here, the phrase “proportional to” implies that one must multiply or divide by a normalizing constant that forces the expression to integrate to one (Feller, 1968; Glickman & van Dyk, 2007). The computation of the integration is shown toward the left side of (5) by $\int p(\theta)p(y|\theta) d\theta$. This integration to a value of one is important because if the

posterior distribution does not integrate to one, then it is considered an improper posterior distribution and an inadmissible solution (Gelman, 2014).

The fifth step was to assess the posterior distribution first for convergence and second for estimates. Since Bayesian inference is dependent upon the formation of this posterior distribution, convergence of the Markov chain Monte Carlo simulations was assessed. This is because the stationary distribution of the Markov chain is the posterior distribution and the lack of convergence means that the parameter space has not been sufficiently explored. A lack of convergence leads to inefficiencies in sampling, as any sampling of the distribution would not approximate the target distribution well. While no one statistic informs the researcher of convergence, typical tests include the Gelman-Rubin, Geweke, and Heidelberger-Welch tests. The Gelman-Rubin test diagnostics rely on parallel chains or simulations to test whether they all converge to the same posterior distribution (SAS Institute, 2016). The Geweke diagnostic compares means from two non-overlapping parts of the chain to see if they come from the same distribution, and the Heidelberger-Welch test calculates a test statistic to test whether the Markov chain is from a stationary distribution (Lam, 2009). To measure the mixing of the chain and dependency among chain samples, correlations between variables and autocorrelation statistics were assessed (Stokes, Chen, & Gunes, 2014). Traceplots were also visually inspected to see if bad mixing occurred at any part of the parameter space (Lam, 2009). Given that the analysis results in a posterior distribution for each model parameter, these metrics were gathered and are provided for each parameter. After confirming

appropriate convergence of the chain, estimates were also reported from the posterior distribution. Typical statistics include the mean, standard deviation, and the 95% credible intervals (van de Schoot & Depaoli, 2014).

The last step in this process was to conduct a sensitivity analysis. This type of analysis assesses the degree to which posterior inferences change when other reasonable probability distributions are used in place of the current prior distribution. As was mentioned briefly in the literature review, comparative model fit is assessed by the DIC statistic. The DIC statistic incorporates both goodness of fit and a penalty term for increasing model complexity. While better model fit results in a larger likelihood value, this is multiplied by -2 which results in an overall smaller value for a better fitting model (Berg, Meyer, & Yu, 2012). This is why the research question was examined with both informative and non-informative prior distributions.

Comparison between the Frequentist and Bayesian Approaches. As mentioned previously, the second set of research questions was interested in formalizing a comparison method between the classical approach and the Bayesian approach to statistics. The first step was to assess the relative strength of the comparison method, evaluating the method on two main points: its applicability to different types of data; and information gained from the comparison, both as it relates to the topic but also to the statistical approach. These two standards were operationalized through a set of indicators: two indicators for the first standard of applicability to different types of data; and five indicators for the second standard of information gained from the comparison.

The two indicators for the first standard were whether the comparison method applies to simulated data and empirical data, and the five indicators for the second standard were as follows: first, whether the comparison method results in a quantifiable component; second, whether the comparison method highlights the meaningful significance of the result; third, whether the comparison method captures the meaning of the approach; fourth, whether the comparison method applies to different types of studies; and fifth, whether the comparison method gathers information which allows for a direct comparison between approaches. After assessing the relative strength by assigning rankings of the comparison methods in this way, the most efficient, or relatively strongest, method was selected.

The next step in this process was to then apply the method to the comparison of critical realignment theory from the classical statistics approach and the Bayesian approach. While this comparison resulted in some kind of measure or statistic, the third step in this process was to then evaluate that statistic for information gained as it relates to critical realignment theory, but also information gained regarding the two approaches to statistics. This third and last step sought to answer the latter two research questions pertaining to the field of research methods and statistics.

CHAPTER THREE: RESULTS

Power Analysis

Prior to running the analysis, a power analysis was conducted for each set of data. The presidential models tested for the average mean difference between the baseline era and the comparison era, while controlling for the other three eras. Consequently, four parameters are estimated in each model. Given four predictors, an alpha level of .05, power set at the recommended level of .80 (Adams-Huet & Ahn, 2009), and assuming a medium effect size of .30, a sample size of 45 is required to detect a significant model, $F(4, 40) = 2.61$. The congressional models also tested for the average mean difference between the baseline era and the comparison era, while controlling for the other three eras and the surge in voting during general election years. Consequently, 5 parameters are estimated in each model. Given five parameters, an alpha level of .05, power set at the recommended level of .80 (Adams-Huet & Ahn, 2009), and assuming a medium effect size of .25, a sample size of 58 is required to detect a significant model, $F(5, 52) = 2.39$. Power was met in both analyses with a sample size of 45 elections for the presidential analysis, and a sample size of 90 elections for the congressional analysis.

Classical Linear Regression Analysis

Presidential Models. Model assumptions were evaluated prior to interpreting model results. First, the presidential dataset was evaluated for outlying and influential elections. This investigation was done by utilizing Cook's D and residual values. By utilizing the Cook's D value, two cases, the 1912 and 1956 elections, were identified as influential. Only the 1912 election had a residual value outside the accepted bounds; consequently, this election was dropped. After dropping this observation, the regressions were rerun and these statistics were again evaluated for additional outliers. After adjusting the Cook's D value for the change in sample size, three additional elections were considered to be influential: the 1936 election, the 1956 election, and the 1964 election. However, only the 1964 election was found to have a residual value outside the accepted bounds. A decision was made to keep this election in the dataset for the following reasons: first, the accepted range for residual values is from -2 to 2, and the residual value of this election was 2.12, only marginally above the cut-off; second, by dropping the 1912 election, the sample size falls from 46 to 45 elections. Given the power analysis, deleting an additional election would create an underpowered study to find a medium effect. Lastly, the 1964 election is one of the elections being tested for a critical realignment. Consequently, it is not surprising that it might appear as an influential data point. Given these reasons, the decision was made to retain this election in the dataset.

After working through an evaluation of influential points, the remaining assumptions of the model were also evaluated. Independence of observations was assumed, as presidential vote returns were recorded at the end of each election and only compiled for this analysis. Autocorrelation was not detected in the presidential data, as indicated by a Durbin-Watson statistic of 1.48. Residual values appeared normally distributed, with a skewness value of -.14 and a kurtosis value of -.39. A linear model was deemed appropriate through inspection of scatterplots of the dependent variable against each independent variable, and the data were found to be homoscedastic by evaluating scatterplots of the residual values against the predicted values. The aforementioned scatterplots can be found in Appendix A. Lastly, there was no indication of multicollinearity, as the tolerance and variance inflation factors fell within accepted bounds.

As previously noted, multiple linear regression analysis was used to develop a model for comparing mean differences in the Democratic percentage of the presidential two-party vote across critical realignments in American electoral history. Basic descriptive statistics and regression coefficients are shown in Tables 1 and 2, respectively. Comparing the first era (1828-1856) to the second era (1860-1892), the mean difference in the Democratic percentage of the presidential two-party vote was found to be statistically significant, $\beta = -0.057$, $t(40) = -2.07$, $p = 0.045$. The model was able to account for 31.3% of the variance in the Democratic percentage of the two-party vote, $F(4,40) = 4.55$, $p = 0.004$, R^2

= .31.¹ Comparing the second era (1860-1892) to the third era (1896-1928), the mean difference in the Democratic share of the presidential vote was not found to be statistically significant, $\beta = -0.052$, $t(40) = -1.89$, $p = 0.07$. Comparing the third era (1896-1928) to the fourth era (1932-1960), the mean difference in the Democratic share of the presidential vote was found to be statistically significant, $\beta = 0.096$, $t(40) = 3.39$, $p = 0.002$. Lastly, comparing the fourth era (1932-1960) to the fifth era (1964-2008), the mean difference in the Democratic percentage of the two-party vote was not found to be statistically significant, $\beta = -0.031$, $t(40) = -1.20$, $p = 0.24$.

The point estimates provided in both the table below and the text above indicate the magnitude of mean difference between the eras noted as measured in percentage points. The sign on the coefficient indicates the direction of the swing in party dominance, with a negative sign indicating a swing toward the conservatives, as it indicates that the Democratic percentage of the vote fell. Likewise, a positive sign on the coefficient indicates a swing toward the liberal side, indicating an increase in the Democratic percentage of the presidential vote. The results above support the conclusion of a critical realignment, based upon statistical significance, in the Republicans' favor in 1860, corroborated by the election of Republican President Abraham Lincoln, and the conclusion of a critical realignment in the Democrats' favor in 1932, corroborated by the election of President Franklin D. Roosevelt of the Democratic Party.

¹ These fit statistics are the same for each presidential model within the classical approach, as only the indicator is changing between runs of the model. Likewise, the sample size, R^2 , adjusted R^2 , the standard error of the estimate, and the Durbin-Watson statistic listed in the table are also consistent across runs of the model.

Table 1

Descriptive Statistics for the Presidential Classical Regression Models

Variables	N	Mean	Standard Deviation
Democratic Percentage of Two-Party Presidential Vote	45	0.49	0.06
Dummy Variables			
1828-1860	45	0.18	0.39
1860-1892	45	0.20	0.40
1896-1928	45	0.18	0.39
1932-1960	45	0.18	0.39
1964-2008	45	0.27	0.45

Table 2

Regression Coefficients for the Presidential Classical Regression Models

Dependent variable: Democratic percentage of two-party vote								
All presidential elections except 1912								
Dummy Variables	(1)		(2)		(3)		(4)	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
1828-1856	-		0.057*	2.07	0.108***	3.85	0.013	0.45
1860-1892	-0.057*	-2.07	-		0.052	1.89	-0.044	-1.60
1896-1928	-0.108***	-3.85	-0.052	-1.89	-		-0.096*	-3.93
1932-1960	-0.013	-0.45	0.044	1.60	0.096**	3.93	-	
1964-2008	-0.044	-1.69	0.013	0.53	0.065*	2.52	-0.031	-1.20
Constant	0.537		0.481		0.429		0.525	
N	45		45		45		45	
R ²	0.313		0.313		0.313		0.313	
Adjusted R ²	0.240		0.240		0.240		0.240	
Standard error of estimate	0.056		0.056		0.056		0.056	
Durbin-Watson statistic	1.480		1.480		1.480		1.480	

* $p < .05$ ** $p < .01$ *** $p < .001$

Congressional Models. The results from the congressional data are addressed next. Similar to the presidential data, the congressional dataset was first evaluated for outliers. An initial evaluation of Cook's D and residual values against the data, using both values as metrics, returned the following elections as outliers: the 1864 election, the 1866 election, the 1890 election, the 1912 election, and the 1936 election. Using only a measure of the Cook's D value added the elections of 1860 and 1894. Since the 1912 election was found to be an outlier in the presidential dataset as well, this observation was first deleted and the analysis was rerun. The other four elections which were found to be influential points remained as such on both indicators when rerunning the analysis.

On a second evaluation of outliers after the deletion of the election of 1912, the 1866 election, the 1864 election, the 1890 election, and the 1936 election remained as influential points on both indicators as mentioned above. At this point, the 1866 election carried the highest absolute residual and Cook's D values; consequently, this election was deleted next from the dataset to see how its removal would affect the other outlying points. After its deletion, the elections of 1864, 1936, and 1890 all remained as influential points as indicated by both residual and Cook's D values. The Cook's D measurement also denoted a series of other elections as influential points; however, these elections did not carry residual values above an absolute value of 2. At this point, the election of 1936 carried the highest residual value and so this election was eliminated to measure the effect on the other outlying points.

A few additional passes were made through the data, but due to concerns of overfitting the model, the decision was made to evaluate outlying and influential points in a different manner. Here, the elections that appeared as influential during the first evaluation of influential points were dummy coded. A dummy variable was created in which the supposed outlying elections were coded as a “1” and all other elections were coded as a “0”. This binary variable was then entered into the regressions and its coefficient was evaluated for statistical significance. In the subsequent regressions, this variable was not found to be a statistically significant predictor, indicating that this group of elections was not statistically significantly affecting the slope of the regression line. Consequently, a decision was made to include these elections for two reasons: first, some of the supposed outlying elections were either located at or near critical juncture points (i.e., the elections of 1864 and 1936); and second, there was some concern regarding statistical power if all influential elections were dropped.

Consequently, after this investigation, assumptions were evaluated. Independence of observations was assumed, as the seat shares were recorded at the end of each election and only compiled for this analysis. Autocorrelation was not detected in the presidential data, as indicated by a Durbin-Watson statistic of 1.22. Residual values appeared normally distributed, with a skewness value of .20 and a kurtosis value of .52. A linear model was deemed appropriate through scatterplots of the dependent variable against each independent variable, and the data were found to be homoscedastic by evaluating scatterplots of the residual values against the predicted values. The aforementioned scatterplots can be found

in Appendix A. Lastly, there was no indication of multicollinearity, as the tolerance and variance inflation factors fell within accepted bounds.

As noted in previous sections, multiple linear regression analysis was used to develop a model for comparing mean differences in the Democratic seat share in the U.S. House across eras of critical realignments in American electoral history. Basic descriptive statistics and regression coefficients are shown in Table 3 and Table 4. Comparing the first era (1828-1858) to the second era (1860-1894), the mean difference in the Democratic seat share in the U.S. House of Representatives was found to be statistically significant, $\beta = -0.112$, $t(85) = -3.33$, $p = 0.001$. The model was able to account for 33.6% of the variance in the Democratic seat share, $F(5,85) = 8.61$, $p < 0.001$, $R^2 = .336$.² Comparing the second era (1860-1894) to the third era (1896-1930), the mean difference in the Democratic seat share was not found to be statistically significant, $\beta = -0.019$, $t(85) = -0.57$, $p = 0.57$. Comparing the third era (1896-1930) to the fourth era (1932-1994), the mean difference in the Democratic seat share was found to be statistically significant, $\beta = 0.15$, $t(85) = 5.19$, $p < .001$. Lastly, comparing the fourth era (1932-1994) to the fifth era (1996-2008), the mean difference in the Democratic seat share was found to be statistically significant, $\beta = -0.114$, $t(85) = -2.80$, $p < 0.001$.

Similar to the presidential models, the point estimates provided both in the table below and the text above indicate the magnitude of mean difference in

² Similar to the presidential models, only the indicator is changing between runs of the congressional model in the classical approach, resulting in consistent fit statistics. This is applicable as well to the aforementioned statistics in the table detailing regression coefficients for the congressional model.

Democratic seat shares in the U.S. House across the specified eras. The sign on the coefficient carries a similar interpretation to that of the presidential models; a negative sign indicates a conservative swing as the average percentage of Democratic seats fell between the two eras, whereas a positive sign indicates the opposite scenario, with a swing toward the liberal side. The results corroborate the presidential analysis, as both the 1860 and 1932 congressional elections demarcated eras of statistically significant mean differences in the Democratic seat share in the U.S. House from the most previous era, and in the same direction as found within the presidential analysis. The one outstanding result is the comparison of the fourth to the fifth era, which did not find statistically significant results in the presidential analysis, but did find such in the congressional analysis.

Table 3

Descriptive Statistics for the Congressional Classical Regression Models

Variables	N	Mean	Standard Deviation
Democratic Percentage of US House Seats	91	0.53	0.12
Midterm Election Surge	91	0.79	11.74
Dummy Variables			
1828-1858	91	0.18	0.38
1860-1894	91	0.20	0.40
1896-1930	91	0.20	0.40
1932-1994	91	0.35	0.48
1996-2008	91	0.08	0.27

Table 4

Regression Coefficients for the Congressional Classical Regression Models

	Dependent variable: Democratic percentage of US House seats
--	---

Dummy Variables	(1)		(2)		(3)		(4)	
	<i>Coefficient</i>	<i>Standard Error</i>	<i>Coefficient</i>	<i>Standard Error</i>	<i>Coefficient</i>	<i>Standard Error</i>	<i>Coefficient</i>	<i>Standard Error</i>
1828-1858	-		0.110**	3.33	0.131***	3.87	-0.019	-0.65
1860-1894	-0.110**	-3.33	-		0.019	0.57	-0.131***	-4.55
1896-1930	-0.130***	-3.87	-0.019	-0.57	-		-0.150***	-5.19
1932-1994	0.020	0.65	0.131***	4.55	0.150***	5.19	-	
1996-2008	-0.100*	-2.15	0.017	0.39	0.035	0.81	-0.114**	-2.8
Midterm Election Surge	-0.001	-1.07	-0.001	-1.07	-0.001	-1.07	-0.001	-1.07
Constant	0.580		0.468		0.449		0.599	
N	91		91		91		91	
R ²	0.336		0.336		0.336		0.336	
Adjusted R ²	0.297		0.297		0.297		0.297	
Standard error of estimate	0.098		0.098		0.098		0.098	
Durbin-Watson statistic	1.221		1.221		1.221		1.221	

* $p < .05$

** $p < .01$

*** $p < .001$

Bayesian Linear Regression Analysis

Specification of Prior Distributions. For the models employing informative prior distributions, the prior distributions for the parameters were assumed to be normal and means and variances were set to the values found in the table below. The means and variances found below were derived from adjusting Burnham's model, as Burnham presupposed an electoral realignment in 1856 and not 1860 and none after 1932. This particularly complicated the variance calculations (M.W. Frank, personal communication, September 16, 2016). For the

models employing non-informative distributions, the default of “COEFFPRIOR=UNIFORM” within the SAS PROC GENMOD command statement was utilized, applying a uniform, or equal probability prior, to all parameters in the model. Code for setting and applying these prior distributions, in addition to all Bayesian analysis, can be found in Appendix B.

Table 5

Means and Variances for the Informative Prior Distributions for the Presidential Bayesian Models

Elections	Mean	Variance
1828-1856	53.0	9.0
1860-1892	48.0	2.5
1896-1928	45.0	30.0
1932-1960	53.0	16.0
1964-2008	50.0	12.0

Table 6

Means and Variances for the Informative Prior Distributions for the Congressional Bayesian Models

Elections	Mean	Variance
1828-1858	55.0	25.0
1860-1894	50.0	30.0
1896-1930	43.0	25.0
1932-1994	62.0	25.0
1996-2008	55.0	25.0

Presidential Models. As previously mentioned, convergence of the Monte Carlo Markov chains is vital for Bayesian analysis. Convergence of the simulation draws ideally achieve a stationary distribution, from which inferences regarding parameters can be made. No one statistic indicates convergence; instead, a series

of diagnostics are evaluated first to assess convergence before interpreting parameters.

For all presidential models, diagnostics provided no evidence that convergence was not achieved. The Gelman-Rubin test, which uses parallel chains with differing initial values to test whether they all converge to the same target distribution, returned similar estimates for each model (i.e., each non-informative model and each informative model) indicating one stationary distribution, respectively. The Geweke test is similar, but evaluates convergence by comparing means from the early and later parts of the Markov chain. Statistically significant results indicate a significant difference in the means, implying a lack of convergence to one stationary distribution. This test returned non-significant results for each presidential model. Visual analysis of the trace plots indicated sufficient burn-in and adequate mixing of the chain. Adequate mixing of the chain was also supported by the autocorrelation graphs and the effective sample size; correlations were low among lagged points and the effective sample size matched precisely the number of Monte Carlo simulations. The Markov chain was deemed long enough through the Heidelberger-Welch test, and accuracy of the percentiles is high, as indicated by the Raftery-Lewis test. This was supported by a dependence factor close to one. Tables detailing these specifics are provided below and the trace plots and posterior distributions for each parameter are provided in Appendix B.

Table 7

Convergence Diagnostics for the Presidential Bayesian Models Using a Non-Informative Prior Distribution

Diagnostic Statistics	Explanation of Test	Accepted Bounds	Model 1: Comparing 1828-1856 to 1860-1892	Model 2: Comparing 1860-1892 to 1896-1928	Model 3: Comparing 1896-1928 to 1932-1960	Model 4: Comparing 1932-1960 to 1964-2008
Gelman-Rubin	Uses parallel chains with differing initial values to assess convergence to same distribution. Failure to do so could indicate a multi-mode posterior distribution.	This is measured by the ratio of within-chain and between-chain variance. A value close to 1 is considered adequate.	Values range from 0.9999 to 1.0003	Values range from 0.9999 to 1.0003	Values range from 0.9999 to 1.0003	Values range from 0.9999 to 1.0003
Geweke	Compares means from early and later parts of Markov chain.	Small p -values indicate rejection.	Lowest p -value is 0.5590	Lowest p -value is 0.5701	Lowest p -value is 0.5701	Lowest p -value is 0.5681
Autocorrelation statistics	Measures dependency among chain samples.	Low correlation between lagged points indicates adequate mixing.	Refer to graphs in appendix.	Refer to graphs in appendix.	Refer to graphs in appendix.	Refer to graphs in appendix.
Effective Sample Size	Similar to autocorrelation; measures mixing of the chain.	Low discrepancy between the effective sample size and the simulation sample size indicates adequate mixing.	Effective sample size is 10,000; simulation sample size equaled 10,000	Effective sample size is 10,000; simulation sample size equaled 10,000	Effective sample size is 10,000; simulation sample size equaled 10,000	Effective sample size is 10,000; simulation sample size equaled 10,000
Heidelberg-Welch	Ensures adequate length of the chain.	Small p -values indicate rejection.	Lowest p -value is 0.1900	Lowest p -value is 0.1800	Lowest p -value is 0.1800	Lowest p -value is 0.1864
Rafferty-Lewis	Evaluates the accuracy of the desired percentiles by reporting the number of samples needed. Failure could indicate the need for a longer chain.	The resulting dependence factor should be close to 1.	Dependence factors range from 0.9904 to 1.0406	Dependence factors range from 0.9904 to 1.0406	Dependence factors range from 0.9904 to 1.0406	Dependence factors range from 0.9904 to 1.0406

**Further mixing of the Markov chains and adequate burn-in are evaluated visually and are presented in the appendix.*

***Table adapted from*

https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_introbayes_sect008.htm#statug_introbayes.bayesess

Table 8

Convergence Diagnostics for the Presidential Bayesian Models Using an Informative Prior Distribution

Diagnostic Statistics	Explanation of Test	Accepted Bounds	Model 1: Comparing 1828-1856 to 1860-1892	Model 2: Comparing 1860-1892 to 1896-1928	Model 3: Comparing 1896-1928 to 1932-1960	Model 4: Comparing 1932-1960 to 1964-2008
Gelman-Rubin	Uses parallel chains with differing initial values to assess convergence to same distribution. Failure to do so could indicate a multi-mode posterior distribution.	This is measured by the ratio of within-chain and between-chain variance. A value close to 1 is considered adequate.	Values range from 1.0001 to 1.0005	Values range from 1.0000 to 1.0005	Values range from 1.0001 to 1.0007	Values range from 1.0001 to 1.0007
Geweke	Compares means from early and later parts of Markov chain.	Small p -values indicate rejection.	Lowest p -value is 0.1603	Lowest p -value is 0.1732	Lowest p -value is 0.1745	Lowest p -value is 0.1553
Autocorrelation statistics	Measures dependency among chain samples.	Low correlation between lagged points indicates adequate mixing, demonstrated graphically	Refer to graphs in appendix.	Refer to graphs in appendix.	Refer to graphs in appendix.	Refer to graphs in appendix.
Effective Sample Size	Similar to autocorrelation; measures mixing of the chain.	Low discrepancy between the effective sample size and the simulation sample size indicates adequate mixing.	Effective sample size ranged from 9,653.10 to 10,268.30; simulation sample size is 10,000	Effective sample size ranged from 9,653.10 to 10,268.30; simulation sample size is 10,000	Effective sample size ranged from 9,653.10 to 10,268.30; simulation sample size is 10,000	Effective sample size ranged from 9,653.10 to 10,268.30; simulation sample size is 10,000

Heidelberger-Welch	Ensures adequate length of the chain.	Small p -values indicate rejection.	Lowest p -value is 0.0771; however, the sample passed the stationary test.	Lowest p -value is 0.0685; however, the sample passed the stationary test.	Lowest p -value is 0.0723; however, the sample passed the stationary test.	Lowest p -value is 0.0767; however, the sample passed the stationary test.
Rafferty-Lewis	Evaluates the accuracy of the desired percentiles by reporting the number of samples needed. Failure could indicate the need for a longer chain.	The resulting dependence factor should be close to 1.	Values range from 0.9586 to 1.0320	Values range from 0.9586 to 1.0320	Values range from 0.9744 to 1.0844	Values range from 0.9744 to 1.0755

**Further mixing of the Markov chains and adequate burn-in are evaluated visually and are presented in the appendix.*

***Table adapted from*

https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_introbayes_sect008.htm#statug_introbayes.bayesess

Since these diagnostics did not reveal any concern with the convergence of the chains, the stationary distributions can be interpreted for parameter estimates. Tables 9 through 16 below summarize the prior and posterior moments of the parameters of each of the models as well as the Deviance Information Criteria (DIC) statistic for each model, which provides a measure of comparative model fit. A comparison of the DIC statistic suggests that the informative model is a better fit when evaluating the first era (1828-1856) to the second era (1860-1892), although the difference in the DIC statistic is small. Consequently, using the informative model, the posterior mean for this difference is -0.052, with a standard deviation of 0.031. Theoretically, this posterior mean indicates that the average Democratic share of the presidential vote fell by 5.22 percentage points from the first to the second era, 1828-1856 to 1860-1892 respectively.

Since model parameters are regarded as random estimates in Bayesian analysis, the conditional probability of parameters carrying a directional effect, given the data, must be estimated from the posterior distribution samples. For example, if one was interested in whether a beta coefficient carried a positive effect on an outcome, one would estimate the conditional probability that the beta coefficient was greater than zero, drawing from the posterior distribution samples. However, in this case, the significance of the covariate is not whether it is greater than zero, but whether it is greater than the average change in Democratic percent of the presidential vote across all U.S. elections. If the average change within the era is significantly greater or lesser than the average change across all U.S. elections, then one may be able to conclude the occurrence of realignment. Consequently, there is a 0.42 probability of the mean difference between the first and second era being greater than the overall average change in the Democratic share of the presidential popular vote.

Comparing the second (1860-1892) era to the third (1896-1928) era for the presidential data, the DIC statistic indicated slightly better model fit for the informative model. The posterior distribution for this difference had a mean of -0.028 with a standard deviation of 0.03. This means that the average Democratic share of the presidential vote fell 2.81 percentage points from the second era to the third era. However, there is only a 0.16 probability that the mean difference for this comparison is greater than the overall average change in the Democratic share of the presidential popular vote.

Comparing the third (1896-1928) era to the fourth (1932-1960) era, again the DIC statistic indicated slightly better model fit for the informative model. The posterior distribution indicated a mean of 0.089 with a standard deviation of 0.03. This means that the average Democratic vote share rose 8.97 percentage points between the third and fourth eras. There is a 0.85 probability, though, that this average mean difference for this comparison is greater than the overall average change in the Democratic share of the presidential popular vote.

Lastly, comparing the fourth (1932-196) era to the fifth (1960-2008) era, the DIC statistic indicated slightly better model fit for the non-informative model. This posterior distribution had a mean of -0.031 and a standard deviation of 0.029. This means that the average Democratic vote share fell 3.05 percentage points between the fourth and fifth eras; however, there is a 0.17 probability that this mean difference is greater than the average change in the Democratic vote share of the presidential popular vote.

Table 9

Prior and Posterior Distribution Information for the Non-Informative Bayesian Presidential Models, Comparing the First and Second Eras

	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1860	-	-	-	-	-	-
1860-1892	0	1.00E-06	-0.0566	0.0302	(-0.1156, 0.0035)	(-0.1172, 0.0013)
1896-1928	0	1.00E-06	-0.0887	0.0305	(-0.1487, -0.0288)	(-0.1477, -0.0280)
1932-1964	0	1.00E-06	-0.0128	0.0316	(-0.0749, 0.0494)	(-0.0730, 0.0509)
1964-2008	0	1.00E-06	-0.0433	0.0286	(-0.0996, 0.0125)	(-0.1011, 0.0107)
Constant	0	1.00E-06	0.5373	0.0220	(0.4940, 0.5798)	(0.4950, 0.5806)

DIC: -118.833

Table 10

Prior and Posterior Distribution Information for the Informative Bayesian Presidential Models, Comparing the First and Second Eras

	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1860	-	-	-	-	-	-
1860-1892	48.00	0.40	-0.0522	0.0306	(-0.1114, 0.0080)	(-0.1132, 0.0060)
1896-1928	45.00	0.03	-0.0843	0.0306	(-0.1442, -0.0235)	(-0.1472, -0.0271)
1932-1964	53.00	0.06	-0.0072	0.0315	(-0.0684, 0.0564)	(-0.0675, 0.0571)
1964-2008	50.00	0.08	-0.0382	0.0287	(-0.0948, 0.0188)	(-0.0952, 0.0177)
Constant	0	1.00E-06	0.5333	0.0221	(0.4898, 0.5771)	(0.4885, 0.5753)

DIC: -119.334

Table 11

Prior and Posterior Distribution Information for the Non-Informative Bayesian Presidential Models, Comparing the Second and Third Eras

	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1860	0	1.00E-06	0.0567	0.0302	(-0.0023, 0.1168)	(-0.0040, 0.1145)
1860-1892	-	-	-	-	-	-
1896-1928	0	1.00E-06	-0.0321	0.0296	(-0.0904, 0.0260)	(-0.0908, 0.0256)
1932-1964	0	1.00E-06	0.0539	0.0307	(-0.0166, 0.1041)	(-0.0146, 0.1058)
1964-2008	0	1.00E-06	0.0133	0.0277	(-0.0410, 0.0676)	(-0.0418, 0.0667)
Constant	0	1.00E-06	0.4807	0.0207	(0.4396, 0.5207)	(0.4412, 0.5218)

DIC: -118.833

Table 12

Prior and Posterior Distribution Information for the Informative Bayesian Presidential Models, Comparing the Second and Third Eras

	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1860	53.00	0.11	0.0606	0.0306	(0.0014, 0.1208)	(-0.0004, 0.1187)

1860-1892	-	-	-	-	-	-
1896-1928	45.00	0.03	-0.0281	0.0297	(-0.0859, 0.0308)	(-0.0859, 0.0307)
1932-1964	53.00	0.06	0.0490	0.0306	(-0.0109, 0.1112)	(-0.0116, 0.1093)
1964-2008	50.00	0.08	0.0180	0.0277	(-0.0368, 0.0731)	(-0.0378, 0.0717)
Constant	0	1.00E-06	0.4772	0.0208	(0.4358, 0.5186)	(0.4339, 0.5158)

DIC: -119.337

Table 13

Prior and Posterior Distribution Information for the Non-Informative Bayesian Presidential Models, Comparing the Third and Fourth Eras

	Prior Information				Posterior Information	
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1860	0	1.00E-06	0.0889	0.0302	(0.0299, 0.1490)	(0.0282, 0.1468)
1860-1892	0	1.00E-06	0.0323	0.0296	(-0.0260, 0.0904)	(-0.0263, 0.0900)
1896-1928	-	-	-	-	-	-
1932-1964	0	1.00E-06	0.0761	0.0307	(0.0157, 0.1363)	(0.0176, 0.1380)
1964-2008	0	1.00E-06	0.0455	0.0277	(-0.0088, 0.0999)	(-0.0096, 0.0989)
Constant	0	1.00E-06	0.4485	0.0207	(0.4074, 0.4885)	(0.4090, 0.4896)

DIC: -118.833

Table 14

Prior and Posterior Distribution Information for the Informative Bayesian Presidential Models, Comparing the Third and Fourth Eras

	Prior Information				Posterior Information	
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1860	53.00	0.11	0.1012	0.0308	(0.0420, 0.1625)	(0.0418, 0.1621)
1860-1892	48.00	0.40	0.0530	0.0303	(-0.0049, 0.1140)	(-0.0051, 0.1136)
1896-1928	-	-	-	-	-	-
1932-1964	53.00	0.06	0.0897	0.0309	(0.0297, 0.1529)	(0.0280, 0.1505)
1964-2008	50.00	0.08	0.0586	0.0281	(0.0041, 0.1149)	(0.0023, 0.1127)
Constant	0	1.00E-06	0.4366	0.0211	(0.3938, 0.4778)	(0.3930, 0.4763)

DIC: -118.872

Table 15

Prior and Posterior Distribution Information for the Non-Informative Bayesian Presidential Models, Comparing the Fourth and Fifth Eras

	Prior Information				Posterior Information	
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1860	0	1.00E-06	0.0128	0.0311	(-0.0479, 0.0746)	(-0.0496, 0.0723)
1860-1892	0	1.00E-06	-0.0438	0.0305	(-0.1039, 0.0162)	(-0.1026, 0.0170)
1896-1928	0	1.00E-06	-0.0761	0.0306	(-0.1365, -0.0155)	(-0.1369, -0.0165)
1932-1964	-	-	-	-	-	-
1964-2008	0	1.00E-06	-0.0305	0.0286	(-0.0869, 0.0252)	(-0.0885, 0.0234)
Constant	0	1.00E-06	0.5246	0.0220	(0.4812, 0.5670)	(0.4830, 0.5685)

DIC: -118.833

Table 16

Prior and Posterior Distribution Information for the Informative Bayesian Presidential Models, Comparing the Fourth and Fifth Eras

	Prior Information				Posterior Information	
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1860	53.00	0.11	0.0257	0.0317	(-0.0352, 0.0886)	(-0.0355, 0.0883)
1860-1892	48.00	0.40	-0.0225	0.0312	(-0.0825, 0.0404)	(-0.0830, 0.0395)
1896-1928	45.00	0.03	-0.0628	0.0309	(-0.1228, 0.0002)	(-0.1235, -0.0011)
1932-1964	-	-	-	-	-	-
1964-2008	50.00	0.08	-0.0171	0.0291	(-0.0737, 0.0409)	(-0.0739, 0.0403)
Constant	0	1.00E-06	0.5121	0.0224	(0.4671, 0.5557)	(0.4684, 0.5568)

DIC: -118.798

Congressional Models. Similar to the presidential models, diagnostics provided no evidence that convergence was not achieved in the congressional models using either a non-informative or informative prior distribution. The Gelman-Rubin test returned similar estimates, again indicating one stationary distribution. Mean differences between the early and later parts of the Markov

chain were not found to be statistically significant by the Geweke test, and visual analysis of the trace plots indicated sufficient burn-in and adequate mixing. Autocorrelation statistics and the effective sample size also indicated adequate mixing of the chain; similar to the presidential models, correlations were low among lagged points and the effective sample size matched the number of Monte Carlo simulations. The Heidelberger-Welch test concluded that a longer Markov chain was not needed, and accuracy of the percentiles was found to be within .005, as indicated by the Raftery-Lewis test. Again, tables detailing these specifics are provided below and the trace plots and posterior distributions for each parameter are provided in Appendix B.

Table 17

Convergence Diagnostics for the Congressional Bayesian Models Using a Non-Informative Prior Distribution

Diagnostic Statistics	Explanation of Test	Accepted Bounds	Model 1: Comparing 1828-1858 to 1860-1894	Model 2: Comparing 1860-1894 to 1896-1930	Model 3: Comparing 1896-1930 to 1932-1994	Model 4: Comparing 1932-1994 to 1996-2008
Gelman-Rubin	Uses parallel chains with differing initial values to assess convergence to same distribution. Failure to do so could indicate a multi-mode posterior distribution.	This is measured by the ratio of within-chain and between-chain variance. A value close to 1 is considered adequate.	Values range from 0.9999 to 1.0002	Values range from 0.9999 to 1.0002	Values range from 0.9999 to 1.0002	Values range from 0.9999 to 1.0003
Geweke	Compares means from early and later parts of Markov chain.	Small p -values indicate rejection.	Lowest p -value is 0.2791	Lowest p -value is 0.2791	Lowest p -value is 0.2791	Lowest p -value is 0.2791
Autocorrelation statistics	Measures dependency among chain samples.	Low correlation between lagged points indicates	Refer to graphs in the appendix	Refer to graphs in the appendix	Refer to graphs in the appendix	Refer to graphs in the appendix

		adequate mixing.				
Effective Sample Size	Similar to autocorrelation; measures mixing of the chain.	Low discrepancy between the effective sample size and the simulation sample size indicates adequate mixing.	Effective sample size ranged from 9,042.50 to 10,000; simulation sample size equaled 10,000	Effective sample size ranged from 9,042.50 to 10,000; simulation sample size equaled 10,000	Effective sample size ranged from 9,042.50 to 10,000; simulation sample size equaled 10,000	Effective sample size ranged from 9,042.50 to 10,000; simulation sample size equaled 10,000
Heidelberger-Welch	Ensures adequate length of the chain.	Small p -values indicate rejection.	Lowest p -value is 0.2736	Lowest p -value is 0.2677	Lowest p -value is 0.2683	Lowest p -value is 0.1835
Rafferty-Lewis	Evaluates the accuracy of the desired percentiles by reporting the number of samples needed. Failure could indicate the need for a longer chain.	The resulting dependence factor should be close to 1.	Values range from 0.9744 to 1.0235	Values range from 0.9744 to 1.0406	Values range from 0.9744 to 1.0406	Values range from 0.9664 to 1.0235

*Further mixing of the Markov chains and adequate burn-in are evaluated visually and are presented in the appendix.

**Table adapted from

https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_introbayes_sect008.htm#statug.introbayes.bayesess

Table 18

Convergence Diagnostics for the Congressional Bayesian Models Using an Informative Prior Distribution

Diagnostic Statistics	Explanation of Test	Accepted Bounds	Model 1: Comparing 1828-1858 to 1860-1894	Model 2: Comparing 1860-1894 to 1896-1930	Model 3: Comparing 1896-1930 to 1932-1994	Model 4: Comparing 1932-1994 to 1996-2008
Gelman-Rubin	Uses parallel chains with differing initial values to assess convergence to same distribution. Failure to do so could indicate a multi-mode posterior distribution.	This is measured by the ratio of within-chain and between-chain variance. A value close to 1 is considered adequate.	Values range from 0.9999 to 1.0003	Values range from 0.9999 to 1.0003	Values range from 0.9999 to 1.0003	Values range from 0.9999 to 1.0004

Geweke	Compares means from early and later parts of Markov chain.	Small p -values indicate rejection.	Lowest p -value is 0.0945	Lowest p -value is 0.0947	Lowest p -value is 0.0948	Lowest p -value is 0.0964
Autocorrelation statistics	Measures dependency among chain samples.	Low correlation between lagged points indicates adequate mixing, demonstrated graphically.	Refer to graphs in the appendix	Refer to graphs in the appendix	Refer to graphs in the appendix	Refer to graphs in the appendix
Effective Sample Size	Similar to autocorrelation; measures mixing of the chain.	Low discrepancy between the effective sample size and the simulation sample size indicates adequate mixing.	Effective sample size ranged from 9,674.10 to 10,218.20; simulation sample size equaled 10,000	Effective sample size ranged from 9,674.10 to 10,218.20; simulation sample size equaled 10,000	Effective sample size ranged from 9,674.10 to 10,218.20; simulation sample size equaled 10,000	Effective sample size ranged from 9,674.10 to 10,218.20; simulation sample size equaled 10,000
Heidelberger-Welch	Ensures adequate length of the chain. Evaluates the accuracy of the desired percentiles by reporting the number of samples needed. Failure could indicate the need for a longer chain.	Small p -values indicate rejection.	Lowest p -value is 0.4462	Lowest p -value is 0.4462	Lowest p -value is 0.4461	Lowest p -value is 0.4469
Rafferty-Lewis		The resulting dependence factor should be close to 1.	Values range from 0.9744 to 1.0152	Values range from 0.9744 to 1.0152	Values range from 0.9744 to 1.0152	Values range from 0.9744 to 1.0320

**Further mixing of the Markov chains and adequate burn-in are evaluated visually and are presented in the appendix.*

***NOTE: Table adapted from*

https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_introbayes_sect008.htm#statug.introbayes.bayesess

Since these diagnostics did not indicate any issue with the convergence of the chains, the stationary distributions can be interpreted for parameter estimates. Similar to the presidential data, Tables 19 through 26 below summarize the prior and posterior moments of the parameters of each of the models. The DIC statistic

is also provided, as it measures comparative model fit. Although the difference in the DIC statistic is small, the DIC statistic suggests that the informative model provides better model fit when evaluating the first era (1828-1858) to the second era (1860-1894). Consequently, using the informative model, the posterior mean for this difference is -0.108, with a standard deviation of 0.03. Theoretically, this posterior mean indicates that the average Democratic share of seats in the U.S. House of Representatives fell by 10.75 percentage points from the first to the second era, 1828-1858 to 1860-1894 respectively. As before, this was compared to the average Democratic seat share across all U.S. elections, and there is a .77 probability that this difference is greater than the average change in Democratic seat share.

Comparing the second (1860-1894) era to the third (1896-1930) era, again, the DIC statistic was slightly smaller for the informative model. The mean of this posterior distribution was -0.014 with a standard deviation of 0.03. This means that the average Democratic seat share in the U.S. House fell by 1.36 percentage points between the second and third eras. There is a .02 probability that this difference is greater than the average change in Democratic seat shares across all U.S. elections.

Comparing the third (1896-1930) era to the fourth (1932-1994) era, the DIC statistic indicated slightly better model fit for the informative model. The posterior mean of this distribution had a value of 0.15 with a standard deviation of 0.03. This means that the average Democratic seat share in the U.S. House grew by 15.48 percentage points between the third and fourth eras. There is a 0.99

probability that this change is greater than the average change in the Democratic seat share.

Lastly, comparing the non-informative and the informative models for the comparison between the fourth (1932-1994) era and the fifth (1996-2008) era, the DIC statistic indicated slightly better model fit for the informative model. The mean of this posterior distribution carried a value of -0.1088 with a standard deviation of 0.04. This indicates that the average change fell by 10.88 percentage points between the fourth and fifth eras, and there is a 0.74 probability that this change exceeds the average change in Democratic seat shares across all U.S. elections.

Table 19

Prior and Posterior Distribution Information for the Non-Informative Bayesian Congressional Models, Comparing the First and Second Eras

Election Years	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1858	-	-	-	-	-	-
1860-1894	0	1.00E-06	-0.1119	0.0342	(-0.1787, -0.0441)	(-0.1756, -0.0421)
1896-1930	0	1.00E-06	-0.1308	0.0342	(-0.1982, -0.0648)	(-0.1982, -0.0649)
1932-1994	0	1.00E-06	0.0192	0.0303	(-0.0398, 0.0787)	(-0.0392, 0.0790)
1996-2008	0	1.00E-06	-0.0952	0.0449	(-0.1844, -0.0076)	(-0.1833, -0.0068)
General Election Surge	0	1.00E-06	-0.0009	0.0009	(-0.0027, 0.0008)	(-0.0027, 0.0009)
Constant	0	1.00E-06	0.5799	0.0248	(0.5314, 0.6278)	(0.5308, 0.6268)

DIC: -156.671

Table 20

Prior and Posterior Distribution Information for the Informative Bayesian Congressional Models, Comparing the First and Second Eras

Election Years	Prior Information	Posterior Information
----------------	-------------------	-----------------------

	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1858	-	-	-	-	-	-
1860-1894	50.00	0.03	-0.1075	0.0343	(-0.1742, -0.0395)	(-0.1675, -0.0422)
1896-1930	43.00	0.04	-0.1251	0.0341	(-0.1924, -0.0577)	(-0.1925, -0.0578)
1932-1994	62.00	0.04	0.0247	0.0306	(-0.0348, 0.0848)	(-0.0344, 0.0852)
1996-2008	55.00	0.04	-0.0840	0.0450	(-0.1715, 0.0050)	(-0.1704, 0.0057)
General Election Surge	0	1.00E-06	-0.0010	0.0009	(-0.0027, 0.0008)	(-0.0027, 0.0008)
Constant	0	1.00E-06	0.5752	0.0249	(0.5256, 0.6232)	(0.5273, 0.6246)

DIC: -156.865

Table 21

Prior and Posterior Distribution Information for the Non-Informative Bayesian Congressional Models, Comparing the Second and Third Eras

Election Years	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1858	0	1.00E-06	0.1120	0.0342	(0.0451, 0.1797)	(0.0482, 0.1818)
1860-1894	-	-	-	-	-	-
1896-1930	0	1.00E-06	-0.0189	0.0331	(-0.0845, 0.0453)	(-0.0860, 0.0434)
1932-1994	0	1.00E-06	0.1312	0.0292	(0.0745, 0.1882)	(0.0745, 0.1882)
1996-2008	0	1.00E-06	0.0167	0.0442	(-0.0713, 0.1027)	(-0.0718, 0.1020)
General Election Surge	0	1.00E-06	-0.0009	0.0009	(-0.0027, 0.0008)	(-0.0027, 0.0009)
Constant	0	1.00E-06	0.4680	0.0234	(0.4219, 0.5131)	(0.4205, 0.5114)

DIC: -156.671

Table 22

Prior and Posterior Distribution Information for the Informative Bayesian Congressional Models, Comparing the Second and Third Eras

Election Years	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1858	55.00	0.04	0.1159	0.0343	(0.0491, 0.1837)	(0.0482, 0.1823)
1860-1894	-	-	-	-	-	-
1896-1930	43.00	0.04	-0.0136	0.0330	(-0.0785, 0.0518)	(-0.0775, 0.0524)
1932-1994	62.00	0.04	0.1362	0.0294	(0.0789, 0.1940)	(0.0797, 0.1947)

1996-2008	55.00	0.04	0.0274	0.0442	(-0.0586, 0.1155)	(-0.0592, 0.1143)
General Election Surge	0	1.00E-06	-0.0010	0.0009	(-0.0027, 0.0008)	(-0.0027, 0.0008)
Constant	0	1.00E-06	0.4637	0.0235	(0.4168, 0.5090)	(0.4161, 0.5078)

DIC: -156.868

Table 23

Prior and Posterior Distribution Information for the Non-Informative Bayesian Congressional Models, Comparing the Third and Fourth Eras

Election Years	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1858	0	1.00E-06	0.1305	0.0342	(0.0636, 0.1984)	(0.0667, 0.2005)
1860-1894	0	1.00E-06	0.0183	0.0331	(-0.0472, 0.0824)	(-0.0474, 0.0820)
1896-1930	-	-	-	-	-	-
1932-1994	0	1.00E-06	0.1497	0.0292	(0.0930, 0.2069)	(0.0931, 0.2069)
1996-2008	0	1.00E-06	0.0353	0.0442	(-0.0528, 0.1213)	(-0.0533, 0.1207)
General Election Surge	0	1.00E-06	-0.0010	0.0009	(-0.0027, 0.0008)	(-0.0027, 0.0009)
Constant	0	1.00E-06	0.4494	0.0235	(0.4031, 0.4947)	(0.4022, 0.4934)

DIC: -156.671

Table 24

Prior and Posterior Distribution Information for the Informative Bayesian Congressional Models, Comparing the Third and Fourth Eras

Election Years	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1858	55.00	0.04	0.1344	0.0344	(0.0675, 0.2024)	(0.0666, 0.2010)
1860-1894	50.00	0.03	0.0234	0.0330	(-0.0415, 0.0887)	(-0.0407, 0.0893)
1896-1930	-	-	-	-	-	-
1932-1994	62.00	0.04	0.1548	0.0294	(0.0974, 0.2127)	(0.0983, 0.2134)
1996-2008	55.00	0.04	0.0460	0.0442	(-0.0400, 0.1341)	(-0.0408, 0.1327)
General Election Surge	0	1.00E-06	-0.0009	0.0009	(-0.0027, 0.0008)	(-0.0027, 0.0009)
Constant	0	1.00E-06	0.4452	0.0235	(0.3982, 0.4906)	(0.3973, 0.4893)

DIC: -156.869

Table 25

Prior and Posterior Distribution Information for the Non-Informative Bayesian Congressional Models, Comparing the Fourth and Fifth Eras

Election Years	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1858	0	1.00E-06	-0.0193	0.0304	(-0.0788, 0.0410)	(-0.0761, 0.0429)
1860-1894	0	1.00E-06	-0.1316	0.0292	(-0.1893, -0.0744)	(-0.1915, -0.0771)
1896-1930	0	1.00E-06	-0.1500	0.0291	(-0.2069, -0.0924)	(-0.2062, -0.0920)
1932-1994	-	-	-	-	-	-
1996-2008	0	1.00E-06	-0.1145	0.0413	(-0.1962, -0.0340)	(-0.1960, -0.0339)
General Election Surge	0	1.00E-06	-0.0009	0.0009	(-0.0027, 0.0008)	(-0.0026, 0.0009)
Constant	0	1.00E-06	0.5993	0.0175	(0.5649, 0.6327)	(0.5652, 0.6328)

DIC: -156.671

Table 26

Prior and Posterior Distribution Information for the Informative Bayesian Congressional Models, Comparing the Fourth and Fifth Eras

Election Years	Prior Information		Posterior Information			
	<i>Mean</i>	<i>Std Dev</i>	<i>Mean</i>	<i>Std Dev</i>	<i>Equal Tail Interval</i>	<i>Highest Posterior Density Interval</i>
1828-1858	55.00	0.04	-0.0178	0.0306	(-0.0772, 0.0425)	(-0.0786, 0.0407)
1860-1894	50.00	0.03	-0.1288	0.0291	(-0.1862, -0.0711)	(-0.1872, -0.0728)
1896-1930	43.00	0.04	-0.1466	0.0292	(-0.2038, -0.0888)	(-0.2044, -0.0900)
1932-1994	-	-	-	-	-	-
1996-2008	55.00	0.04	-0.1088	0.0413	(-0.1894, -0.0264)	(-0.1910, -0.0287)
General Election Surge	0	1.00E-06	-0.0010	0.0009	(-0.0027, 0.0008)	(-0.0027, 0.0008)
Constant	0	1.00E-06	0.5974	0.0175	(0.5626, 0.6310)	(0.5639, 0.6316)

DIC: -156.91

Sensitivity Analysis. A sensitivity analysis was also run to assess the robustness of both the non-informative and informative presidential and congressional models. This is done by assuming a second prior distribution for each parameter. Given the model adjustments noted by the expert opinion and its

potential influence on variance, the means were held constant and the variances were set to the informative prior values to investigate this effect. The results of this analysis are shown in Appendix C.

Looking at the presidential models, the DIC statistic indicated slightly better model fit for this alternative set of models, but the difference in the DIC statistic was small. Changes in the mean of the posterior distribution were contained to less than 0.005 for all but one; the outlier was the comparison between the third and fourth eras, where the informative model returned a mean estimate of 0.089 and the alternative model returned an estimate of 0.076. The DIC statistics were -118.87 and -119.38, respectively.

The congressional models returned similar results, with the alternative models demonstrating slightly better model fit. However, differences between all mean comparisons across the two sets of models remained within 0.005. Practically, this translates to half of a percentage point. This result, and the closeness of the DIC statistic, demonstrates the robustness of the informative model results for both sets of data.

Formal Comparison

This next section addresses the formal comparison between the classical and Bayesian statistical approaches. To aid in understanding, some definitions are provided. There are three parts to this comparison: first, the five methods of comparison that are noted in the literature review are used in an attempt to compare the classical and Bayesian approaches to statistics. These five methods will be referred to as the “comparison methods.” Second, two standards were

developed for the purpose of the comparison: the first one relating to types of data, and the second one relating to information gained from the comparison. These will be referred to as “standards.” Third, to clarify and operationalize the standards, the standards are broken down into indicators. The first standard has two indicators, and the second standard has five indicators. Overall, each of the five comparison methods was measured against all seven indicators. A table detailing this information follows the text to help clarify this formulation.

As mentioned previously, the literature discusses five comparison methods for comparing the classical and Bayesian approach to statistics: a comparison to an underlying known estimate, a comparison in terms of bias, a comparison of frequentist confidence intervals to Bayesian credible intervals, a comparison done through kappa scores, and a qualitative comparison. As outlined in the methods section, each one of these comparison methods was measured against two standards: the general applicability of the method, and as well as a comparison of the information gained from the method. From this, a rank was assigned to each comparison method through a point system and the top ranked method applied to the case at hand.

This next section further explains the standards utilized, followed by an evaluation of the different comparison methods against these standards. Lastly, the top ranked method was applied to the case at hand.

Explanation of Standards. The first standard that the comparison method was measured against was whether the method is applicable to different types of data. Since Bayesian methods utilize simulated data at times, this is an important

consideration concerning the applicability of the method to the social sciences, and for this study in particular. Types of data were grouped into two main groups: empirical and simulated. If the method is applicable to both, the comparison method received two points. Otherwise, one point was awarded for the type of data to which the comparison method applies.

This second standard is focused particularly on information gained from the comparison method, but operationalized through five indicators: first, whether the comparison method resulted in a quantifiable component; second, whether the comparison method carried a component by which to determine the meaningful significance of the result; third, whether the comparison method captures the meaning of the model or statistical approach; fourth, whether the comparison method carries application to different types of studies, which is to be differentiated from the different types of data mentioned in the first standard; and fifth, whether information that allows a direct comparison is gained.

Evaluation of Comparison Methods. As mentioned by Betti, Cazzaniga, and Tornatore (2011) and Guikema (2005), one way to compare the effectiveness of the two statistical approaches is to compare each approach to an underlying known estimate. However, using known estimates limits the data type to simulations, and consequently, resulted in the comparison method receiving only one point on the first standard. Regarding the indicators for the second standard, this comparison method met four of the five indicators. This comparison method resulted in a quantifiable component, namely the distance from the resulting point estimate to the known point estimate (though not a true underlying value), and

this measure of distance can be used to differentiate results and thus produce some level of meaningful significance concerning the results. This comparison method also captured the meaning of each statistical approach, as it measures distance from a resulting point estimate or the mean of a distribution to the known underlying point estimate. It also resulted in congruent information across statistical approaches, as the unit of measurement is distance and this can be compared across approaches. However, this comparison method did not meet the fourth indicator of applicability across different types of studies, as it does not work when the underlying point estimate is not known.

Bennett, Crowe, Price, Stamey, and Seaman, Jr. (2013) and Price (2012) built on the first comparison method, describing a comparison method in which the amount of bias present in the model parameters is calculated. As this comparison method is also based on a known underlying value, it only works with simulated data. With regard to the second standard, this comparison method met three out of the five indicators. Since the outcome of this comparison method would be the amount of bias present in model parameters, this method resulted in a quantifiable component that provides a level of meaningful significance to the result. This comparison method also reflected the meaning of the statistical approach, as it measures the amount of influence on model parameters. However, it was not found to be applicable to different types of studies, as the underlying point estimate must be known, and congruent information was not presented between models because model parameters are treated fundamentally differently across statistical approaches.

The third comparison method between the two statistical approaches was a comparison of frequentist confidence intervals to Bayesian credible intervals. Concerning the first standard, this comparison method was applicable to both empirical and simulated data. Regarding the second standard, this comparison method met four out of the five indicators. Since both intervals can be represented numerically, the outcome of the comparison would also be numeric and thus quantifiable in nature. While these intervals did not meet the fifth indicator of providing congruent information, traditionally these intervals are seen as counterpart measures across approaches and so some level of meaningful significance as it pertains to the models and approaches could be discerned. This comparison method is applicable to different types of studies and does capture the meaning of the appropriate statistical approach as each interval measures the interval for the presumed underlying point estimate.

Although not as prevalent in the literature as the first three comparison methods, this fourth comparison method utilized a standard measure across statistical approaches. This fourth comparison method utilized kappa scores, a method of inter-rater agreement for categorical variables, and compared the level of percent agreement. As for the first standard, kappa scoring can be utilized with both empirical and simulated data. As for the second standard, this method met three out of the five indicators. This method carries a quantifiable component, but also provides a level of significance of results across models. While congruent information is gained, as the percent agreement is directly comparable across models, this comparison method is not applicable to all studies, as it deals mainly

with classification of categorical variables, and does not ultimately capture the meaning of the statistical approach as it does not deal with the underlying point estimate.

The final comparison method is only indirectly discussed in the literature, mainly through a discussion of results from comparison studies or simulations between the classical and Bayesian approaches to statistics. This final comparison method was a qualitative comparison, trying to provide a more shaded picture of the information gained from the models. As for the first standard, this method can be applied to both empirical and simulated data. As for the second standard, this comparison method met four out of the five indicators. There is no quantifiable component, but the qualitative description of the results of the model does speak to the approach to statistics and also provides some level of meaningful significance regarding the model. A qualitative comparison is applicable to any type of model and could easily provide congruent information, in addition to specific points of differences that might not be highlighted in a quantitative comparison.

Given this evaluation of the standards and as displayed in the table below, two comparison methods were found to be most efficient as they pertain to this comparison process. Consequently, those two methods were applied to the present example.

Table 27

Rankings of Comparison Methods for Comparisons between the Classical and Bayesian Statistical Approaches

Comparison Method	First Standard: Applicability to different types of data		Second standard: Measurement of information gained from comparison					Total Points
	<i>Simulated Data</i>	<i>Experimental Data</i>	<i>Has a quantifiable component</i>	<i>Has an ability to differentiate results in terms of importance</i>	<i>Captures the meaning of the model</i>	<i>Carries application to different types of studies</i>	<i>Whether information that allows a direct comparison is gained</i>	
Comparison 1: Comparison to an underlying known estimate	Yes	No	Yes	Yes	Yes	No; this method is only applicable to simulated studies	Yes	5
Comparison 2: Comparison in terms of bias present in parameters	Yes	No	Yes	Yes	Yes	No; this method is only applicable for studies in which one can induce bias and thus limited to simulated studies	No; this method does not provide comparable information because model parameters are understood differently across approaches	4

<p>Comparison 3: Comparison between frequentist confidence intervals and Bayesian credible intervals</p>	<p>Yes</p>	<p>Yes</p>	<p>Yes</p>	<p>Yes</p>	<p>Yes</p>	<p>Yes</p>	<p>No; this method does not provide directly comparable information because confidence intervals and credible intervals represent different concepts</p>	<p>6</p>
<p>Comparison 4: Comparison of kappa scores between the two methods</p>	<p>Yes</p>	<p>Yes</p>	<p>Yes</p>	<p>Yes</p>	<p>No; this method looks at only the congruence between percentage correctly classified, which does not reveal anything particular about the overall statistical approach</p>	<p>No; this method only works with classification studies</p>	<p>Yes</p>	<p>5</p>
<p>Comparison 5: Comparison completed qualitatively, focusing on interpretation</p>	<p>Yes</p>	<p>Yes</p>	<p>No; this method is focused on discussing the differences in interpretation</p>	<p>Yes</p>	<p>Yes</p>	<p>Yes</p>	<p>Yes</p>	<p>6</p>

Application of Comparison Methods. The comparison numerically between confidence intervals and credible intervals is shown in the table below. Two additional columns are part of this table; one column indicates statistical significance, and the other table indicates the comparison probability to the baseline utilized earlier in this section. With this view, a few findings can be seen. First, the results across the models were quite similar. Looking more closely at a comparison between the means and the point estimates, differences in the congressional comparisons ranged from 0.0045 to 0.0054. The presidential models contained a much larger range between the two models: differences in corresponding means and point estimates ranged from 0.0005 to 0.02.

Second, statistical significance appeared to correlate with the Bayesian probabilities. This was more pronounced in the congressional results; however, it appeared that high probabilities correlate with statistical significance or low p -values. One interesting result was the comparison between the first and second presidential eras: the classical approach returned a p -value of .045, which would fall under the standard alpha level of 0.05. However, the associated probability was 0.42, which indicates that this mean was only approximately as likely as the average change in the Democratic two-party presidential vote. A second interesting result was the saliency of this pattern in the congressional results. Of course, the p -values are more extreme, but so are the probabilities.

Table 28

Differences in Point Estimates and Frequentist Confidence Intervals and Bayesian Credible Intervals for the Presidential Models

Comparison Eras	Difference in Point Estimate	Differences in Interval Bounds		Differences in Interval Range	Associated <i>p</i> -value	Associated Probability
		<i>Minimum Value</i>	<i>Maximum Value</i>			
1828-1856 to 1860-1892	-0.0048	-0.0006	-0.0090	-0.0084	0.0450	0.4179
1860-1892 to 1896-1928	-0.0239	-0.0211	-0.0268	-0.0057	0.0660	0.1556
1896-1928 to 1932-1960	0.0063	0.0093	-0.0009	-0.0102	0.0020	0.8458
1932-1960 to 1964-2008	-0.0005	0.0039	-0.0042	-0.0081	0.2380	0.1667

Table 29

Differences in Point Estimates and Frequentist Confidence Intervals and Bayesian Credible Intervals for the Congressional Models

Comparison Eras	Difference in Point Estimate	Differences in Interval Bounds		Differences in Interval Range	Associated <i>p</i> -value	Associated Probability
		<i>Minimum Value</i>	<i>Maximum Value</i>			
1828-1858 to 1860-1894	-0.0045	-0.0048	-0.0055	-0.0007	0.0010	0.7664
1860-1894 to 1896-1930	-0.0054	-0.0045	-0.0058	-0.0013	0.7180	0.0214
1896-1930 to 1932-1994	-0.0048	-0.0054	-0.0057	-0.0003	0.0000	0.9920
1932-1994 to 1996-2008	-0.0052	-0.0066	-0.0066	0.0000	0.0070	0.7405

Next, the two statistical approaches were evaluated qualitatively as they pertained to these results. Looking through this second comparative lens, two main points can be made, but with one major caveat. First, and again, the results are similar. This is seen in closeness of the means and point estimates and in the closeness of the confidence intervals and credible intervals. Second, the closeness of these results could lead one to conclude that with some margin of error, both

approaches would arrive at the same conclusion. In this case, this is true.

However, there is one major caveat to this viewpoint: the interpretations of each mean, point estimate, credible interval, and confidence interval are very different between these two approaches. First, the mean references the mean of a distribution, which means that it is the average point for this random parameter. However, the point estimate has a very different interpretation: this estimate is understood as the underlying true value. Consequently, while these values may appear to be the same, their interpretation is very different.

This same argument applies to the credible intervals and confidence intervals. While these measures are considered counterparts in the literature, these two intervals do not represent the same concept. Confidence intervals, with an alpha level of .05, are typically explained as 95% of the intervals contain the true estimate. However, this means the answer as it pertains to any one confidence interval is binary: either the interval contains the point estimate or it does not. This is differentiated from the Bayesian credible interval, where a 95% Bayesian credible interval is interpreted as 95% of the posterior distribution lies within that particular region. Consequently, while the actual numeric values are close, the interpretation and lens across these measures is again very different.

The main point of the comparison of results from a frequentist and Bayesian perspective is: while the results are similar, the interpretation of the results leads to very different conclusions because they stem from very different philosophical viewpoints. With the current comparison methods available, a comparison done only numerically may not necessarily highlight this point. While

a qualitative comparison as this may be included in a discussion of results, it is important to recognize that this comparison still may be the most efficient way to compare these statistical approaches.

CHAPTER FOUR: DISCUSSION

The results presented in the previous chapter underscore the importance of continued analysis and discussion of critical elections, from both a political science perspective and a statistical perspective. This chapter extends that discussion, paying particular attention to the research questions posed at the beginning of this study. This section begins with a review of the results, explains their significance, and then turns toward the implications and limitations of this research before posing avenues for future research.

Summary of Results

Utilizing the classical statistical approach, the beta coefficients of two of the four election years tested were found to be statistically significant in the presidential analysis, which was also found for the congressional analysis. The first year, 1860, saw a reduction in the Democratic two-party percentage of the presidential vote of almost 6 percentage points with the election of Abraham Lincoln. Likewise, Congress saw an 11 percentage point reduction in Democratic seat share within the same election cycle. The second election year, 1932, elected Franklin D. Roosevelt to the presidency with almost a 10 point swing in the Democratic percentage of the presidential vote. Again, Congress confirmed this pattern with a rise in the Democratic seat share of 15 percentage points when compared to the average Democratic seat share of the previous era.

The congressional analysis also returned a statistically significant result when comparing the fourth and fifth eras. This was not seen in the presidential analysis; however, a number of reasons might explain this result. First, the demarcation for the fifth era was different between the presidential and congressional analysis. This is due to literature surrounding the context of election cycles in Congress and the presumption of a gradual shift instead of a strong swing in one election cycle (Campbell, 2006). Second, while both outcomes are national indicators, the presidential vote may mask a national swing in party dominance due to the level of aggregation, whereas enough disaggregation may exist in seat shares such that a swing may still be noticeable. Third, in spite of the different demarcations of the congressional eras, the timing of the presidential election may still not capture the swing in political climate and dominance. Given that congressional elections occur more frequently, this may better display the political climate and geography at the time.

Generally, the results of the Bayesian analysis yielded similar conclusions. For the congressional analysis, there was a 76.64% probability that the change seen between the first and second era was greater than the average change seen in Democratic seat share between any two congressional election years, and there was a 99.20% probability that the change seen between the third and fourth eras was greater than the average change seen between any two congressional election years. The presidential results also returned a high probability that the change between the third and fourth eras was greater than the average change seen in the Democratic two-party percentage of the presidential vote at 84.58%. The only

exception to this statement was the presidential election of 1860. This presidential election had an only 41.79% probability that the change seen here between eras was greater than the average change in the Democratic two-party percentage of presidential vote seen between any two election years. However, this finding did coincide with the results of the classical presidential analysis; while the resultant p -value was statistically significant, the p -value fell right at the cut-off for significance with a value of 0.045.

The last portion of the results section sought to formally compare the results of the two statistical approaches. This was done by applying a set of standards uniformly across five different comparison methods that were noted in the literature. The five comparison methods were comparisons to known estimates, comparisons in terms of bias present in parameters, comparisons between frequentist confidence intervals and Bayesian credible intervals, comparisons between kappa scores, and a qualitative comparison between the two approaches. After ranking these five comparison methods on a set of standards, two comparison methods were found to be most efficient: first, the comparison between confidence intervals and credible intervals; and second, the qualitative comparison. Each of these comparison methods has their strengths. The interval comparison carries a quantifiable component whereas the qualitative comparison allows for all differences to be highlighted and discussed at length. Applying these comparison methods to the topic at hand, the comparison between intervals was very close numerically, but did not highlight easily that the two point

estimates demonstrate very different concepts. The latter point was the primary objective of the second qualitative comparison.

Importance of Findings

Consequently, the results of this study yield two main findings: first, critical elections were found in 1860 and 1932; and second, the most efficient methods of comparison between the classical and Bayesian statistical approaches are a comparison of confidence and credible intervals and a qualitative comparison. The first main finding is discussed, followed by a discussion of the second finding. Uncovering critical elections in 1860 and 1932 is important for a couple of different reasons. First, the identification of critical elections indicates shifts in the American electorate and political climate. This means that the American electorate is not static and is responsive to different stimuli within the political climate, whether that be economic change, a change in demographics, or a shift in the ideological stances of the parties (Lodge, Steenbergen, & Brau, 1995).

Further, this identification of critical elections and the ability to replicate some of the results discussed by Campbell (2006) discredits the argument made by Mayhew (2002). Mayhew (2002) argued that the empirical validity of some previous work on the topic was questionable, as he and others were not able to replicate prior results. The study completed in this work utilized more electoral history than Campbell's 2006 work, which could explain a lack of complete congruence in results; however, the partial replication of this work defies Mayhew's validity claim. Mayhew also claimed that creating a typology of

elections only served that purpose--that there is no greater reason for such classification, and in fact, the classification creates a narrow perspective through which to view American electoral history (Mayhew, 2002). One could argue, however, that the identification of such elections actually carries the opposite effect. By identifying critical elections and then closely evaluating the historical context, factors which carry a motivating effect on the electorate might be realized and then utilized to forecast future eras of realignment within American electoral history (Carmines & Wagner, 2006).

Continuing in this vein, Mayhew's claims were not discredited by only one statistical approach. The closeness of the results across approaches serve as confirmation for the elections that were identified as realigning. Obviously differences in interpretation between approaches exist; however, the same substantial result was achieved. Extending this point further, not only are results confirmed across approaches, but this comparison also indicates that the identification of critical elections is not dependent upon the classical approach to statistics. The ability to apply a secondary approach to the same topic and find the same results is not only confirmation of those results, but adds credence to the theory.

Other contributions of this study to the field exist. One contribution of the classical approach is the addition of election years to previous work, such as Campbell's 2006 work, and partial confirmation of his results. A second contribution of the classical approach is the accessibility of the approach and subsequent results, as Bayesian methods, until recently, were limited due to a lack

of inclusion in standard statistical packages (Peck, 2015) and the dominance of frequentist methods in statistics programs (Bolstad, 2002). Alternatively, Bayesian methods are able to incorporate a more direct expression of uncertainty through the use of prior distributions (van de Schoot et al., 2013) and also produce a more precise result, providing a numeric probability instead of a statement concerning statistical significance (Lilford & Braunholtz, 1996). While not all researchers present results in this way, the focus on statistical significance can lead to incorrect conclusions (Kirk, 1996) and a dichotomous view that may be better explained through probabilistic terms (Rothman, 2014). However, while both approaches offer contributions, one could argue that Bayesian methods perform better at identifying critical elections for two reasons: first, Bayesian methods better capture the social science situation under study; and second, Bayesian methods provide a probability which is perhaps more easily interpretable than a p -value (Kruschke, 2011).

The second main finding of the work presented here concerns comparisons between the classical and Bayesian statistical approaches. The results discussed above identified a comparison between confidence and credible intervals and a comparison done qualitatively as the most effective. This finding is important for several reasons. First, there is no formalized way to compare the classical and Bayesian approaches. Some theorists have attempted a compromise between the two approaches, but none have been adopted due to a lack of justification from either perspective (Berger, Boukai, & Wang, 1997). Consequently, this work

contributes to the field by drawing attention to this need, illustrating its importance, and also providing a starting point for the conversation.

Tangential to this first point, the comparison between frequentist and Bayesian approaches highlights their philosophical differences. Some comparison methods captured the meaning of the model, or accounted for the philosophical difference, but then information was either not directly comparable or the method only worked on simulated data. The only exception to this rule was the qualitative comparison. This demonstrates that not only has a formal comparison not been developed, but a solid formal comparison needs to account for the differences in perspective across approaches.

One other main contribution of this study is the systematic review of comparison methods and their relative strengths. As previously mentioned, the five methods utilized were the following: first, a comparison to a known, underlying point estimate; second, a comparison in terms of bias present in the models; third, a comparison of frequentist confidence intervals to Bayesian credible intervals; fourth, a comparison of kappa scores; and fifth, a qualitative comparison. Beginning with a comparison to a known, underlying estimate, the relative strength of this method is the differential amount between the result of the model and true point estimate. One is able to know with complete certainty the true distance between the two values. Likewise, the relative strength for the comparison in terms of bias is that the true amount of bias in each of the parameter estimates is known. The actual effect of inducing bias in the model is known, again, with complete certainty. The fourth method of comparison noted

above is the comparison of kappa scores. The relative strength of this method is that information on the same scale and with the same implications is compared. This is obviously the highest standard; however, this method is only applicable to classification studies. The comparison between the two intervals is quantifiable and easily reported from analyses using the models. It can also be applied in empirical settings, and carries applicability to different types of studies. Lastly, comparing the two approaches qualitatively also carries a few relative strengths. First, incorporating a qualitative comparison allows for complexities pertaining to the models to be identified and discussed in full, such as qualifying statements that might not be immediately noticed or incorporated into other methods of comparison. Likewise, a qualitative comparison also allows for the discussion to reflect the uniqueness of the statistical approach--again, a characteristic that may not be available with other comparisons.

Implications for Theory and Practice

A few implications exist for the field as a result of this study. First, the presence of statistically significant results not only demonstrates that the electorate is not static, but also supports continued study of the theory. Carmines and Wagner (2006) have begun this conversation, acknowledging the critiques concerning the usefulness of the theory, but purport that looking beneath the elections themselves to the evolution of issues over time may further enhance the perspective of realignment theory.

A second implication for both theory and practice stems one main difference between the two statistical approaches: the treatment of the parameters

as either fixed or random estimates. This study highlights the need for researchers to be aware of this particular difference in the treatment of parameters. It not only asks that researchers be aware and knowledgeable about this difference, but to also ensure that the method applied appropriately captures the targeted understanding of the final estimate. Likewise, the choice of approach is not only realized in the final result, but can also be evident in the statement and operationalization of the research question. Bayesian methods allow for other questions to be asked of the data (Austin, Naylor, & Tu, 2000; Kruschke, 2011), which influences model development and operationalization of the question.

A final point of consideration for theory and practice is the prior distribution utilized in the model. While Bayesian analysis is useful as it makes uncertainty explicit (Coory, Wills, & Barnett, 2009), one point of contention is the use of prior distributions in Bayesian models. Some researchers note the difficulty in setting prior probabilities (Chang & Boral, 2008), and proponents of the frequentist approach argue that misspecified prior distributions can alter results and induce bias in the model (Bennett, Crowe, Price, Stamey & Seaman, Jr., 2013). Proponents of the Bayesian approach counter that ignoring prior information on the topic also biases results. Consequently, the use and specification of prior distributions is not only a contentious topic of discussion for researchers between the two camps, but is also an area that the Bayesian researcher should treat carefully.

Study Limitations

Limitations exist to the results presented here. First, national datasets were used. Although the decision to use national datasets was based on other research on this topic, the use of such data presumes a national realignment. Realignments could occur regionally, by state, or even on a local level. The use of national data may mask the presence of lower level realignments, and a realignment, for instance, in the North may not be realized by a direct and opposite reversal in the South. Consequently, a more full investigation should be done, and would involve testing for realignments at other levels of government.

A second limitation to the results presented here is the particular definition for “critical election” utilized within this study. This study utilized a definition by Brunell, Grofman, and Merrill, III (2012) of a decisive and durable shift in party dominance. Other definitions of critical elections exist, and the operationalization of these definitions could also affect the outcomes of this study.

A third and intertwined limitation is the demarcation of critical elections utilized within this study. One major assumption made within this study is that critical elections have occurred only at the elections tested in this study. While the choice of 1860, 1896, 1932, and 1964 were based on previous research, this study is limited in testing only those elections and did not seek to uncover any other potential elections. Related to this discussion is the demarcation of the fifth congressional era. Instead of spanning the same timeframe as the presidential analysis, the fourth congressional era was expanded, consequently reducing the number of elections in the fifth congressional era. This decision was based on

literature and the situational understanding of a gradual change in Congress; however, this delineation of eras could potentially alter results.

A fourth limitation to the study presented here relates to the data and methods utilized. This study utilized secondary data, collected from national sources. However, any data misreported in those sources, or any subsequent miscoding by the researcher, could influence results. Likewise, not only the use of the particular definition of critical election discussed here, but also its operationalization and how that affects the choice of method could also alter results. This study was highly influenced by Campbell's 2006 work, as it was the most comprehensive work on the topic of critical elections. However, choices made by the researcher as it pertains to data sources, years included, and methods utilized could also influence results.

The operationalization of the research question also affects the Bayesian analysis undertaken here. Since critical realignment theory requires a durable shift in party power (Brunell, Grofman, & Merrill III, 2012), this limits analyses to a retrospective look. This perspective, combined with the current operationalized definition of critical elections, greatly hinders the application of Bayesian methods. While still applicable, the strength of Bayesian approach is seen in the method's ability to account for prior information. However, due to the operationalization presented here, this strength is not seen to its fullest extent.

Avenues for Future Research

Given these limitations, many avenues for future research exist. This study assumed a certain definition of critical elections; one could investigate the effect

of operationalizing that definition differently. Likewise, this study also tested only specific, presupposed critical elections; one could test other elections to investigate whether other realigning eras exist. This study also used national data; consequently, looking at a different level of government might also influence results. The perspective of issue evolution within and across elections could also be investigated. However, the major avenue for further research stemming from this study is the development of a metric through which to compare the classical and Bayesian statistical approaches. As was denoted in this study, comparison methods exist between the two approaches; however, many of these methods do not directly compare both approaches, taking into account the basis of the approach. More research is needed to explore whether such a comparison is possible, given the different treatment of probability.

CHAPTER FIVE: CONCLUSION

This study began from one main point of difference: the treatment of probability between the classical approach and the Bayesian approach to statistics. The difference is substantial: in the classical approach, probability is treated as a long-term frequency of a particular event occurring, whereas in the Bayesian approach, probability is viewed as a degree of uncertainty. In theory, this is simply a difference in definition; however, in practice, the implications of that difference require more thought regarding the applicability of the approach. The classical approach, due to its definition of a long-term frequency, asks that conditions remain precisely the same among random replicates of the measured phenomenon. Social science situations rarely meet this requirement, and elections are no different. The change in voters, issues, and candidates impede the actualization of this requirement, leading to an inadequate application of the classical approach to statistics in this setting. This situation is then only compounded by the controversy that exists over set significance levels. As these levels are set arbitrarily and in particular, are interpreted dichotomously (i.e., either the result is statistically significant or not), the results of these analyses are limited in their interpretation and, thus, in practical significance. Consequently, this main difference between statistical approaches requires a reevaluation of

social science phenomena, and in this particular case, critical elections, from a new perspective.

Based on this difference and its implications for practice, this study sought to understand the question of critical elections through the Bayesian statistical approach. The purpose of this was two-fold: first, prior to this study, Bayesian modeling had not been applied to the study of critical elections in American electoral history; and second, critical elections had not been evaluated in this way, meaning that new information pertaining to critical elections might be uncovered. This study sought to expand Bayesian methods to this topic, but also see what further information could be gained regarding critical elections through this viewpoint. This purpose resulted in the statement of six research questions: the first three relating to the field of political science and the latter three relating to the field of research methods and statistics. The first three questions looked at the occurrence of critical elections from the classical approach and the Bayesian approach, evaluating the contribution of each approach to the identification of such elections. The latter three questions focused on a comparison between approaches, discussing the relative strengths of comparison methods derived from literature in addition to information gained from the application of the most efficient method to the comparison at hand.

The scope of this study was limited to an evaluation of the traditionally accepted critical elections: 1860, 1896, and 1932. The election of 1964 was added to the analysis based on evidence found in literature. The study as a whole took a national perspective, and elections between 1828 and 2008 were evaluated. This

section will first provide a summary of the study in its entirety, followed by some concluding thoughts.

For many researchers, V.O. Key (1955) is credited with deriving the base of critical realignment theory. He perceived elections as acts of “collective decision”, and perceived critical elections as ones where the result is a sharp change in party lines which persists for subsequent elections (Key, 1955). Other influential researchers in the field, such as Schattschneider (1960), Burnham (1970), Sundquist (1973), and Kleppner (1987), followed in his footsteps, each providing their unique contribution to the field of critical realignment theory. In the end, we find that critical elections are characterized by highly concerned and polarized voters (Key, 1955) at times when the ideological distance between parties grows (Burnham, 1970). Oftentimes critical elections result in new party lines, splicing the electorate at a national level along current, highly politicized issues (Key, 1955; Schattschneider, 1960).

Such elections can occur for a variety of reasons, but those reasons are generally grouped into three main categories: the conversion of voters, the mobilization of voters, or the demobilization of voters (Darmofal & Nardulli, 2010). The conversion thesis posits that a critical election will occur if enough voters change their party attachment from one party to the opposing party. This can happen for a number of reasons, such as the strength of the local and state parties, group membership, or the rise of divisive issues (Burnham, 1970; Darmofal & Nardulli, 2010; Kleppner, 1987; Sundquist, 1973; Zingher, 2014). The mobilization thesis explains a rise in voter turnout, as previous non-voters

become inclined to vote. Oftentimes the high intensity of the political climate, potentially due to divisive issues, encourages political participation of this group which subsequently changes the current balance of parties and can cause a critical election (Andersen, 1979; Beck, 1982; Wanat & Burke, 1982). Lastly, the demobilization thesis posits that a critical election can occur based on the alienation of once active, partisan voters. This occurs most predominantly through the disillusionment with one's party platform (Kleppner, 1987).

Arguably, through these theses, critical elections have occurred at specific timepoints throughout American electoral history. The traditionally accepted elections are 1860, 1896, and 1932. Given literature on the political climate at the time, the election of 1964 was also added to the analysis presented in this study. These elections were the ones evaluated throughout this work. However, not all researchers agree with the inclusion of these specified elections, or with a theory of a cyclical realignment throughout American electoral history. Mayhew (2002) argues against the entire genre on three main points: first, the validity of the empirical work completed; second, the added benefit of the genre; and third, the lack of relevancy to the present day. He takes issue with some of the empirical work, as he and others have not been able to replicate it, but more so he argues that there is little benefit to identifying critical elections. He argues that not all elections are equal—some are more important than others, but to classify elections in such a way as this creates a useless dichotomy. This dichotomy does nothing to propel the genre forward and instead constrains more interesting work on the topic. His third point is that critical elections have not followed the pattern

prescribed; consequently, while there may have been a cyclical look earlier in history, the lack of a known critical election in recent history illustrates the lack of relevancy of the genre.

As stated previously, these elections were evaluated through both the classical statistical approach and the Bayesian statistical approach. The classical approach is more known, and is also referred to as the frequentist approach. Although an older technique, the Bayesian approach has laid dormant for a number of years due to a lack of computing power. Bayesian methods are predicated on the idea of conditional probability, and uses those probabilities, conditional on data, to express beliefs about unknown quantities. There are three main components to Bayesian methods--the prior, the likelihood, and the posterior, and it is the prior distribution, or previously known information concerning an event, that is combined with the evidence at hand to create the posterior distribution. For this study, the posterior distribution was obtained via Markov chain Monte Carlo simulations, and it is the mean and standard deviation of that posterior distribution that serves as the point estimate.

To assess whether Bayesian methods provide any improvement over the classical approach as it pertains to the identification of critical elections, the elections of 1860, 1896, 1932, and 1964 were tested through both the classical approach and the Bayesian approach. Two sets of analyses were run; one set of models utilized the Democratic two-party percentage of the presidential vote as the outcome variable, and the second set of models utilized the Democratic seat share within the U.S. House of Representatives as the outcome variable.

Secondary data were used, and were gathered from the Office of the Clerk within the U.S. House of Representatives, the Historical Statistics of the United States, and the CQ Press' Guide to U.S. Elections. In order to test the hypothesized critical elections, eras were formed with the hypothesized critical election beginning the next era. For example, within the presidential analysis, the first era spanned from 1828 to 1856, the second era spanned from 1860 to 1892, the third era spanned from 1896 to 1928, the fourth era spanned from 1932 to 1960, and the fifth era spanned from 1964 to 2008. It should be noted that the fifth congressional era began with 1996 instead of 1964. From here, the data were coded to represent the appropriate era and regressions were run to assess the mean difference between the current era and the most previous era.

To further understand any improvements of the Bayesian approach over the frequentist approach, a formal comparison between the two approaches was also conducted. A review of the literature identified five methods of comparison used by other researchers. These methods included a comparison to a known estimate, a comparison in terms of bias, a comparison of frequentist confidence intervals and Bayesian credible intervals, a comparison of kappa scores, and a qualitative comparison. These different methods were ranked based on developed criteria and the highest ranked methods were applied to the question at hand.

The analysis of the presidential data found two critical elections, one in 1860 and one in 1932. The sign on the coefficient was of the expected direction in both cases, corroborated by the election of Republican President Abraham Lincoln in 1860 and the election of Democratic President Franklin D. Roosevelt

in 1932. The congressional analysis supported these results, and also found a critical election in 1996. Furthermore, the Bayesian models served to confirm these results, as most of the probabilities of these results being greater than the average change in that outcome variable were relatively high. The one outlying result was the first critical election in the presidential analysis. For this 1860 election, there was a 41.79% probability that the change seen between the first and second eras was greater than the average change in the Democratic two-party percentage of the presidential vote. However, there was an 84.58% probability that the change seen between the third and fourth eras was greater than the average change in the same outcome variable. For the congressional analysis, there was a 76.64% probability that the change seen between the first and second eras was greater than the average change in Democratic seat share, and there was a 99.20% probability that the change seen between the third and fourth eras was greater than the average change seen in the same outcome variable.

As discussed, methods of comparing the frequentist and Bayesian statistical approaches were also evaluated and ranked according to a developed set of standards. This evaluation resulted in two comparison methods being applied to the topic at hand: first, the comparison between frequentist confidence intervals and Bayesian credible intervals; and second, the qualitative comparison. The first comparison method resulted in very similar point estimates, but did not account for the difference in understanding between the intervals developed with the frequentist and Bayesian analyses. The second comparison method was able to account for the differences between these two methods, but did not provide a

numerical comparison. Overall, results were very similar, both numerically and in the identification of critical elections; however, it is to be remembered that the interpretation of these intervals is very different.

Returning to the research questions, the first three questions were focused on the contribution of each statistical approach to the genre of critical realignment theory as well as noting any improvements of the Bayesian approach over the classical approach as it pertains to the identification of critical elections. From this study, the contributions of the classical approach to the field of critical realignment theory are the expansion and extension of Campbell's work with the addition of nearly 50 years of elections, but more so the accessibility of results. Due to the familiarity of the classical approach and the presentation of results in typical studies, the accessibility of such results is high and is certainly a benefit and contribution of the classical approach. Turning toward the Bayesian approach, the main contribution of Bayesian methods to the theory is the precision with which results are presented. Instead of a dichotomous statement regarding statistical significance in the classical approach, Bayesian methods yield a probability that represents a comparison between the simulated point estimates from the posterior distribution and a predetermined baseline. However, as to whether this aids in the identification of critical elections is a moot point; the classical approach does offer a more concise method.

The latter three research questions focused on the contribution of this study to the field of research methods and statistics, and more specifically, on the formal comparison method between the two approaches. As discussed previously,

five comparison methods were found in the literature, and the fourth research question evaluated each comparison method for its relative strength. After ranking these comparison methods by their relative strength, the fifth research question surrounded the application of that method to the question at hand. Through the comparison of confidence and credible intervals with an additional qualitative comparison, it was noted that the results, both numerically and substantively, were very similar. However, it is important to note that while similar numbers were returned for each set of intervals, the interpretation of these intervals is very different.

The last research question was focused on gathering further information regarding the two statistical approaches from the completion of these comparisons. This question highlighted the increased precision from the Bayesian approach. This is seen in both the comparison between confidence and credible intervals and the probability associated with the mean estimates. In the classical approach, 95% of confidence intervals contain the true parameter. This means that with any one interval, either the true parameter falls within the interval or it does not. In the Bayesian approach, the 95% credible interval consistently demarcates a region that contains 95% of the probability distribution. This leads to a statement of greater certainty in the Bayesian approach. Likewise, instead of using an arbitrarily set significance level in the classical approach where the answer is typically interpreted in a dichotomous fashion, the Bayesian approach has the flexibility to return a probability associated with a certain point estimate. Regardless, the overarching contribution of this study is to be aware, informed,

and cognizant of the differences in interpretation and more so the treatment of parameters between the two methods. This carries strong implications for researchers and practitioners, as choosing a method which does not describe the practical situation being modeled may lead to inadequate conclusions.

Based on this summary, a few conclusions can be reached. First, while philosophical differences are apparent between the Bayesian and classical statistical approaches, the conclusions of this study result in the same substantive outcome. Interpretations differ between the two approaches, but it is still considered as confirmatory that both methods resulted in the same critical junctures. Second, while comparisons between statistical approaches can be made, much room exists for work to be done as it pertains to the development of a formal, comparative statistic. As presented earlier, some methods currently exist, but these do not capture the philosophical differences that exist between the two approaches. This carries great implications for interpretation.

Finally, the last main conclusion of this work is that critical elections are a part of the United States' electoral history. As Darmofal and Nardulli (2010) note, these types of elections are important as they serve to hold political elites accountable. Consequently, understanding their occurrence and frequency becomes a near necessity. While Mayhew (2002) may argue that critical elections have lost relevancy and are no longer a part of the United States' electoral cycle, the election of President Obama in 2008 particularly followed by the election of President Trump in 2016 may provide evidence counter to that claim. While it is too early to fully analyze these elections utilizing the methods discussed here,

preliminary evidence may suggest a shift. Ball's comment that the "old [Republican] party establishment went into exile, perhaps never to return" (2016) may suggest a new political cleavage, and his comments regarding Republicans leaving the party could provide further evidence of conversion or demobilization. Regardless of the occurrence of a 'true' critical election, the prospect is certainly intriguing for political scientists and demands further study.

BIBLIOGRAPHY

- Abramowitz, A. I. & Saunders, K. L. (1998). Ideological realignment in the U.S. electorate. *The Journal of Politics*, 60(3), 634-652.
- Adams-Huet, B. & Ahn, C. (2009). Bridging clinical investigators and statisticians: Writing the statistical methodology for a research proposal. *Journal of Investigative Medicine*, 57(8), 818-824.
- American Statistical Association. (2016). *American statistical association releases statement on statistical significance and p-values* [Press release]. Retrieved from <https://www.amstat.org/newsroom/pressreleases/P-ValueStatement.pdf>.
- Andersen, K. (1979). *The creation of a Democratic majority, 1928-1936*. Chicago, IL: University of Chicago Press.
- Austin, P. C., Naylor, C. D., & Tu, J. V. (2000). A comparison of a Bayesian vs. a frequentist method for profiling hospital performance. *Journal of Evaluation in Clinical Practice*, 7(1), 35-45.
- Ball, M. (2016, May 4). The day the Republican party died. *The Atlantic*. Retrieved May 21, 2016 from <https://www.theatlantic.com/politics/archive/2016/05/the-day-the-republican-party-died/481176/>.
- Beck, P. A. (1982). Realignment begins?: The Republican surge in Florida. *American Politics Quarterly*, 10(4), 421-438.
- Bennett, M. M., Crowe, B. J., Price, K. L., Stamey, J. D., & Seaman, Jr., J. W. (2013). Comparison of Bayesian and frequentist meta-analytical

- approaches for analyzing time to event data. *Journal of Biopharmaceutical Statistics*, 23, 129-145.
- Berg, A., Meyer, R., & Yu, J. (2004). Deviance information criterion for comparing stochastic volatility models. *Journal of Business & Economic Statistics*, 22, 107-120.
- Berger, J. O., Boukai, B., & Wang, Y. (1997). Unified frequentist and Bayesian testing of a precise hypothesis. *Statistical Science*, 21(3), 133-160.
- Betti, B., Cazzaniga, N. E., Tornatore, V. (2011). Deformation assessment considering an a priori functional model in a Bayesian framework. *Journal of Surveying Engineering*, 137(4), 113-119.
- Black, E., & Black, M. (1992). *The vital South: How presidents are elected*. Cambridge, MA: Harvard University Press.
- Bolstad, W. M. (2002). *Teaching Bayesian statistics to undergraduates: Who, what, where, when, why, and how*. Paper presented at the Sixth International Conference on Teaching Statistics: Developing a Statistically Literate Society, Cape Town, South Africa. Cape Town, South Africa: International Conference on Teaching Statistics.
- Brooks, D. (2016, September 9). Time for a realignment. *The New York Times*, p. A23.
- Brunell, T. L., Grofman, B., & Merrill III, S. (2012). Magnitude and durability of electoral change: Identifying critical elections in the U.S. Congress 1854-2010. *Electoral Studies*, 31, 816-828.

- Buchanan, S. E. (2002). The realignment of 1964?. *Politics & Polity*, 30(1), 140-158.
- Bullock, C. S., Hoffman, D. R., & Gaddie, R. K. (2006). Regional variations in the realignment of American politics, 1944-2004. *Social Science Quarterly*, 87(3), 494-518.
- Burnham, W. D. (1970). *Critical elections and the mainsprings of American politics*. New York, NY: WW Norton and Company.
- Campbell, J. E. (2006). Party systems and realignments in the United States, 1868-2004. *Social Science History*, 30(3), 359-386.
- Carmines, E. G., Huckfeldt, R., & McCurley, C. (1995). Mobilization, counter-mobilization and the politics of race. *Political Geography*, 14(7), 601-619.
- Carmines, E. G., & Stimson, J. A. (1989). *Issue evolution: Race and the transformation of American politics*. Princeton, NJ: Princeton University Press.
- Carmines, E. G., & Wagner, M. W. (2006). Political issues and party alignments: Assessing the issue evolution perspective. *Annual Review of Political Science*, 9, 67-81.
- Chang, M., & Boral, A. (2008). ABC of Bayesian approaches to drug development. *Pharmaceutical Medicine*, 22(3), 141-150.
- Coory, M. D., Wills, R. A., Barnett, A. G. (2009). Bayesian versus frequentist statistical inference for investigating a one-off cancer cluster reported to a health department. *BMC Medical Research Methodology*, 9(30), 1-7.
- CQ Press. (2010). *Guide to U.S. elections*. Washington, D.C.: CQ Press.

- Darmofal, D., & Nardulli, P. F. (2010). The dynamics of critical realignments: An analysis across time and space. *Political Behavior*, 32, 255-283.
- Darmofal, D. (2008). The political geography of the new deal realignment. *American Politics Research*, 36(6), 934-961.
- Downey, A. B. (2012). *Think Bayes: Bayesian statistics made simple*. Needham, MA: Green Tea Press.
- Falk, R., & Greenbaum, C. W. (1995). Significance tests die hard: The amazing persistence of a probabilistic misconception. *Theory & Psychology*, 5(1), 75-98.
- Feinstein, B. D., & Schickler, E. (2008). Platforms and partners: The Civil Rights realignment reconsidered. *Studies in American Political Development*, 22, 1-31.
- Feller, W. (1968). *An introduction to probability theory and its applications*. New York, NY: John Wiley and Sons, Inc.
- Gans, D. J. (1985). Persistence of party success in American presidential elections. *The Journal of Interdisciplinary History*, 16(2), 221-237.
- Gelman, A. B. (2014, February 12). How to think about “identifiability” in Bayesian inference? [Web log comment]. Retrieved from <http://andrewgelman.com/2014/02/12/think-identifiability-bayesian-inference/>.
- Gelman, A. B., & Stern, H. S. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *The American Statistician*, 60(4), 328-331.

- Gelman, A. B., Carlin, J. S., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. New York, NY: Chapman & Hall.
- Gigerenzer, S. L. (1998). Precis of statistical significance: rationale, validity, and utility. *Behavioral and Brain Sciences*, *21*, 169-239.
- Gill, J. (1999). The insignificance of null hypothesis significance testing. *Political Research Quarterly*, *52*(3), 647-674.
- Giroux, G. (2012, December 12). Voters throw bums in while holding Congress in disdain. *Bloomberg*. Retrieved from <https://www.bloomberg.com/news/articles/2012-12-13/voters-throw-bums-in-while-disdaining-congress-bgov-barometer>.
- Glickman, M. E., & van Dyk, D. A. (2007). Basic Bayesian methods. In W. T. Ambrosius (Ed.), *Methods in Molecular Biology, vol. 404: Topics in Biostatistics* (319-338). Totowa, NJ: Humana Press Inc.
- Gliner, J. A., Leech, N. L., Morgan, G. A. (2002). Problems with null hypothesis significance testing (NHST): What do the textbooks say?. *The Journal of Experimental Education*, *71*(1), 83-92.
- Goodman, S. N. (1999). Toward the evidence-based medical statistics. 1. The P value fallacy. *Annals of Internal Medicine*, *130*, 995-1004.
- Gross, J. H. (2015). Testing what matters (if you must test at all): A context-driven approach to substantive and statistical significance. *American Journal of Political Science*, *59*(3), 775-788.
- Guikema, S. D. (2005). A comparison of reliability estimation methods for binary systems. *Reliability Engineering and System Safety*, *87*, 365-376.

- Hawley, G., & Sagarzazu, I. (2012). Where did the votes go? Reassessing American party realignments via vote transfers between major parties from 1860 to 2008. *Electoral Studies*, 31, 726-739.
- Historical Statistics of the United States Millennial Edition Online. (2006). *Political party affiliations in Congress and the presidency: 1789-2002* [Data file]. Available from <http://hsus.cambridge.org/HSUSWeb/HSUSEntryServlet>.
- Holmes, C. (n.d.). *Markov chain Monte Carlo and applied Bayesian statistics: A short course* [PDF document]. Retrieved from http://www.stats.ox.ac.uk/~cholmes/Courses/BDA/bda_mcmc.pdf.
- Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach*. Chicago, IL: Open Court Publishing Company.
- Kantor, S., Fishback, P. V., & Wallis, J. J. (2013). Did the New Deal solidify the 1932 democratic realignment?. *Explorations in Economic History*, 50, 620-633.
- Key, V. O. (1955). A theory of critical elections. *The Journal of Politics*, 17(1), 3-18.
- Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and psychological measurement*, 56(6), 746-759.
- Kleppner, P. (1987). *Continuity and change in electoral politics, 1893-1928*. New York, NY: Greenwood Press.

- Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, 6(3), 299-312.
- Lam, P. (2009). *Convergence diagnostics* [Powerpoint slides]. Retrieved from http://www.stat.missouri.edu/~dsun/8640/convergence_print.pdf.
- Lichtman, A. J. (1976). Critical election theory and the reality of American presidential politics. *The American Historical Review*, 81(2), 317-351.
- Lilford, R. J., & Braunholtz, D. (1996). The statistical basis of public policy: A paradigm shift is overdue. *British Medical Journal*, 313(7057), 603-607.
- Liu, P., Yang, H., Qiang, L., Xiao, S., & Shi, Z. (2012). Estimation of the sensitivity and specificity of assays for screening antibodies to HIV: A comparison between the frequentist and Bayesian approaches. *Journal of Virological Methods*, 186, 89-93.
- Lodge, M., Steenbergen, M., & Brau, S. (1995). The responsive voter: Campaign information and the dynamics of candidate evaluation. *American Political Science Review*, 89(2), 309-326. doi:10.2307/2082427
- Mayhew, D. R. (2002). *Electoral realignments: a critique of an American genre*. New Haven, CT: Yale University Press.
- McGrayne, S. B. (2012). *The theory that would not die: How Bayes' rule cracked the enigma code, hunted down Russian submarines, & emerged triumphant from two centuries of controversy*. New Haven, CT: Yale University Press.

- Milkis, S. M. (2012). The transformation of American democracy: Teddy Roosevelt, the 1912 election, and the Progressive party. *First Principles*, 43. Retrieved from <http://www.heritage.org/political-process/report/the-transformation-american-democracy-teddy-roosevelt-the-1912-election>.
- Minium, E. W., King, B. M., & Bear, G. (1993). *Statistical reasoning in psychology and education* (3rd ed.). New York, NY: Wiley.
- Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46(4), 806-834.
- New York Daily News. (2016). May 4, 2016: Republican party, 1854-2016 [Online image]. Retrieved May 5, 2016 from <http://www.nydailynews.com/news/new-york-daily-news-front-pages-presidential-election-gallery-1.2512941?pmSlide=1.2624119>.
- Norpoth, H., & Rusk, J. G. (2007). Electoral myth and reality: Realignments in American politics. *Electoral Studies*, 26, 392-403.
- Peck, J. (2015, October 7). New Bayesian extension commands for SPSS statistics [Web log comment]. Retrieved from <https://developer.ibm.com/predictiveanalytics/2015/10/07/new-bayesian-extension-commands-for-spss-statistics/>.
- Price, L. R. (2012). Small sample properties of Bayesian multivariate autoregressive time series models. *Structural Equation Modeling: A Multidisciplinary Journal*, 19(1), 51-64.

- Rothman, K. J. (2014). Six persistent research misconceptions. *Journal of General Internal Medicine, 29*(7), 1060-1064.
- SAS Institute Inc. (2009). *SAS/STAT® 9.2 user's guide, second edition*. Cary, NC: SAS Institute Inc.
- Saunders, K. L., & Abramowitz, A. I. (2004). Ideological realignment and active partisans in the American electorate. *American Politics Research, 32*(3), 285-309.
- Schattschneider, E. E. (1956). United States: The functional approach to party government. In S. Neumann (Ed.), *Modern Political Parties: Approaches to Comparative Politics* (194-215). Chicago, IL: University of Chicago Press.
- Schattschneider, E. E. (1960). *The semi-sovereign people: A realist's view of democracy in America*. New York, NY: Holt, Rinehart, and Winston.
- Schofield, N., Miller, G., & Martin, A. (2003). Critical elections and political realignments in the USA: 1860-2000. *Political Studies, 51*, 217-240.
- Shafer, B. E. (1991). The notion of an electoral order: The structure of electoral politics at the accession of George Bush. In B. E. Shafer (Ed.), *The end of realignment? Interpreting American electoral eras* (37-84). Madison, WI: The University of Wisconsin Press.
- Shelley, F. M., Zerr, K. J., & Proffer, A. M. (2007). The Civil Rights movement and recent electoral realignment in the South. *Southeastern Geographer, 47*(1), 13-26.

- Shively, W. P. (1992). Differential abstention to conversion: A change in electoral change, 1864-1988. *American Journal of Political Science*, 36(2), 309-330.
- Silbey, J. H. (1991). Beyond realignment and realignment theory: American political eras, 1789-1989. In B. E. Shafer (Ed.), *The end of realignment? Interpreting American electoral eras* (37-84). Madison, WI: The University of Wisconsin Press.
- Silver, N. (2012). *The signal and the noise: Why so many predictions fail- but some don't*. New York, NY: The Penguin Press.
- StataCorp. (2015). *Stata Bayesian analysis reference manual release 14*. College Station, TX: Stata Press.
- Stegmueller, D. (2013). How many countries for multilevel modeling? A comparison of frequentist and Bayesian approaches. *American Journal of Political Science*, 57(3), 748-761.
- Stokes, M., Chen, F., & Gunes, F. (2014). An introduction to Bayesian analysis with SAS/STAT software. SAS Institute Inc, 1-25.
- Stonecash, J. M., & Silina, E. (2005). The 1896 realignment: A reassessment. *American Politics Research*, 33(1), 3-32.
- Sundquist, J. L. (1973). *Dynamics of the party system: Alignment and realignment of political parties in the United States*. Washington, DC: The Brookings Institution.
- The Royal Society. (2017). *Fellow details: Bayes; Thomas (1701-1761)*. Retrieved from

<https://collections.royalsociety.org/Dserve.exe?dsqIni=Dserve.ini&dsqApp=Archive&dsqDb=Persons&dsqSearch=Code==%27NA8121%27&dsqCmd=Show.tcl>.

- United States House of Representatives. (2016). *Party divisions of the House of Representatives, 1789-present* [Data file]. Retrieved from <http://history.house.gov/Institution/Party-Divisions/Party-Divisions/>.
- van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. G. (2013). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child Development*, 1-19.
- van de Schoot, R., & Depaoli, S. (2014). Bayesian analyses: where to start and what to report. *The European Health Psychologist*, 16(2), 75-84.
- Wanat, J., & Burke, K. (1982). Estimating the degree of mobilization and conversion in the 1890s: An inquiry into the nature of electoral change. *The American Political Science Review*, 76(2), 360-370.
- Wilkinson, M. (2014). Clinical and practical importance vs statistical significance: Limitations of conventional statistical inference. *International Journal of Therapy and Rehabilitation*, 21(10), 488-495.
- Zingher, J. N. (2014). An analysis of the changing social bases of America's political parties: 1952-2008. *Electoral Studies*, 35, 272-282.

APPENDIX A

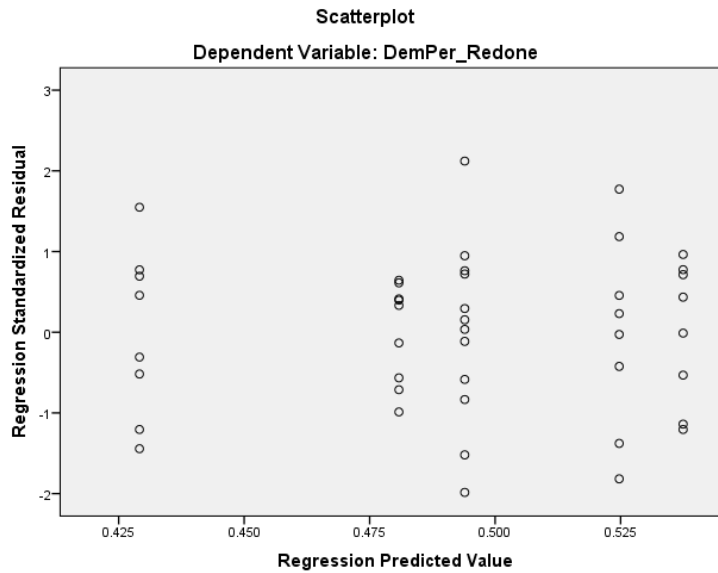


Figure A1. Scatterplot demonstrating the homoscedasticity of the presidential models.

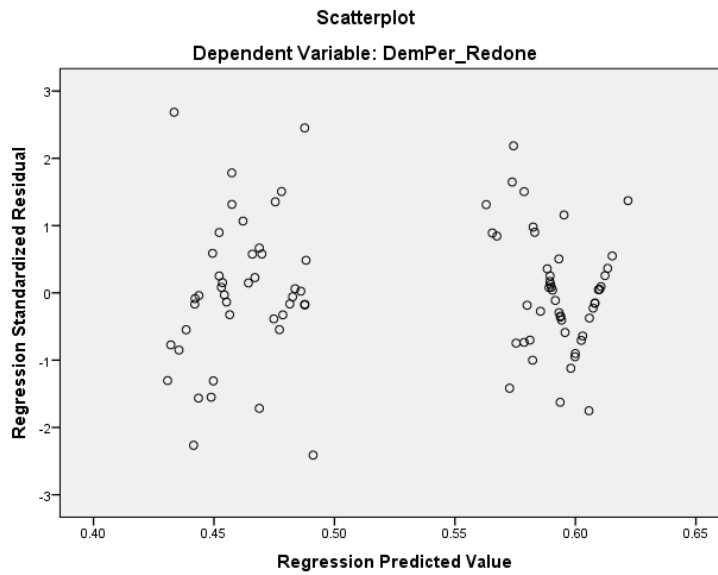


Figure A2. Scatterplot demonstrating the homoscedasticity of the congressional models.

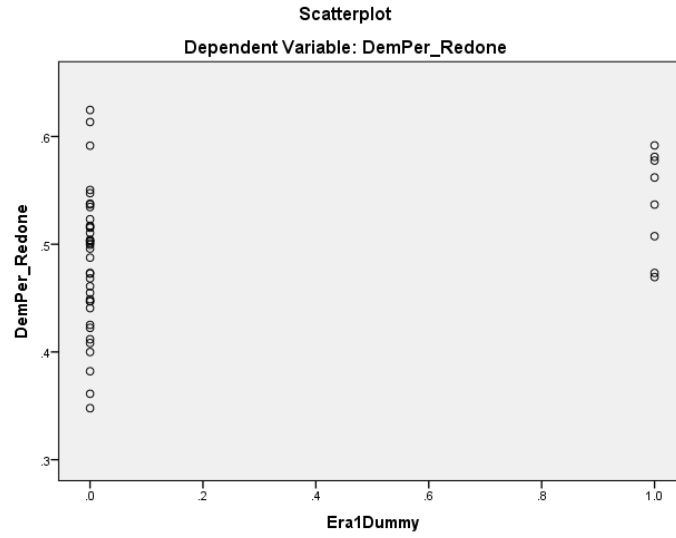


Figure A3. Scatterplot demonstrating the linearity between the Democratic two-party presidential vote and the first era, 1828-1856.

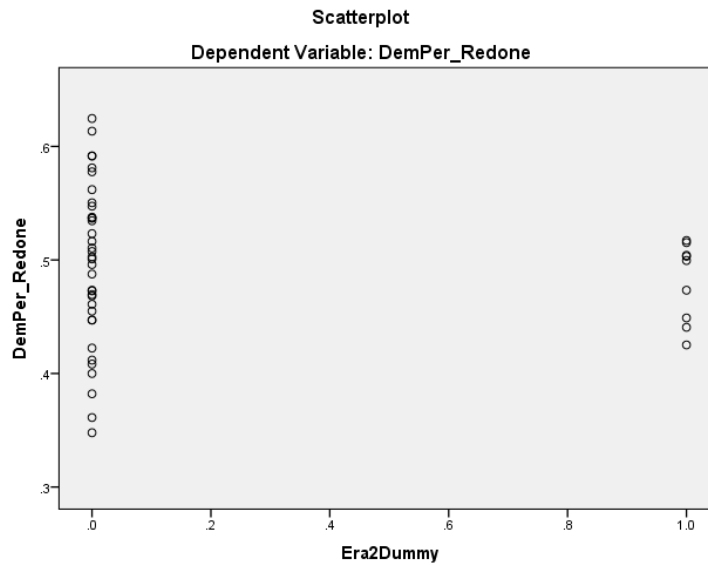


Figure A4. Scatterplot demonstrating the linearity between the Democratic two-party presidential vote and the second era, 1860-1892.

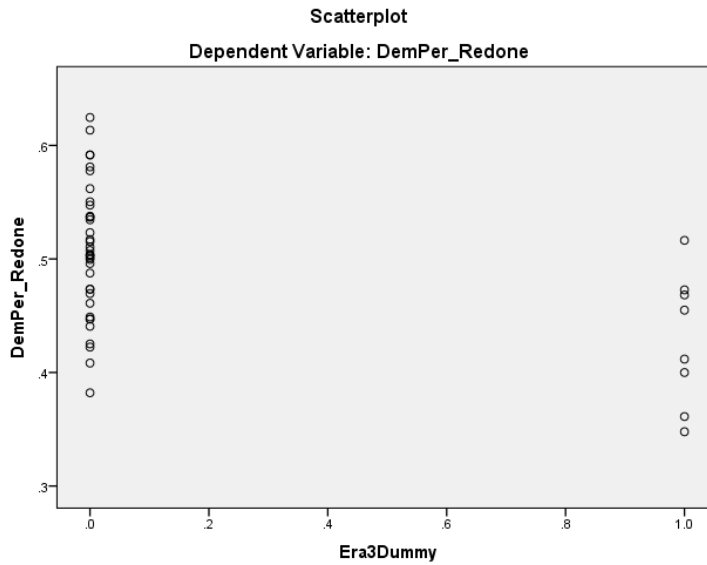


Figure A5. Scatterplot demonstrating the linearity between the Democratic two-party presidential vote and the third era, 1896-1928.

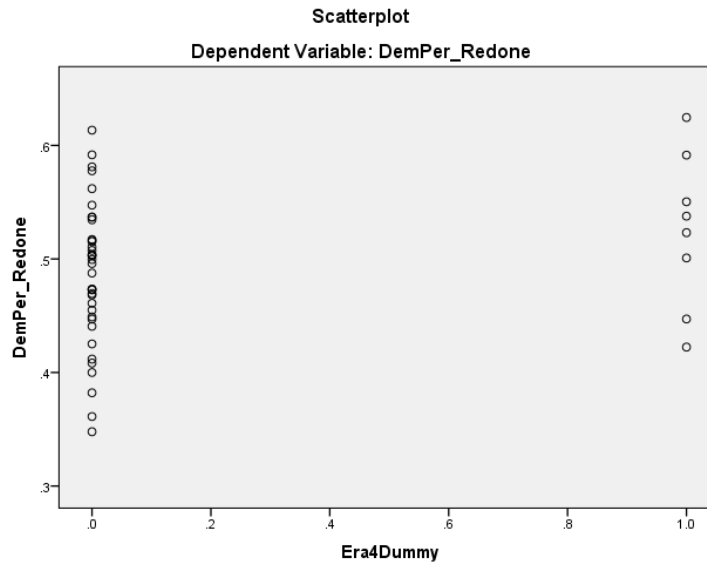


Figure A6. Scatterplot demonstrating the linearity between the Democratic two-party presidential vote and the fourth era, 1932-1960.

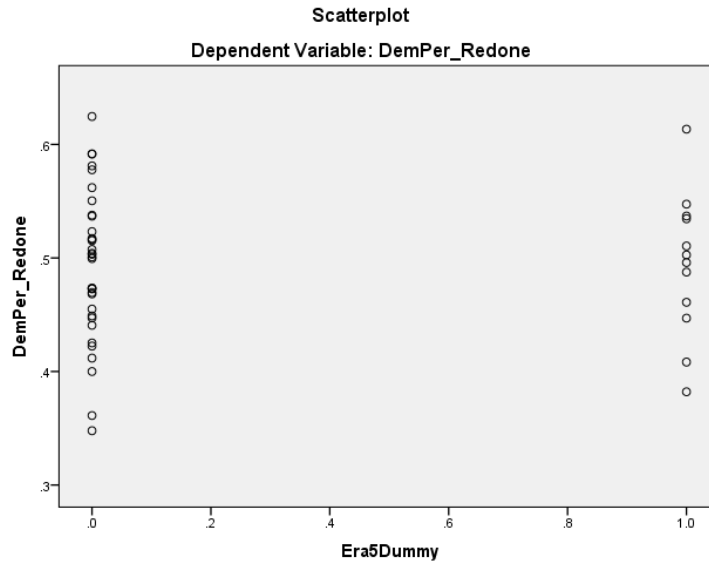


Figure A7. Scatterplot demonstrating the linearity between the Democratic two-party presidential vote and the fifth era, 1964-2008.

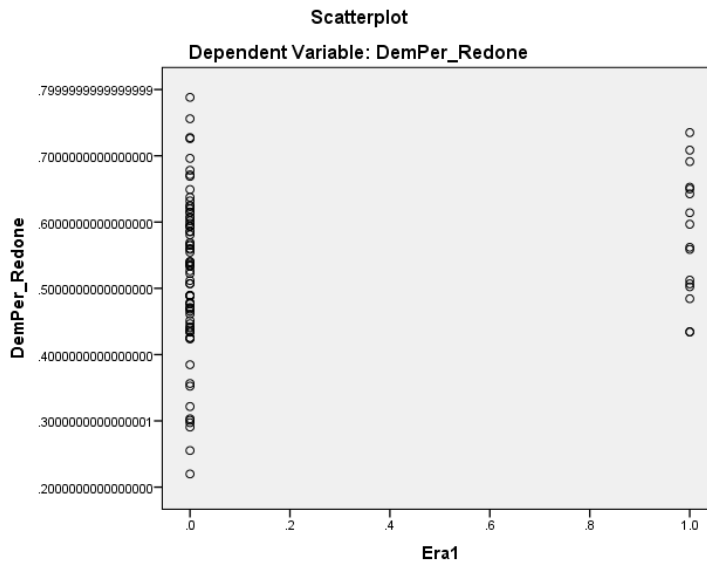


Figure A8. Scatterplot demonstrating the linearity between the U.S. House between Representative seat share and the first era, 1828-1858.

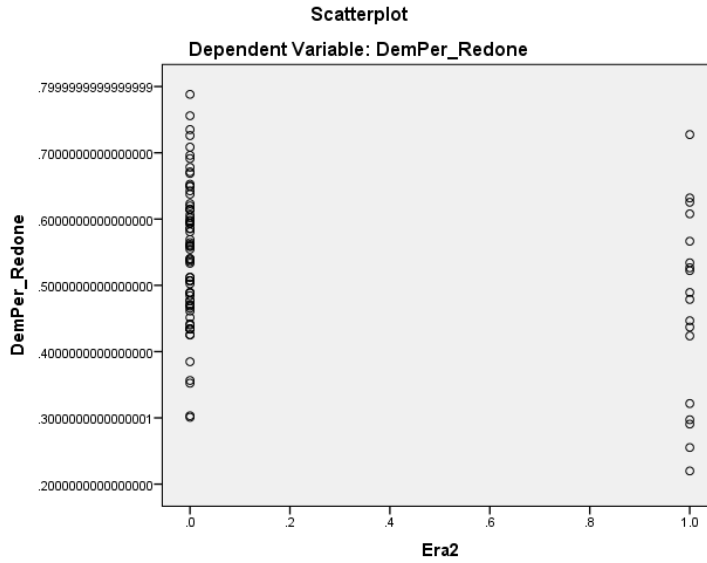


Figure A9. Scatterplot demonstrating the linearity between the U.S. House between Representative seat share and the second era, 1860-1894.

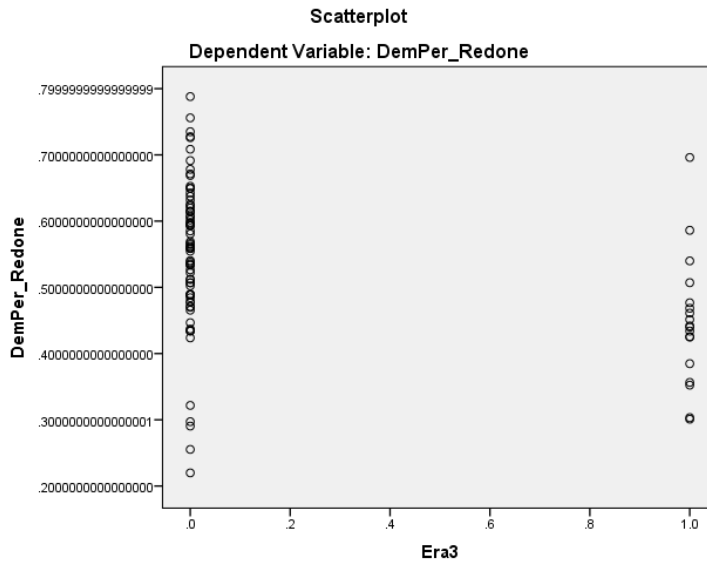


Figure A10. Scatterplot demonstrating the linearity between the U.S. House between Representative seat share and the third era, 1896-1930.

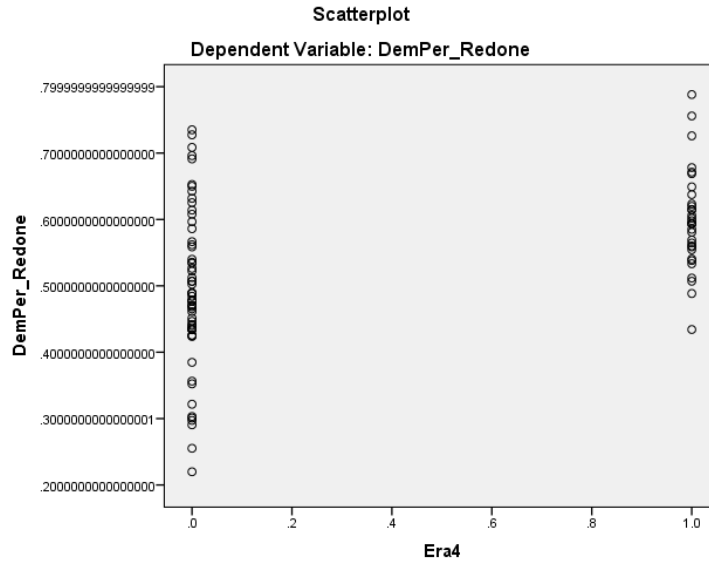


Figure A11. Scatterplot demonstrating the linearity between the U.S. House between Representative seat share and the fourth era, 1932-1994.

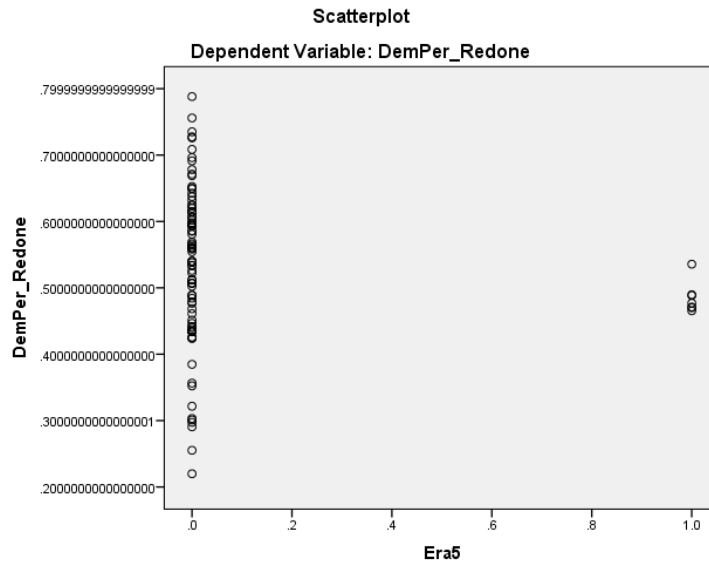


Figure A12. Scatterplot demonstrating the linearity between the U.S. House between Representative seat share and the fifth era, 1996-2008.

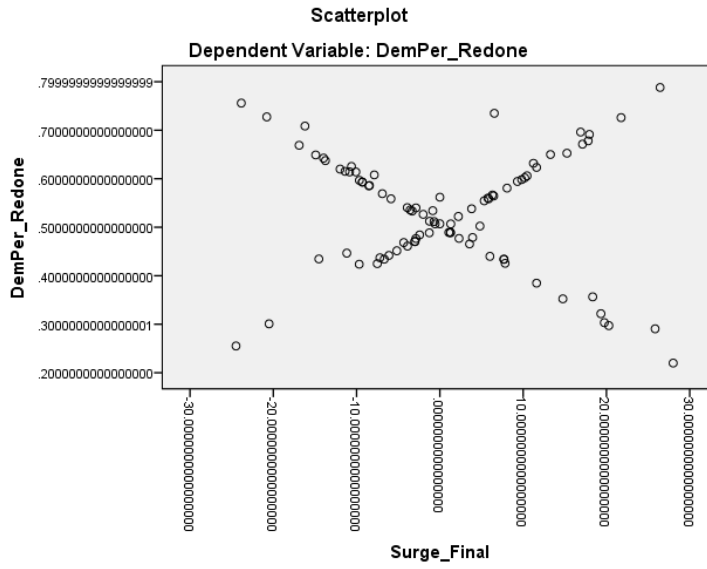


Figure A13. Scatterplot demonstrating the linearity between the U.S. House between Representative seat share and the general election surge.

APPENDIX B

```
%include "T:\FHF\Research\Data\Analysis\Include\SAS Setup and
Help\Setup\TaraAutoExec.sas";
%include "T:\FHF\Research\Data\Analysis\Include\SAS Setup and
Help\Setup\setup.Inc.sas";
***SAS program for dissertation***
Author: TAR
Date: 8/27/16
Program: Results for Bayesian Non-Informative
*****;

*step 1: import data file;
proc import out=data1
  datafile="T:\FHF\Research\Tara's
Documents\School\Dissertation\Presidential Data2.sav"
  dbms= sav replace;
run;

ods rtf text="Reg1";
proc genmod data=data1;
  model demper_redone=era1constant era2dummy era3dummy era4dummy
era5dummy / dist=normal;
  bayes seed=08212016 outpost=PostReg1 diagnostics=(gelman
geweke heidelberger raftery);
run;

ods rtf text="Reg2";
proc genmod data=data1;
  model demper_redone=era2constant era1dummy era3dummy era4dummy
era5dummy / dist=normal;
  bayes seed=08212016 outpost=PostReg2 diagnostics=(gelman
geweke heidelberger raftery);
run;

ods rtf text="Reg3";
proc genmod data=data1;
  model demper_redone=era3constant era1dummy era2dummy era4dummy
era5dummy / dist=normal;
  bayes seed=08212016 outpost=PostReg3 diagnostics=(gelman
geweke heidelberger raftery);
run;

ods rtf text="Reg4";
proc genmod data=data1;
  model demper_redone=era4constant era1dummy era2dummy era3dummy
era5dummy / dist=normal;
  bayes seed=08212016 outpost=PostReg4 diagnostics=(gelman
geweke heidelberger raftery);
run;

ods rtf text="Reg5";
```

```

proc genmod data=data1;
    model demper_redone=era5constant era1dummy era2dummy era3dummy
era4dummy / dist=normal;
    bayes seed=08212016 outpost=PostReg5 diagnostics=(gelman
geweke heidelberger raftery);
run;

****Assessing Posterior Densities;
proc means data=postreg1; var era2dummy; run;
proc kde data=postreg1; var era2dummy; run;
proc kde data=postreg1; univar era2dummy; run;

proc contents data=postreg1; run;
data postreg1a;
    set postreg1;
Abs=abs(era2dummy);
Indicator=(abs>0.058401);
run;
proc means data=postreg1a (keep=indicator); run;

proc means data=postreg2; var era3dummy; run;
proc kde data=postreg2; var era3dummy; run;
proc kde data=postreg2; univar era3dummy; run;

proc contents data=postreg2; run;
data postreg2a;
    set postreg2;
Abs=abs(era3dummy);
Indicator=(abs>0.058401);
run;
proc means data=postreg2a (keep=indicator); run;

proc means data=postreg3; var era4dummy; run;
proc kde data=postreg3; var era4dummy; run;
proc kde data=postreg3; univar era4dummy; run;

proc contents data=postreg3; run;
data postreg3a;
    set postreg3;
Abs=abs(era4dummy);
Indicator=(abs>0.058401);
run;
proc means data=postreg3a (keep=indicator); run;

proc means data=postreg4; var era5dummy; run;
proc kde data=postreg4; var era5dummy; run;
proc kde data=postreg4; univar era5dummy; run;

proc contents data=postreg4; run;
data postreg4a;

```

```
set postreg4;
Abs=abs(era5dummy);
Indicator=(abs>0.058401);
run;
proc means data=postreg4a (keep=indicator); run;

proc means data=postreg5; var era4dummy; run;
proc kde data=postreg5; var era4dummy; run;
proc kde data=postreg5; univar era4dummy; run;
```

Figure B1. SAS programming used for the presidential Bayesian models, using a non-informative prior distribution.


```

%include "T:\FHF\Research\Data\Analysis\Include\SAS Setup and
Help\Setup\TaraAutoExec.sas";
%include "T:\FHF\Research\Data\Analysis\Include\SAS Setup and
Help\Setup\setup.Inc.sas";
***SAS program for dissertation***
Author: TAR
Date: 8/21/16
Program: Results for Bayesian Non-Informative
*****;

*step 1: import data file;
proc import out=data1
    datafile="T:\FHF\Research\Tara's
Documents\School\Dissertation\Assumptions\SPSS\US House Data3.sav"
    dbms= sav replace;
run;

ods rtf text="Reg1";
proc genmod data=data1;
    model demper_redone=average1 era2 era3 era4 era5 surge_final /
    dist=normal;
    bayes seed=08212016 outpost=PostReg1 diagnostics=(gelman
geweke heidelberger raftery);
run;

ods rtf text="Reg2";
proc genmod data=data1;
    model demper_redone=average2 era1 era3 era4 era5 surge_final /
    dist=normal;
    bayes seed=08212016 outpost=PostReg2 diagnostics=(gelman
geweke heidelberger raftery);
run;

ods rtf text="Reg3";
proc genmod data=data1;
    model demper_redone=average3 era1 era2 era4 era5 surge_final /
    dist=normal;
    bayes seed=08212016 outpost=PostReg3 diagnostics=(gelman
geweke heidelberger raftery);
run;

ods rtf text="Reg4";
proc genmod data=data1;
    model demper_redone=average4 era1 era2 era3 era5 surge_final /
    dist=normal;
    bayes seed=08212016 outpost=PostReg4 diagnostics=(gelman
geweke heidelberger raftery);
run;

ods rtf text="Reg5";

```

```

proc genmod data=data1;
    model demper_redone=average5 era1 era2 era3 era4 surge_final /
dist=normal;
    bayes seed=08212016 outpost=PostReg5 diagnostics=(gelman
geweke heidelberger raftey);
run;

***Assessing Posterior Densities;
proc means data=postreg1; var era2; run;
proc kde data=postreg1; var era2; run;
proc kde data=postreg1; univar era2; run;

proc contents data=postreg1; run;
data postreg1a;
    set postreg1;
Abs=abs(era2);
Indicator=(abs>0.08268);
run;
proc means data=postreg1a (keep=indicator); run;

proc means data=postreg2; var era3; run;
proc kde data=postreg2; var era3; run;
proc kde data=postreg2; univar era3; run;

proc contents data=postreg2; run;
data postreg2a;
    set postreg2;
Abs=abs(era3);
Indicator=(abs>0.08268);
run;
proc means data=postreg2a (keep=indicator); run;

proc means data=postreg3; var era4; run;
proc kde data=postreg3; var era4; run;
proc kde data=postreg3; univar era4; run;

proc contents data=postreg3; run;
data postreg3a;
    set postreg3;
Abs=abs(era4);
Indicator=(abs>0.08268);
run;
proc means data=postreg3a (keep=indicator); run;

proc means data=postreg4; var era5; run;
proc kde data=postreg4; var era5; run;
proc kde data=postreg4; univar era5; run;

proc contents data=postreg4; run;
data postreg4a;

```

```
set postreg4;
Abs=abs(era5);
Indicator=(abs>0.08268);
run;
proc means data=postreg4a (keep=indicator); run;

proc means data=postreg5; var era4; run;
proc kde data=postreg5; var era4; run;
proc kde data=postreg5; univar era4; run;
```

Figure B2. SAS programming used for the congressional Bayesian models, using a non-informative prior distribution.

```

%include "T:\FHF\Research\Data\Analysis\Include\SAS Setup and
Help\Setup\TaraAutoExec.sas";
%include "T:\FHF\Research\Data\Analysis\Include\SAS Setup and
Help\Setup\setup.Inc.sas";
***SAS program for dissertation***
Author: TAR
Date: 10/17/16
Program: Results for Bayesian Informative
*****;
*pre-step to set prior distributions;
data normalprior1;
    input _type_ $ era1constant era2dummy era3dummy era4dummy
era5dummy;
    datalines;
Var 9 2.5 30 16 12
Mean 53 48 45 53 50
;
run;

data normalprior2;
    input _type_ $ era2constant era1dummy era3dummy era4dummy
era5dummy;
    datalines;
Var 2.5 9 30 16 12
Mean 48 53 45 53 50
;
run;

data normalprior3;
    input _type_ $ era3constant era1dummy era2dummy era4dummy
era5dummy;
    datalines;
Var 30 9 2.5 16 12
Mean 45 53 48 53 50
;
run;

data normalprior4;
    input _type_ $ era4constant era1dummy era2dummy era3dummy
era5dummy;
    datalines;
Var 16 9 2.5 30 12
Mean 53 53 48 45 50
;
run;

data normalprior5;
    input _type_ $ era5constant era1dummy era2dummy era3dummy
era4dummy;
    datalines;

```

```

Var 12 9 2.5 30 16
Mean 50 53 48 45 53
;
run;

*step 1: import data file;
proc import out=data1
    datafile="T:\FHF\Research\Tara's
Documents\School\Dissertation\Presidential Data2.sav"
    dbms= sav replace;
run;

ods rtf text="Reg1";
proc genmod data=data1;
    model demper_redone=era1constant era2dummy era3dummy era4dummy
era5dummy / dist=normal;
    bayes seed=08212016 outpost=PostReg1 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior1);
run;

ods rtf text="Reg2";
proc genmod data=data1;
    model demper_redone=era2constant era1dummy era3dummy era4dummy
era5dummy / dist=normal;
    bayes seed=08212016 outpost=PostReg2 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior2);
run;

ods rtf text="Reg3";
proc genmod data=data1;
    model demper_redone=era3constant era1dummy era2dummy era4dummy
era5dummy / dist=normal;
    bayes seed=08212016 outpost=PostReg3 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior3);
run;

ods rtf text="Reg4";
proc genmod data=data1;
    model demper_redone=era4constant era1dummy era2dummy era3dummy
era5dummy / dist=normal;
    bayes seed=08212016 outpost=PostReg4 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior4);
run;

ods rtf text="Reg5";
proc genmod data=data1;
    model demper_redone=era5constant era1dummy era2dummy era3dummy
era4dummy / dist=normal;
    bayes seed=08212016 outpost=PostReg5 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior5);

```

```

run;

***Assessing Posterior Densities;
proc means data=postreg1; var era2dummy; run;
proc kde data=postreg1; var era2dummy; run;
proc kde data=postreg1; univar era2dummy; run;

proc contents data=postreg1; run;
data postreg1a;
    set postreg1;
Abs=abs(era2dummy);
Indicator=(abs>0.058401);
run;
proc means data=postreg1a (keep=indicator); run;

proc means data=postreg2; var era3dummy; run;
proc kde data=postreg2; var era3dummy; run;
proc kde data=postreg2; univar era3dummy; run;

proc contents data=postreg2; run;
data postreg2a;
    set postreg2;
Abs=abs(era3dummy);
Indicator=(abs>0.058401);
run;
proc means data=postreg2a (keep=indicator); run;

proc means data=postreg3; var era4dummy; run;
proc kde data=postreg3; var era4dummy; run;
proc kde data=postreg3; univar era4dummy; run;

proc contents data=postreg3; run;
data postreg3a;
    set postreg3;
Abs=abs(era4dummy);
Indicator=(abs>0.058401);
run;
proc means data=postreg3a (keep=indicator); run;

proc means data=postreg4; var era5dummy; run;
proc kde data=postreg4; var era5dummy; run;
proc kde data=postreg4; univar era5dummy; run;

proc contents data=postreg4; run;
data postreg4a;
    set postreg4;
Abs=abs(era5dummy);
Indicator=(abs>0.058401);
run;
proc means data=postreg4a (keep=indicator); run;

```

```
proc means data=postreg5; var era4dummy; run;  
proc kde data=postreg5; var era4dummy; run;  
proc kde data=postreg5; univar era4dummy; run;
```

Figure B3. SAS programming used for the presidential Bayesian models, using an informative prior distribution.

```

%include "T:\FHF\Research\Data\Analysis\Include\SAS Setup and
Help\Setup\TaraAutoExec.sas";
%include "T:\FHF\Research\Data\Analysis\Include\SAS Setup and
Help\Setup\setup.Inc.sas";
***SAS program for dissertation***
Author: TAR
Date: 11/05/2016
Program: Results for Congressional Bayesian Informative
*****;
*pre-step to set prior distributions;
data normalprior1;
    input _type_ $ average1 era2 era3 era4 era5 surge_final;
    datalines;
Var 25 30 25 25 25 1e6
Mean 55 50 43 62 55 0
;
run;

data normalprior2;
    input _type_ $ average2 era1 era3 era4 era5 surge_final;
    datalines;
Var 30 25 25 25 25 1e6
Mean 50 55 43 62 55 0
;
run;

data normalprior3;
    input _type_ $ average3 era1 era2 era4 era5 surge_final;
    datalines;
Var 25 25 30 25 25 1e6
Mean 43 55 50 62 55 0
;
run;

data normalprior4;
    input _type_ $ average4 era1 era2 era3 era5 surge_final;
    datalines;
Var 25 25 30 25 25 1e6
Mean 62 55 50 43 55 0
;
run;

data normalprior5;
    input _type_ $ average5 era1 era2 era3 era4 surge_final;
    datalines;
Var 25 25 30 25 25 1e6
Mean 55 55 50 43 62 0
;
run;

```



```

*step 1: import data file;
proc import out=data1
  datafile="T:\FHF\Research\Tara's
Documents\School\Dissertation\Assumptions\SPSS\US House Data3.sav"
  dbms= sav replace;
run;

ods rtf text="Reg1";
proc genmod data=data1;
  model demper_redone=average1 era2 era3 era4 era5 surge_final /
  dist=normal;
  bayes seed=08212016 outpost=PostReg1 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior1);
run;

ods rtf text="Reg2";
proc genmod data=data1;
  model demper_redone=average2 era1 era3 era4 era5 surge_final /
  dist=normal;
  bayes seed=08212016 outpost=PostReg2 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior2);
run;

ods rtf text="Reg3";
proc genmod data=data1;
  model demper_redone=average3 era1 era2 era4 era5 surge_final /
  dist=normal;
  bayes seed=08212016 outpost=PostReg3 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior3);
run;

ods rtf text="Reg4";
proc genmod data=data1;
  model demper_redone=average4 era1 era2 era3 era5 surge_final /
  dist=normal;
  bayes seed=08212016 outpost=PostReg4 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior4);
run;

ods rtf text="Reg5";
proc genmod data=data1;
  model demper_redone=average5 era1 era2 era3 era4 surge_final /
  dist=normal;
  bayes seed=08212016 outpost=PostReg5 diagnostics=(gelman
geweke heidelberger raftery) coeffprior=normal(input=NormalPrior5);
run;

***Assessing Posterior Densities;
proc means data=postreg1; var era2; run;
proc kde data=postreg1; var era2; run;

```

```

proc kde data=postreg1; univar era2; run;

proc contents data=postreg1; run;
data postreg1a;
    set postreg1;
Abs=abs(era2);
Indicator=(abs>0.08268);
run;
proc means data=postreg1a (keep=indicator); run;

proc means data=postreg2; var era3; run;
proc kde data=postreg2; var era3; run;
proc kde data=postreg2; univar era3; run;

proc contents data=postreg2; run;
data postreg2a;
    set postreg2;
Abs=abs(era3);
Indicator=(abs>0.08268);
run;
proc means data=postreg2a (keep=indicator); run;

proc means data=postreg3; var era4; run;
proc kde data=postreg3; var era4; run;
proc kde data=postreg3; univar era4; run;

proc contents data=postreg3; run;
data postreg3a;
    set postreg3;
Abs=abs(era4);
Indicator=(abs>0.08268);
run;
proc means data=postreg3a (keep=indicator); run;

proc means data=postreg4; var era5; run;
proc kde data=postreg4; var era5; run;
proc kde data=postreg4; univar era5; run;

proc contents data=postreg4; run;
data postreg4a;
    set postreg4;
Abs=abs(era5);
Indicator=(abs>0.08268);
run;
proc means data=postreg4a (keep=indicator); run;

proc means data=postreg5; var era4; run;
proc kde data=postreg5; var era4; run;
proc kde data=postreg5; univar era4; run;

```

Figure B4. SAS programming used for the congressional Bayesian models, using an informative prior distribution.

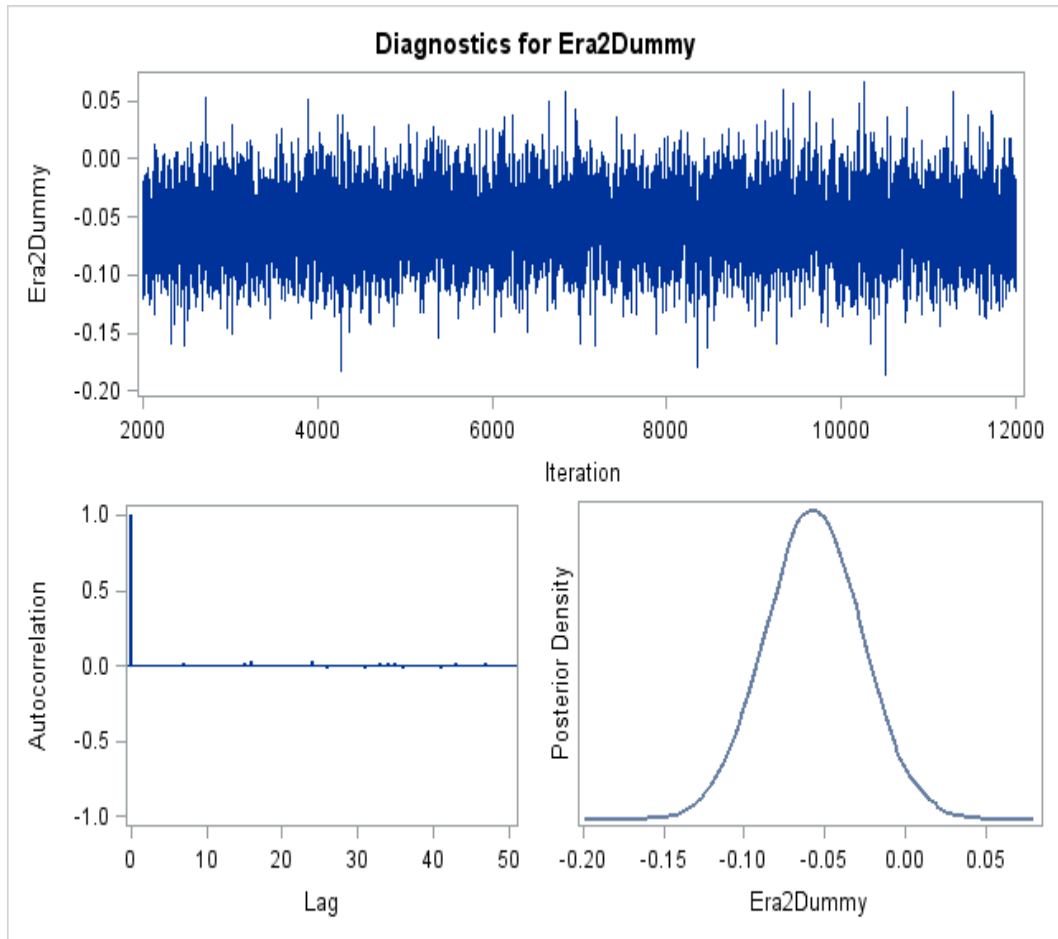


Figure B5. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1860-1892 era, when testing the difference between the 1828-1856 and 1860-1892 eras.

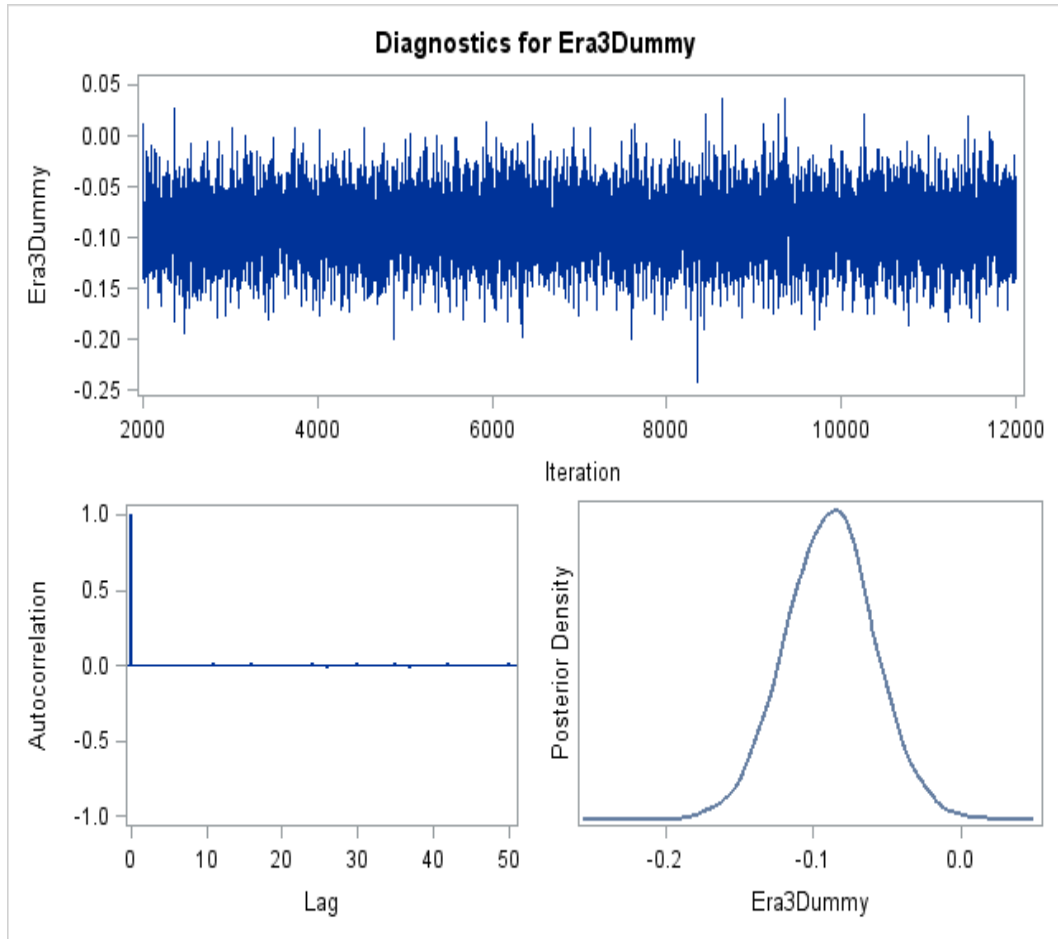


Figure B6. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1896-1928 era, when testing the difference between the 1828-1856 and 1860-1892 eras.

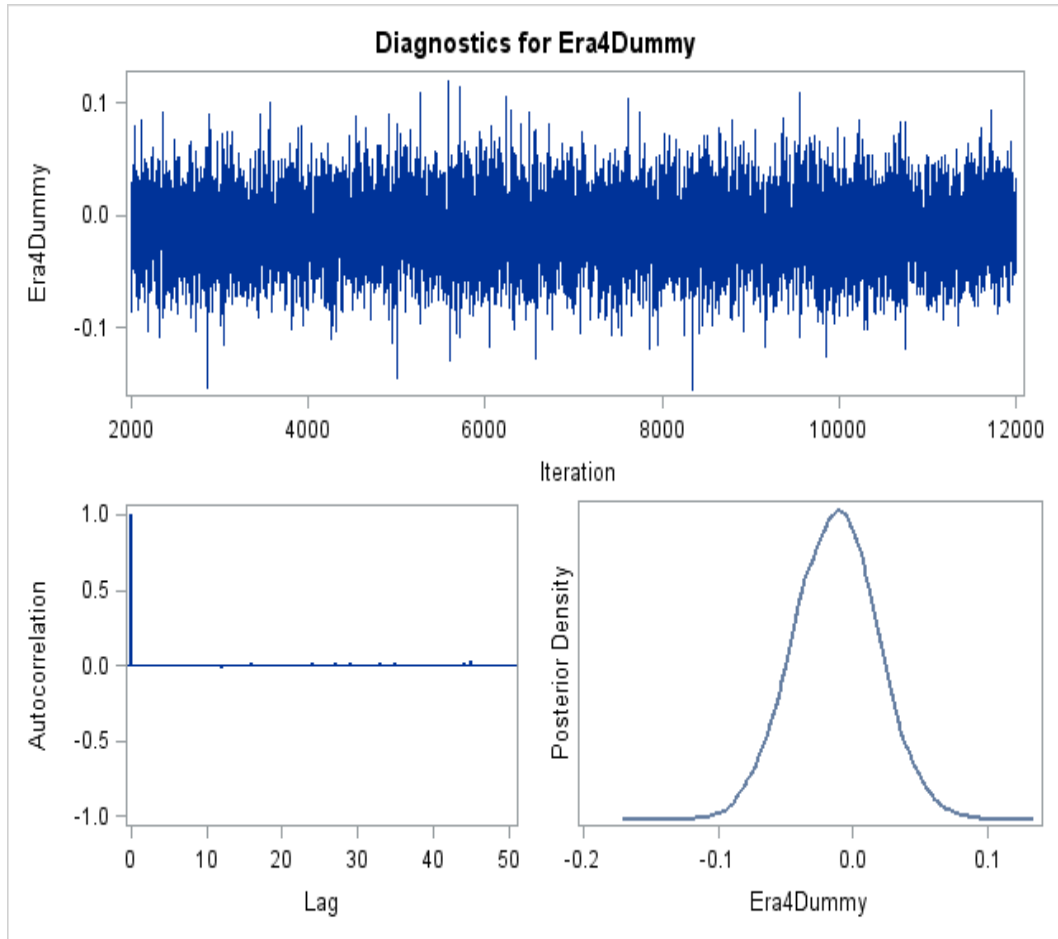


Figure B7. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1932-1960 era, when testing the difference between the 1828-1856 and 1860-1892 eras.

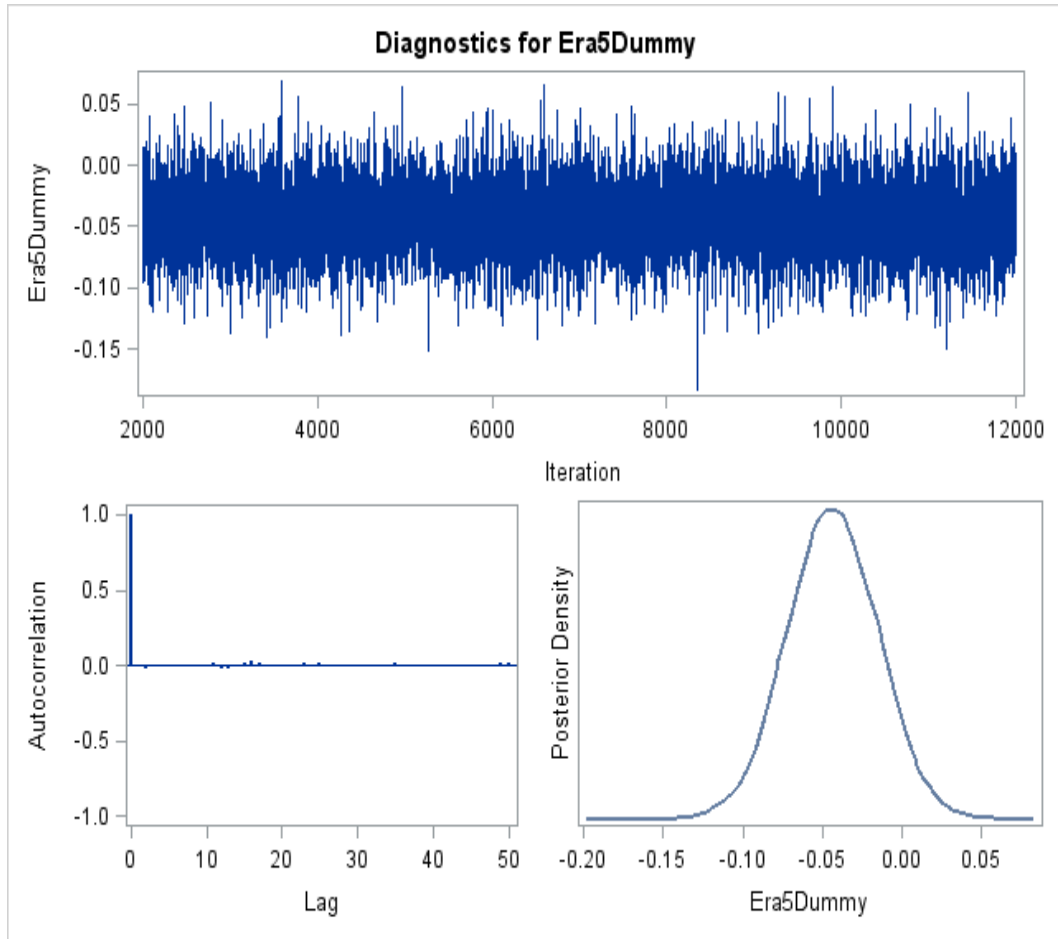


Figure B8. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1964-2008 era, when testing the difference between the 1828-1856 and 1860-1892 eras.

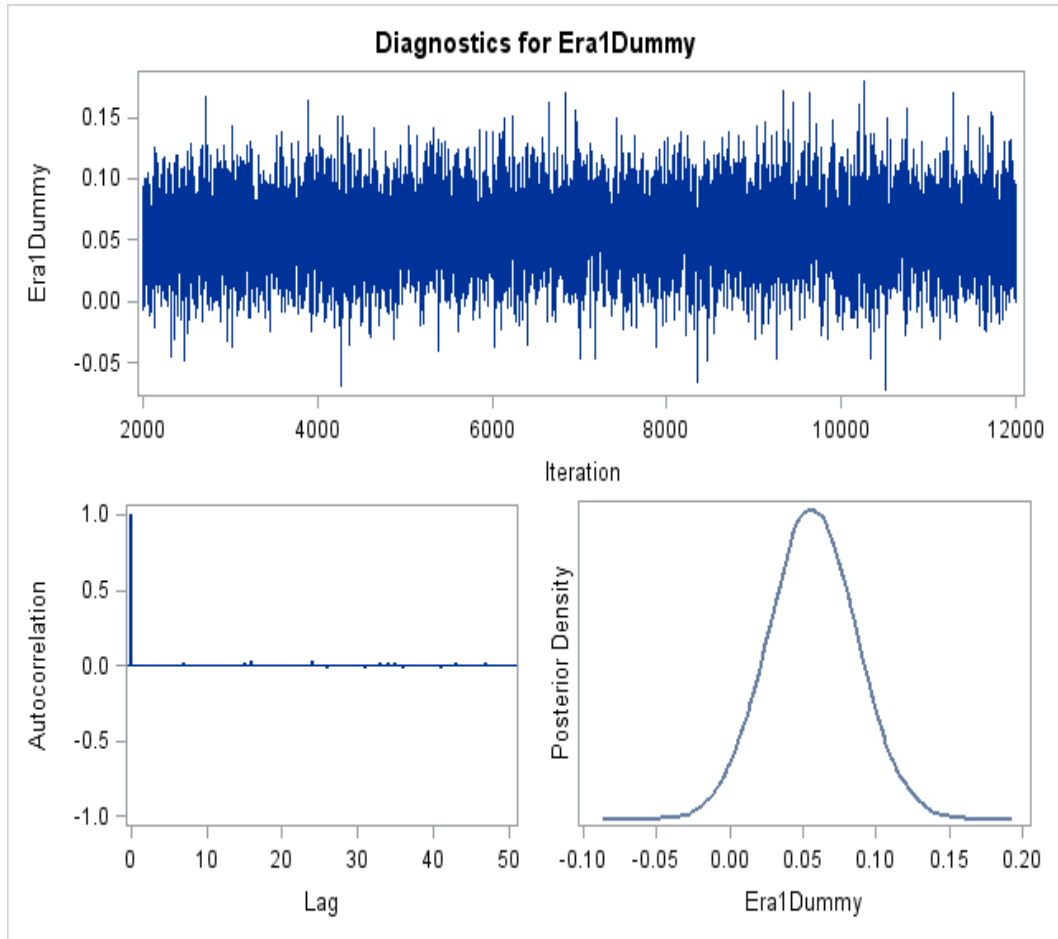


Figure B9. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1828-1856 era, when testing the difference between the 1860-1892 and 1896-1928 eras.

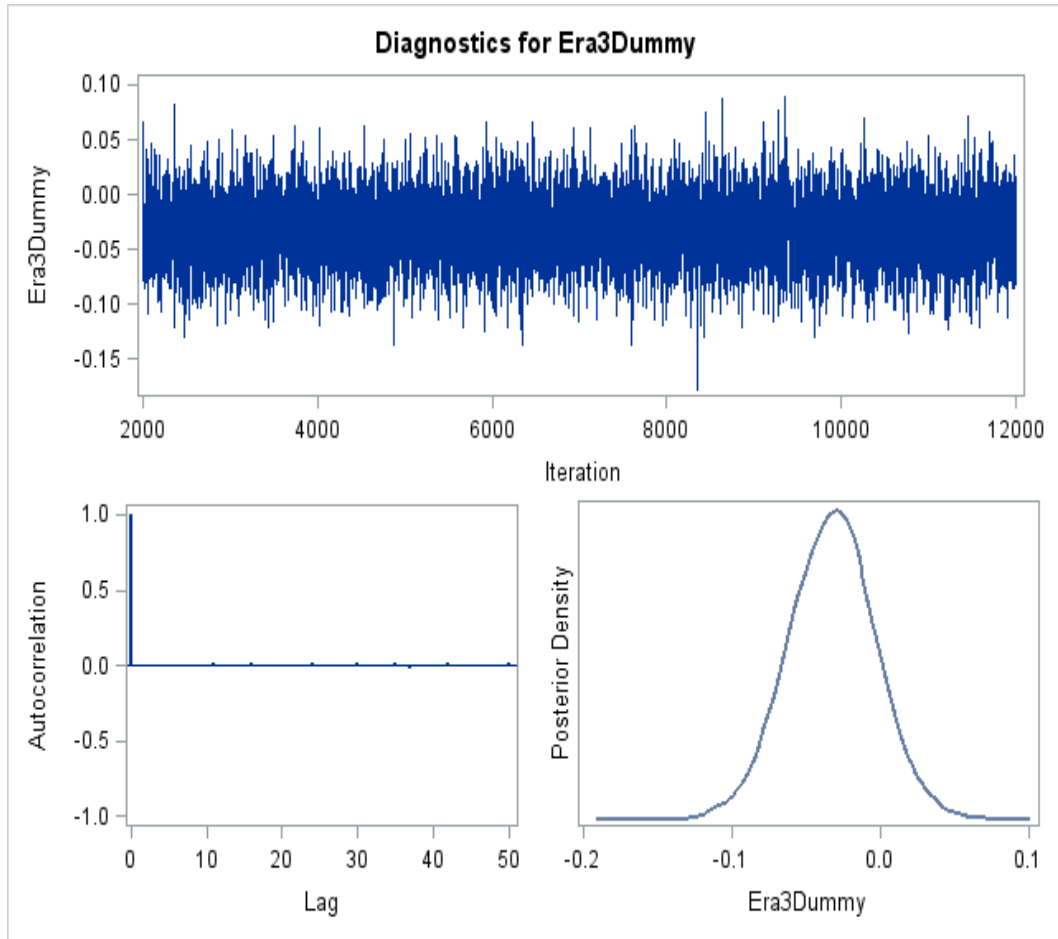


Figure B10. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1896-1928 era, when testing the difference between the 1860-1892 and 1896-1928 eras.

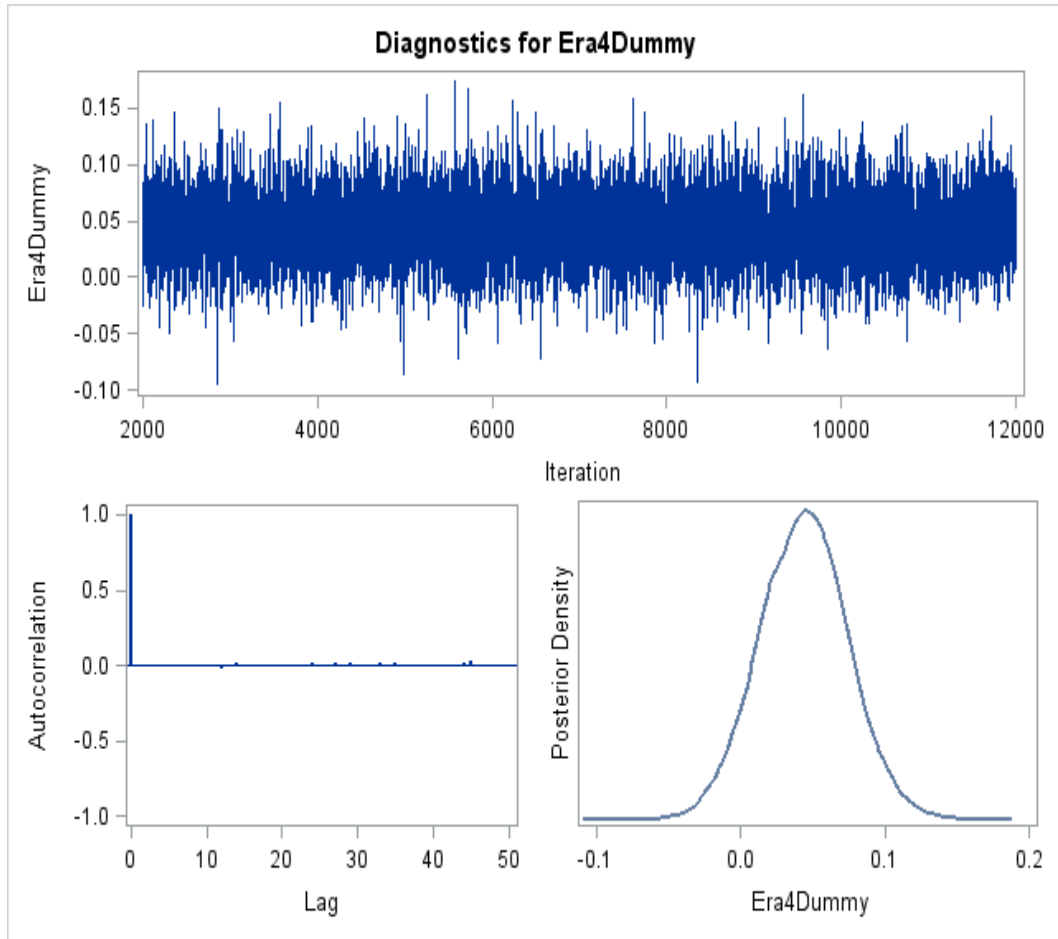


Figure B11. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1932-1960 era, when testing the difference between the 1860-1892 and 1896-1928 eras.

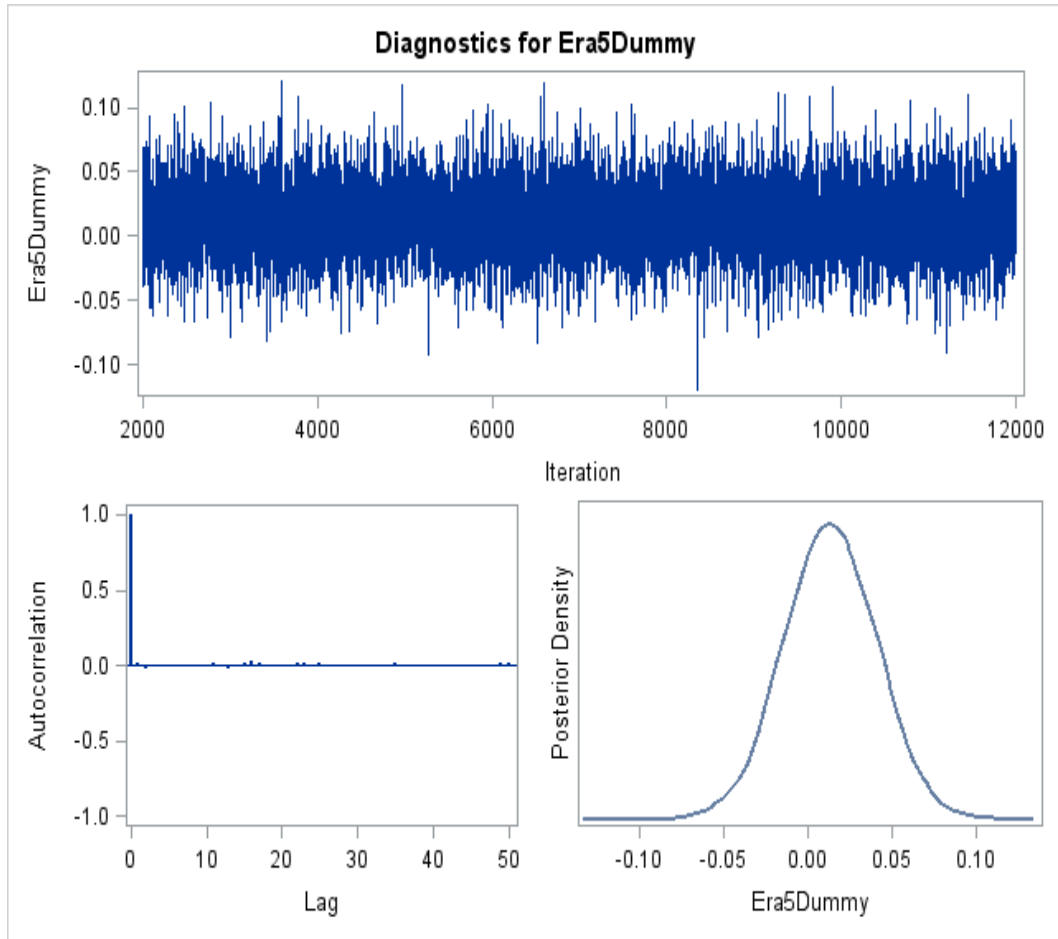


Figure B12. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1964-2008 era, when testing the difference between the 1860-1892 and 1896-1928 eras.

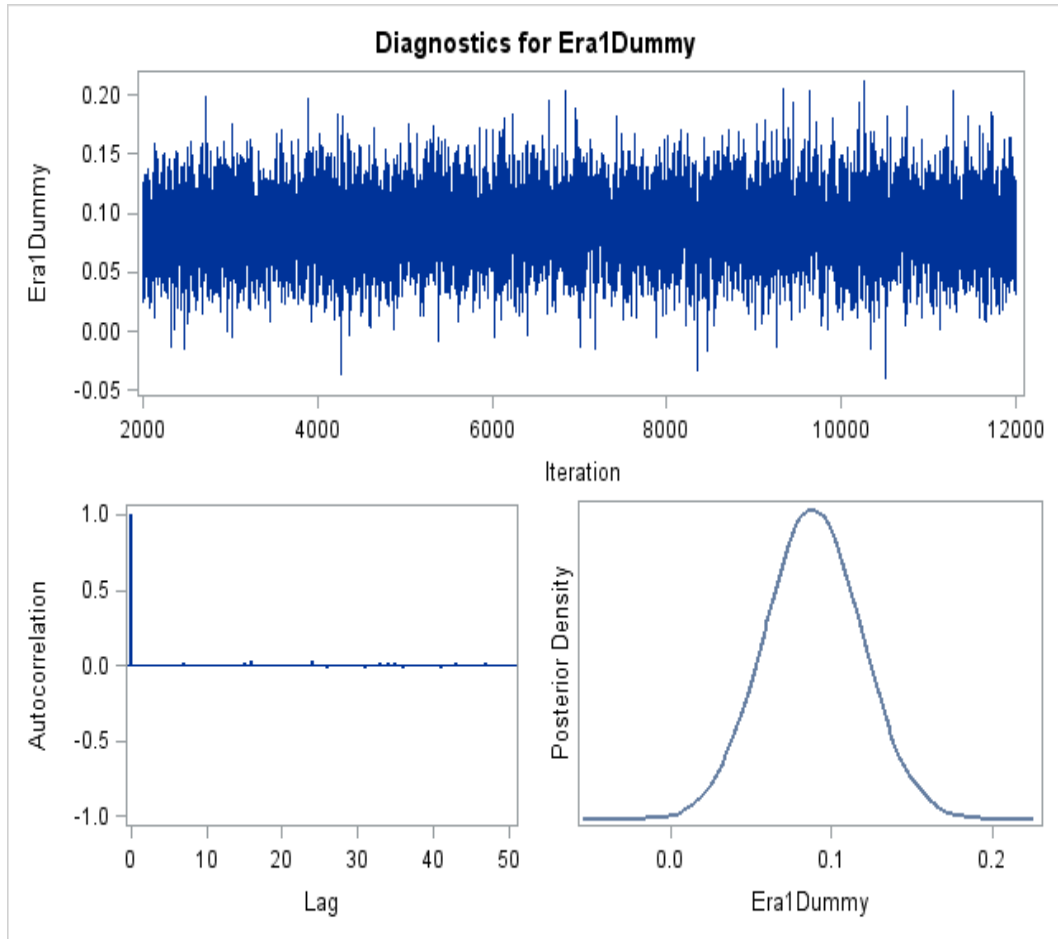


Figure B13. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1828-1856 era, when testing the difference between the 1896-1928 and 1932-1960 eras.

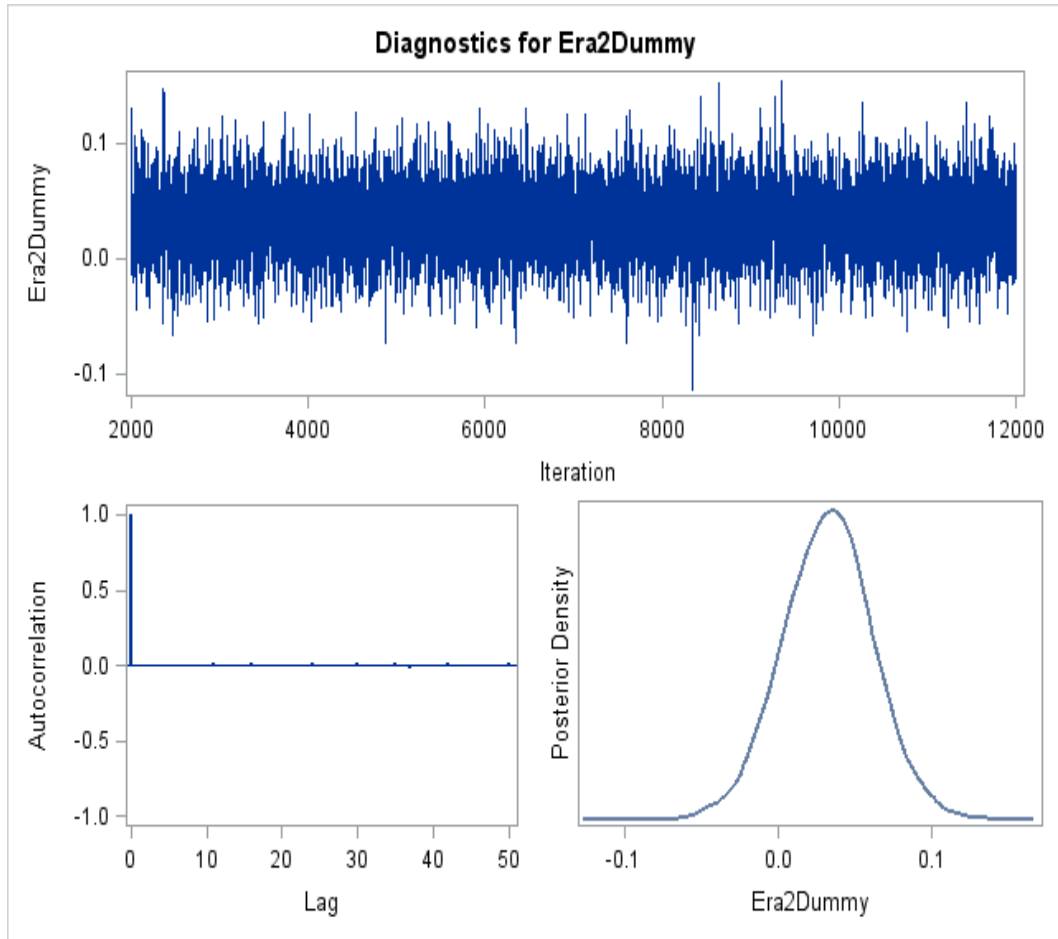


Figure B14. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1860-1892 era, when testing the difference between the 1896-1928 and 1932-1960 eras.

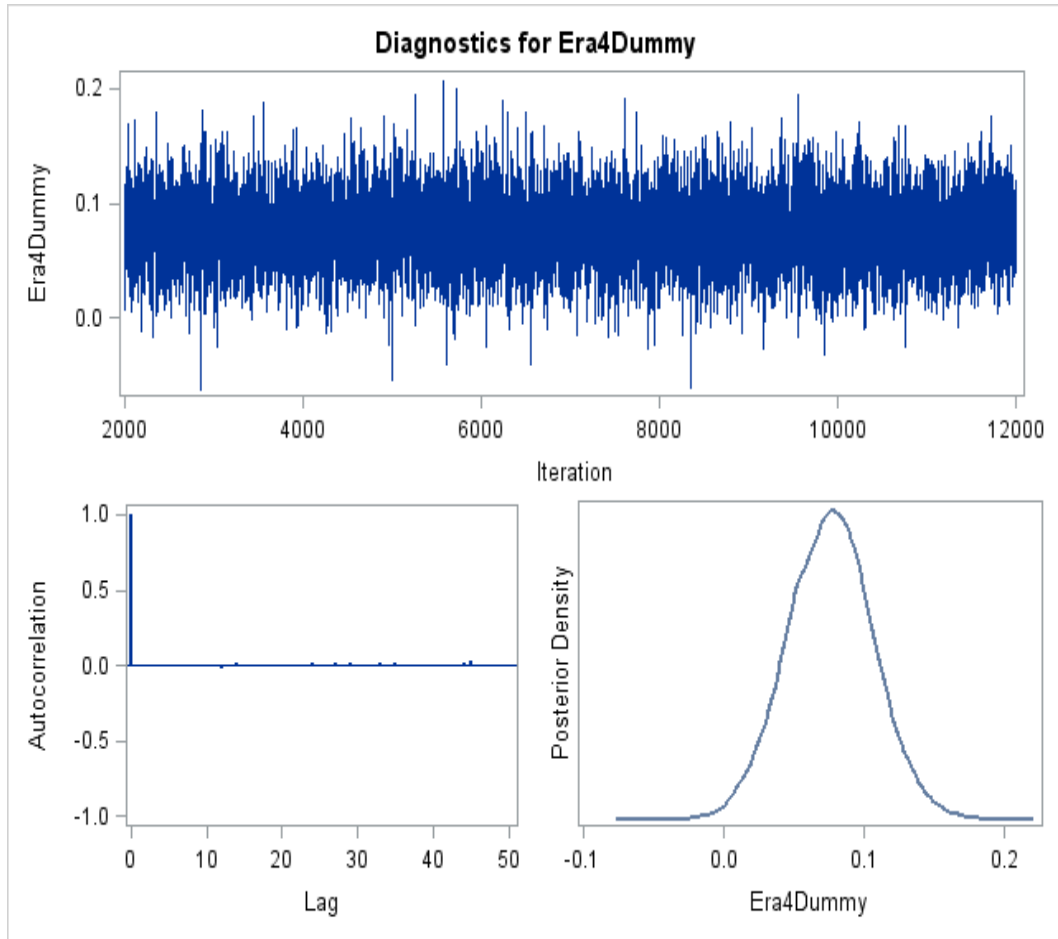


Figure B15. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1932-1960 era, when testing the difference between the 1896-1928 and 1932-1960 eras.

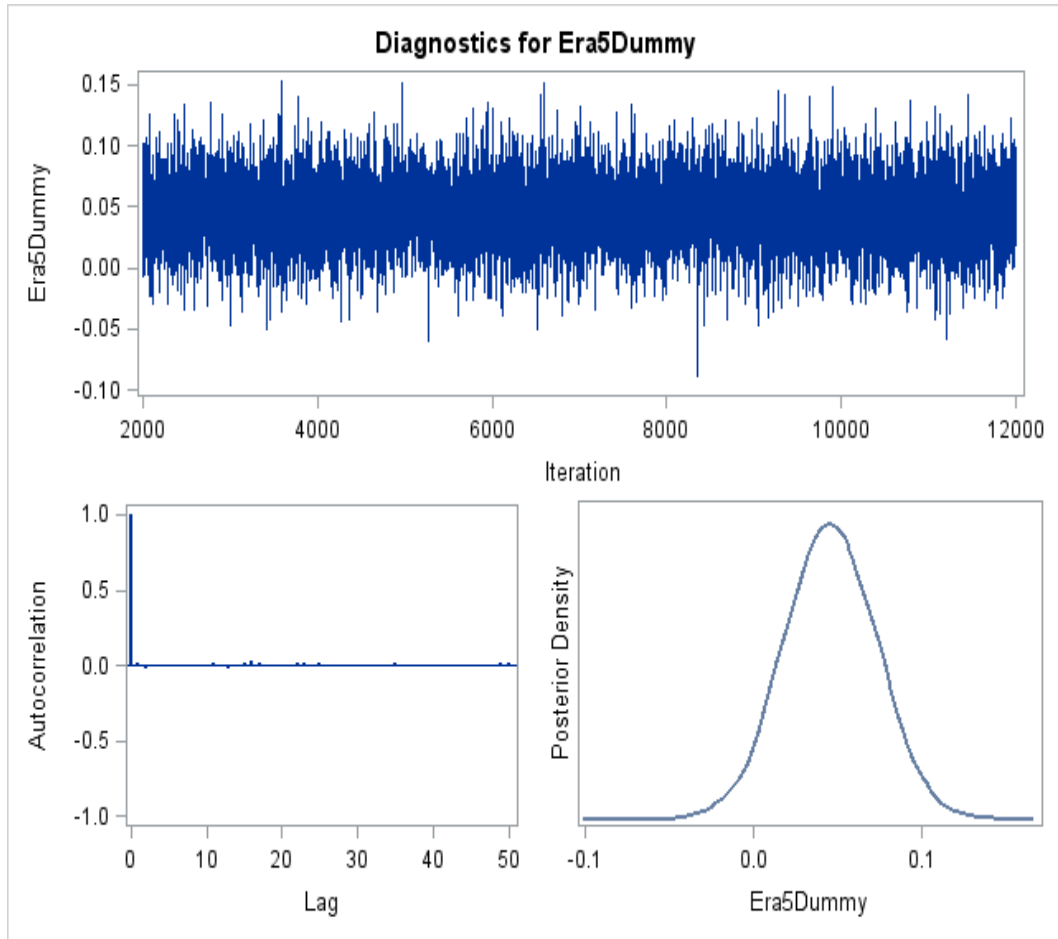


Figure B16. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1964-2008 era, when testing the difference between the 1896-1928 and 1932-1960 eras.

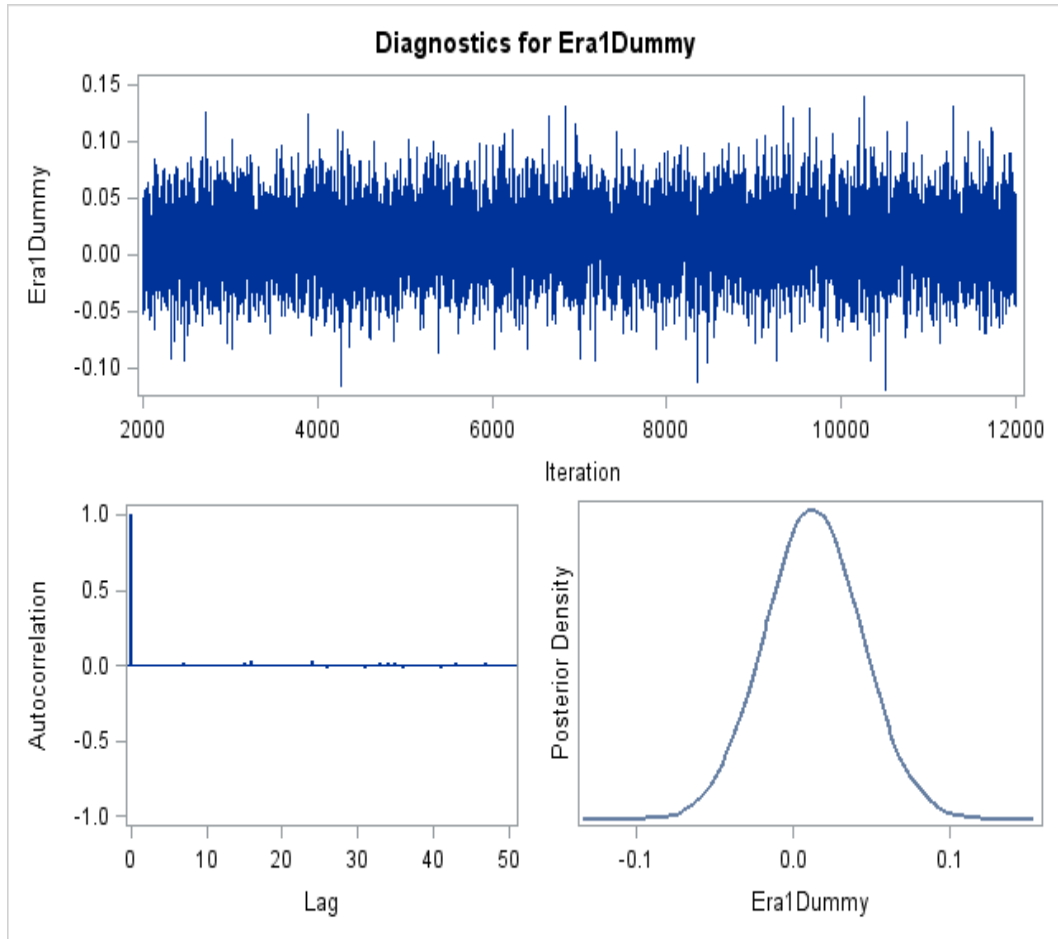


Figure B17. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1828-1856 era, when testing the difference between the 1932-1960 and 1964-2008 eras.

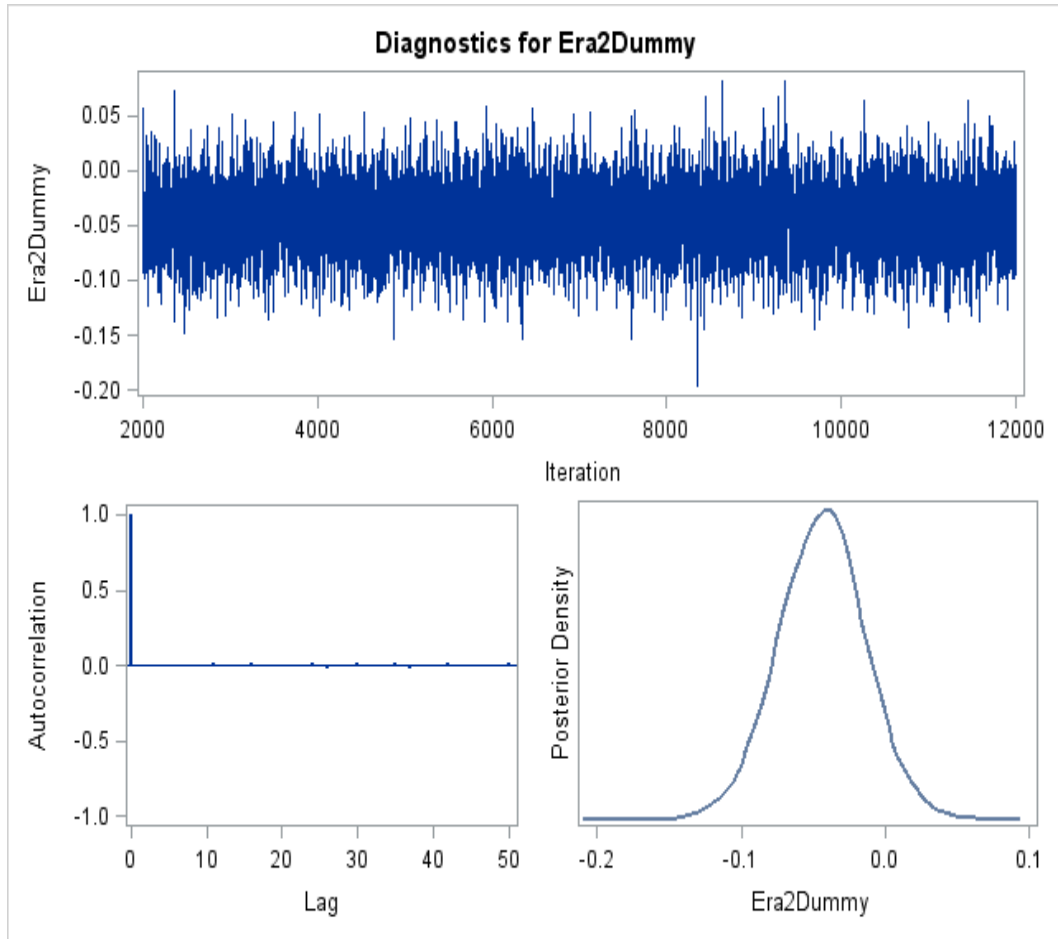


Figure B18. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1860-1892 era, when testing the difference between the 1932-1960 and 1964-2008 eras.

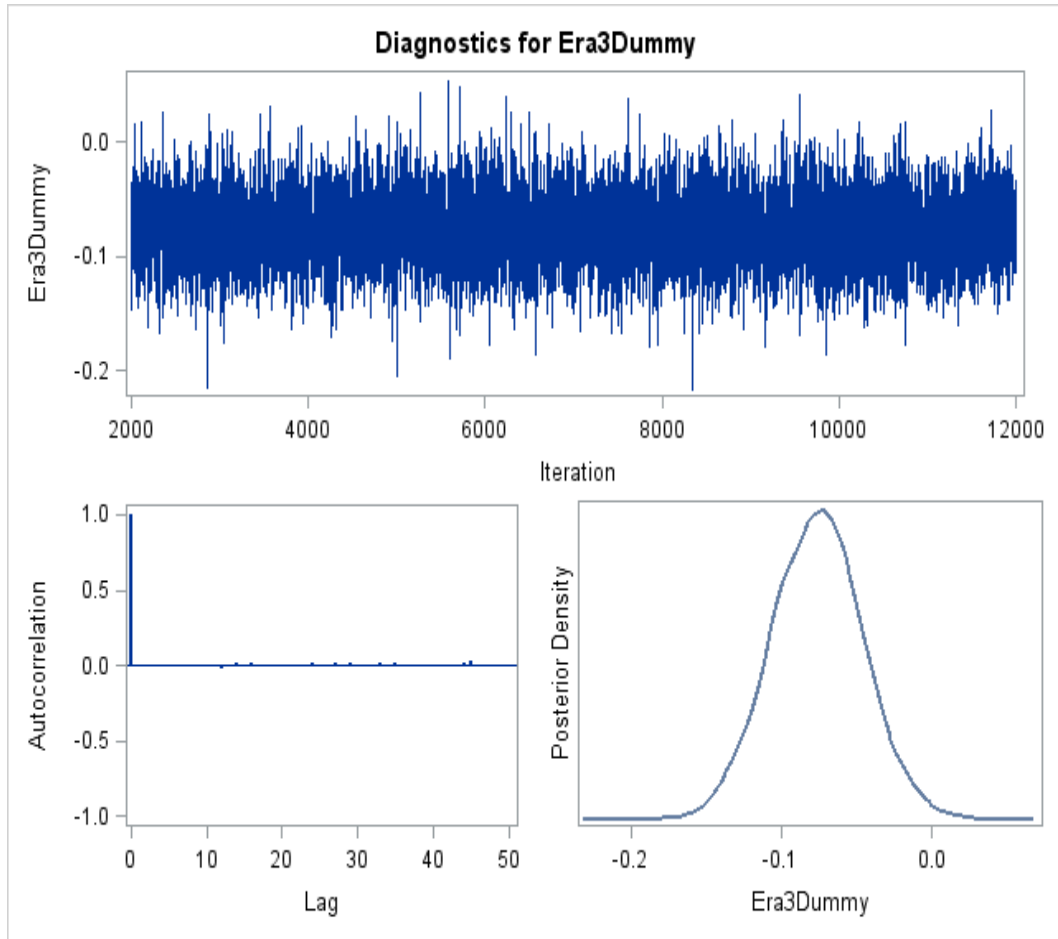


Figure B19. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1896-1928 era, when testing the difference between the 1932-1960 and 1964-2008 eras.

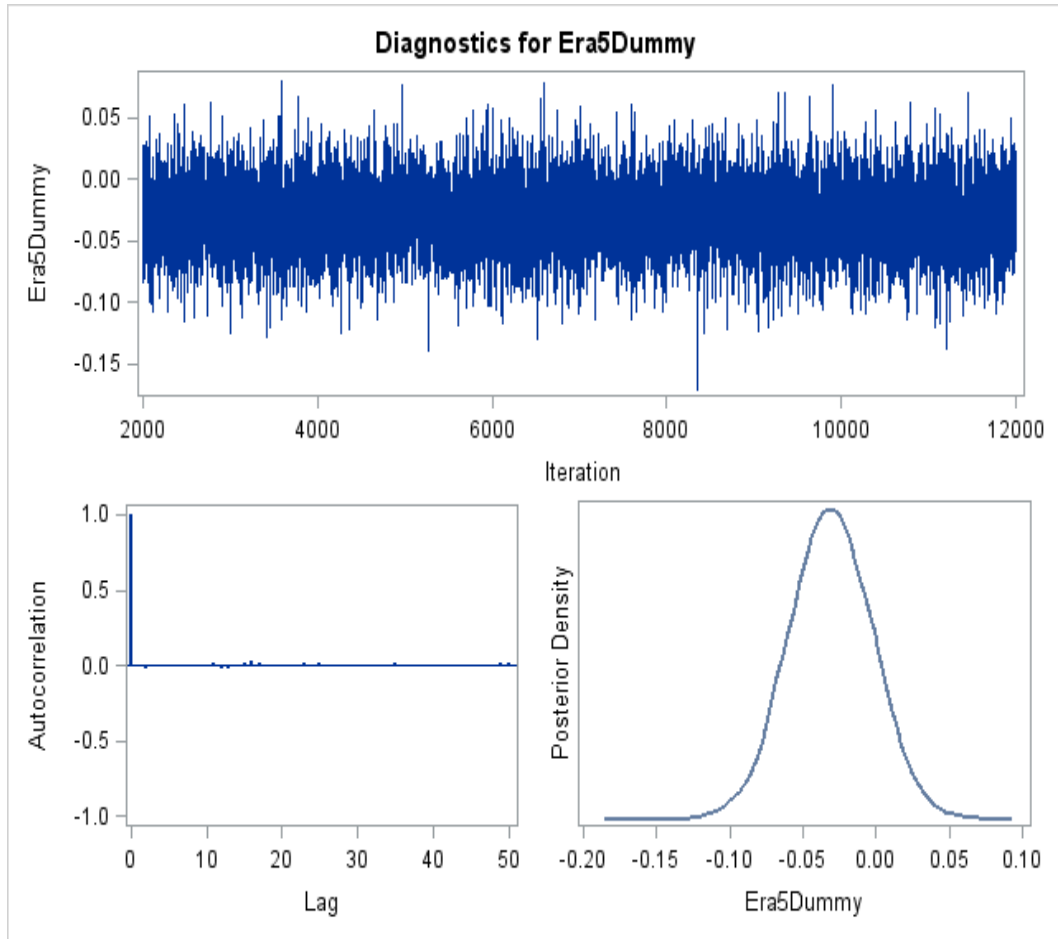


Figure B20. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative presidential model parameter of the 1964-2008 era, when testing the difference between the 1932-1960 and 1964-2008 eras.

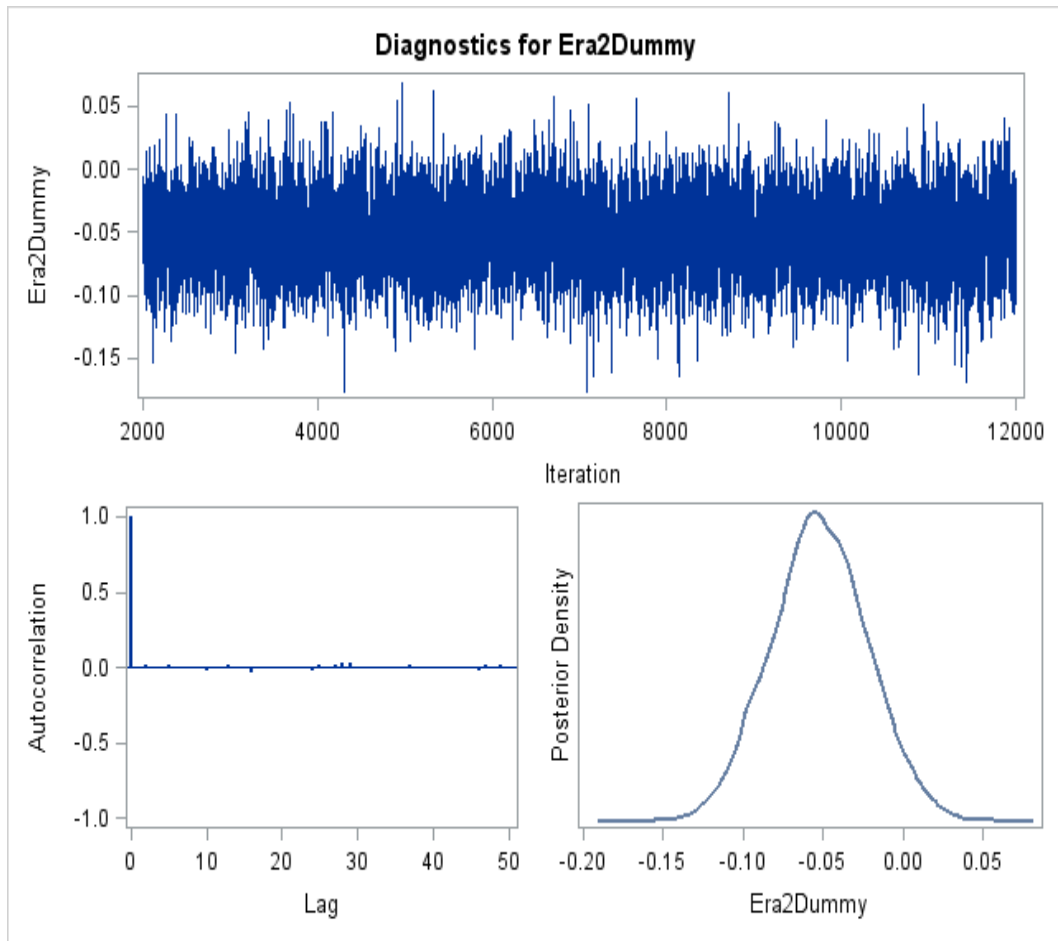


Figure B21. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1860-1892 era, when testing the difference between the 1828-1856 and 1860-1892 eras.

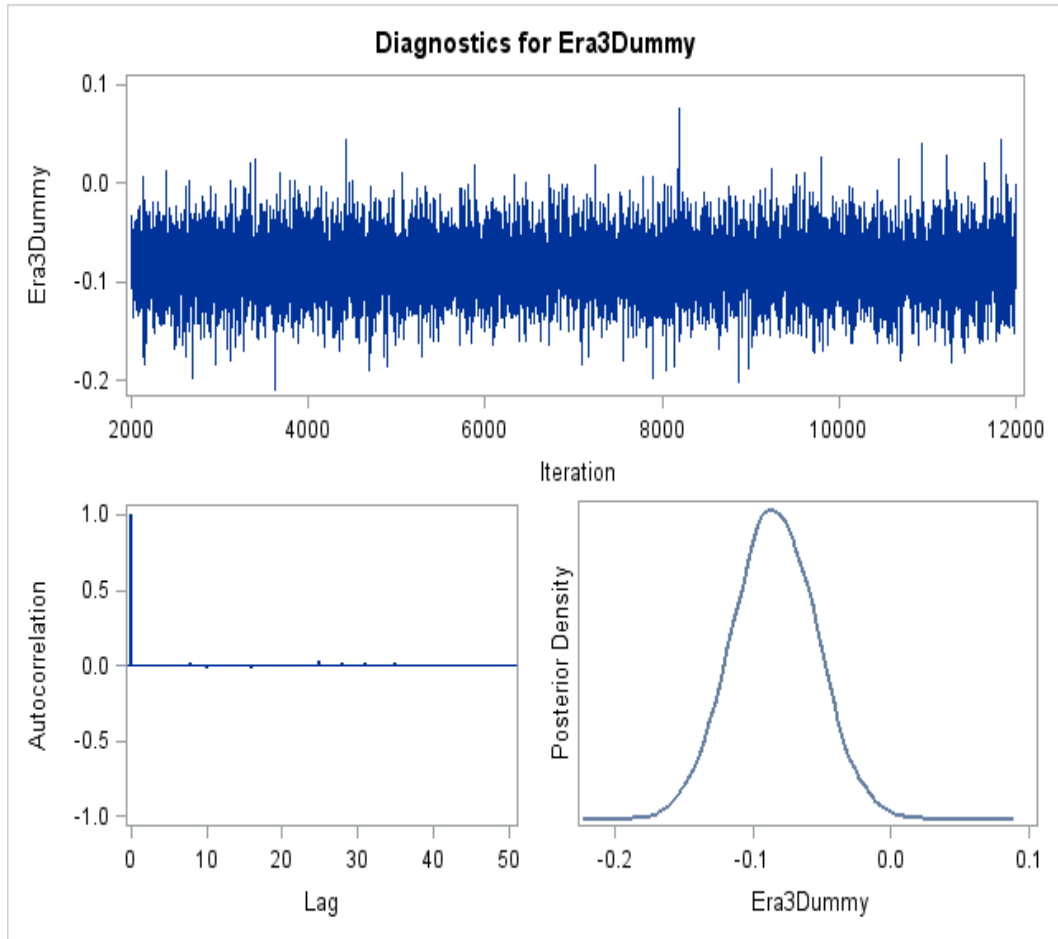


Figure B22. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1896-1928 era, when testing the difference between the 1828-1856 and 1860-1892 eras.

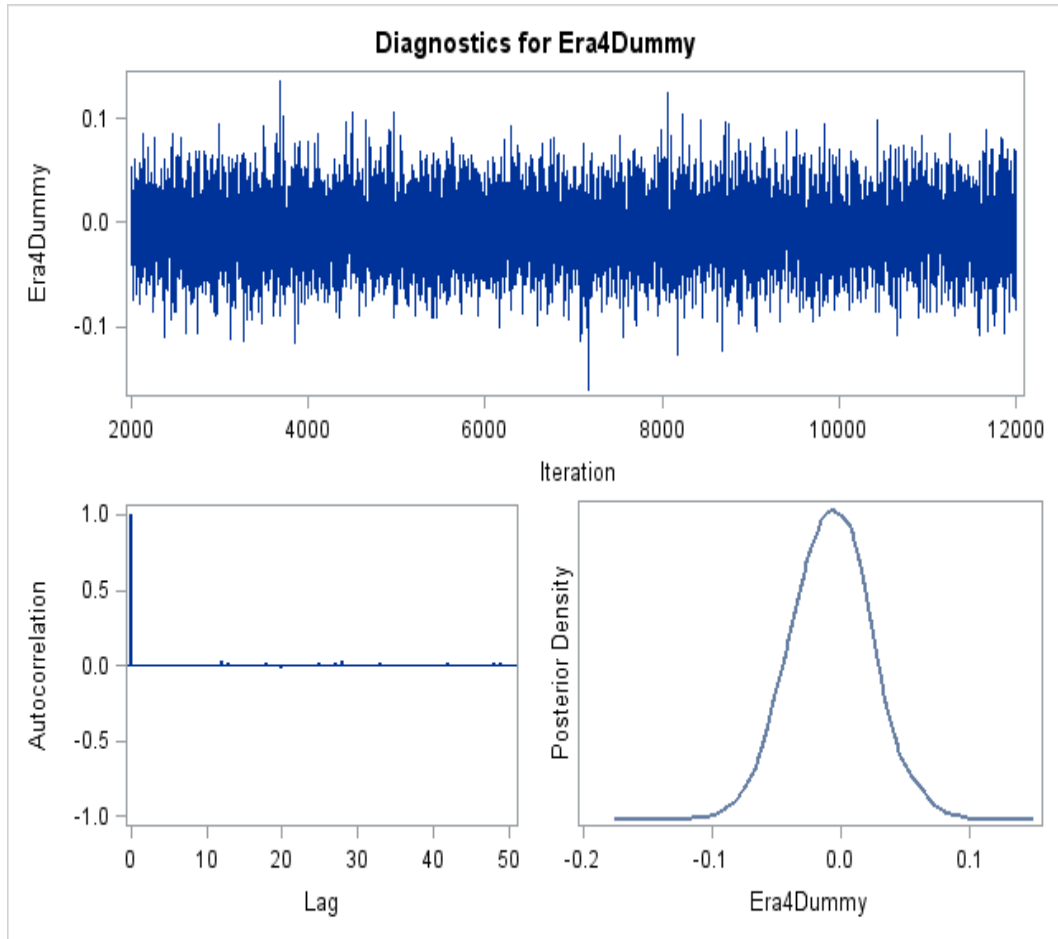


Figure B23. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1932-1960 era, when testing the difference between the 1828-1856 and 1860-1892 eras.

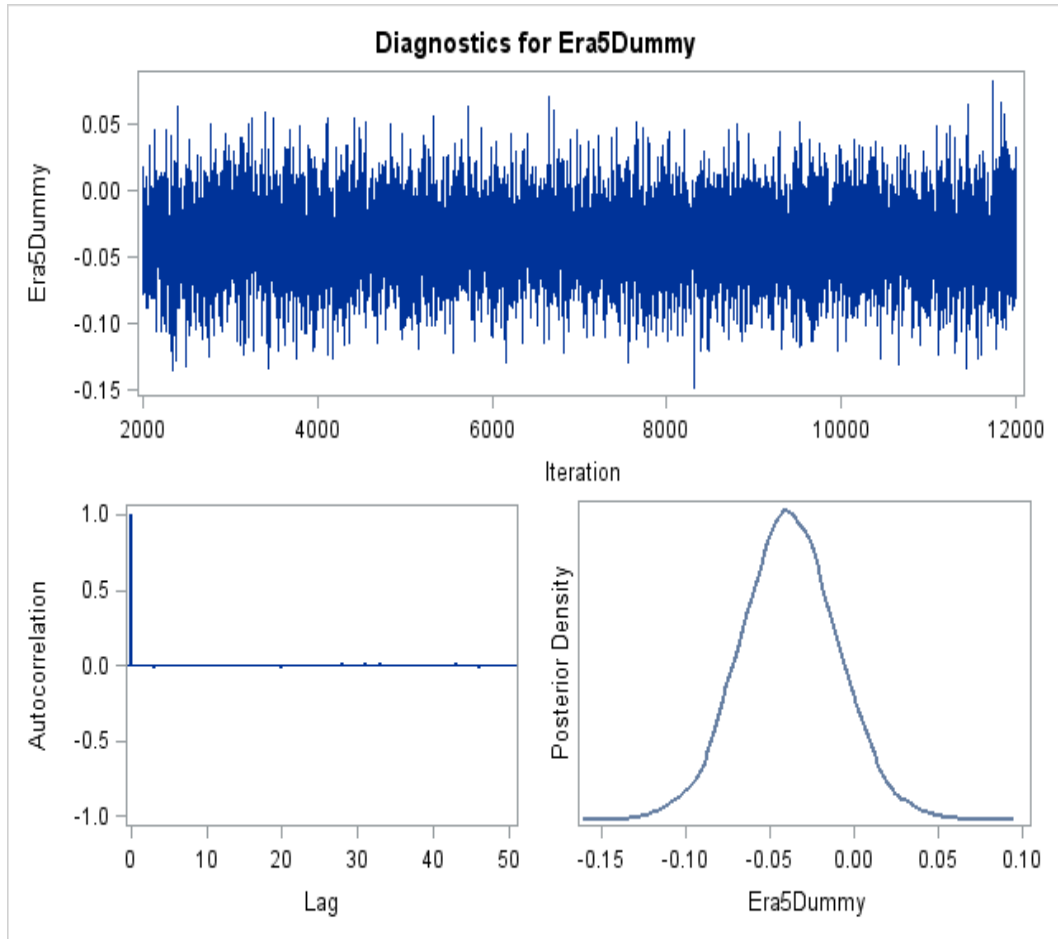


Figure B24. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1964-2008 era, when testing the difference between the 1828-1856 and 1860-1892 eras.

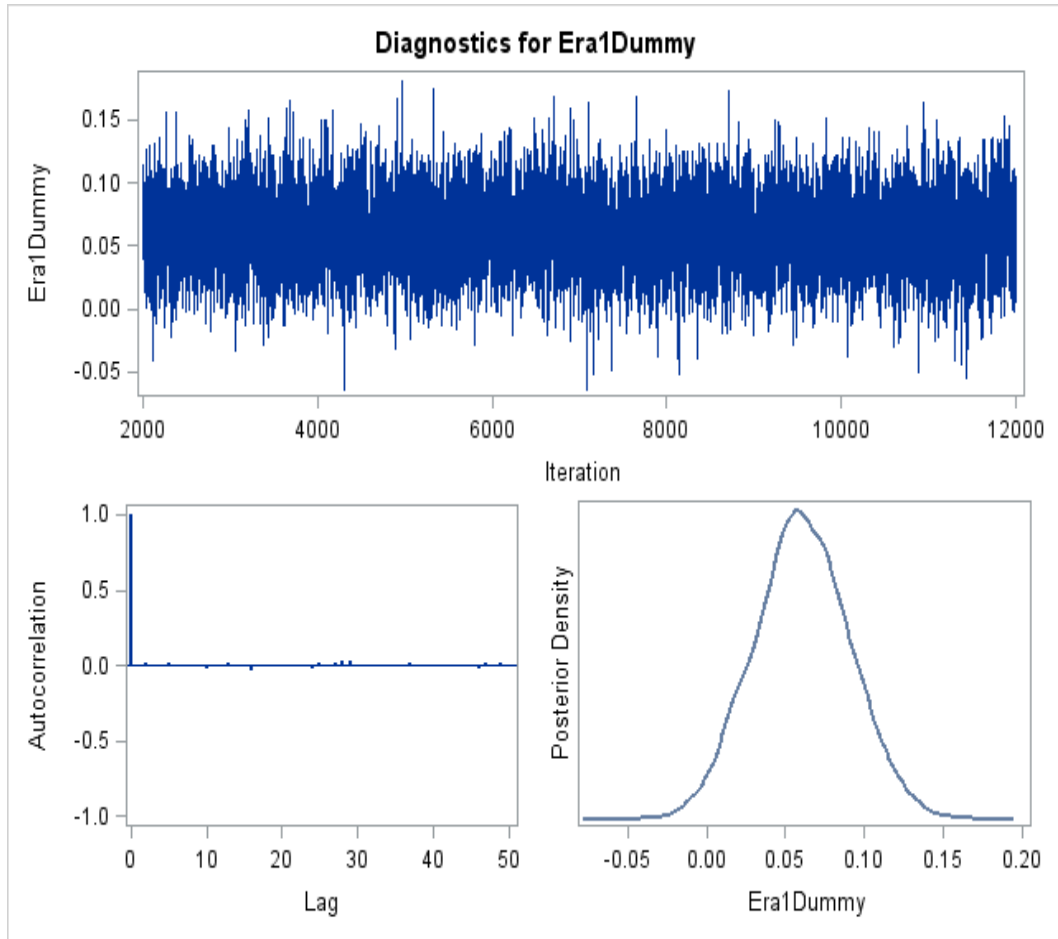


Figure B25. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1828-1856 era, when testing the difference between the 1860-1892 and 1896-1928 eras.

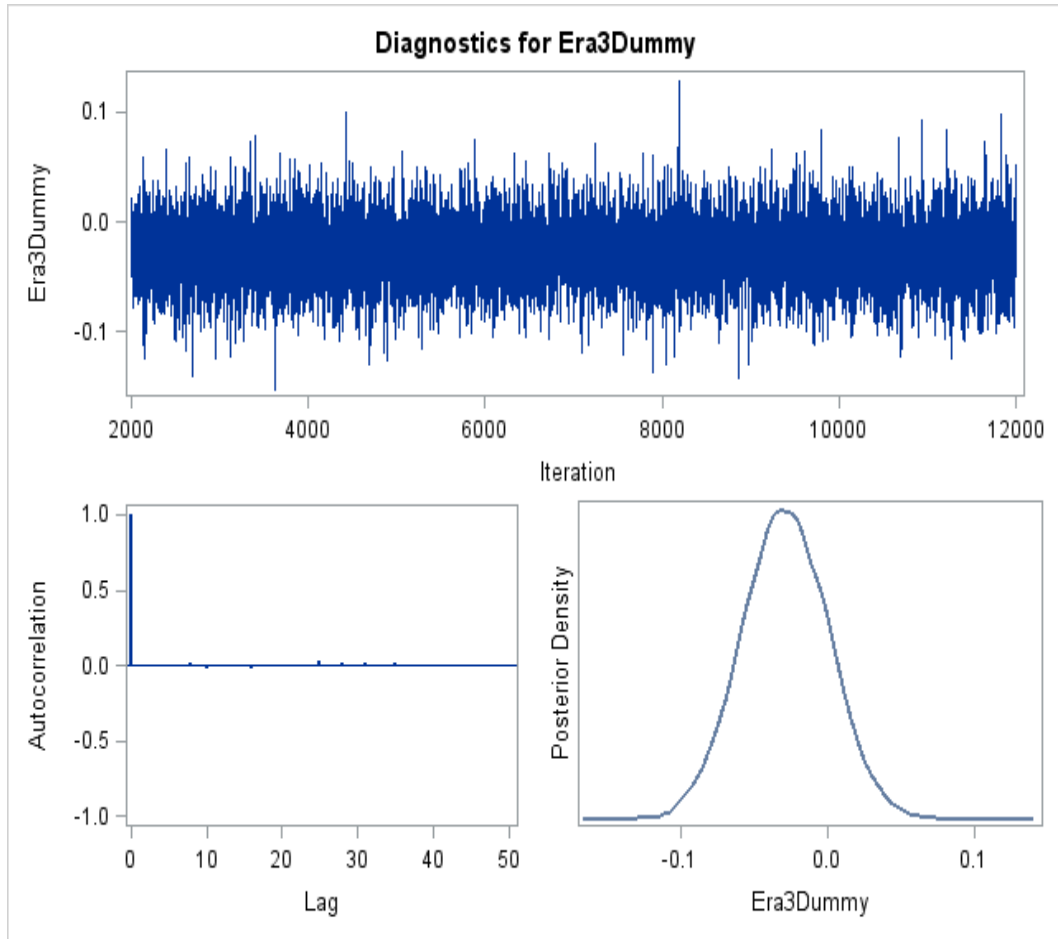


Figure B26. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1896-1928 era, when testing the difference between the 1860-1892 and 1896-1928 eras.

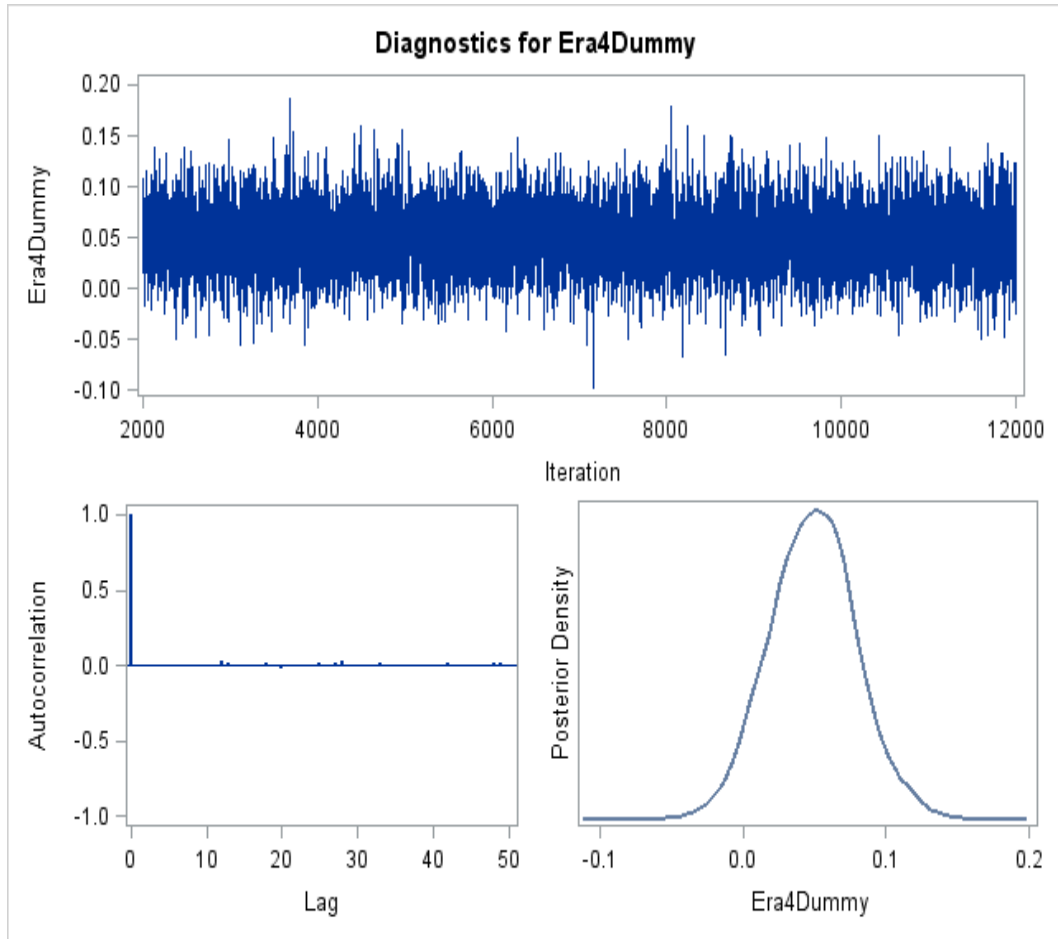


Figure B27. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1932-1960 era, when testing the difference between the 1860-1892 and 1896-1928 eras.

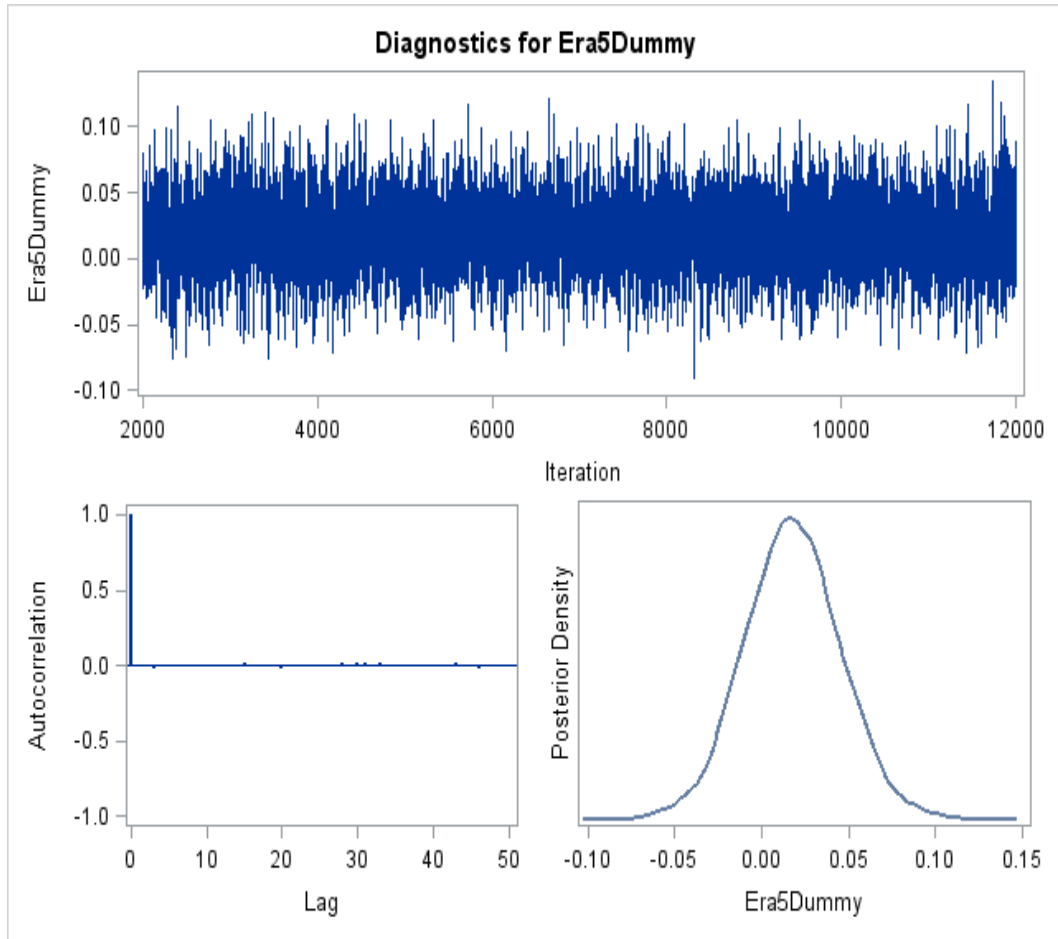


Figure B28. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1964-2008 era, when testing the difference between the 1860-1892 and 1896-1928 eras.

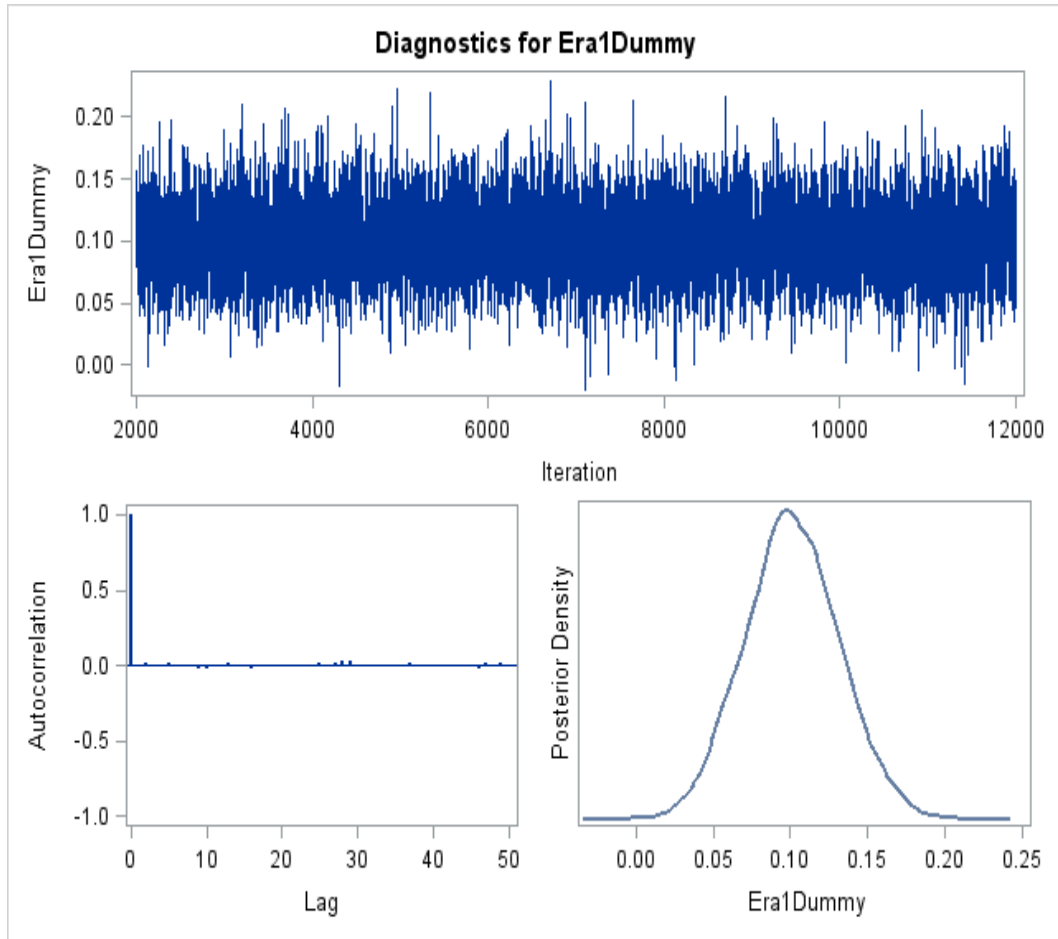


Figure B29. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1828-1856 era, when testing the difference between the 1896-1928 and 1932-1960 eras.

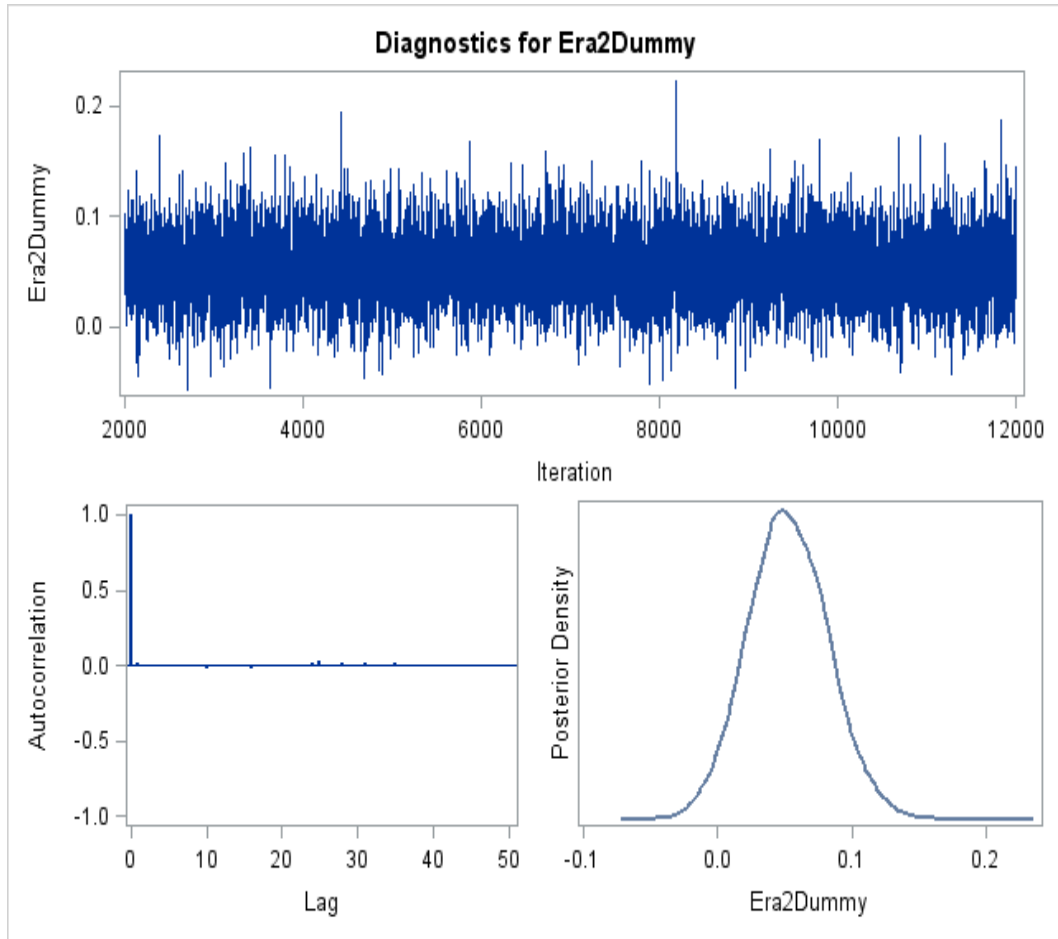


Figure B30. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1860-1892 era, when testing the difference between the 1896-1928 and 1932-1960 eras.

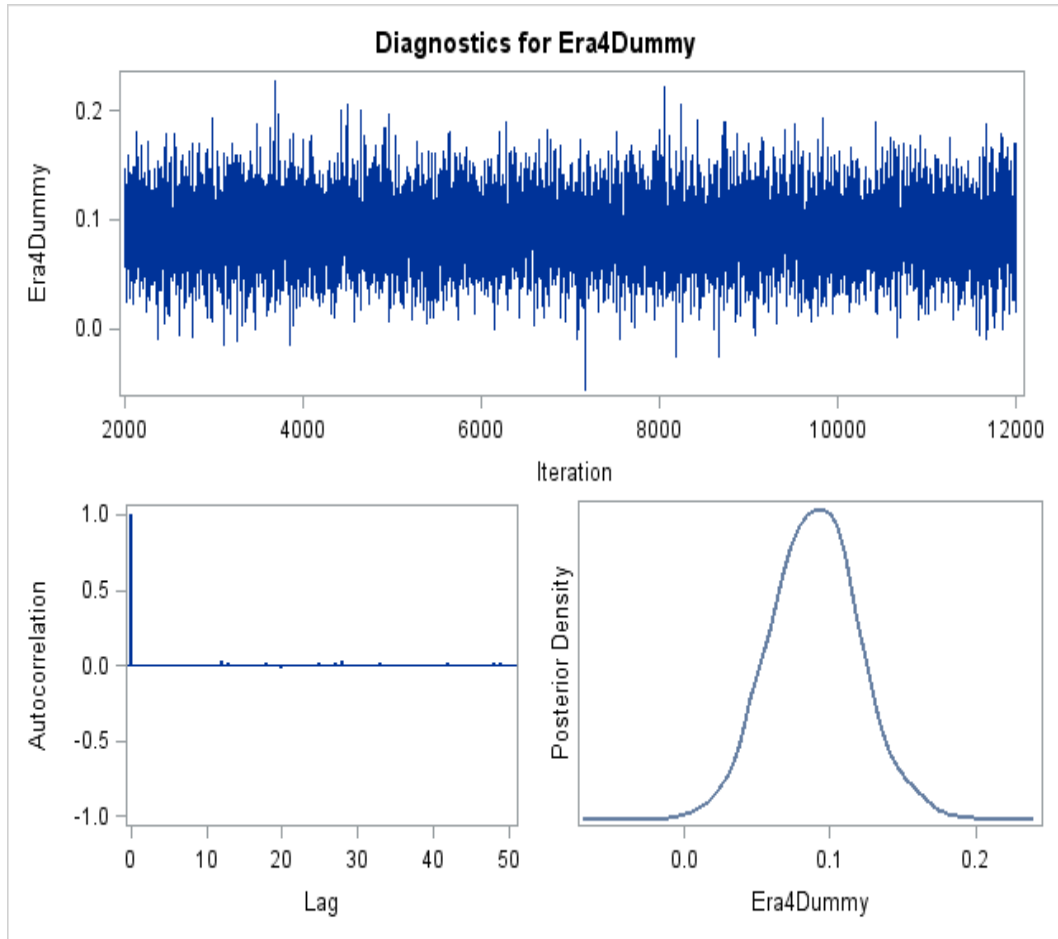


Figure B31. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1932-1960 era, when testing the difference between the 1896-1928 and 1932-1960 eras.

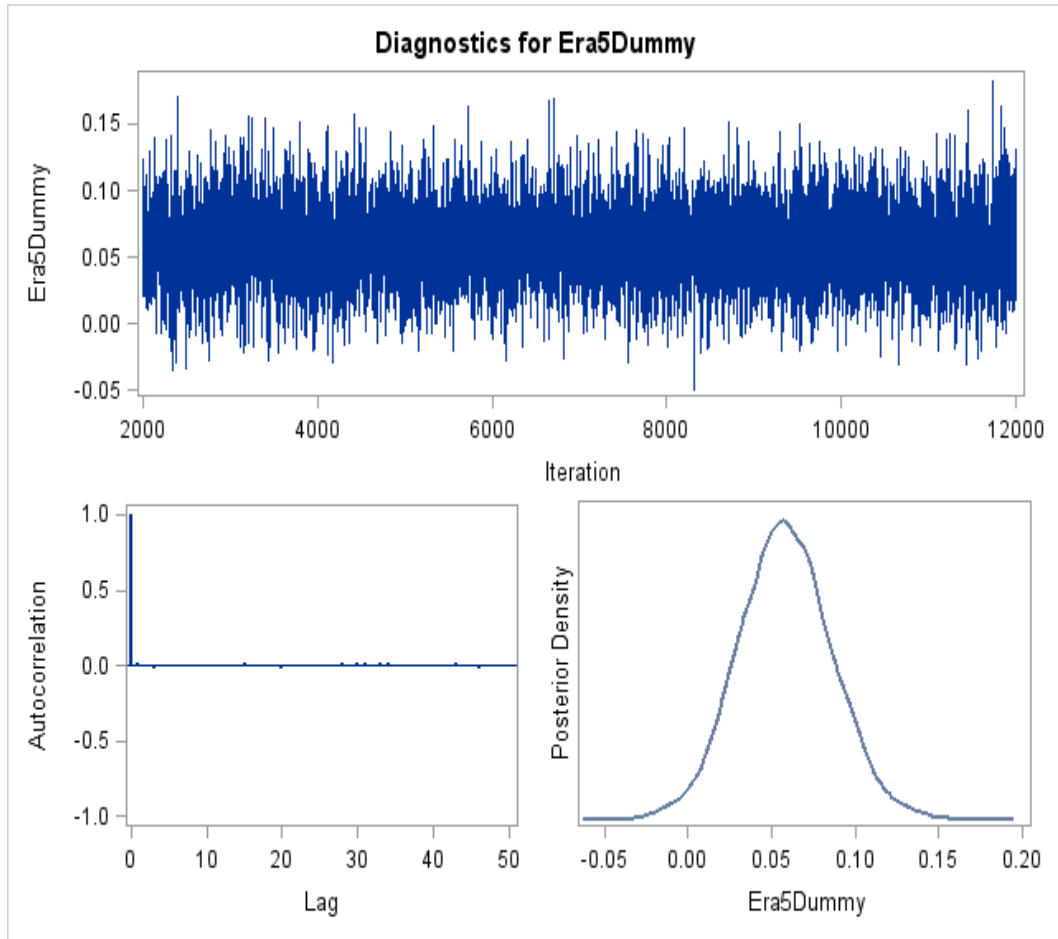


Figure B32. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1964-2008 era, when testing the difference between the 1896-1928 and 1932-1960 eras.

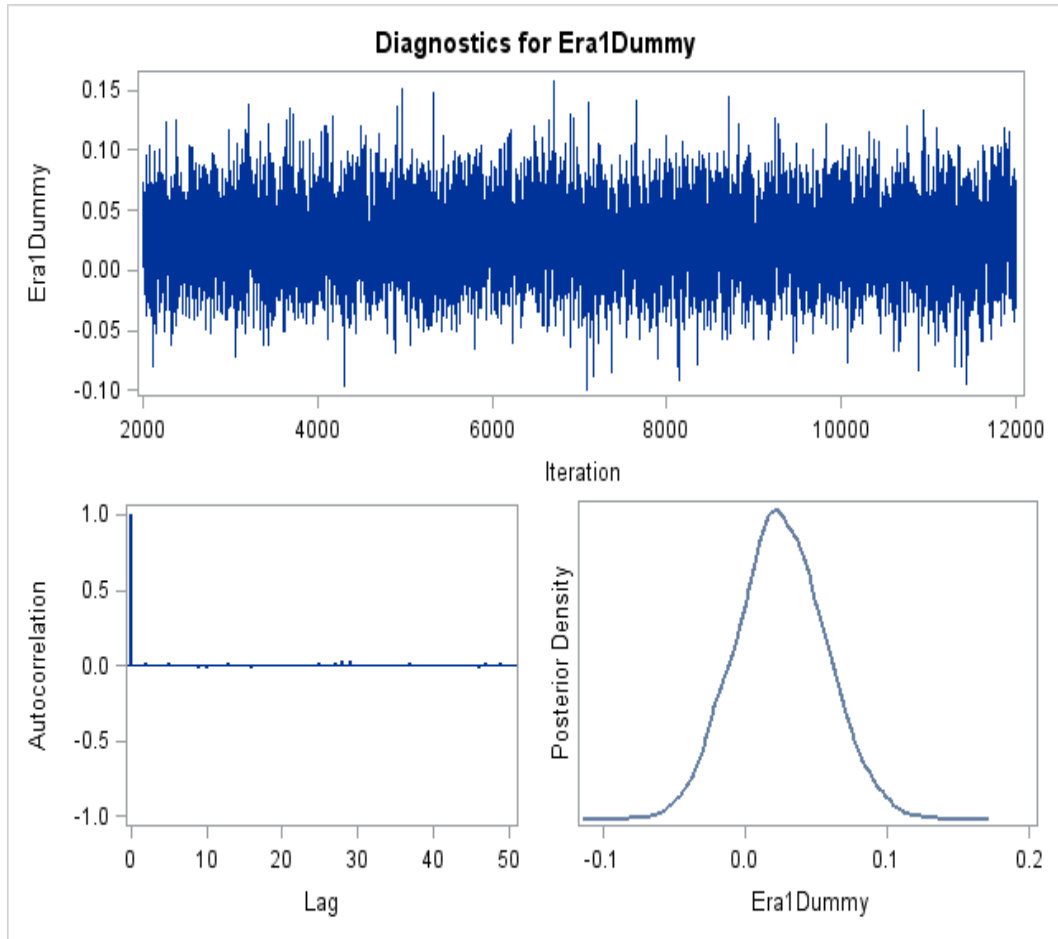


Figure B33. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1828-1856 era, when testing the difference between the 1932-1960 and 1964-2008 eras.

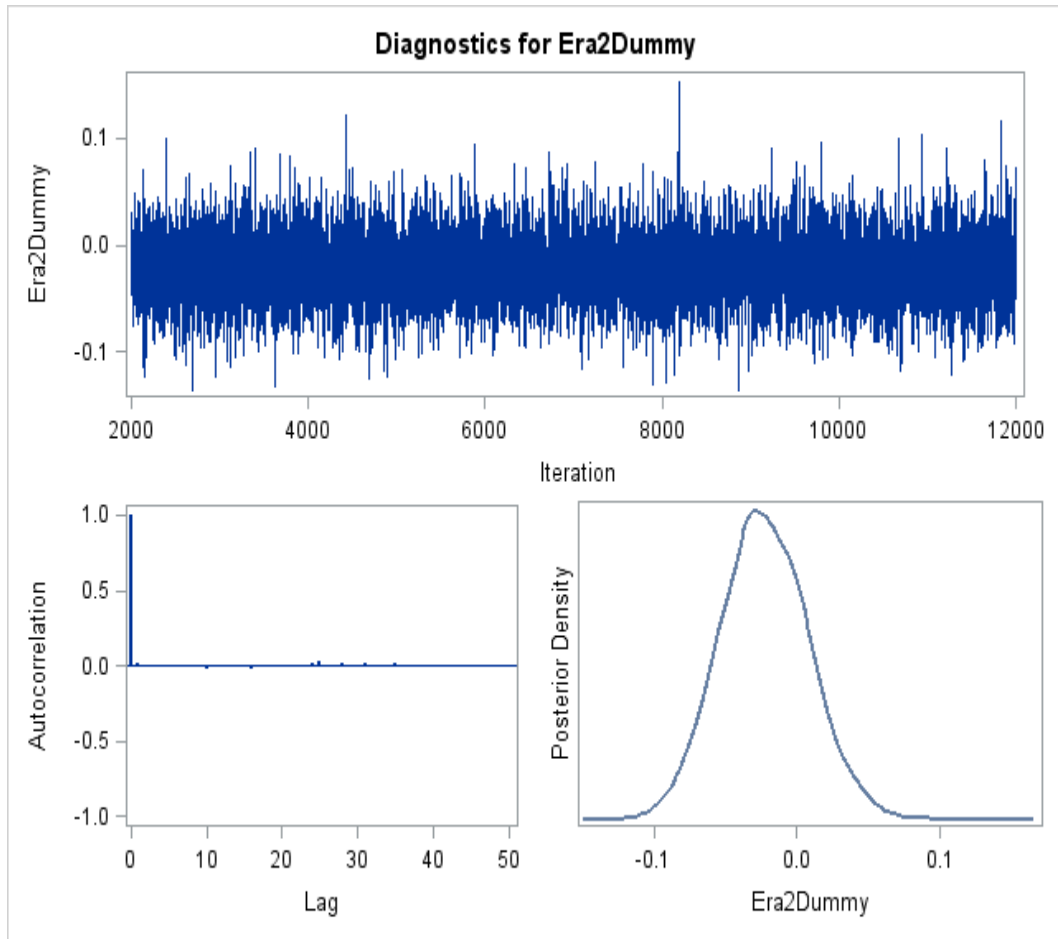


Figure B34. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1860-1892 era, when testing the difference between the 1932-1960 and 1964-2008 eras.

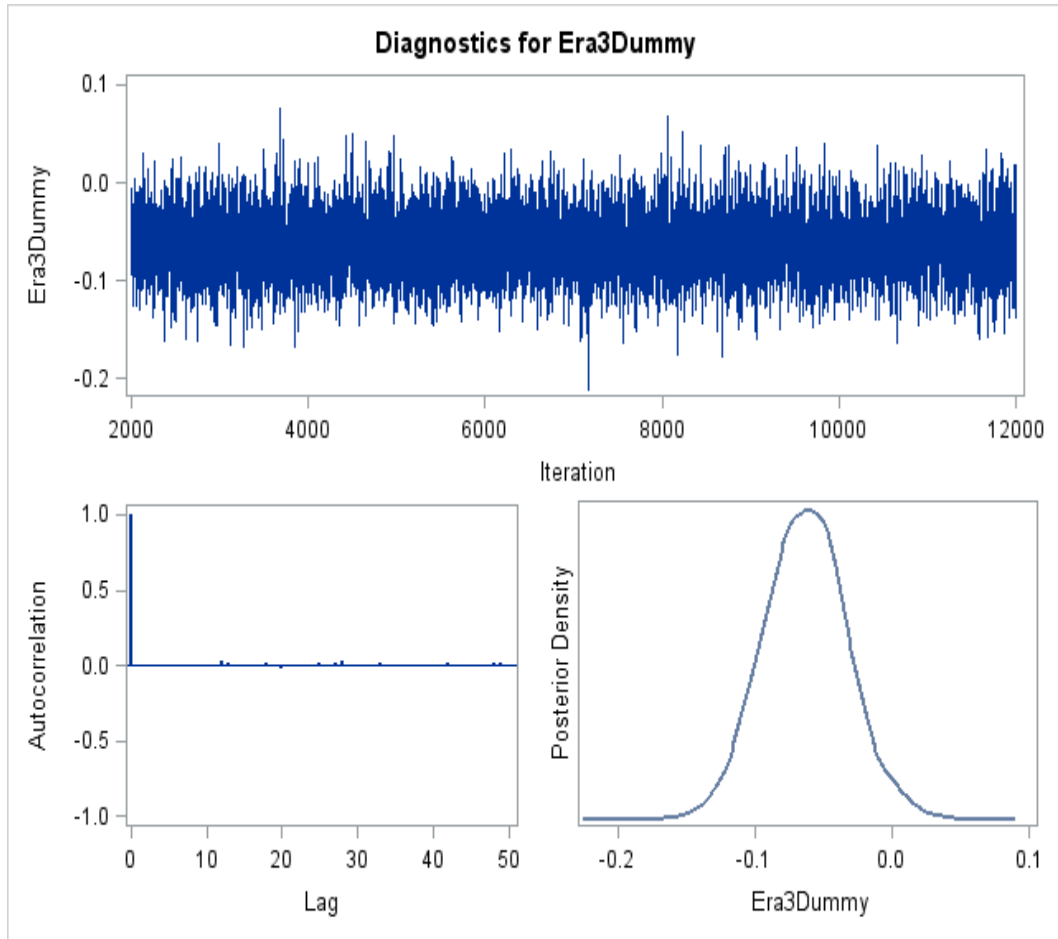


Figure B35. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1896-1928 era, when testing the difference between the 1932-1960 and 1964-2008 eras.

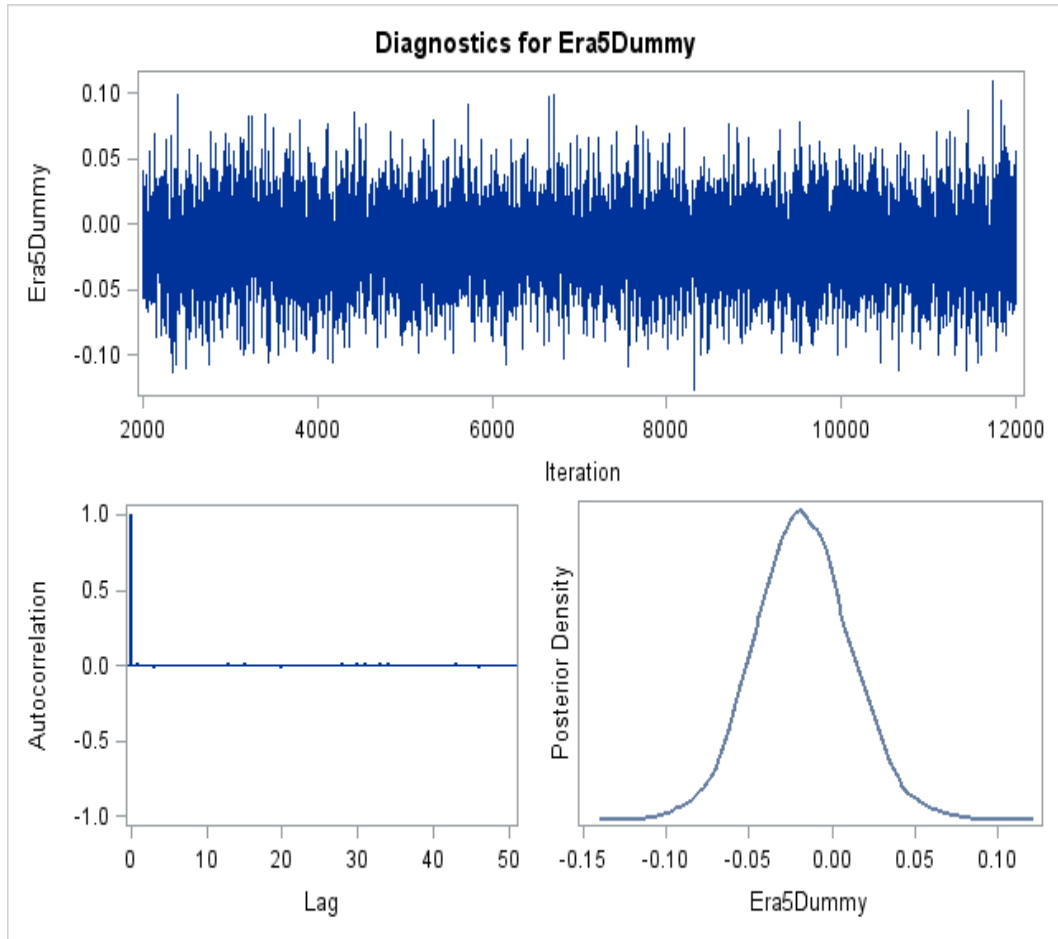


Figure B36. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative presidential model parameter of the 1964-2008 era, when testing the difference between the 1932-1960 and 1964-2008 eras.

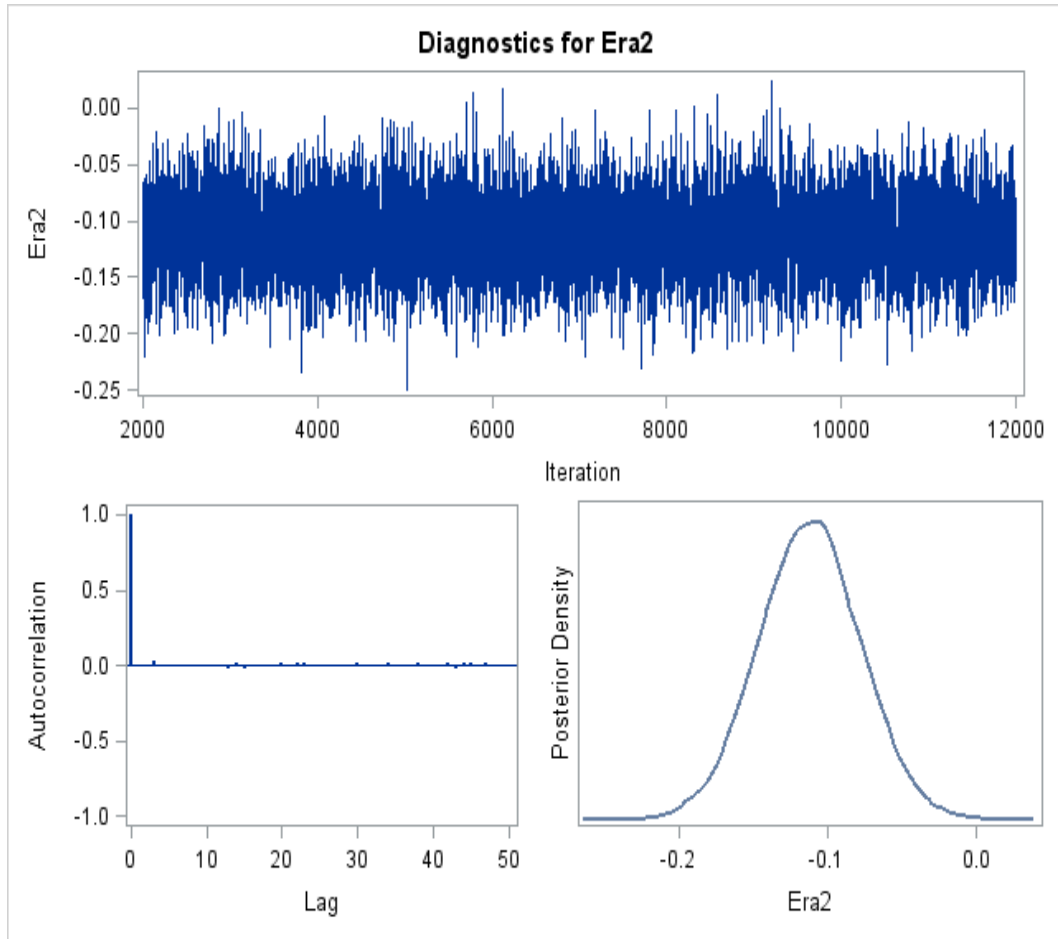


Figure B37. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1860-1894 era, when testing the difference between the 1828-1858 and 1860-1894 eras.

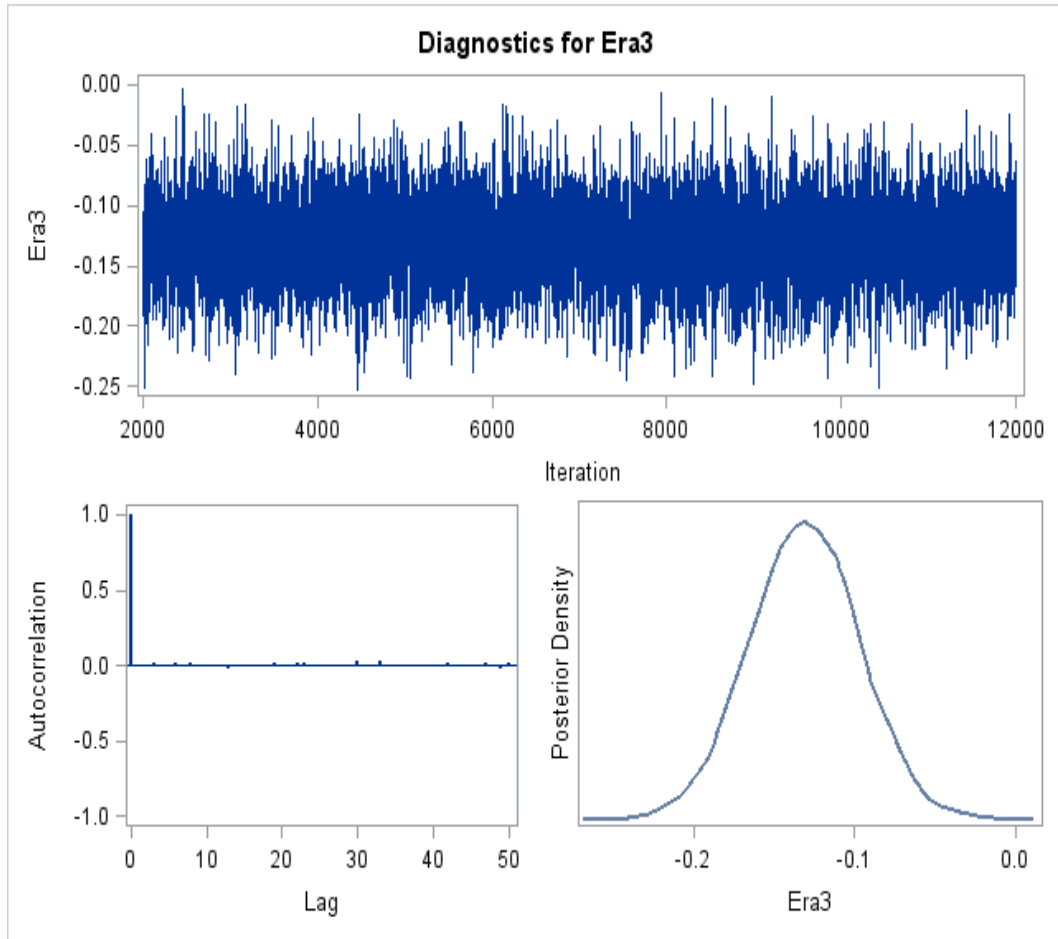


Figure B38. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1896-1930 era, when testing the difference between the 1828-1858 and 1860-1894 eras.

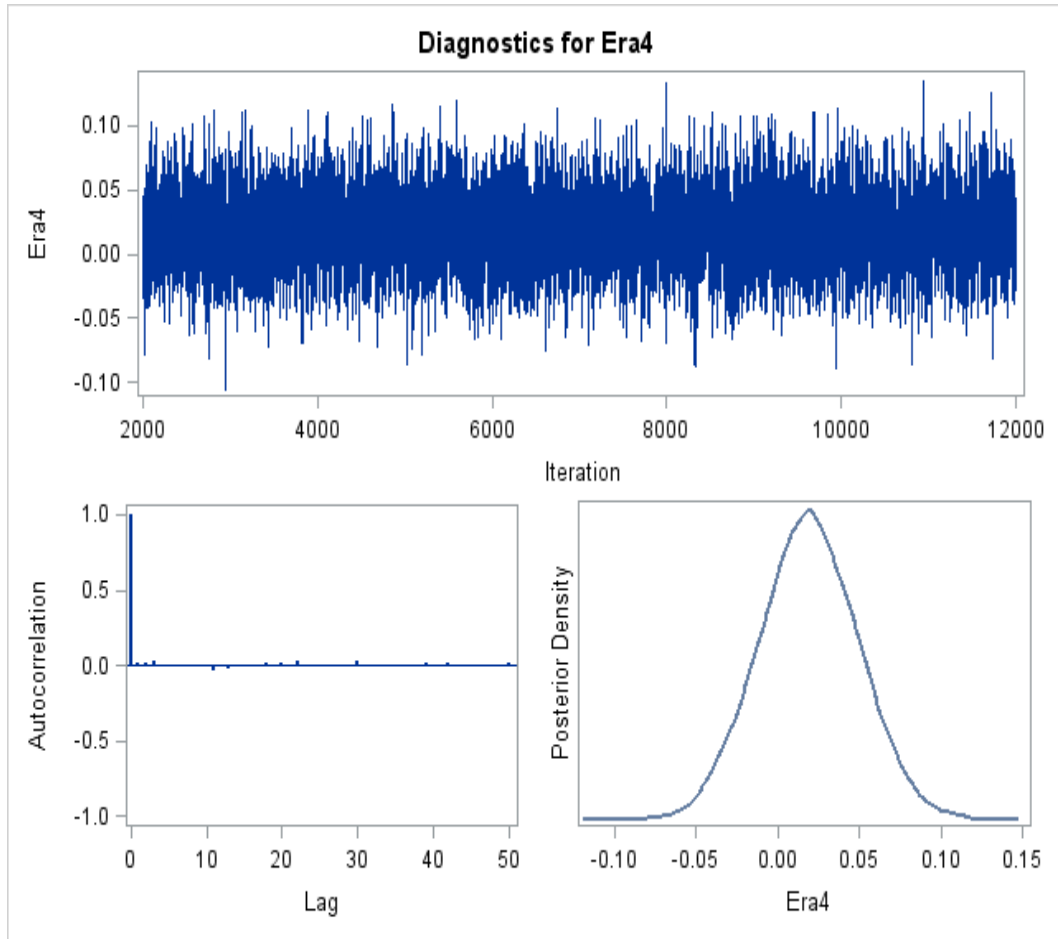


Figure B39. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1932-1994 era, when testing the difference between the 1828-1858 and 1860-1894 eras.

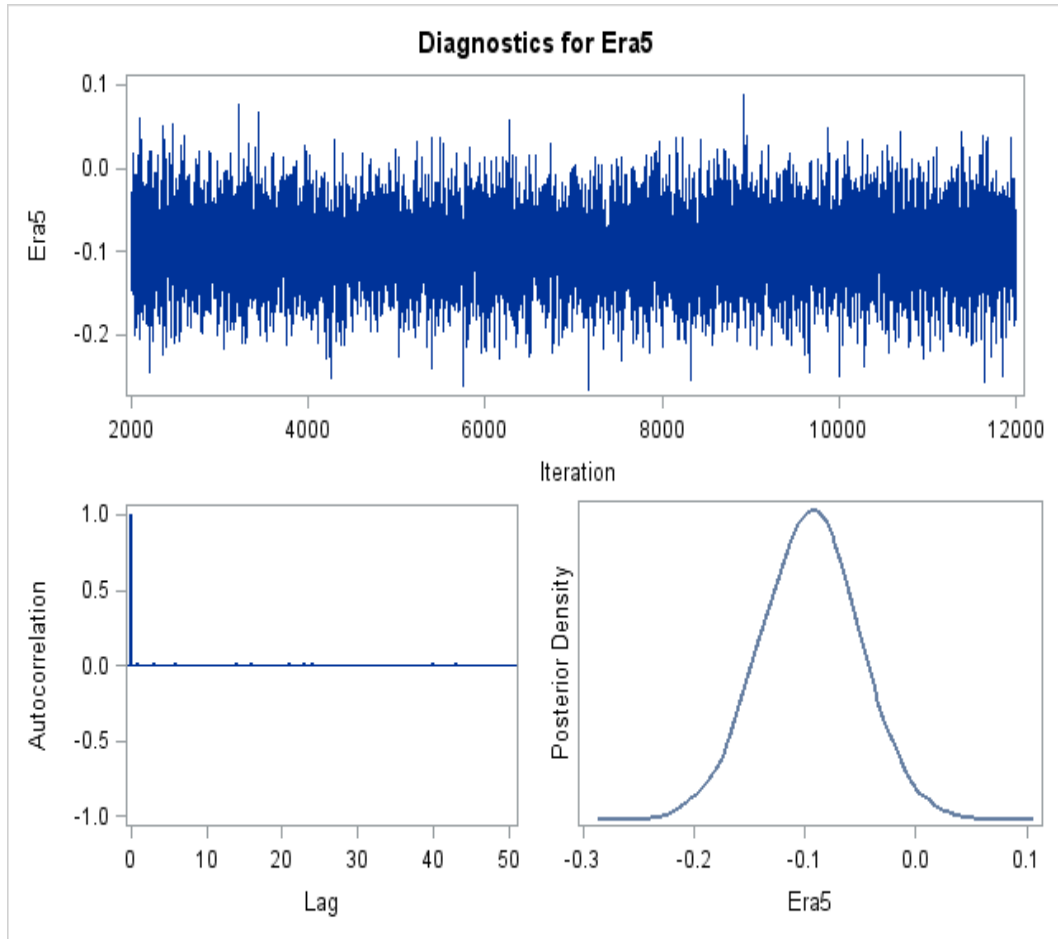


Figure B40. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1996-2008 era, when testing the difference between the 1828-1858 and 1860-1894 eras.

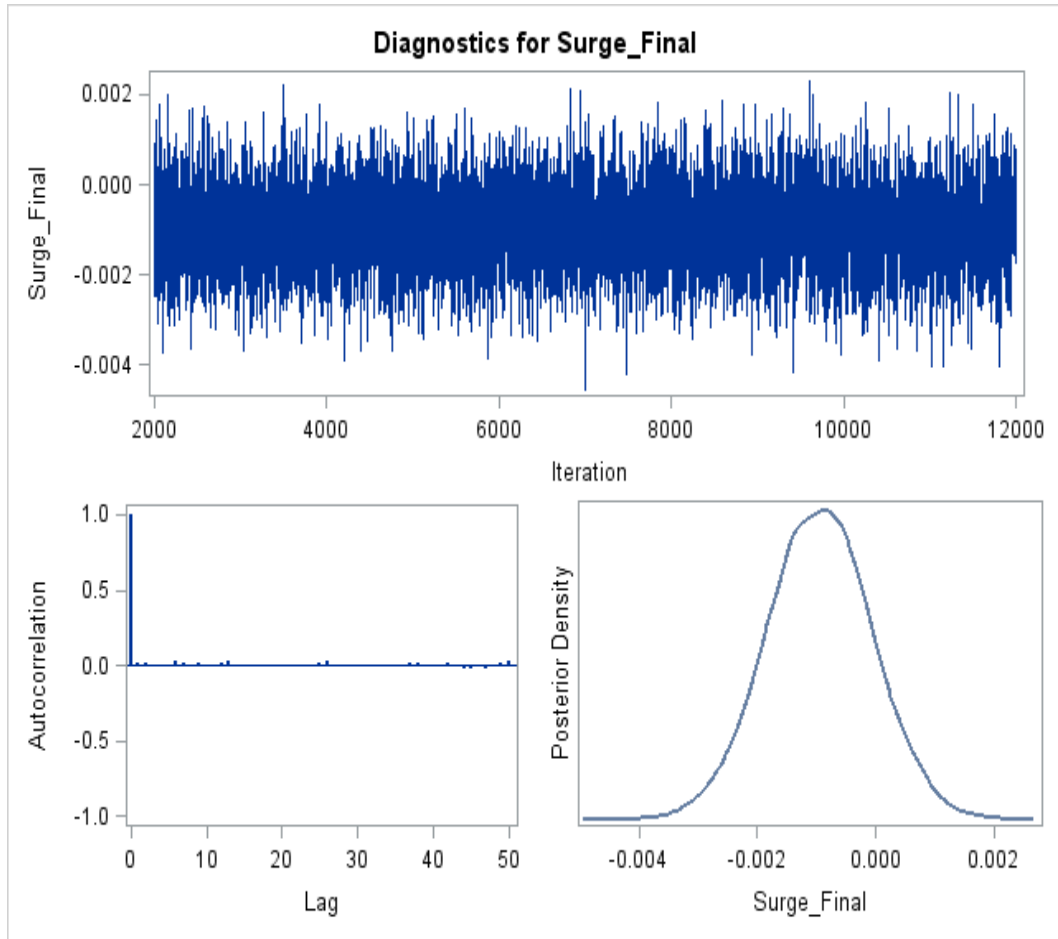


Figure B41. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter controlling for general election surge, when testing the difference between the 1828-1858 and 1860-1894 eras.

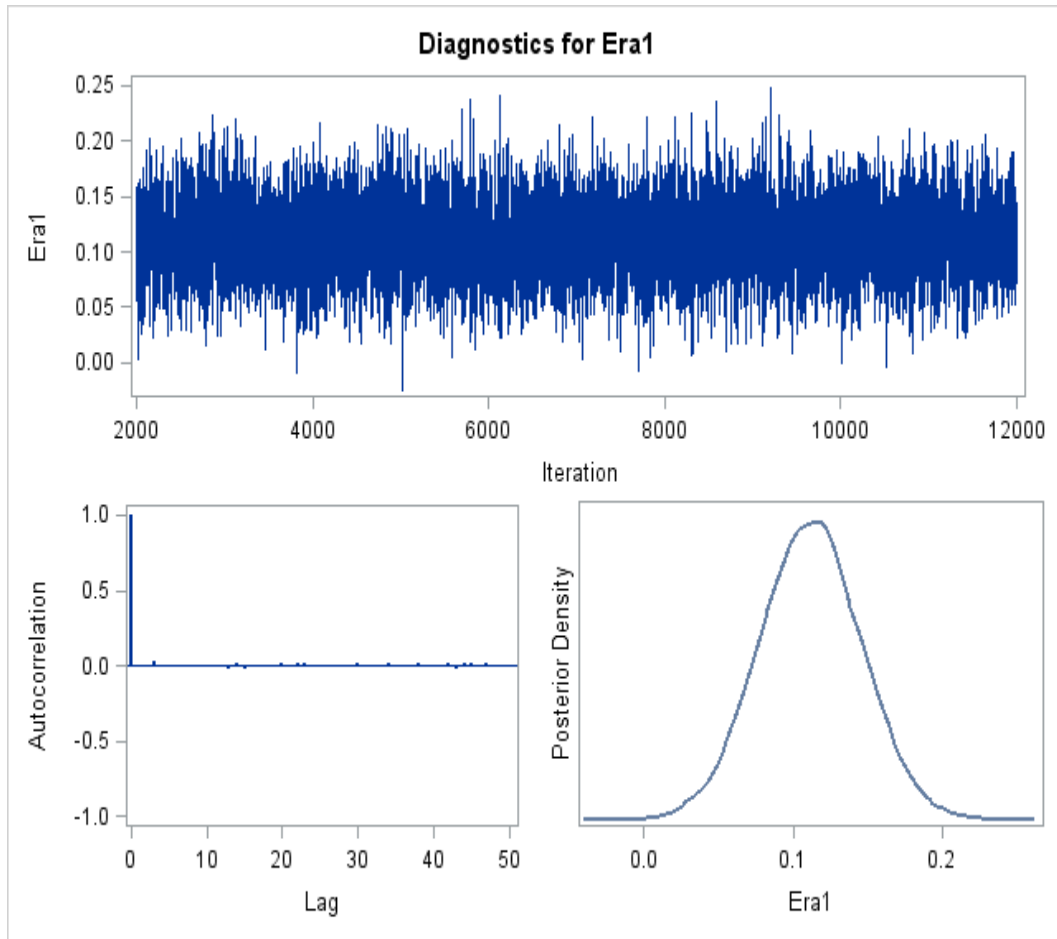


Figure B42. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1828-1858 era, when testing the difference between the 1860-1894 and 1896-1930 eras.

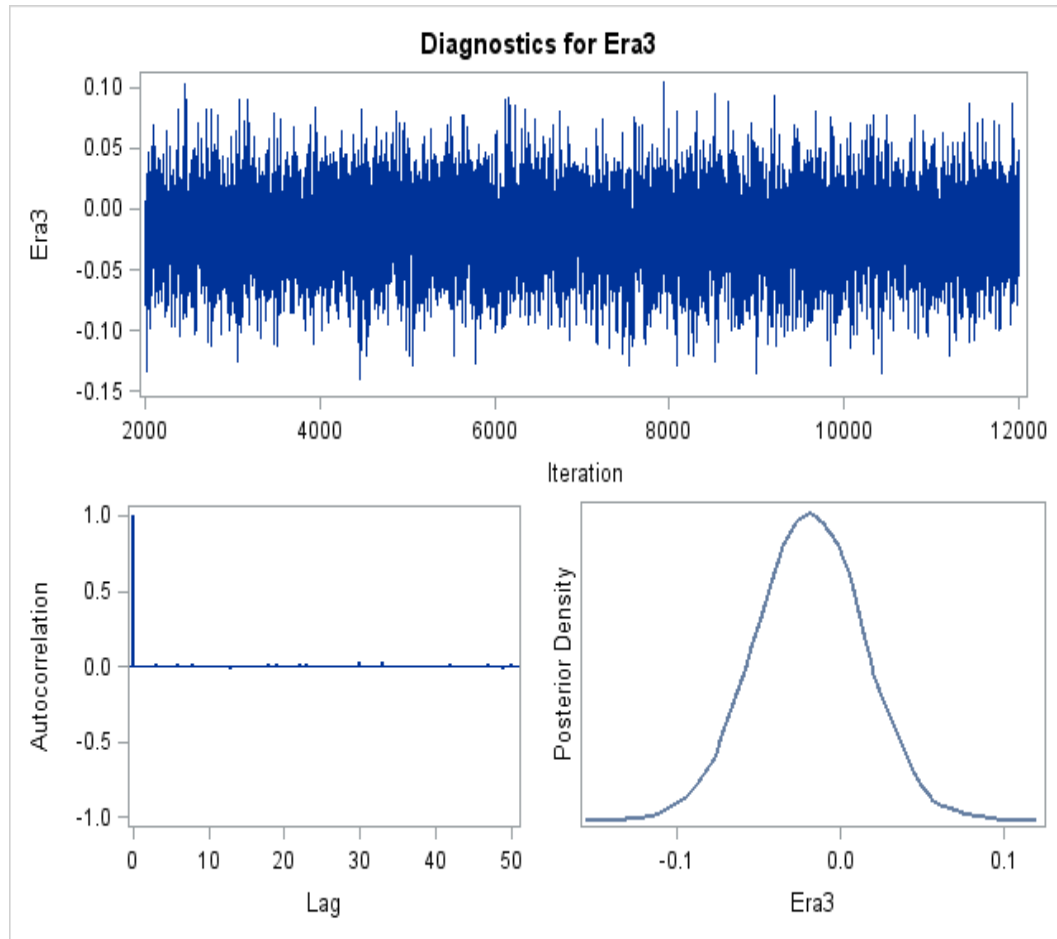


Figure B43. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1896-1930 era, when testing the difference between the 1860-1894 and 1896-1930 eras.

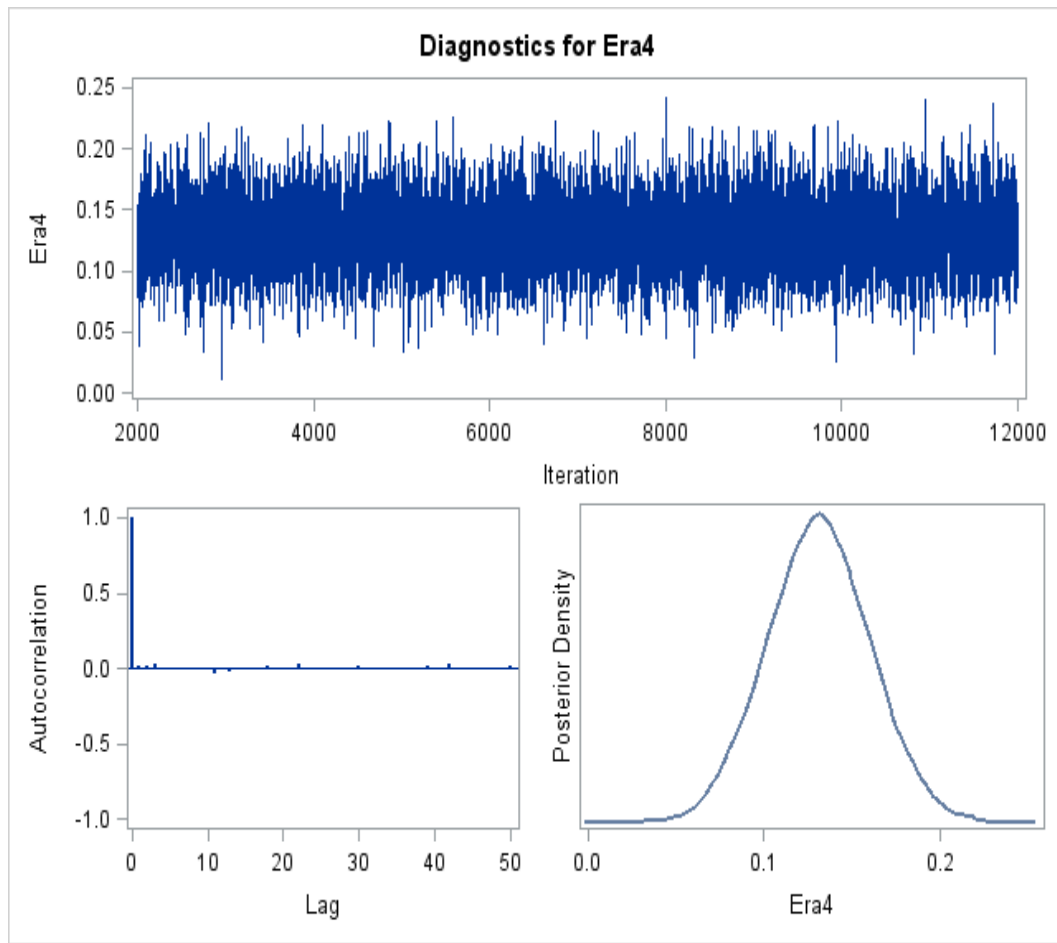


Figure B44. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1932-1994 era, when testing the difference between the 1860-1894 and 1896-1930 eras.

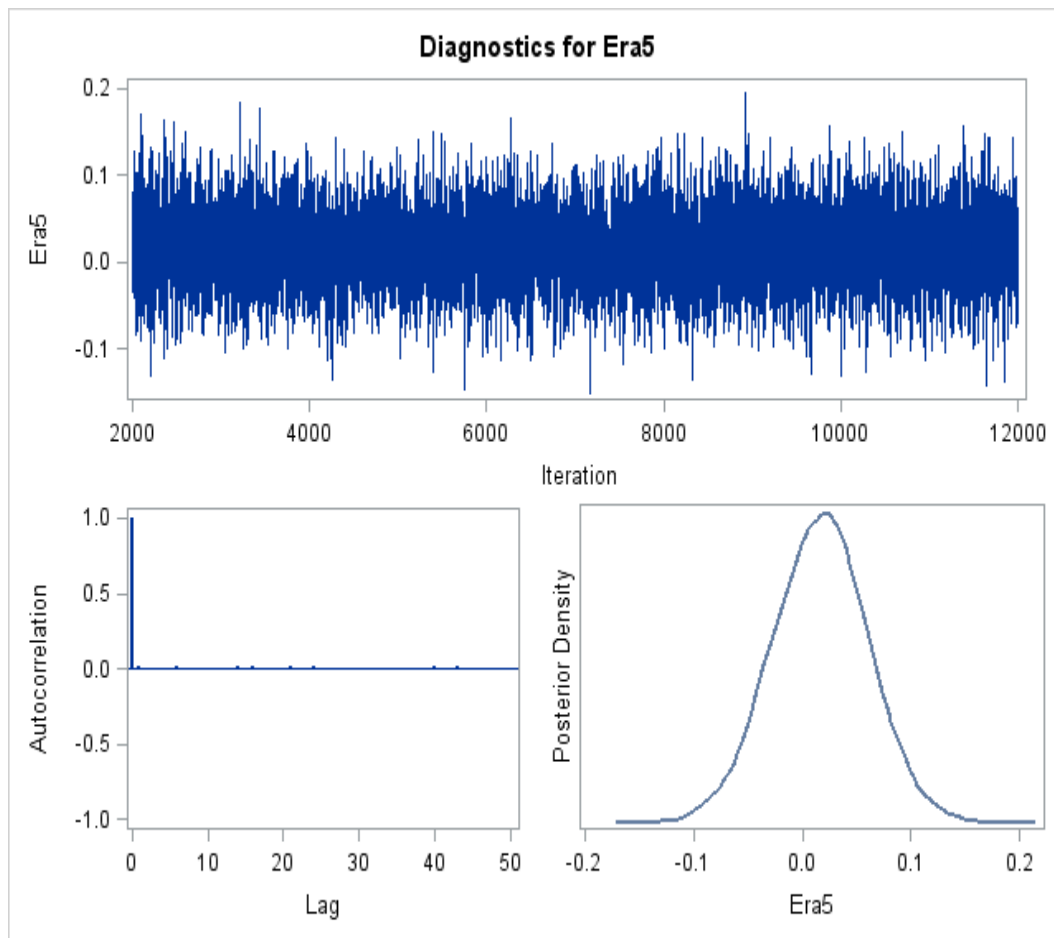


Figure B45. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1996-2008 era, when testing the difference between the 1860-1894 and 1896-1930 eras.

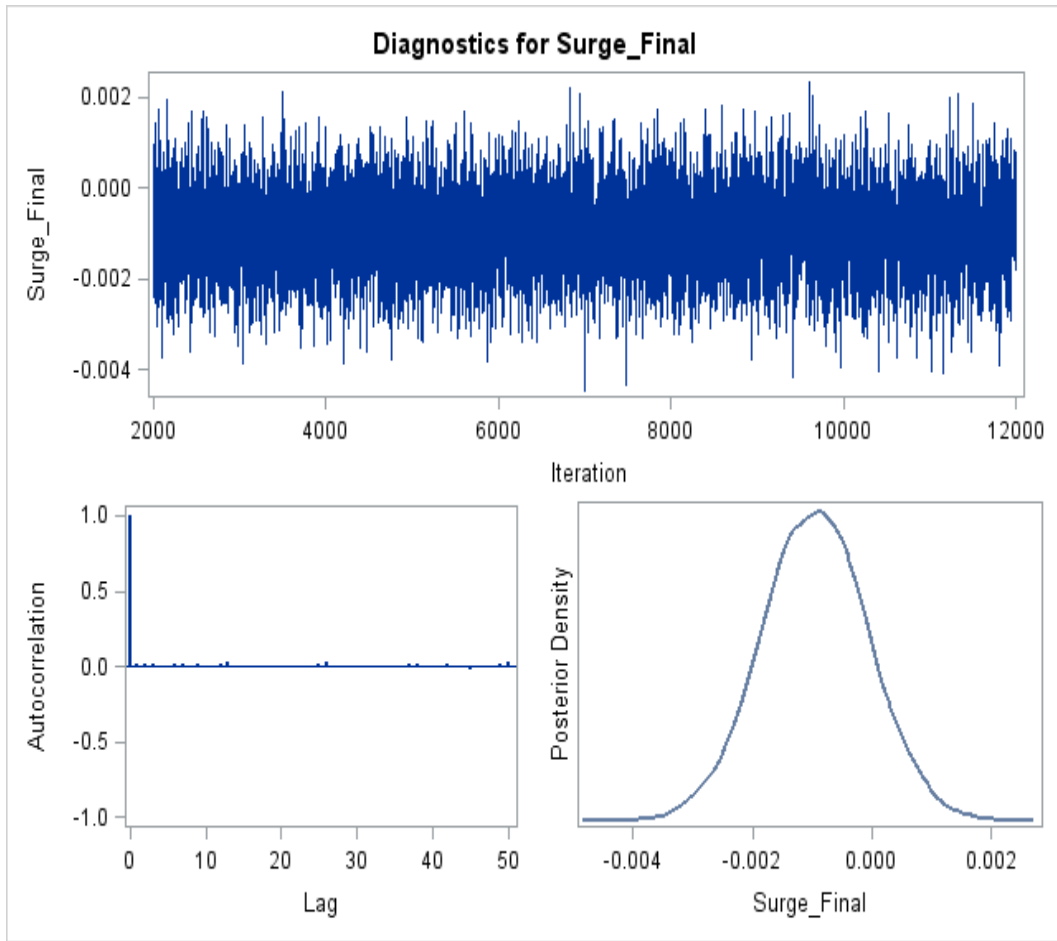


Figure B46. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter controlling for general election surge, when testing the difference between the 1860-1894 and 1896-1930 eras.

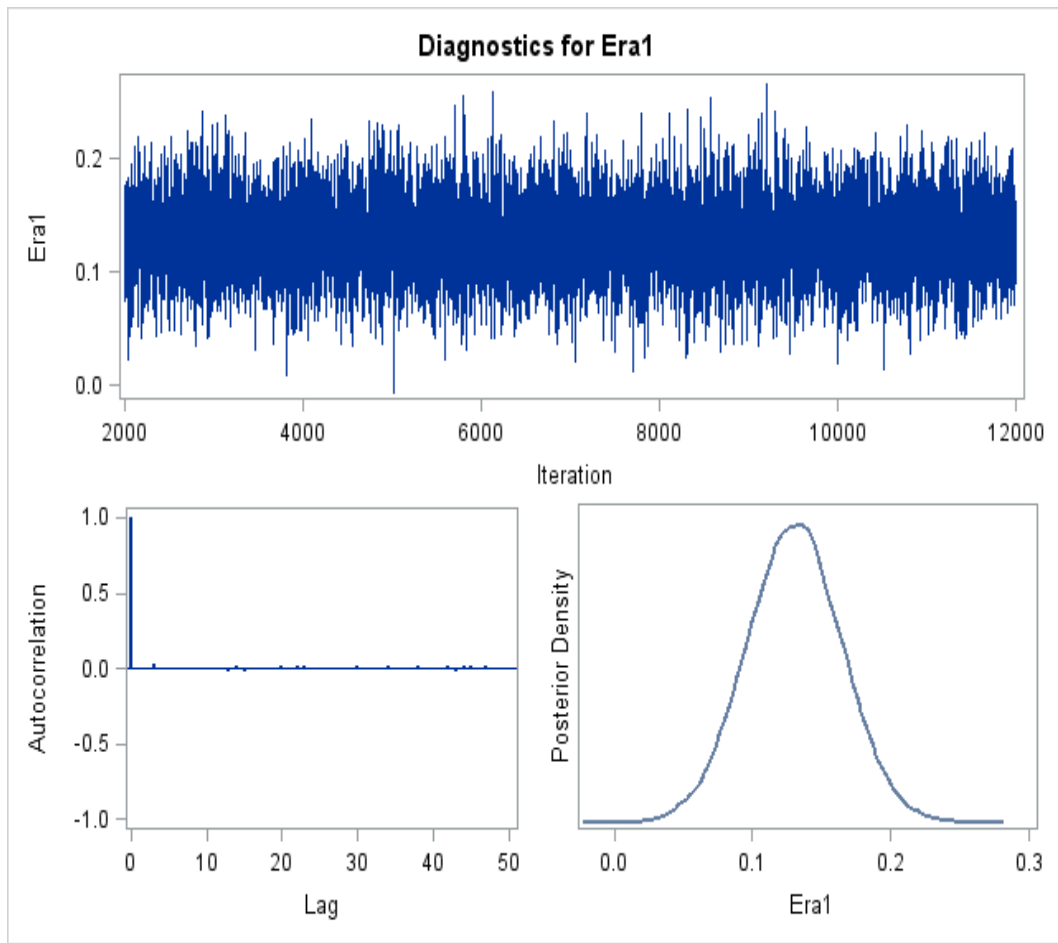


Figure B47. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1828-1858 era, when testing the difference between the 1896-1930 and 1932-1994 eras.

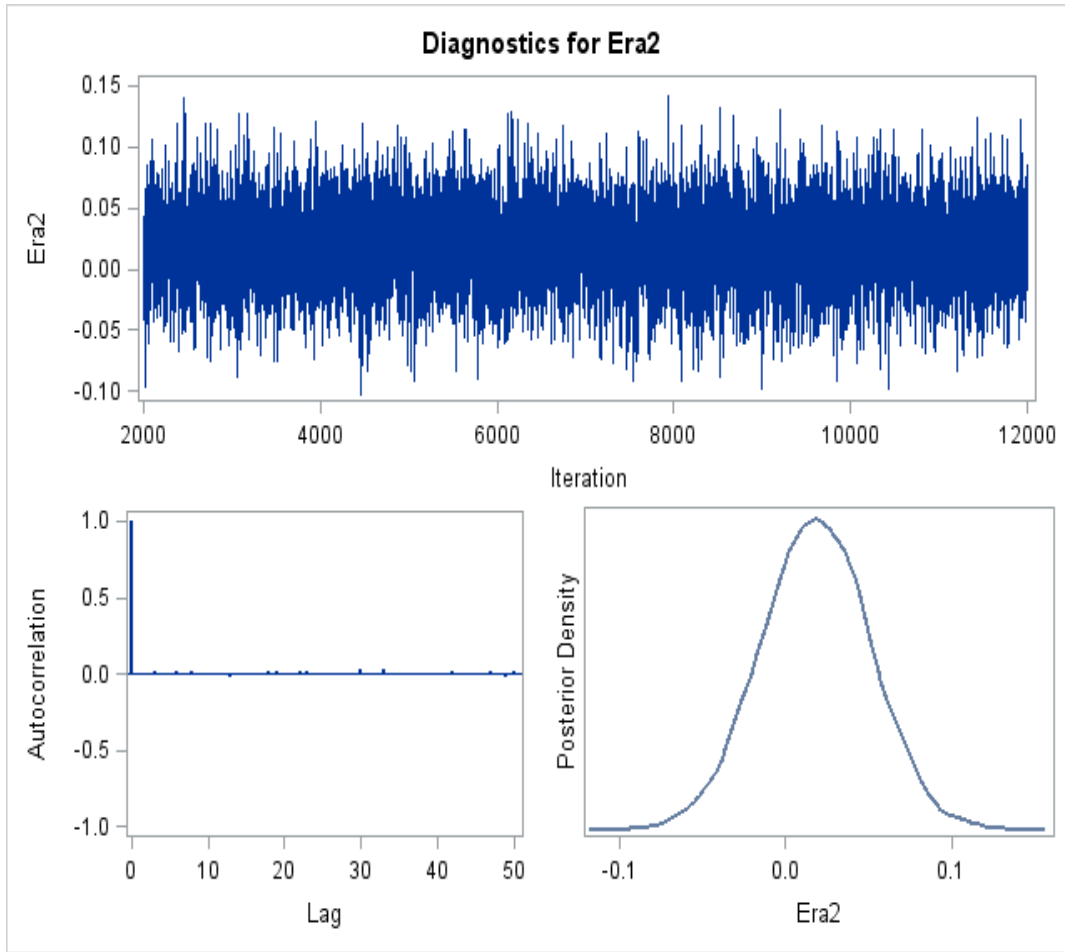


Figure B48. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1860-1894 era, when testing the difference between the 1896-1930 and 1932-1994 eras.

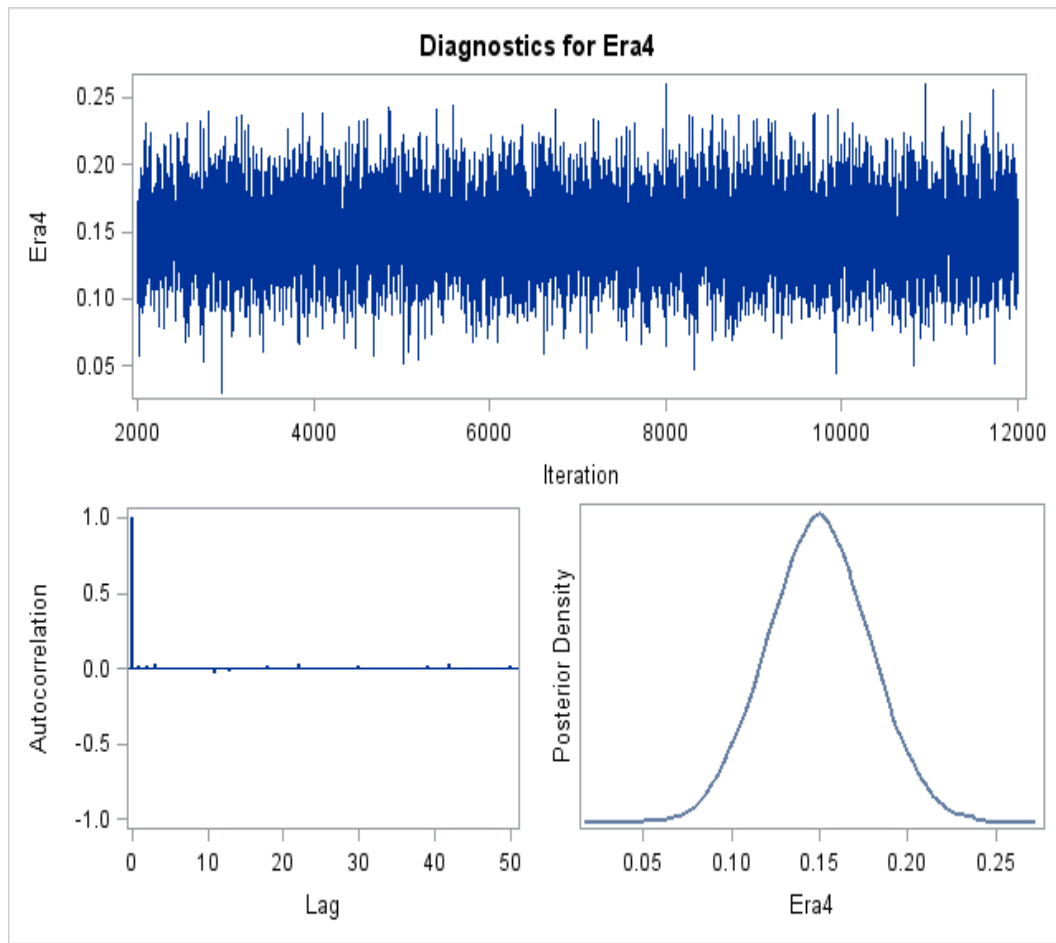


Figure B49. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1932-1994 era, when testing the difference between the 1896-1930 and 1932-1994 eras.

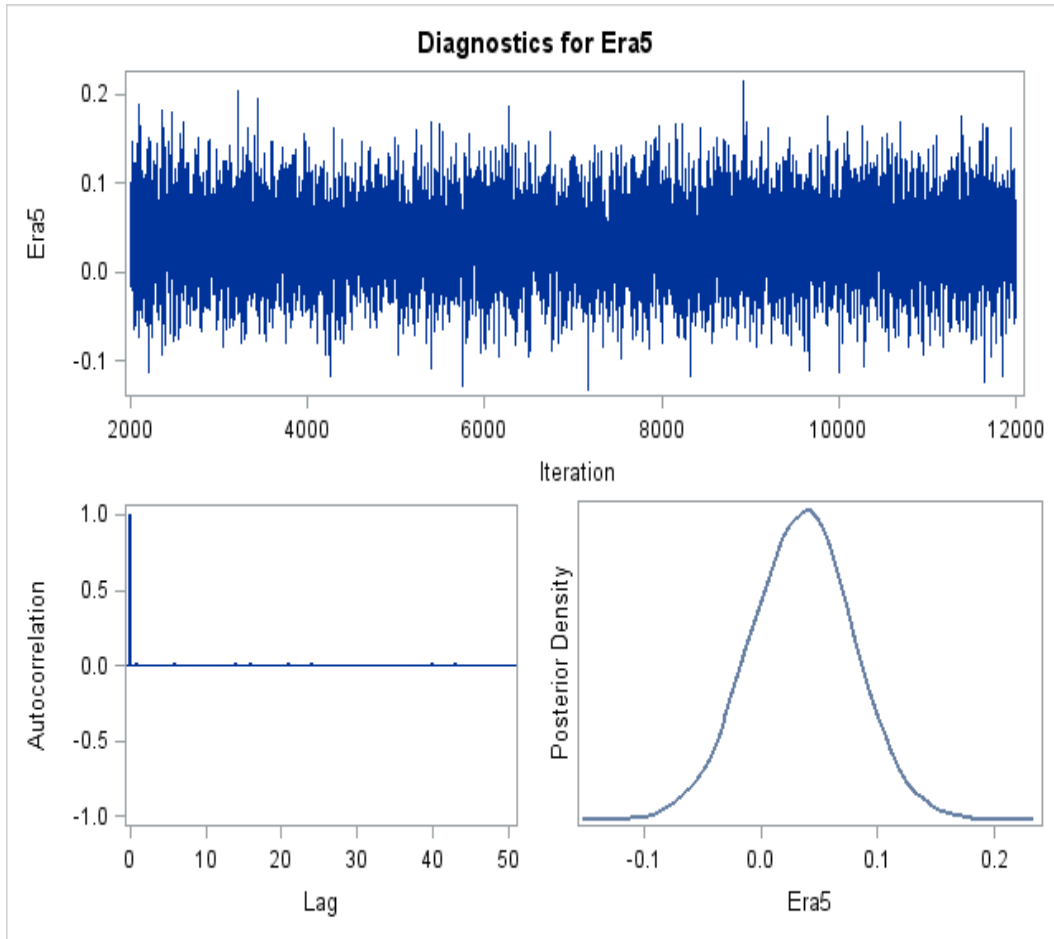


Figure B50. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1996-2008 era, when testing the difference between the 1896-1930 and 1932-1994 eras.

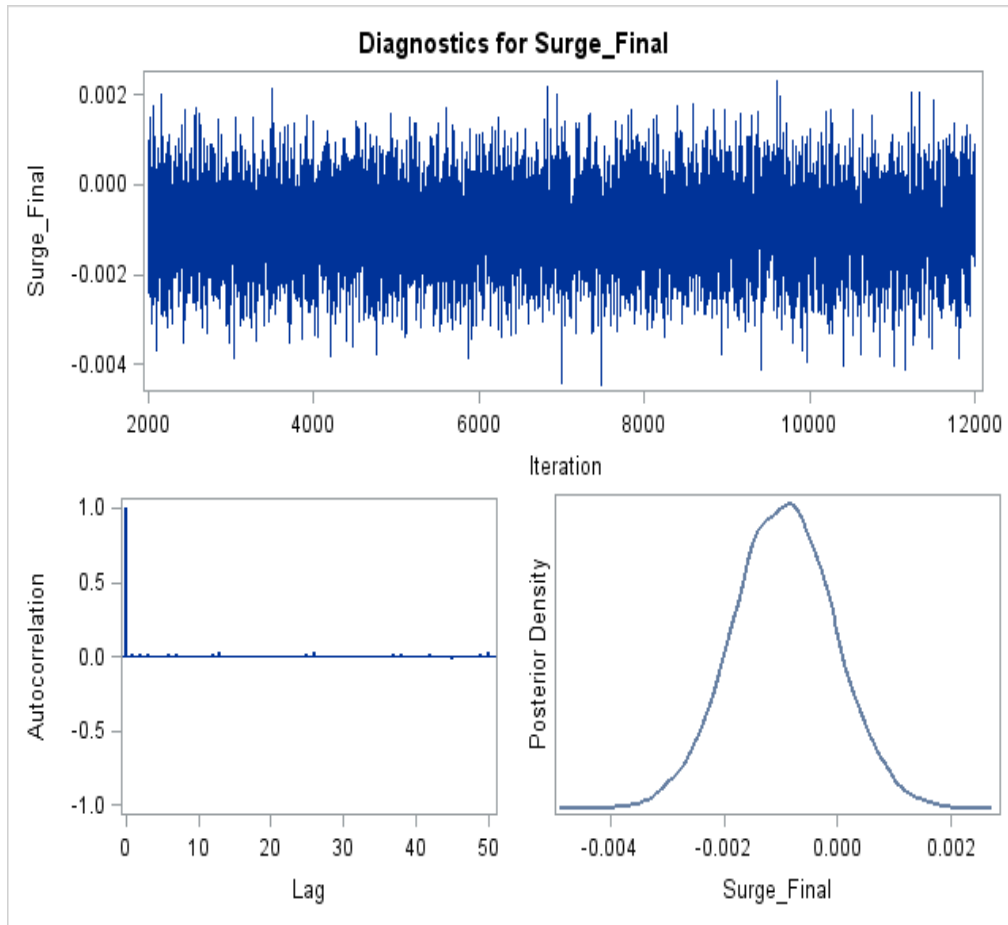


Figure B51. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter controlling for general election surge, when testing the difference between the 1896-1930 and 1932-1994 eras.

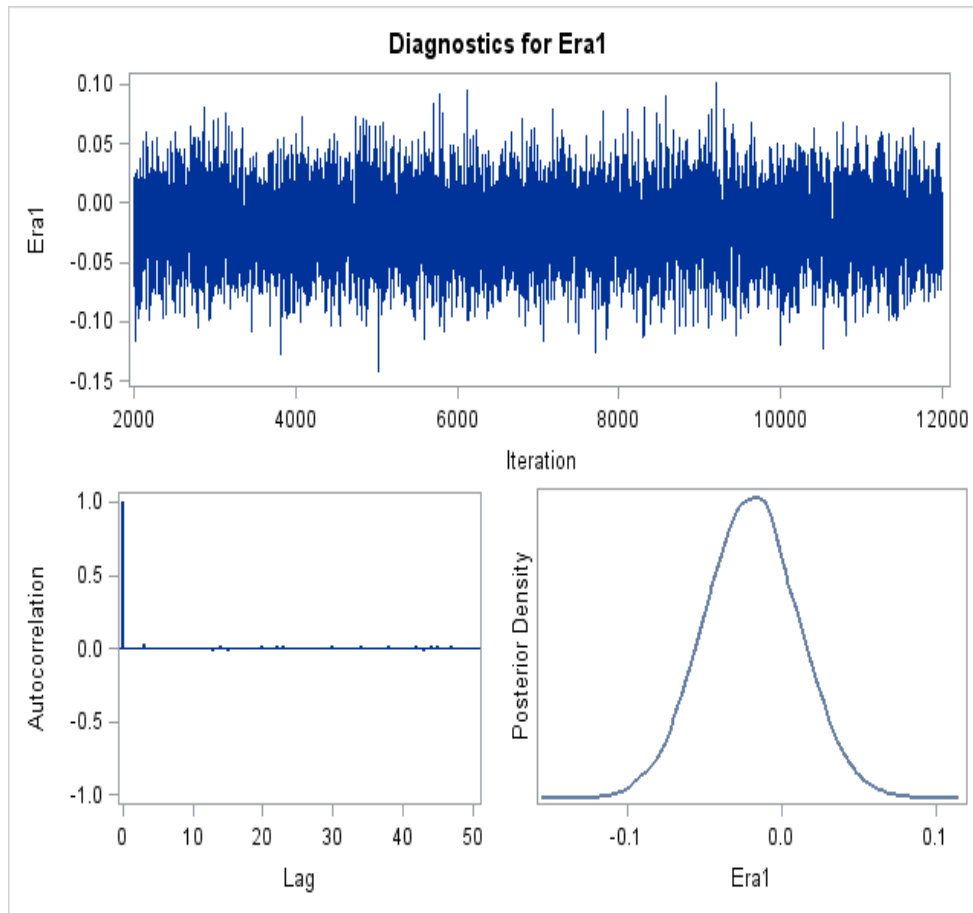


Figure B52. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1828-1858 era, when testing the difference between the 1932-1994 and 1996-2008 eras.

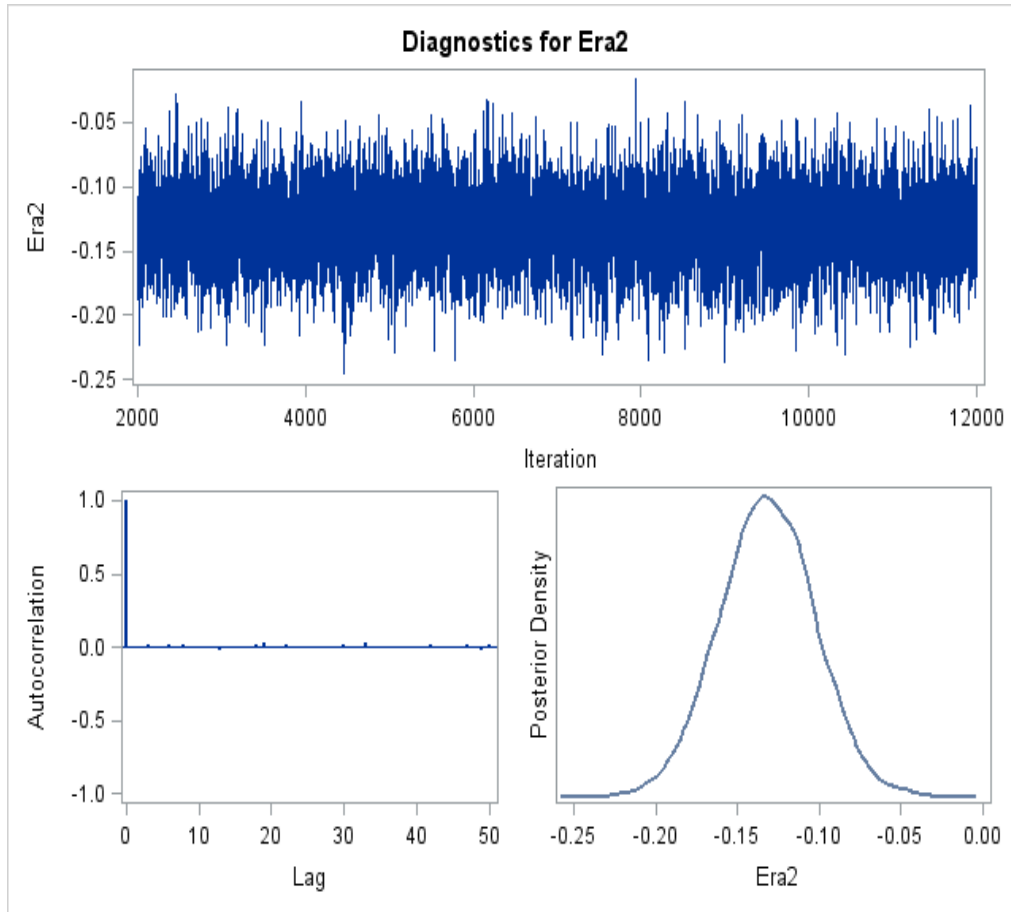


Figure B53. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1860-1894 era, when testing the difference between the 1932-1994 and 1996-2008 eras.

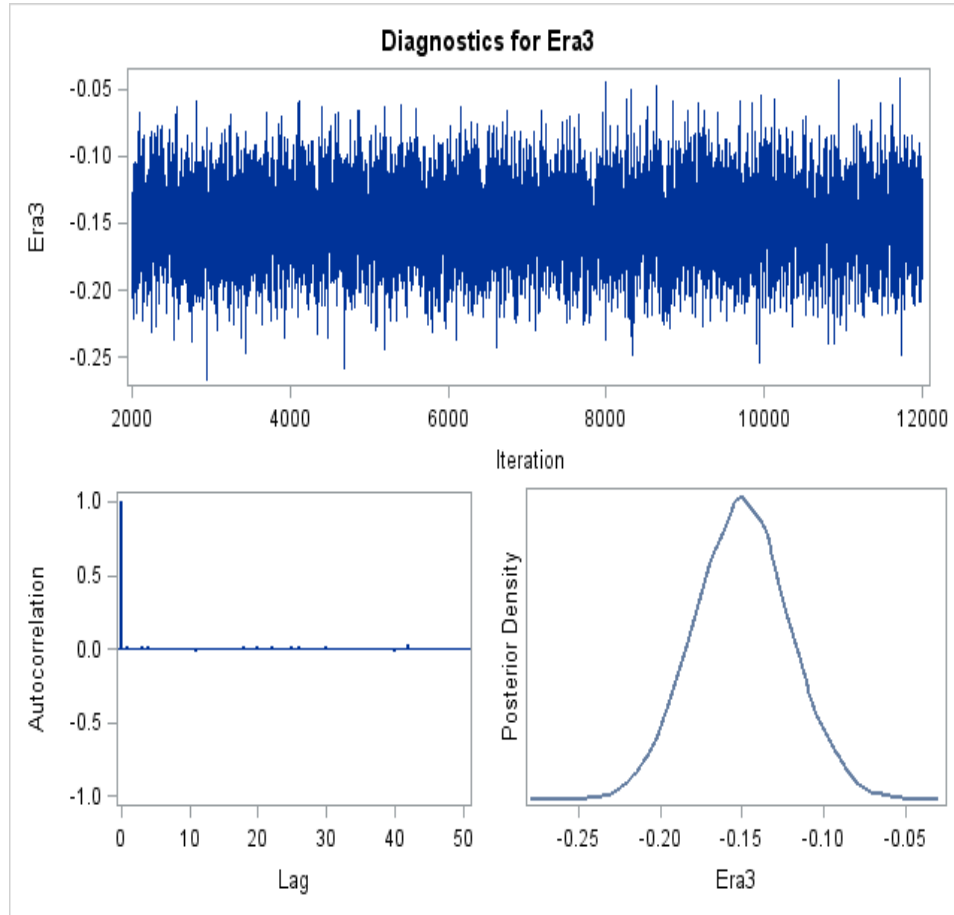


Figure B54. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1896-1930 era, when testing the difference between the 1932-1994 and 1996-2008 eras.

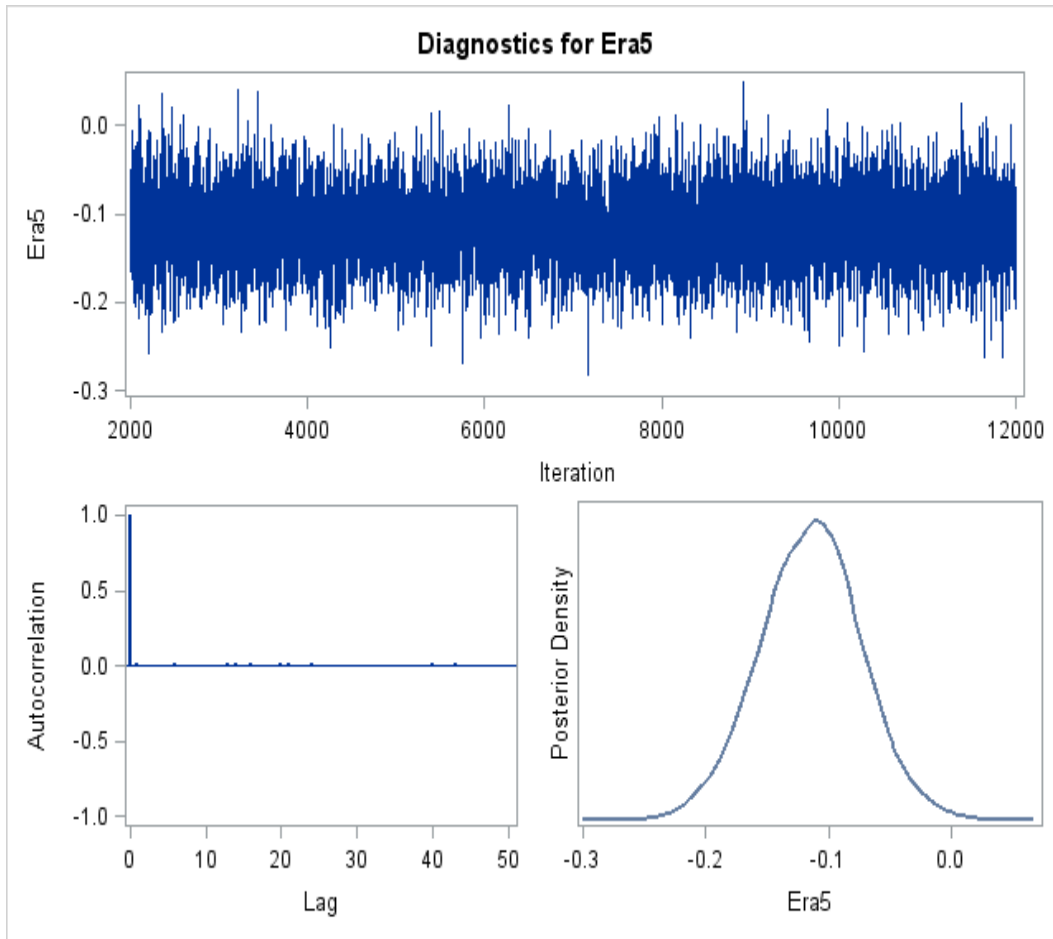


Figure B55. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter of the 1996-2008 era, when testing the difference between the 1932-1994 and 1996-2008 eras.

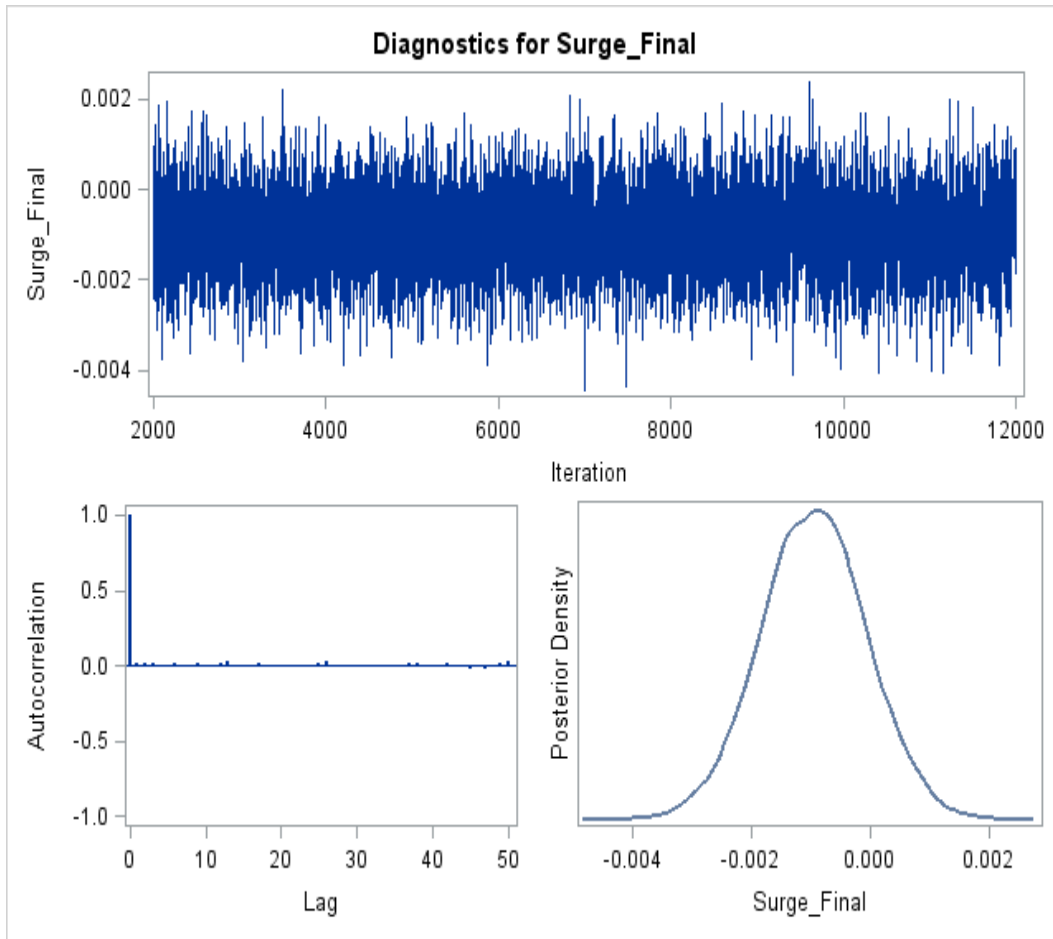


Figure B56. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the non-informative congressional model parameter controlling for general election surge, when testing the difference between the 1932-1994 and 1996-2008 eras.

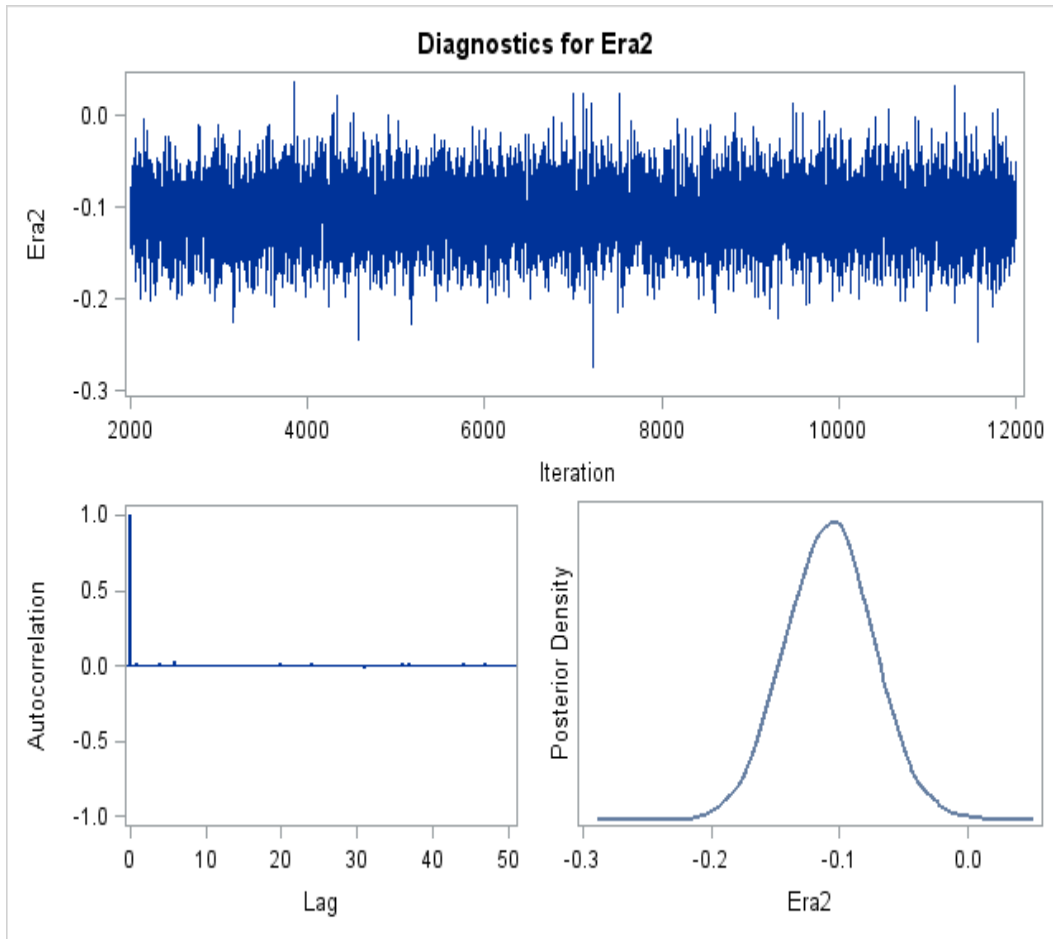


Figure B57. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1860-1894 era, when testing the difference between the 1828-1858 and 1860-1894 eras.

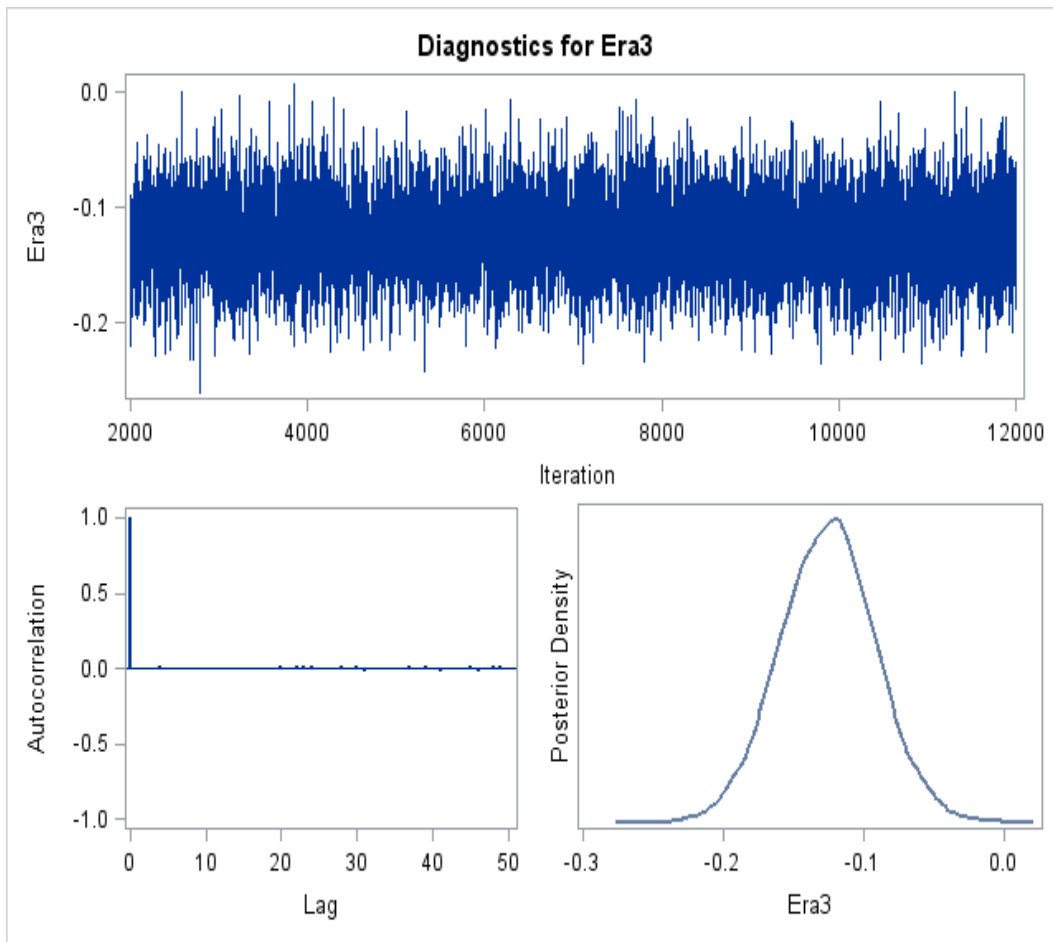


Figure B58. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1896-1930 era, when testing the difference between the 1828-1858 and 1860-1894 eras.

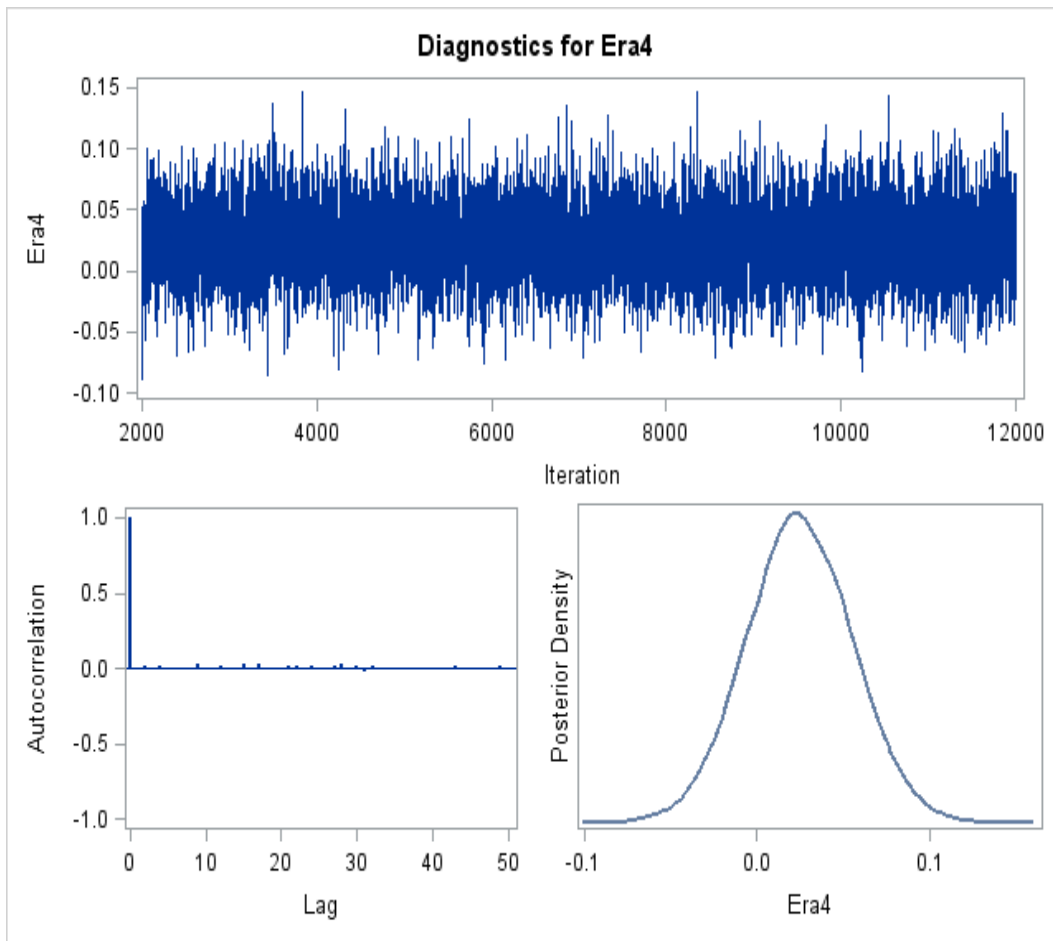


Figure B59. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1932-1994 era, when testing the difference between the 1828-1858 and 1860-1894 eras.

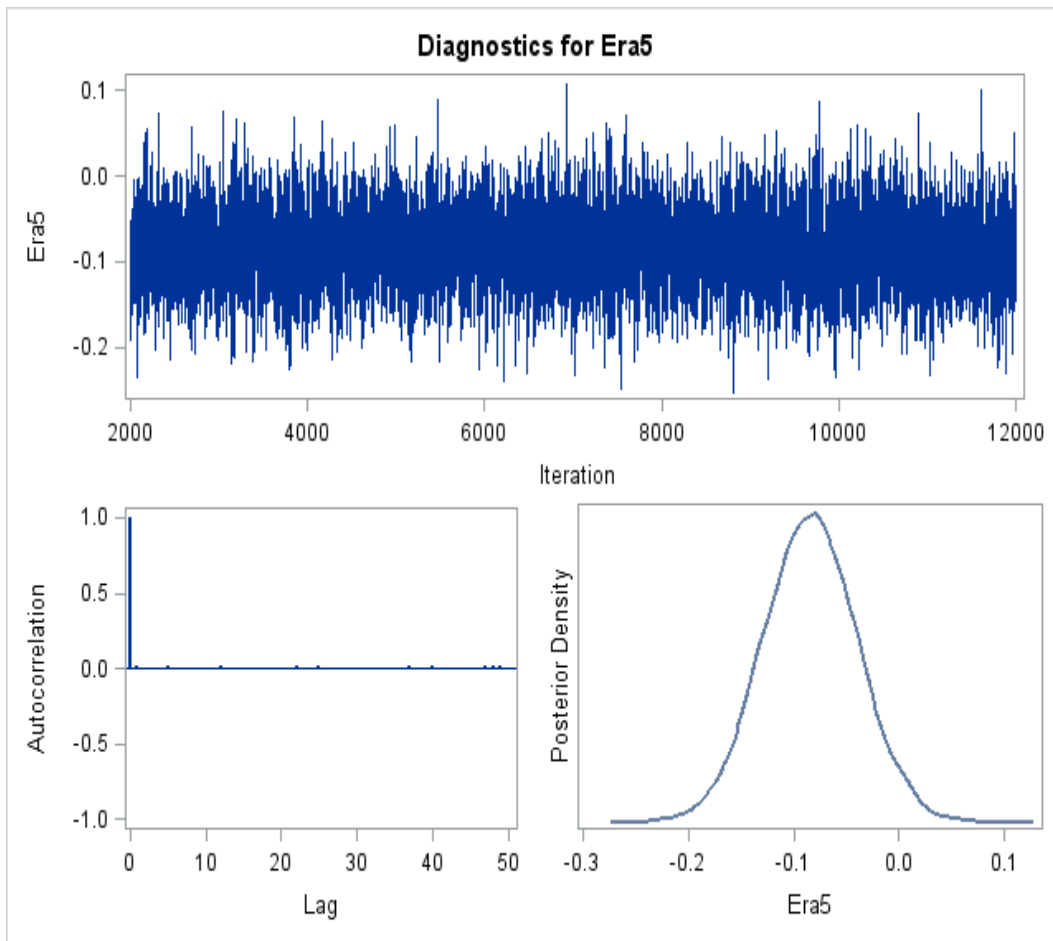


Figure B60. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1996-2008 era, when testing the difference between the 1828-1858 and 1860-1894 eras.

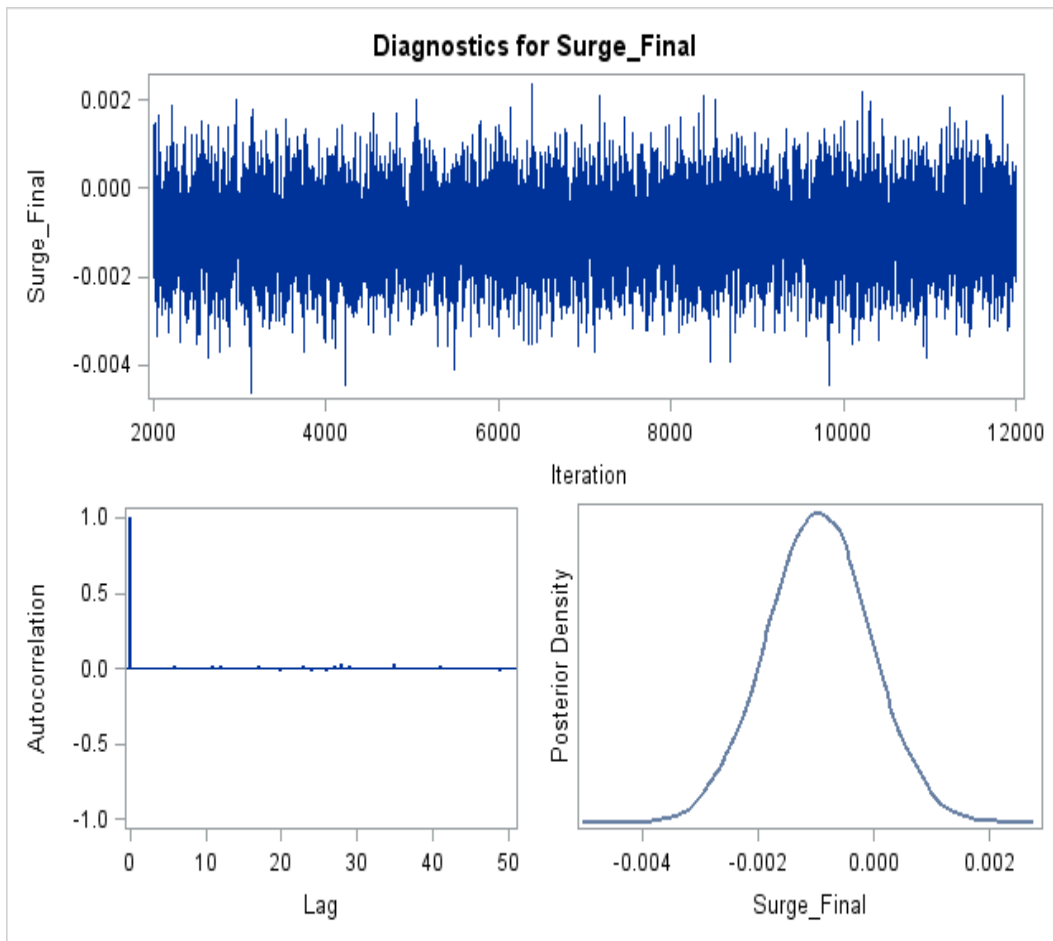


Figure B61. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter controlling for general election surge, when testing the difference between the 1828-1858 and 1860-1894 eras.

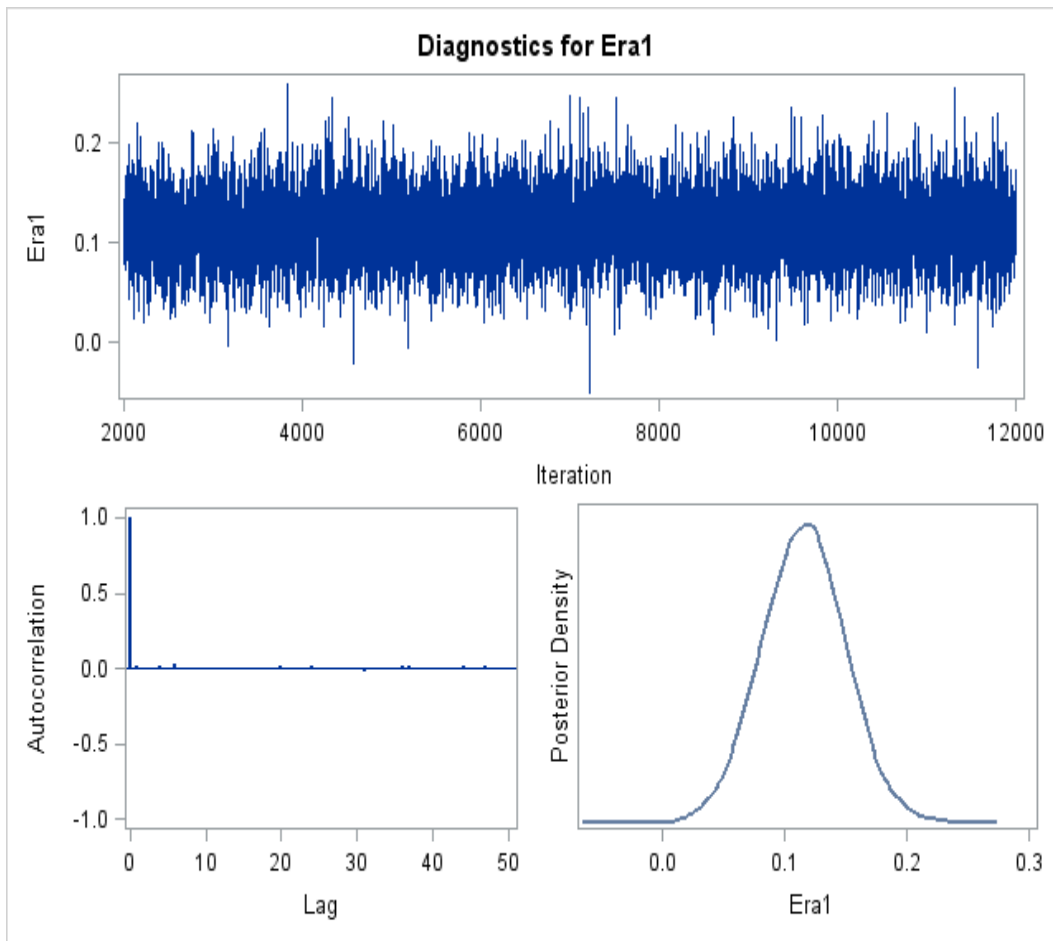


Figure B62. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1828-1858 era, when testing the difference between the 1860-1894 and 1896-1930 eras.

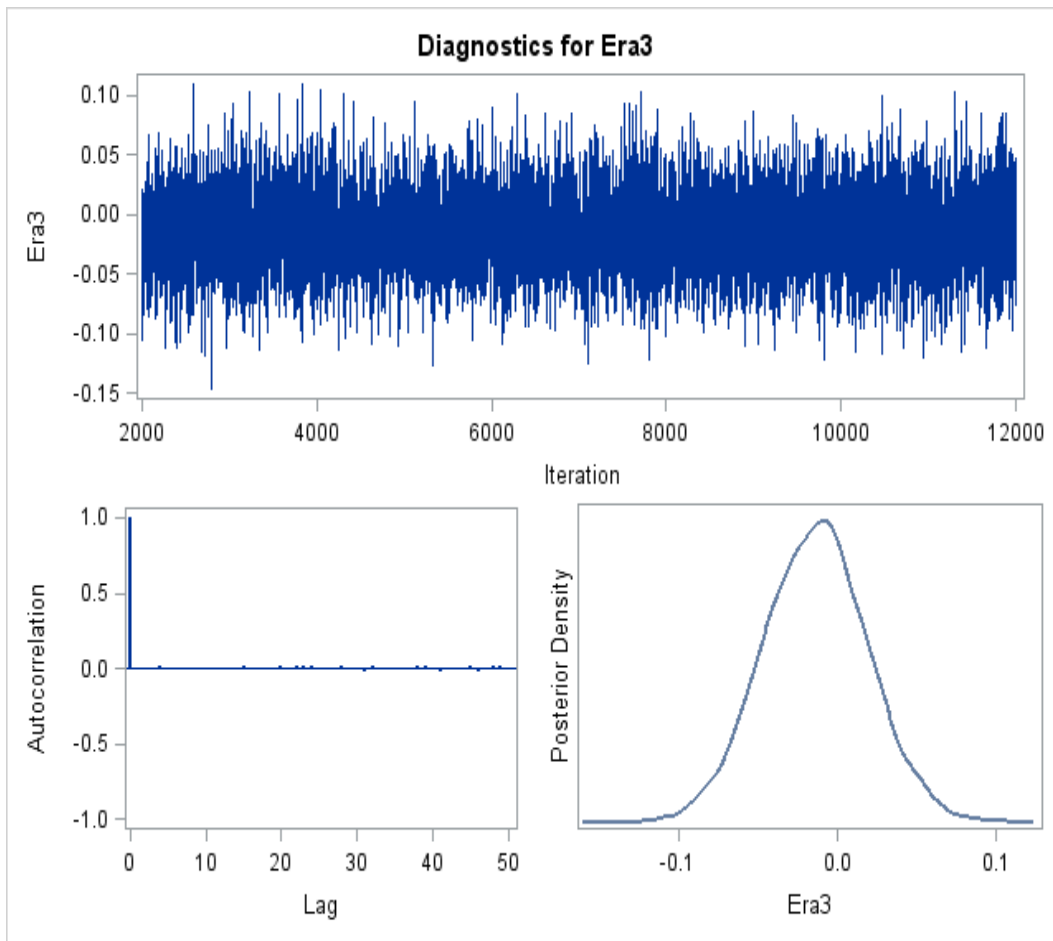


Figure B63. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1896-1930 era, when testing the difference between the 1860-1894 and 1896-1930 eras.

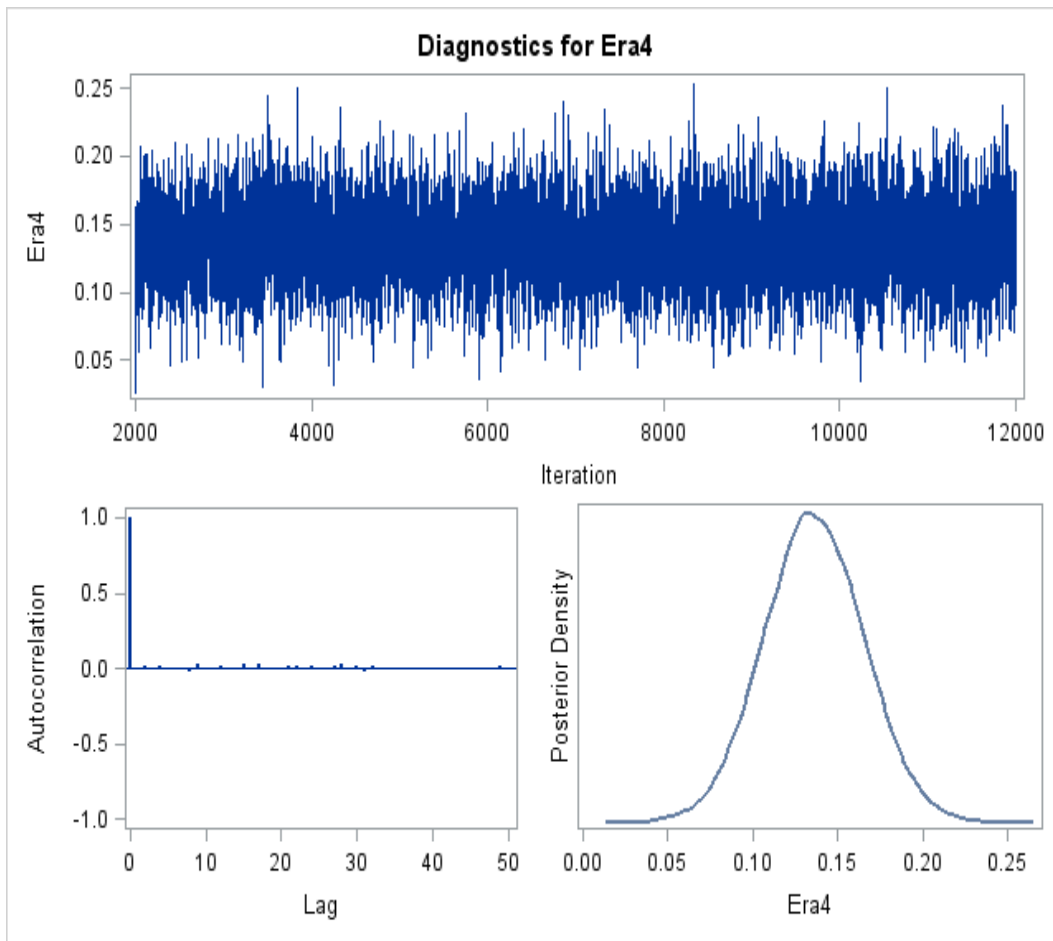


Figure B64. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1932-1994 era, when testing the difference between the 1860-1894 and 1896-1930 eras.

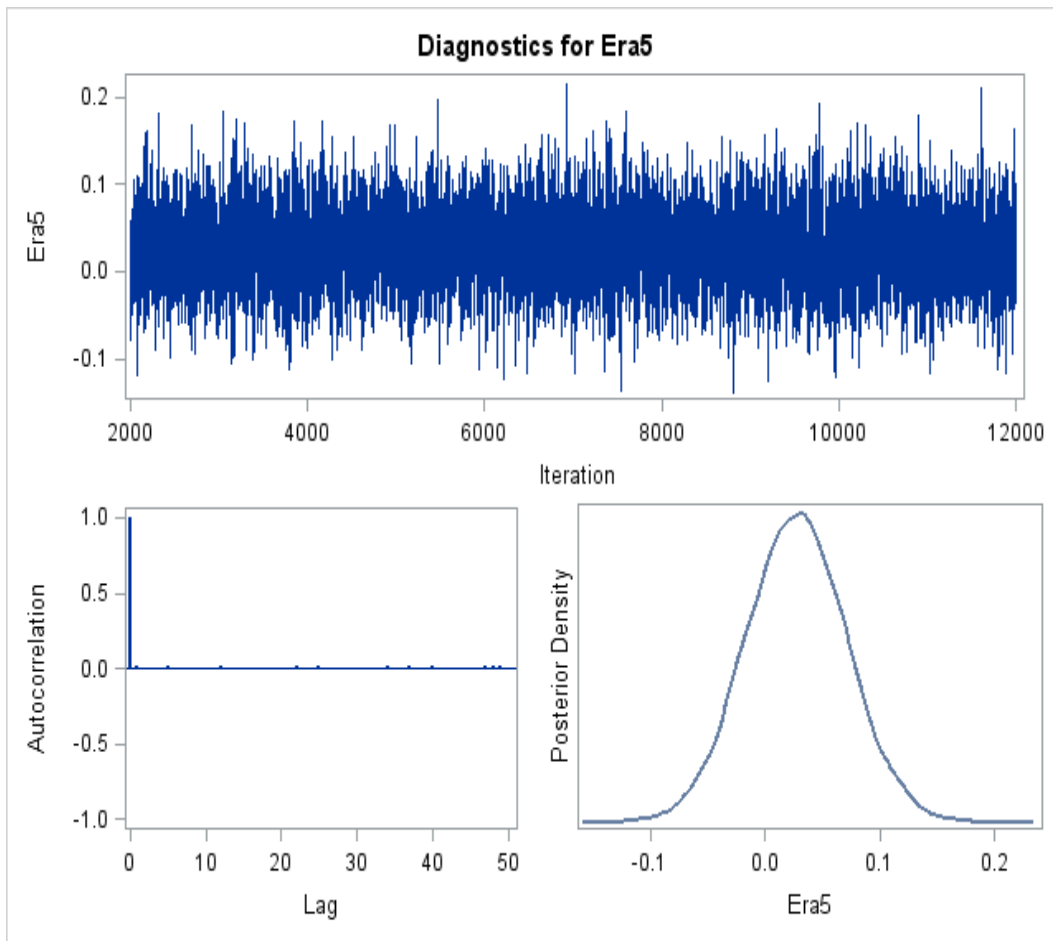


Figure B65. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1996-2008 era, when testing the difference between the 1860-1894 and 1896-1930 eras.

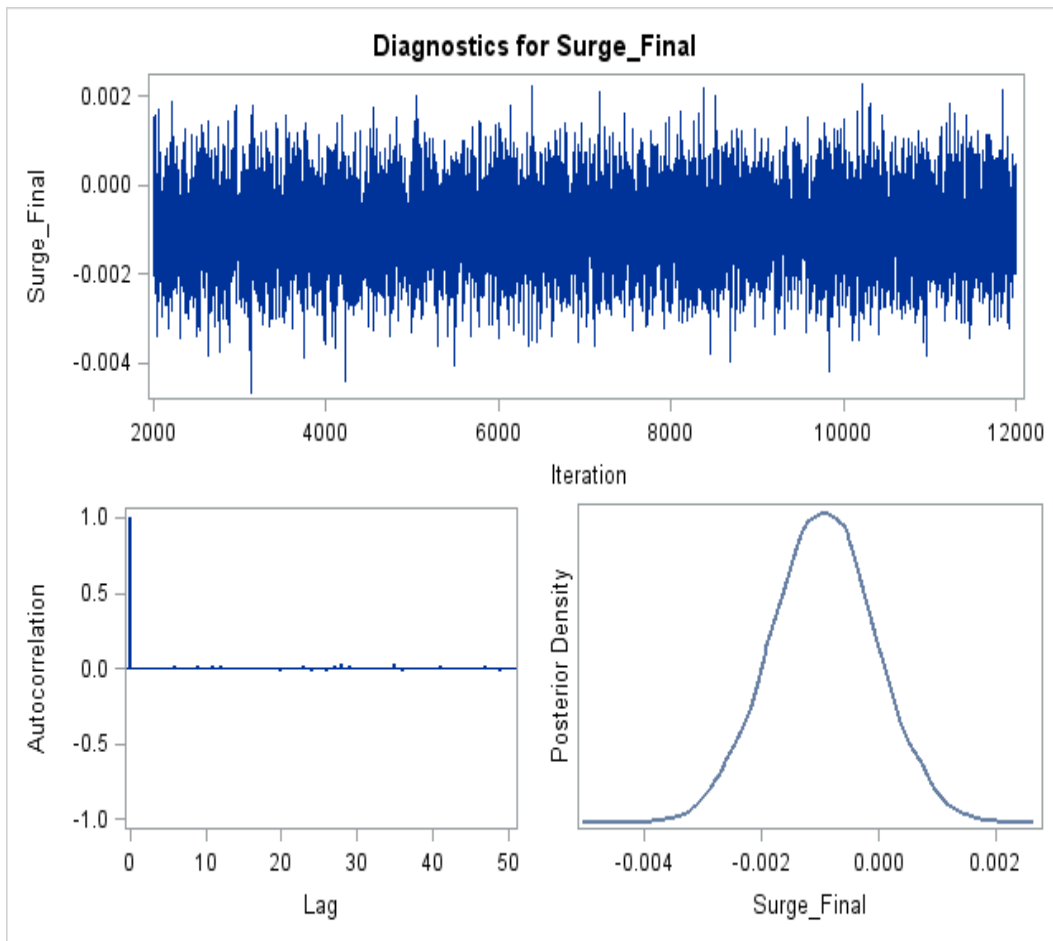


Figure B66. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter controlling for general election surge, when testing the difference between the 1860-1894 and 1896-1930 eras.

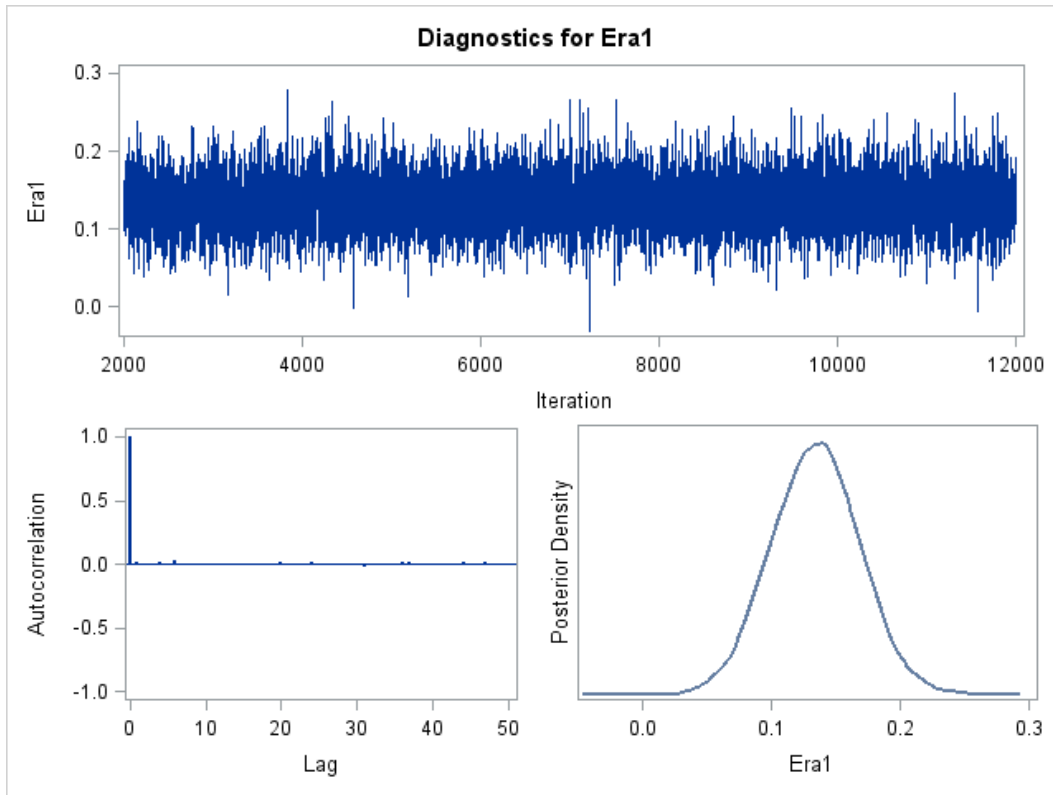


Figure B67. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1828-1858 era, when testing the difference between the 1896-1930 and 1932-1994 eras.

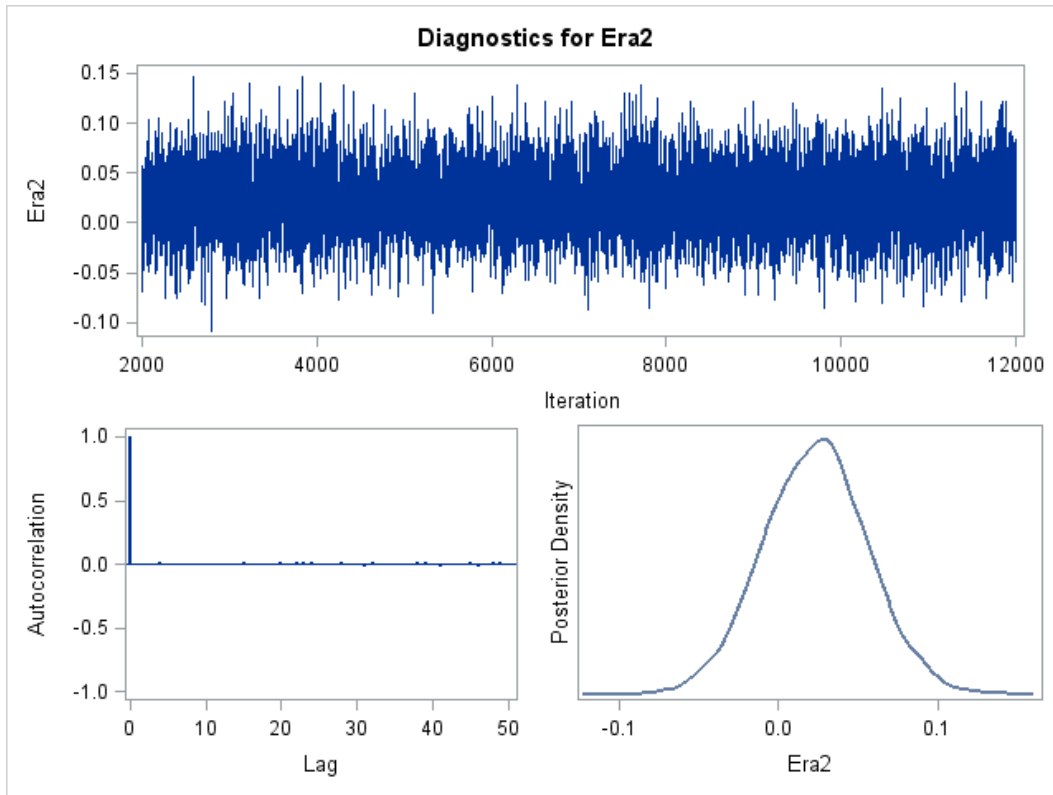


Figure B68. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1860-1894 era, when testing the difference between the 1896-1930 and 1932-1994 eras.

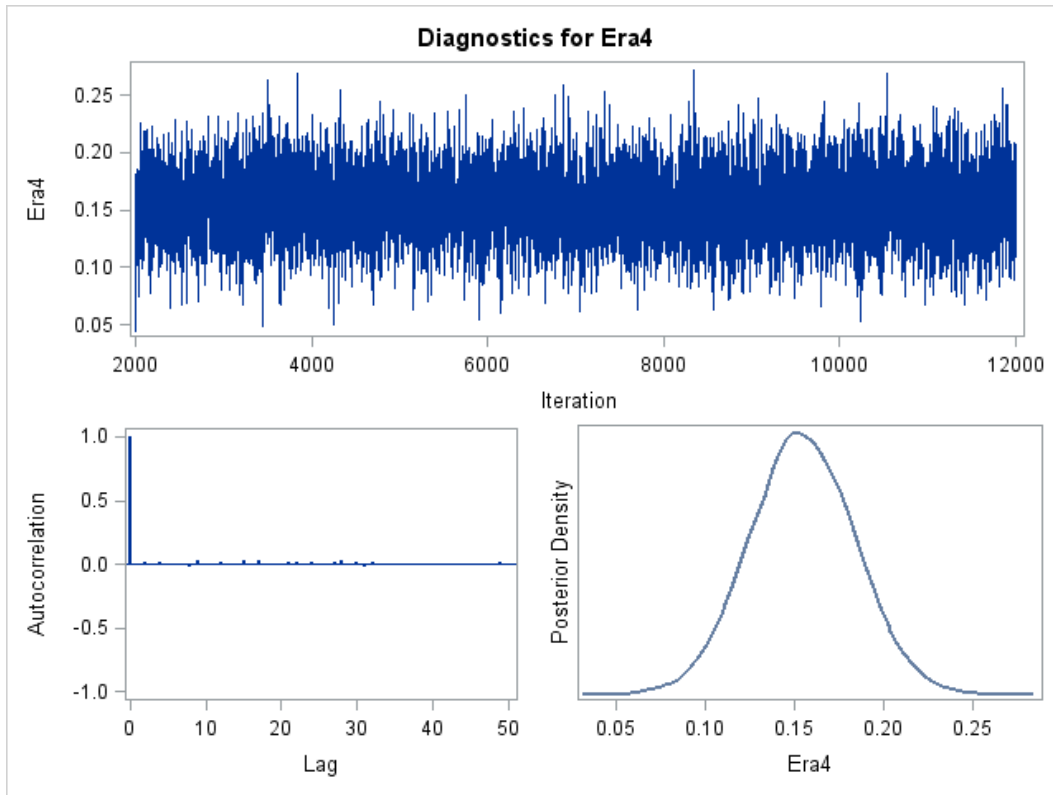


Figure B69. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1932-1994 era, when testing the difference between the 1896-1930 and 1932-1994 eras.

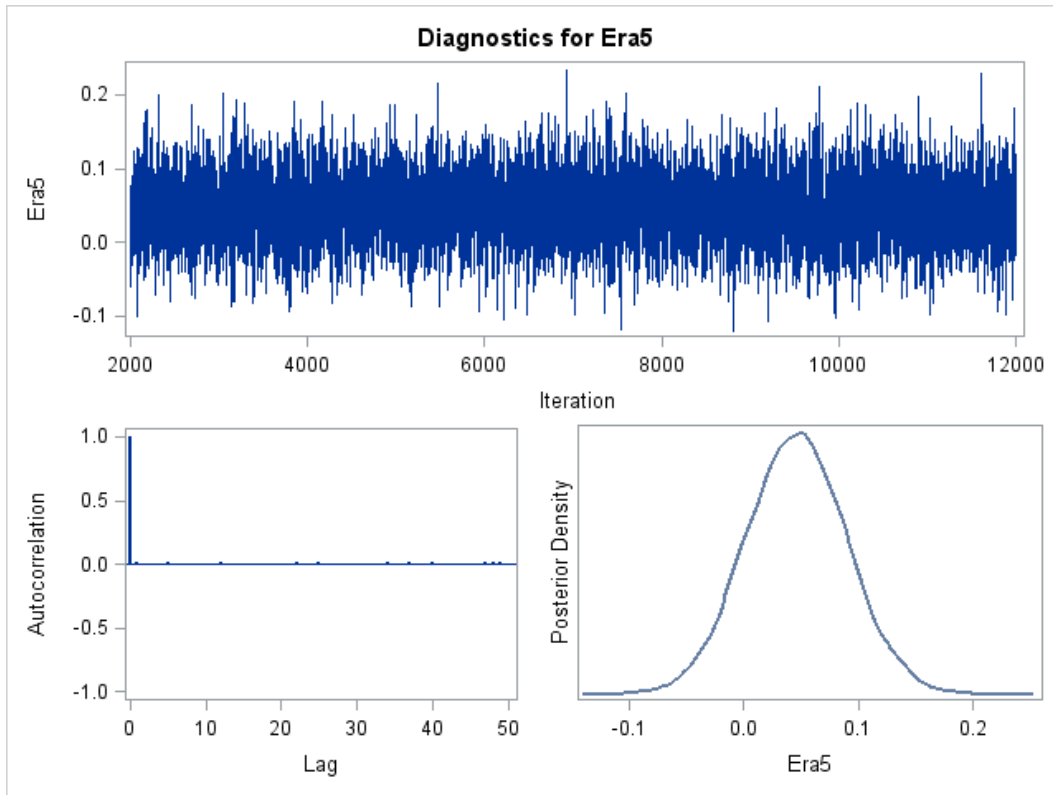


Figure B70. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1996-2008 era, when testing the difference between the 1896-1930 and 1932-1994 eras.

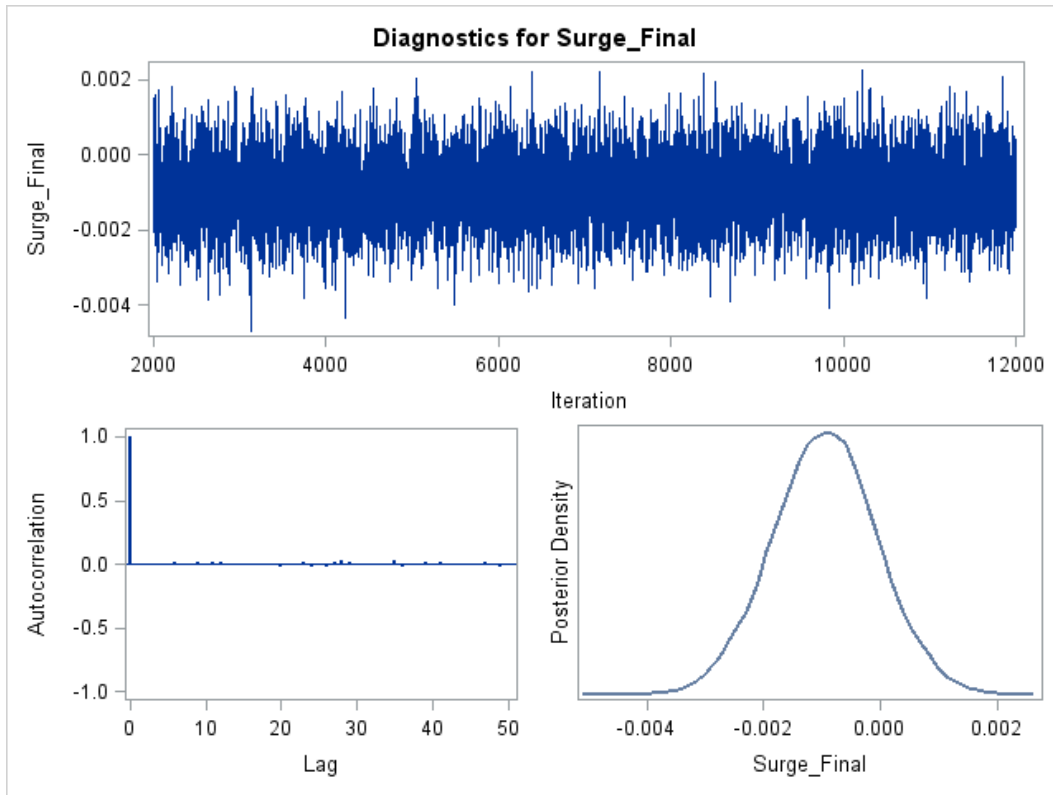


Figure B71. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter controlling for general election surge, when testing the difference between the 1896-1930 and 1932-1994 eras.

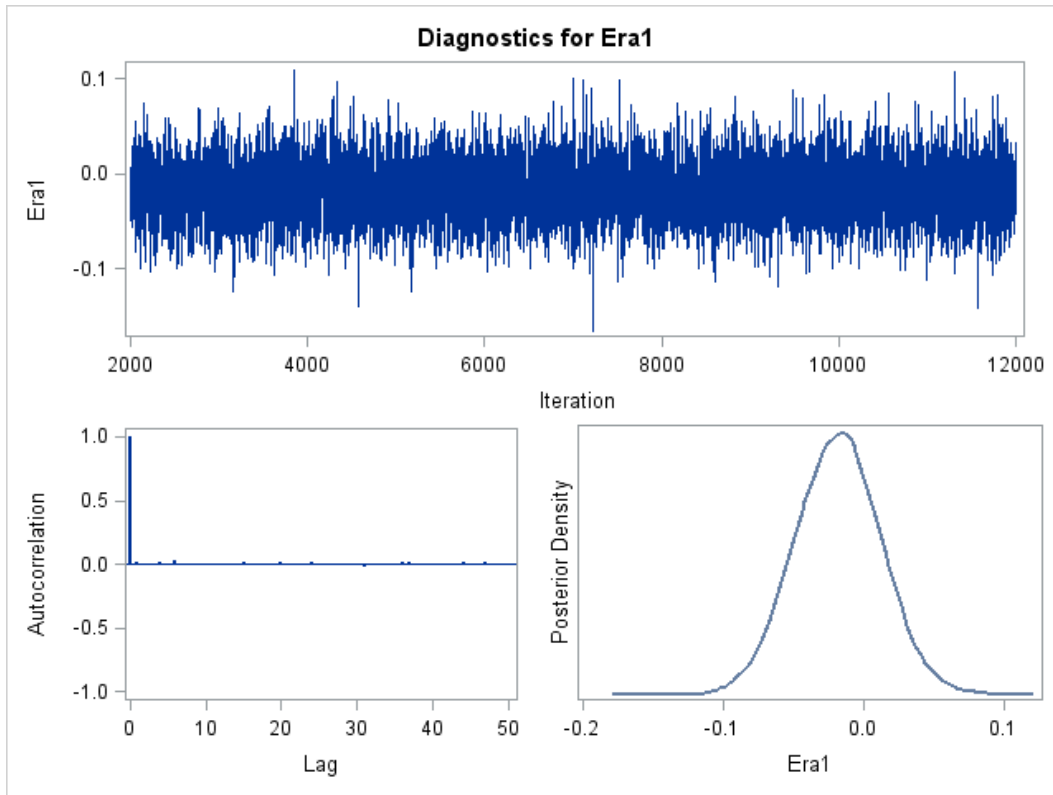


Figure B72. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1828-1858 era, when testing the difference between the 1932-1994 and 1996-2008 eras.

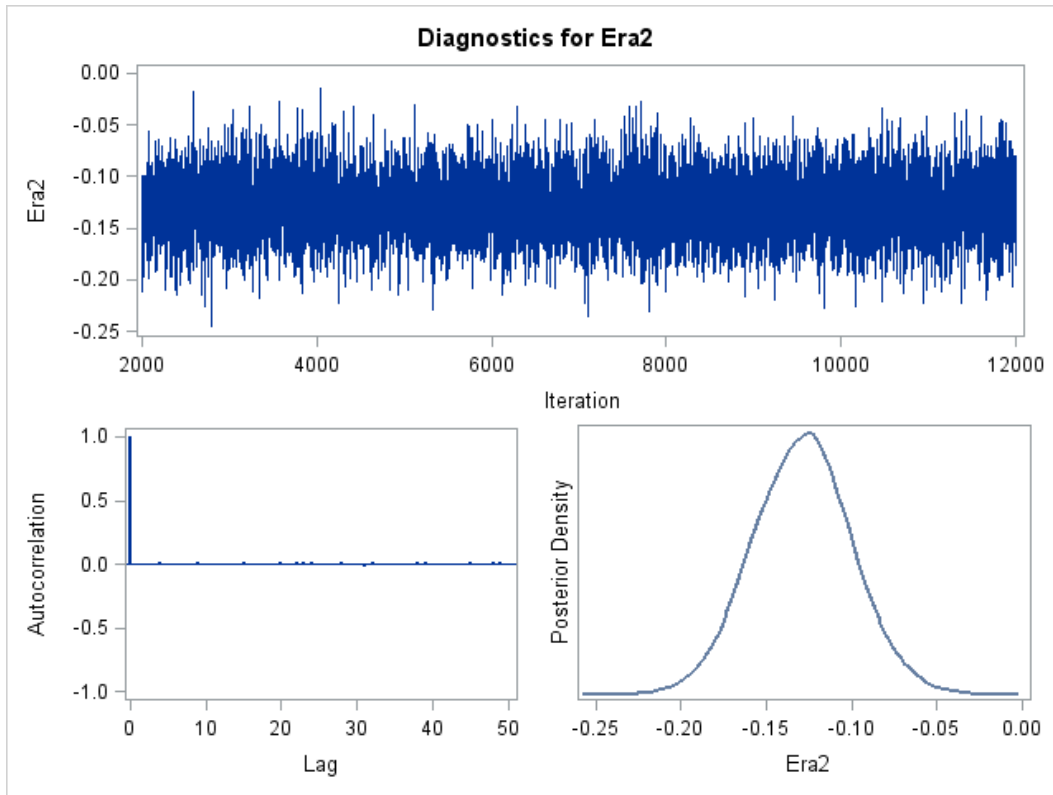


Figure B73. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1860-1894 era, when testing the difference between the 1932-1994 and 1996-2008 eras.

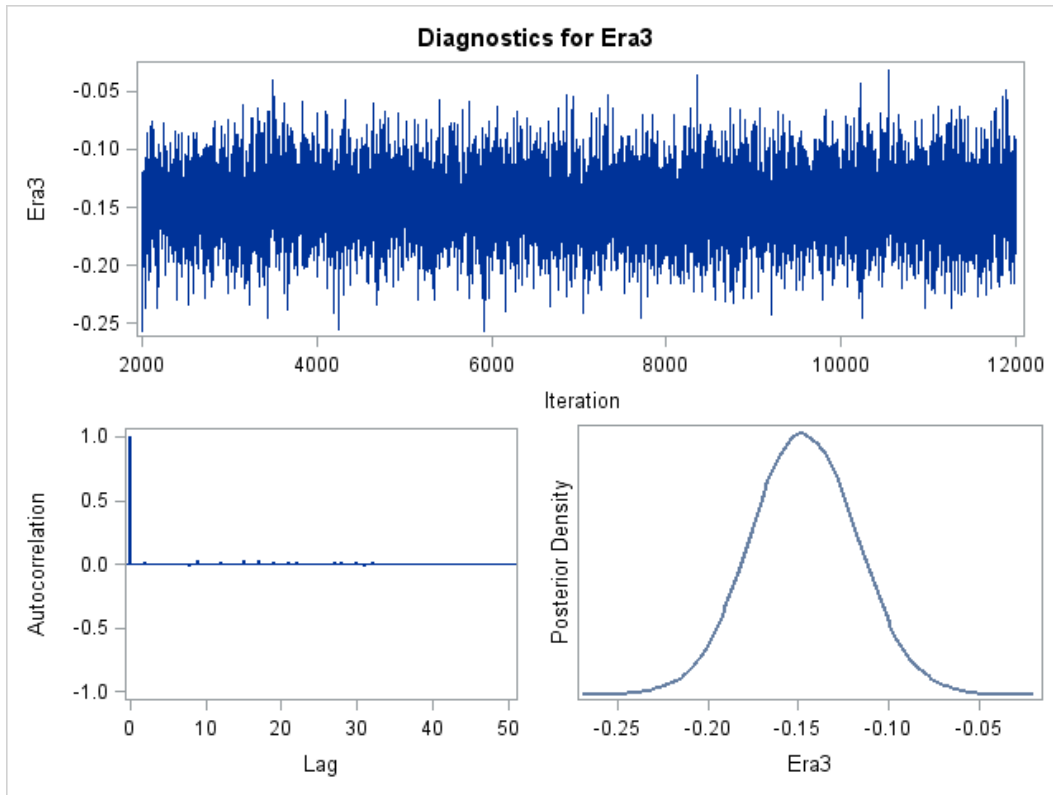


Figure B74. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1896-1930 era, when testing the difference between the 1932-1994 and 1996-2008 eras.

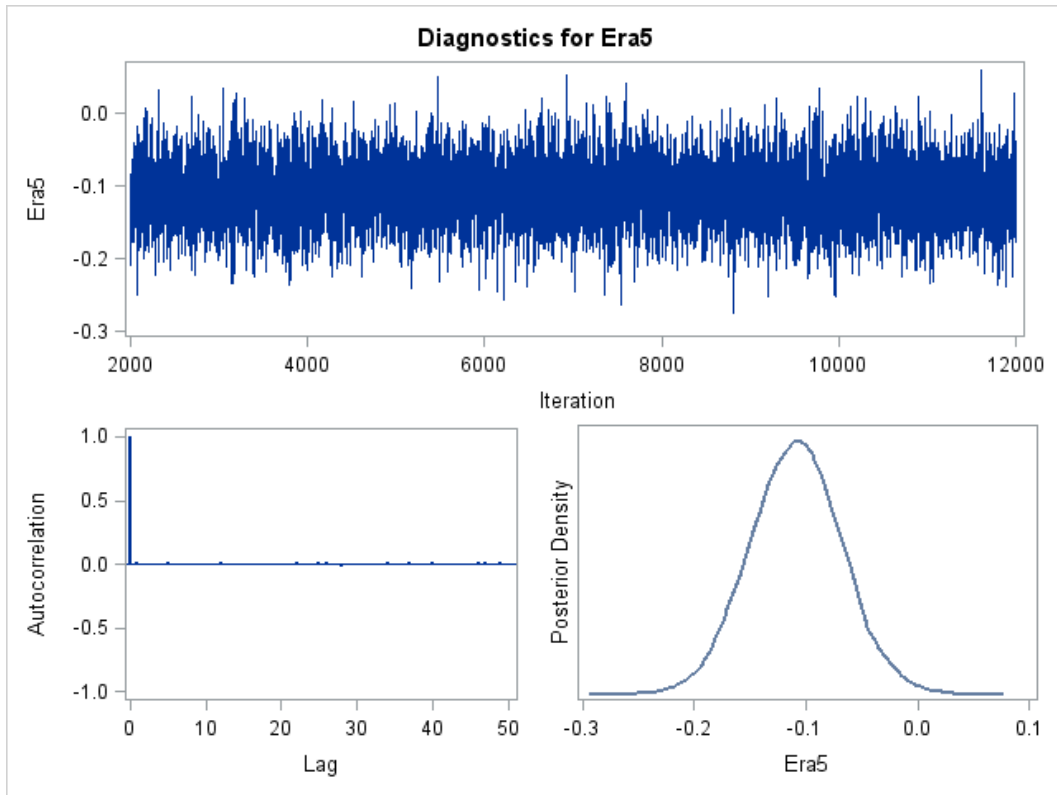


Figure B75. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter of the 1996-2008 era, when testing the difference between the 1932-1994 and 1996-2008 eras.

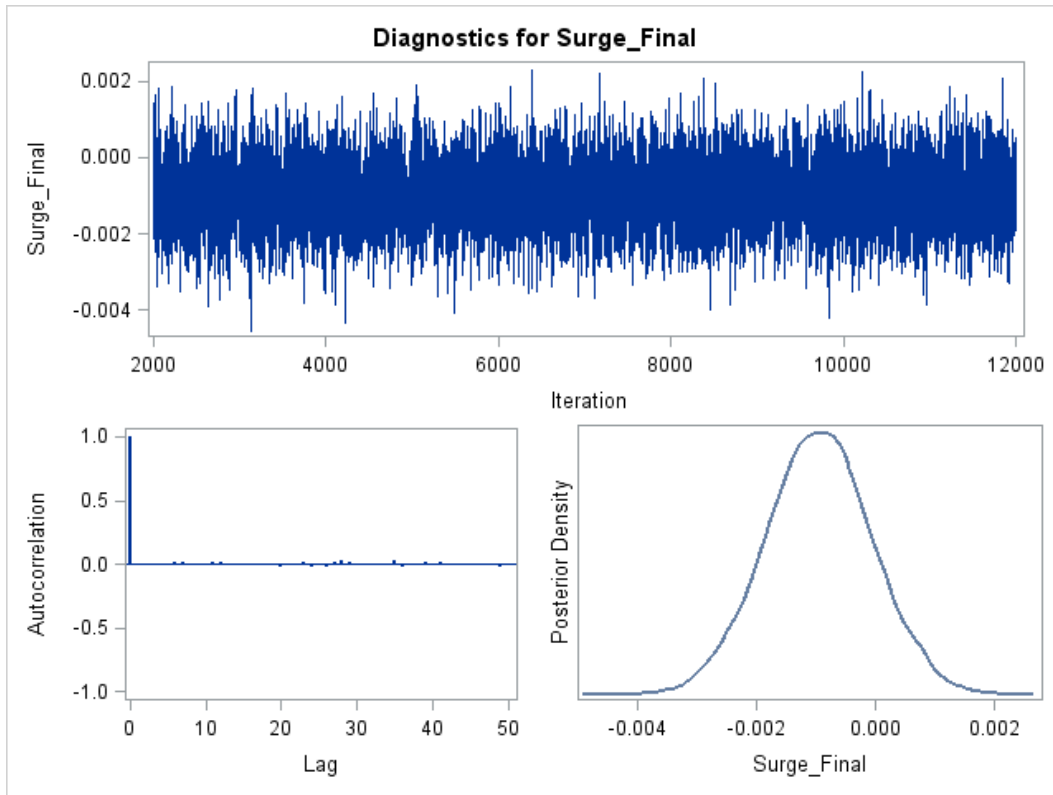


Figure B76. Graphs of autocorrelation, adequate mixing of the chain, and the posterior distribution for the informative congressional model parameter controlling for general election surge, when testing the difference between the 1932-1994 and 1996-2008 eras.

APPENDIX C

Table C1. Results of the sensitivity analysis for the presidential data, comparing 1828-1856 to 1860-1892.

Variable	Prior Information		Posterior Information			
	Mean	Standard Deviation	Mean	Standard Deviation	Equal Tail Interval	Highest Posterior Density
1828-1856	-	-	-	-	-	-
1860-1892	0	0.4000	-0.0563	0.0305	(-0.1158, 0.0037)	(-0.1162, 0.0028)
1896-1928	0	0.0333	-0.0891	0.0306	(-0.1496, -0.0288)	(-0.1517, -0.0317)
1932-1960	0	0.0625	-0.0127	0.0315	(-0.0740, 0.0503)	(-0.0738, 0.0504)
1964-2008	0	0.0833	-0.0436	0.0287	(-0.1004, 0.0130)	(-0.1020, 0.0112)
Constant	0	1.00E-06	0.5375	0.022	(0.4943, 0.5816)	(0.4938, 0.5805)
DIC: -119.385						

Table C2. Results of the sensitivity analysis for the presidential data, comparing 1860-1892 to 1896-1928.

Variable	Prior Information		Posterior Information			
	Mean	Standard Deviation	Mean	Standard Deviation	Equal Tail Interval	Highest Posterior Density
1828-1856	0	0.1111	0.0569	0.0306	(-0.0026, 0.1170)	(-0.0030, 0.1160)
1860-1892	-	-	-	-	-	-
1896-1928	0	0.0333	-0.0325	0.0297	(-0.0911, 0.0262)	(-0.0921, 0.0246)
1932-1960	0	0.0625	0.0439	0.0306	(-0.0158, 0.1059)	(-0.0173, 0.1032)
1964-2008	0	0.0833	0.0130	0.0277	(-0.0423, 0.0675)	(-0.0394, 0.0701)
Constant	0	1.00E-06	0.4809	0.0208	(0.4400, 0.5222)	(0.4376, 0.5192)
DIC: -119.384						

Table C3. Results of the sensitivity analysis for the presidential data, comparing 1896-1928 to 1932-1960.

Variable	Prior Information		Posterior Information			
	Mean	Standard Deviation	Mean	Standard Deviation	Equal Tail Interval	Highest Posterior Density
1828-1856	0	0.1111	0.0891	0.0306	(0.0296, 0.1492)	(0.0293, 0.1483)
1860-1892	0	0.4000	0.0319	0.0297	(-0.0267, 0.0906)	(-0.0277, 0.0890)
1896-1928	-	-	-	-	-	-
1932-1960	0	0.0625	0.0761	0.0306	(0.0164, 0.1381)	(0.0149, 0.1354)

1964-2008	0	0.0833	0.0452	0.0277	(-0.0101, 0.0997)	(-0.0072, 0.1023)
Constant	0	1.00E-06	0.4487	0.0208	(0.4078, 0.4900)	(0.4054, 0.4870)
DIC: -119.384						

Table C4. Results of the sensitivity analysis for the presidential data, comparing 1932-1960 to 1964-2008.

Variable	Prior Information		Posterior Information			
	Mean	Standard Deviation	Mean	Standard Deviation	Equal Tail Interval	Highest Posterior Density
1828-1856	0	0.1111	0.0131	0.0314	(-0.0482, 0.0749)	(-0.0486, 0.0739)
1860-1892	0	0.4000	-0.0442	0.0306	(-0.1046, 0.0162)	(-0.1066, 0.0134)
1896-1928	0	0.0333	-0.0761	0.0306	(-0.1357, - 0.0148)	(-0.1379, -0.0171)
1932-1960	-	-	-	-	-	-
1964-2008	0	0.0833	-0.0309	0.0287	(-0.0880, 0.0256)	(-0.0875, 0.0257)
Constant	0	1.00E-06	0.5247	0.0221	(0.4813, 0.5686)	(0.4803, 0.5669)
DIC: -119.385						

Table C5. Results of the sensitivity analysis for the congressional data, comparing 1828-1858 to 1860-1894.

Variable	Prior Information		Posterior Information			
	Mean	Standard Deviation	Mean	Standard Deviation	Equal Tail Interval	Highest Posterior Density
1828-1858	-	-	-	-	-	-
1858-1894	0	0.0330	-0.1123	0.0343	(-0.1794, - 0.0449)	(-0.1802, -0.0464)
1896-1930	0	0.0400	-0.1310	0.0341	(-0.1983, - 0.0636)	(-0.1974, -0.0631)
1932-1994	0	0.0400	0.0194	0.0305	(-0.0405, 0.0792)	(-0.0342, 0.0773)
1996-2008	0	0.0400	-0.0951	0.0449	(-0.1829, - 0.0064)	(-0.1812, -0.0051)
General Election Surge	0	1.00E-06	-0.0010	0.0009	(-0.0027, 0.0008)	(-0.0026, 0.0008)
Constant	0	1.00E-06	0.5800	0.0249	(0.5306, 0.6281)	(0.5317, 0.6286)
DIC: -156.938						

Table C6. Results of the sensitivity analysis for the congressional data, comparing 1860-1894 to 1896-1930.

Variable	Prior Information		Posterior Information			
	Mean	Standard Deviation	Mean	Standard Deviation	Equal Tail Interval	Highest Posterior Density
1828-1858	0	0.0400	0.1115	0.0343	(0.0445, 0.1790)	(0.0436, 0.1775)
1858-1894	-	-	-	-	-	-
1896-1930	0	0.0400	-	0.0330	(-0.0839, 0.0460)	(-0.0827, 0.0471)
1932-1994	0	0.0400	0.1314	0.0293	(0.0735, 0.1889)	(0.0725, 0.1876)
1996-2008	0	0.0169	0.0441	0.0449	(-0.0698, 0.1041)	(-0.0703, 0.1032)
General Election Surge	0	1.00E-06	-	0.0009	(-0.0027, 0.0008)	(-0.0027, 0.0008)
Constant	0	1.00E-06	0.4681	0.0235	(0.4215, 0.5136)	(0.4223, 0.5139)
DIC: -156.938						

Table C7. Results of the sensitivity analysis for the congressional data, comparing 1896-1930 to 1932-1994.

Variable	Prior Information		Posterior Information			
	Mean	Standard Deviation	Mean	Standard Deviation	Equal Tail Interval	Highest Posterior Density
1828-1858	0	0.0400	0.1301	0.0344	(0.0629, 0.1977)	(0.0621, 0.1962)
1858-1894	0	0.0333	0.0181	0.0330	(-0.0468, 0.0831)	(-0.0455, 0.0843)
1896-1930	-	-	-	-	-	-
1932-1994	0	0.0400	0.1500	0.0294	(0.0920, 0.2076)	(0.0911, 0.2063)
1996-2008	0	0.0400	0.0354	0.0442	(-0.0513, 0.1227)	(-0.0519, 0.1217)
General Election Surge	0	1.00E-06	-	0.0009	(-0.0027, 0.0008)	(-0.0026, 0.0009)
Constant	0	1.00E-06	0.4495	0.0235	(0.4028, 0.4951)	(0.4052, 0.4970)
DIC: -156.938						

Table C8. Results of the sensitivity analysis for the congressional data, comparing 1932-1994 to 1996-2008.

Variable	Prior Information		Posterior Information			
	Mean	Standard Deviation	Mean	Standard Deviation	Equal Tail Interval	Highest Posterior Density
1828-1858	0	0.0400	-	0.0306	(-0.0794, 0.0404)	(-0.0802, 0.0390)
1858-1894	0	0.0333	-	0.0290	(-0.1891, -0.0739)	(-0.1903, -0.0761)
1896-1930	0	0.0400	-	0.0292	(-0.2068, -0.0920)	(-0.2072, -0.0930)
1932-1994	-	-	-	-	-	-

1996-2008	0	0.0400	-	0.0413	(-0.1954, -	(-0.1964, -0.0342)
			0.1144		0.0322)	
General Election	0	1.00E-06	-	0.0009	(-0.0027,	(-0.0027, 0.0008)
Surge			0.0009		0.0008)	
Constant	0	1.00E-06	0.5993	0.0175	(0.5647, 0.6330)	(0.5658, 0.6337)
DIC: -156.938						
