WORKERS' COMPENSATION PREDICTION MODELING USING MULTIPLE REGRESSION

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Muhammad Arsalan Raza Azmi

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Construction Management and Engineering

July 2018

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Workers' Compensation Prediction Modeling Using Multiple Regression

By

Muhammad Arsalan Raza Azmi

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Eric Asa
Chair

Dr. Matthew Stone

Dr. Derek Lehmberg

Approved:

July 6th, 2018 — Date

Dr. Jerry Gao — Department Chair

ABSTRACT

Workers' compensation insurance (WCI) is the highest cost to an employer following accidents. It is needed to predict the benefits value without taking into account the past records of an employee, which is not readily available in most cases. Employment and workers' compensation data were acquired from the Bureau of Labor Statistics and the National Academy of Social Insurance, respectively. The statistical model was developed with SAS using multiple regression and the process was simplified using analysis of covariance (ANCOVA). The model predicted future values of workers compensation given a known number of covered workers for all U.S. states. The model is statistically proven to be fit for all states. The states were compared on the basis of percentage deviation from the actual values. By using this model, insurance companies and policymakers can have better understanding of workers' compensation trend and they can quotes premiums and develop policies more accurately.

ACKNOWLEDGEMENTS

First and foremost all praise is for ALLAH the Almighty, who gave me the strength, commitment, persistence, and confidence to complete the entire thesis successfully.

My sincere regards to my advisor, Dr. Eric Asa. Dr. Asa accepted to advise me when I was halfway through my degree and had to change my academic advisor. The topic of workers' compensation I was working on was relatively challenging because the data was too sensitive and difficult to obtain. It was a great challenge for me to complete my thesis but Dr. Asa was very helpful and kind to demonstrate the slightest details in the research process. I would also like to thank my committee members Dr. Matthew Stone and Dr. Derek Lehmberg. I presented my final exam in summer semester when both of them were traveling. I am really thankful to them that they made themselves available even when they were not supposed to. I would also like to express my deepest gratitude to Ingrid Scarski and Ann Denney for the administrative support throughout my studies. Extending my acknowledgment, I want to thank the wonderful on-campus resources that I took benefit of being an international and graduate student. My humble regards to Student activities office, International students and study abroad services, Student Government, Centre for writers, Statistical consultants, and PAK, the organization of Pakistani students.

Special thanks to my parents and siblings who encouraged me to get a master's degree and supported me throughout these years. My mother was not very happy to send me overseas but she was wise enough to know the benefits in the long run. I first heard about NDSU from Syed Owais Ahmed who is an NDSU alumnus. He told me about the potential assistantship opportunities that cover tuition costs. But of course, getting an assistantship is not guaranteed. I asked my father for $27k to cover all the expenses of traveling to NDSU for the first year. He

told me that this was a huge amount and he could not afford it. Since I had already passed my GRE and TOEFL exams, I applied for admission and was pleased to be accepted. I told my father and he worked very hard to gather the necessary funds for my studies. I am thankful to God that I got the assistantship in my very first semester and thus could save a lot of money. This was the first time I asked my father for something real and the way he responded and supported me was amazing. I took a great risk in coming to NDSU. I knew that if I did not get assistantship I would probably have to go back to Pakistan or if I did manage to pay for the tuition, my family would be under heavy debt. I just wanted to jump into this completely new situation and I was determined to get through it. Besides this, I experienced so much love and connection with my father which I was not expecting. Now I believe that my parents are always there to support me and I have my own reasons to say that my parents are the best in the world.

Last but not the least I would thank my Fargo family and friends. Greg, Fowzia, and the Purdins made me feel at home and gave me warm wishes in the coldest weather. It was worth going out of my comfort zone to meet these wonderful people. The Pakistani community in Fargo is very strong. Though there are only a handful of people, our connection is robust. Everyone knows each other. When I landed in Fargo for the first time in August 2015, I went to live with some Pakistani students in Moorhead who were studying at MSUM. They were very welcoming. They showed me the city and taught me the American culture and a few initial English lessons. I came to University Village at NDSU to live on-campus after one month and there I met Adnan, Bilal, Asif Arshid, Awais, Faiz, Basim, Asif Abubakar, Asad Abbas, Farhan, Mughees, Naveed, and many other Pakistanis. I was so surprised to see the love and friendship between them. We met almost every day. I can never forget the fun, laughter, and teasing. Especially the crazy road trips we had. In two years I had already driven to 36 states of the

United States and 4 provinces of Canada. I don’t know anyone yet who has visited as many places as I have in the U. S. Best thing about this group was that they showed tremendous support in the hard times. I remember how my friends reacted and support me when I was going through financial crises.

I thank the whole NDSU community to make a positive impact in reshaping my personality for good. As I move on having the graduate degree in hand and entering the professional world I hope I represent my country, Pakistan and my institution, NDSU with pride. I pray for everyone who is mentioned here and all those who are not. May God bless everyone and may they continue to spread love forever.

DEDICATION

To my mother.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A_e .... Actual excess losses

AK .... Alaska

AL .... Alabama

ANCOVA .... Analysis of Covariance

ANOVA .... Analysis of Variance

A_p .... Actual primary losses

AR .... Arkansas

AZ .... Arizona

B .... Ballast Value

B .... Before Christ

BLS .... Bureau of Labor Statistics

BR .... Business Roundtable

CA .... California

CCI .... Claim Control Incentive

CO .... Colorado

CT .... Connecticut

DC .... District of Columbia

DE .... Delaware

E .... Expected excess losses

EMR .... Experience Modification Rate

FL .... Florida

FOIA .... Freedom of Information Act

GA .... Georgia

HI .... Hawaii

IMIS ....Integrated Management Information System

KBS..Knowledge Based System

LEED ....Leadership in Energy and Environmental Design

NASI ....National Academy of Social Insurance

NCCI..National Council on Compensation Insurance

OCIP ....Owner-Controlled Insurance Program

OSHA...Occupational Safety and Health Administration

OWCP ....Office of Workers' Compensation Program

RIR...Recordable Incident Rate

SMD...Site Monitoring Discount

SOII..Survey of Occupational Injury and Illness

UK..United Kingdom

USA..United States of America

W..Weighting value

WCI..Workers' Compensation Insurance

1. INTRODUCTION

1.1. Problem Statement

The objective of workers' compensation is to provide coverage to employees for work related injuries and diseases, security against disruption for income, and providing suitable and necessary medical treatment and rehabilitation. Hence the workers' compensation programs provide critical support to workers who get sick or injured during work. There have been attempts to make the workers' compensation system better. Various tools and techniques proposed by the researchers have been adopted by the workers' compensation insurance system but there are still loop holes in the system. Therefore, there is a need to present the authorities and policy makers of the workers' compensation insurance system with new techniques and ideas for them to better understand the trends in the past as well as the upcoming trends of the future.

1.2. Research Questions

This research was carried out to answer the following questions. The main questions are answered by addressing various sub-questions. The main questions are:

How can the future values of total benefits in workers' compensation be predicted using statistical analysis? How can the different states, in a country like the United States of America (USA), be compared on the basis of the workers' compensation prediction model?

The sub-questions are:

1. What sources are available to get the employment and total benefits paid data?
2. What research has been done on the topic of workers' compensation?
3. What is the trend of employment and total benefits paid? How are the two sets of data correlated?

1.3. Background

There is a comparison between BLS (Survey of Occupational Injury and Illness) with the 3 state provided databases: trauma registry, hospital discharge, worker compensation databases and OSHA citations. Finger amputation were considered mainly. Result shows that SOII (Survey of Occupational Injury and Illness) presented 3,984 amputations out of which 94% were related to finger injuries, whereas in reality there were 3,637 amputations with 80% being related to fingers. The databases were linked together by probabilistic model (Friedman 2013).

Another study is present by Ikpe (2012) which shows the cost benefits analysis of preventing construction accidents. The questionnaire based survey was carried out among different contractor sizes (small, medium and large) in the construction industry of United Kingdom. The ratio analysis of the data on costs and benefits of accident prevention is shown in this study. The ratio is calculated by dividing the turnover by costs or benefits. The result showed that £1 spent in accident prevention results in generating the benefits equal to £3. The benefits include insurance, medical, litigation savings etc. Small contractor spend more on accident prevention in total than medium or large contractors. Therefore, a small contractor gain more benefits than that gained by large contractors (Ikpe 2012).

Heinrich's postulate states that the indirect cost of construction is four times more than the direct cost of construction. Hinze and Lytle validated this postulate by an analysis on direct and indirect cost based on a questionnaire survey. A grape (ratio of indirect to direct cost vs. direct cost of injuries) was presented. The grape showed that the ratio decreases with the increase in amount of direct cost (Hinze and Lytle 1991). However, this thesis will be mainly concerned about the direct costs of construction accidents.

1.4. Workers' Compensation

Workers' compensation is a kind of insurance which provides cash benefits and medical benefits to the workers when they undergo an injury during their work. The cash benefits are usually their wage replacement for the days they are away from work and the medical benefit bears all the expenses which are needed for a proper medical treatment to allow the worker return to work as soon as possible. However, the workers' compensation benefits asks the employee to give up his right to sue his employer for the negligence and unsafe work conditions (Strunin and Boded 2004). To be able to provide the workers' compensation insurance to the employers, the employer can get insurance from private carriers, state funds or the employer can be self-insured. Every state in the United States of America has its own laws for workers' compensation. The state laws determines what kind of insurance the employer is obligated to buy. For example, the state of North Dakota only allows the employer to buy insurance from state fund. (Strunin and Boded 2004)

1.5. North Dakota Workforce Safety and Insurance

North Dakota is a monopolistic state for example North Dakota has special legislation that recommends the coverage of workers' compensation be provided exclusively by the workers compensation program designated by the state. It is not allowed to get the insurance through any other private insurance company.

2. BACKGROUND AND LITERATURE REVIEW

2.1. History of Workers' Compensation

Guyton (1999) has illustrated the complete scenario of advent and implementation of workers' compensation from ancient times to the boom of industry in mid 1900s in the developed countries of the world. The first evidences of workers' compensation dates back to 2050 B.C. when labor was compensated for a lost body part or injury. No rules and regulations were set at that time and that was considered as an act of courtesy (Guyton 1999).

When the massive industrialization began in the late 18th century the need of workers' compensation grew. The employers at that time figured out ways of retreating from paying the workers for their loss. They would make workers sign contracts which reduced the employer's responsibility to compensate for the bill by half. For example, the injury occurred by the ignorance or carelessness of the worker himself or a fellow worker then it would not be compensated. These contracts were named as "worker's right to die" or "death contracts" because it extricated the employer from providing safe work environment and extricated the employer from paying the compensation as well. Later the British took steps to pass an Employer Liability Act. This law had little effect on the earlier stages but it went on polishing until a few years ago in 1990 the Americans with Disabilities Act was passed which can be considered as a prime step for giving the workers' life more importance (Guyton 1999).

2.2. Workers' Compensation in the US

United States of America adopted the workers' compensation rule after the Europe. The labor regulation in the United States which is decentralized in nature caused some hindrance in implementation of the law. Nevertheless, the first workers' compensation law was passed in 1911 in the state of Wisconsin, to be followed by forty seven other states until 1920. Mississippi was the last state to accept the comprehensive legislation which occurred in 1948 (Hinze et al. 1995).

Workers' compensation insurance in the United States is determined by a simple formula shown by National Council on Compensation Insurance (NCCI 2016). The final value of standard premium is acquired by the multiplication of manual rates, payroll units and experience modification rating. Some factors involved in this formula are not controllable by the company but most of them are. Manual rates are determined by the state or the insurance company from which the employer is getting the workers' compensation insurance. Payroll units are determined by dividing the wages paid by $100. Final and most important factor which is absolutely in control of the company is the experience modification rating. This represents the safety record of the company over the last 4 years. Last year is not included but the three years before the last years are incorporated (Hinze et al. 1995).

Brahmasrene & Smith (2008) has provided evidence that safety training activities play a vital role in lowering the experience modification rating of a company. They have also shown that the revenue that company generates also counts largely towards the determination of experience modification rate.

2.3. Workers' Compensation around the World

Liao and Chiang (2015) discussed the significant factors of construction accident compensation using ANOVA and correlation coefficients. The significant factors were identified after analyzing the 574 fatalities in the construction industry of Taiwan between years 1999 to 2011. The compensation procedure in Taiwan is not fair for workers, hence the families of deceased workers are forced to take the amount offered by the employers and not step in court proceedings. Categorical factors were analyzed through ANOVA and numerical factors were analyzed through correlation regression. The factors which comply with the safety management regulations are provided higher compensation than the factors which do not comply with safety management regulations. Projects with subcontracting have higher compensation which encourages subcontracting influence in the construction. (Liao and Chiang 2015)

Since the global economy is making the competition more intense, it is very important to realize cost cutting in business operations. Losing a lot of money by paying higher premium rates of workers' compensation insurance will cause nothing but problems in running the business and competing against the other international companies (Brahmasrene 2008).

There are different methods adopted by different countries for calculating the workers' compensation insurance. Imriyas et al. (2008) explains the new method adopted in Singapore which replaces the experience modification element. Their claim is that EMR values do not perfectly describe about the safety record of the particular company hence it is difficult for the owner to make a decision in selecting a contractor for a job. The new method proposed in (Imriyas et al. 2008) is the collection of a specific amount from the contractor as well as the owner and if there is no accident the whole amount will be returned to them as it is after keeping the fees. Contractor's responsibility is to safely carry out site operations whereas owner's responsibility is to keep an eye on the contractor's safety practices.

2.4. Importance of Workers' Compensation

Workers' compensation insurance was implemented in every state in the United States in 1948. These laws were made to provide benefit, cure and income to injured workers or their family members. In a case of a work related accident, the worker's compensation provides secuirty against loss, injury, and financial burden for the employee and allows the employer to not be held responsible for the accident. Delays, workload and cost arising from litigation are decreased. Money wasted in lawyer fees and payment to witnesses is eliminated and time is saved because the element of appeals and trials in court is eradicated. A safe and friendly work environment is maintained under the umbrella of workers' compensation where employees feel a sense of security of their health, wealth and future (Everett 1995).

Loss frequency is given relatively more importance than loss severity in setting up a premium rate. For example, a company with 5 accidents of $10,000 each will have the standard premium much greater than the company with 1 accident of $50,000. All the other factors are kept constant which may be applicable for calculating Workers' compensation insurance. The company with several small accidents is considered more risky than the company with only one big accident is because of the safety practices they have adopted. No one can tell that any one of those several small accidents could turn out to be a bigger one with some minute on site changes. (ABCs of Experience Rating 2016)

Accidents cost can highly impact any small or starting contractor. The calculation of breakeven point of an organization is done which shows that an organization can achieve its break-even point much earlier if the direct cost of accidents are prevented. The direct cost of accidents are often hidden in the variable cost portion of financial reports. They are often neglected. But the economic impact of an accident is not to be ignored. Break-even cost was calculated including the direct cost of accidents and then the break-even cost was calculated excluding the direct cost of accidents. The result showed that the revenue required to cover the direct cost of accident was 14.5% of the gross income (Veltri 1990).

2.5. Workers' Compensation Insurance Calculation

Everett & Thompson (1995) have given a detailed methodology for calculating workers' compensation insurance. There are many factors included in the calculation but these factors are not constant in every state in the United States of America. The information in the paper cited above is sourced from National Council on Compensation Insurance (NCCI). The basic formula of calculating the standard premium is as follows:

Standard Premium = Manual Rate x Payroll Units x Experience Modification Rating

Starting with the explanation of manual rates which is a key factor in determining the standard premium. Manual rates are called standard premium because they are published in a state issued manual for each work class. This factor is not controllable by any company since the regulatory authority determines the value by the past actual losses occurred. The manual rate of banks and clerical job will be far less than the manual rate of building construction. These rates are revised each year since new safety practices are introduced and the workers' compensation claim data changes from year to year. These rates are expressed in dollar per $100 of straight time payroll. The values are usually smaller than $100 but they can be greater.

When the straight time direct labor costs of any employer is divided by $100, the resulting value is called payroll units. This makes sure that the worker is given compensation coverage for his services regardless of any time of the day or the number of hours. By seeing this definition one thing will come into mind that the employer paying more against the wages will be paying more in insurance than the employer paying less. This is contradicted by the regulatory authorities to encourage the employer to pay more to their worker. In the final calculation of WCI the employer paying higher wages is compensated in his insurance premium value.

The most critical component in calculating the final amount of insurance premium is Experience Modification Rating or EMR. This is entirely controllable by the employer by being safe and having a clean accident record. The accident record sticks with the employer for three years. This means that an employer eventually ends up paying for its own accidents in higher premiums in future. Everett and Thompson (1995) have given a formula for calculating the EMR value that is:

$$Experience\ Modification\ Rating = \frac{A_p + \mathrm{W}A_e + (1 - \mathrm{W})E_e + B}{E + B}$$

In this equation,

A_p = Actual primary losses (total of costs below $16500 accident)

W = Weighting value

A_e = Actual excess losses (total of costs above $16500 per accident)

E = Expected excess losses

B = Ballast Value (from state provided manuals)

2.6. EMR and Workers' Compensation Relation

Experience modification rating is an important factor when it comes to determination of the standard premium in workers' compensation insurance (Hinze et al. 1995). Although assuming a company's safety record only on the basis of experience modification rating alone is not justified. If there are two companies doing a similar job but having different revenues or even different number of employees, their value of EMR would be different and they will end up paying different amounts for workers' compensation insurance. As discussed above, loss frequency is given relatively more weight than loss severity in setting up a premium rate. A_e in the above equation or Actual excess losses (total of costs above $16,500 per accident) indicates the loss severity of the employer. Whereas, E or expected excess losses in the equation determines the loss frequency of the employers. Finally, there is W (weighting value) which is responsible of giving more emphasis to loss frequency than to loss severity.

There is another factor called Ballast value which is added in both the numerator as well as the denominator of the equation. This makes sure that the EMR values stays closer to 1.00 and does not change drastically with one large accident cost. Another term in the numerator is (1-W) which has the same effect as ballast value. Both these terms help smoothing the changing effect in EMR values over the course of the years.

2.7. New Techniques in Workers' Compensation Phenomenon

Imriyas et al. (2008) states about the concept of the Knowledge Based System (KBS) which revolves around the basic formula of:

$$\text{WCI net premium} = \text{Risk fee} - \text{CCI} - \text{SMD}$$

In this formula, the risk fee is the total fee which the contractor pays to the insurance company which the contractor can get it reimbursed by the client or owner. CCI or Claim Control Incentive is the amount paid by the insurance company to the contractor at the expiration of the period for controlling the claims. The more claims there would be on site the less CCI would be paid to the contractor. SMD or Site Monitoring Discount is the amount paid to the client by the insurance company at the expiration of the agreement for keeping an eye the site operation. Less accidents on site will lead to greater amount of SMD paid to the client in return for his responsible services during site operation.

2.8. Predictive Modeling Using ANCOVA

Analysis of Covariance (ANCOVA) has been used in designing the predictive models. The example of applying ANCOVA to the construction industry is rare. However, people from other fields of study have been using this statistical technique to design the predictive models. ANCOVA vastly used in the field of phycology, Murray (2010) has used this technique to predict the future test scores of students based on their previous performances.

Another use of this technique is addressed in the field of biology. Lessard (2000) designed a predictive model for the species living near a dam to perform the effects of temperature on the living organisms in Michigan. This was done to see the pattern of the habitat parameter and the density. (Lessard 2000)

2.9. Literature Review Database

A comprehensive literature review has been done from seven different databases available to the students from North Dakota State University Library. The Databases and the number of research articles extracted from each database are shown in the table below.

Table 2.1. List of Databases for Literature Review

S. No.	Distribution By Database	No. of Articles
1	Academic Search Premier	3
2	America: History & Life	2
3	American Society of Civil Engineers	12
4	Business Source Premier	6
5	Google Scholar	4
6	Science Direct	5
7	Web of Science	3

Furthermore, the research papers are distributed by the country in which they are focusing and the data used for those particular countries. The table below shows the various countries and the number of publications discussing them.

Table 2.2. Distribution by Countries

S. No.	Distribution By Countries	No. of Articles
1	Australia	2
2	Canada	1
3	*Multiple*	3
4	Singapore	4
5	Taiwan	1
6	UK	2
7	USA	22

There are a mix of articles according to their age. For the introduction and importance of workers' compensation the old articles are used because they contain the basic definition whereas new articles are cited for discussing the analysis on workers' compensation. In the last 25 years, 36 articles have been selected with a mix of topics. Topics breakdown will be discussed later. Below is the table 2.3 and the figure 2.1 is illustrating the articles selection from each particular year.

Table 2.3. Distribution by Year

Distribution By Year	No. of Articles
1990	1
1991	1
1992	0
1993	1
1994	2
1995	3
1996	1
1997	3
1998	0
1999	2
2000	0
2001	0
2002	1
2003	0
2004	1
2005	1
2006	0
2007	2
2008	2
2009	1
2010	1
2011	2
2012	5
2013	2
2014	1
2015	2

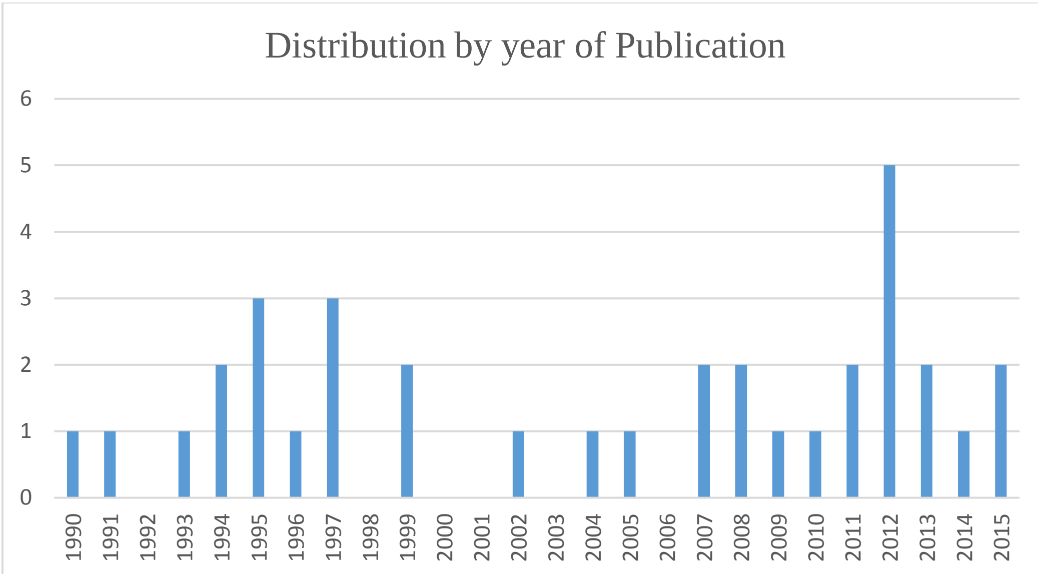


Figure 2.1. Distribution of Articles by Year

Finally the most important distribution is the distribution of topics. Four topics are selected which seemed relevant to the workers' compensation. They are 1. Safety 2. Introduction or Importance of Workers' Compensation 3. Study on Workers' Compensation and 4. Accident Cost. The number of articles are shown in the table below.

Table 2.4. Distribution by Topics

S. No.	Distribution By Topic	No. of Articles
1	Safety	6
2	Intro/Importance of Workers' Comp.	7
3	Study on Workers' Comp.	18
4	Accident Cost	3

The master list showing all the research article reviewed is present below. It shows the name of paper, author, year of publication, country, research findings, and the methodology used.

Table 2.5. Master List of Literature Review

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
1	Workers' Compensation In Construction: Workers' Benefits Under Alternative Dispute Resolution Systems	Robert D. Emerson ; R. Edward Minchin Jr., and Stephen Gruneberg	2013	USA	Alternative dispute Resolution (ADR), workers get benefit through ADR much quicker but they have to compromise on the amount. ADR usually deals with just pain and suffering.	Comparison of methods in Maryland state.
2	Costs Of Accidents And Injuries To The Construction Industry	John G. Everett and Peter B. Frank	1996	USA	Accidents costs have risen from 6.5% in 1979 to 15% in 1996 of total cost, cost of workers comp has increased, indirect cost is not included.	Data was collected from Business Rountable (BR) and analyzed, calculations were compared of BR and present paper.
3	Identification Of Safety Risks For High Performance Sustainable Construction Projects	Bernard Fortunato, Matthew Hallowell, Michael Behm and Katie Dewlaney	2012	USA	LEED certified buildings are more prone to accidents because workers have to work at higher altitude and near machinery for long period of time.	Six case studies were done and two validation cases.

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
4	Construction Safety Risk Mitigation	Matthew Hallowell and John Gambatese	2009	USA	13 safety program elements were identified and divided into 4 tiers. Tier distribution was from upper management to record keeping.	Safety risk classification system was generated, safety program elements were identified and then analyzed using Delphi method.
5	Analysis Of Fatalities Recorded By OSHA	Jimmie Hinze and Debra Russell	1995	USA	States were divided into regions and it was seen which states have more number of OSHA violations resulting in accidents. Fall is most common accident.	OSHA IMIS data of fatalities for 3 years was analyzed.
6	Cost-Benefit Analysis For Accident Prevention In Construction Projects	Elias Ikpe, Felix Hammon and David Oloke	2012	UK	1£ spent in accident prevention saves 3£. Small contractors spend more so they get more benefit too.	Questionnaire survey among contractors. Ratio is calculated by turnover/costs or benefits.

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
7	Experience Modification Rating For Workers Compensation Insurance	John G. Everett and Willard Thompson	1995	USA	Shows how WCI premiums are calculated using Payroll, EMR and manual rates. Employers pay for their accident expenses eventually. Loss frequency is more imp than loss severity.	Premium = Manual rates x EMR x Payroll. EMR has a whole different and complex formula of its calculation.
8	Transportation Agency Use Of Owner-Controlled Insurance Programs	Cliff Schexnayder, Sandra Weber and Scott David	2004	USA	Benefits of OCIP: 1 only one company dealing with WCI. 2 one point of communication 3 provides safer jobsites. Disadvantages: 1 Administrative burden increases 2 there are some gaps in liability.	Questionnaires were sent to DOT of all 50 states and then they were interviewed too.
9	Premium-Rating Model For Workers' Compensation Insurance	K. Imriyas, S. Low, A. Teo and S. Chan	2008	Singapore	Construction projects are unique so EMR method of insurance calculation is ineffective. KBS is presented to calculate insurance premiums.	A/c to KBS: WCI premium = Risk Fee - CCI - SMD

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
10	An Accident Cost Impact Model: The Direct Cost Component	Anthony Veltri	1990	USA	Business can achieve its breakeven much earlier if accident costs are prevented. Accidents costed 14.5% on that project.	Data from marine cost study of major self-insured west coast port was analyzed.
11	Cost Of Construction Injuries	Jimmie Hinze and Lisa Lytle Appelgate	1991	USA	Validated Heinrich's postulate i.e. indirect cost of accident is four times the direct cost.	Questionnaire survey was done and direct and indirect costs were analyzed.
12	The Examination Of Workers' Compensation For Occupational Fatalities In The Construction Industry	CW Liao and TL Chiang	2015	Taiwan	Tells about the procedures of workers' compensation in Indonesia, japan, UK, Mexico, Germany and Lebanon. Projects with subcontracting have higher compensation.	574 fatalities (in Taiwan b/w 1999-2011) were analyzed and factors of accident compensation discussed using ANOVA and correlation coefficients.
13	Occupational Amputations In Illinois 2000-2007: BLS Vs Data Linkage Of Trauma Registry, Hospital Discharge, Worker Compensation Databases And OSHA Citations	L Friedman, C Krupczak, S Brandt-Rauf, L Forst	2013	USA	Out of 3948 amputations 94% were of fingers shown by SOII but in real there were 3637 and 80% finger amputation.	Comparison b/w BLS's SOII and trauma registry, hospital discharge and workers' compensation. Dat a were linked by probabilistic model.

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
14	Strategies For Construction Contractors To Reduce Workers' Compensation Costs	Pankaj Agarwal and John G. Everett	1997	USA	Contractors can lower their WCI premium by state legislative reforms and effective safety programs.	Groups of like-minded employers were made and then the reforms were lobbied for same cause.
15	Comparison Of Construction Safety Codes In United States And Honduras	Guillermo Arturo Recarte Suazo and Edward J. Jaselskis	1993	Multiple	Compensation paid is very less in Honduras as compared to USA. WCI premium is Honduras is very less too. Many workers are not compensated because they are temporary.	Safety laws of both countries were reviewed and Honduran construction managers were interviewed.
16	Improving Workers' Compensation Management In Construction	Donn E. Hancher, Jesus M. de la Garza,z and Gregory K. Eckere and Gregory K. Eckert	1997	USA	Many contractors do not educate their employees about WC, they consider this cost as overhead and do nothing about its management practices.	CII research project's findings were studied which showed improved methods of management for WCI program. Contractors were surveyed.

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
17	Workers' Comp. Premiums: Disparities In Penalties For Identical Losses	y John G. Everett and I.Thng Yang	1997	USA	WCI has fixed portion and a variable portion. Variable portion is controllable and all the companies are paying this portion differently on the basis of their safety record and experiences. WCI calculation is shown.	The method of calculating WCI is looked in depth and the results are inferred.
18	Experience Modification Rating As Measure Of Safety Performance	J Hinze, DC Bren, N Piepho	1994	USA	Injury frequency has greater effect than injury severity. There should be caution while comparing two firms on EMR basis. EMR means differently for large and small contractors.	EMR calculation is shown and then 3 scenarios are discussed that how EMR means different for all the different cases.
19	Overview And Analysis Of Safety Management Studies In The Construction Industry	Z Zhou, YM Goh, Q Li	2015	Singapore	Human error should be minimized. 29 safety topics are researched b/w 2011-2013. Innovative technology is used. Communication and information flow is getting better.	439 papers were selected b/w 1978-2013. Publications were distributed by countries, year, project type, project phase and technology used.

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
20	A Framework Of Computing Workers' Compensation Insurance Premium In Construction	K Imriyas, LS Pheng, EAL Teo	2007	Singapore	8 out of 17 factors were found important. New premium rating framework was established to get optimal WCI premium rates. Risk control strategy is also established for clients and contractors.	17 factors for WCI premium were identified and classified into 4 categories. Questionnaire was designed to assess the significance of each factor.
21	A Fuzzy Knowledge-Based System For Premium Rating Of Workers' Compensation Insurance For Building Projects	K Imriyas, LS Pheng, EAL Teo	2007	Singapore	KBS incorporates real time assessment of project hazard, safety, market condition and insurers' internal factors for premium rating.	Data from interviews and past WC claim was adopted to develop fuzzy KBS.
22	Predictors Of Sustained Return To Work After Work-Related Injury Or Disease: Insights From Workers' Compensation Claims Records	Janneke Berecki-Gisolf, Fiona J. ClayAlex CollieRoderick J. McClure	2011	Australia	94% of claimants had at least 1 return to work (RTW). of those 37% had at least one recurrence. Work disability effects on sustained RTW.	Income compensation and payment data was acquired form Australian WorkSafe. Regression models were used for demographic, occupational, workplace and injury characteristics.

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
23	The Lack Of Correspondence Between Work-Related Disability And Receipt Of Workers' Compensation Benefits	EA Spieler, JF Burton	2012	USA	Many worker do not get compensation because of the strict state rules.	People were surveyed who got disability on job site.
24	Preserving Workers' Dignity In Workers Compensation Systems: An International Perspective	K Lippel	2012	Multiple	Dignity of claimants can be promoted by reducing the adversarial interactions.	Literature review and analysis of legal methods. Accident compensation in New Zealand, Netherlands and Canada is looked at.
25	Contributing Factors In Construction Accidents	R.A. Haslama, S.A. Hide, A.G.F. Gibb, D.E. Gyi, T.Pavitt, S.Atkinson, A.R. Duff	2005	UK	Key factors of accidents were found to be: 1. problems from work team 2. Workplace issues 3. PPE absence 4. Problems with material 5. lack of risk management.	Findings are connected from previous researches and 100 construction accidents. Site staff was interviewed and investigation was authenticated with offsite personnel.

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
26	How Many Injured Workers Do Not File Claims For Workers' Compensation Benefits?	Harry S. Shannon and Graham S. Lowe	2002	Canada	40% of eligible injured worker did not file claim because of low injury severity.	Claim submission was predicted by questioning labor and whether claim was submitted.
27	Frequency And Cost Of Claims By Injury Type From A State Workers' Compensation Fund From 1998 Through 2008	T. M.Mroz, A. R.Carlini, K. R.Archer, S. T.Wegener, J. I.Hoolachan, W. Stiers, R. A.Shore, R. C.Castillo	2014	USA	Shoulder, knee and back injuries were most expensive because of their number of occurrences.	Data from Maryland worker's compensation insurer 1998-2008 was analyzed.
28	Empirical Evidence Of Factors Affecting Experience Modification Rate Used By The U.S. Insurance Industry	Tantatape Brahmasrene and Sarah Sanders Smith	2008	USA	EMR is inversely proportional to cost and effort put in safety training. EMR calculation benefits large companies more than the companies with lower revenue.	U.S. National survey was conducted.

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
29	A Brief History Of Workers' Compensation	GP Guyton	1999	Multiple	1837 was when concept of workers comp emerged. Initially it had many restrictions but by the time it evolved to compensate workers as much as possible. In US first comprehensive WC law was passed in 1911.	Literature review
30	Compensation Of Residential And Nonresidential Construction Workers	Thomas Moehrle	2010	USA	This paper talks about employment trends, compensation pattern and the influential factors on compensation.	Data is gathered from National Compensation Survey which is reported in U.S. Bureau of Labor Statistics.
31	Workers Compensation Reform Past And Present: An Analysis Of Issues And Changes In Benefits	Lawrence W. Boyd	1999	USA	In 1970s benefits were raised and in 1990s the cost of WC was reduced. These events are not inter-related but the recession of 1990-1991 drove the timing of the cost cutting reform.	Literature review

Table 2.5. Master List of Literature Review (Continued)

S. No.	Paper Name	Author	Year	Country	Attributes/ Findings	Methodology
32	Workers' Compensation Laws: Significant Changes In 1993	Charles A. Berreth	1994	USA	State by state description of changes in laws to reduce frauds, manage healthcare plan and improve workplace safety.	*Government issued document*
33	Repeat Workers' Compensation Claims: Risk Factors, Costs And Work Disability	Rasa Ruseckaite and Alex Collie	2011	Australia	37% of workers filed more than one claim who were male and mostly there work condition was not changed.	Data was gathered for a period of five years (1996-2000). Repeat claims were identified. Days away from work and financial impact was compared between single claimants and repeat ones.
34	Analysis Of Ethnic Disparities In Workers' Compensation Claims Using Data Linkage.	LS Friedman, P Ruestow, L Forst	2012	USA	Non-Hispanic white workers were given more privilege after any accident in terms of amount compensated or days off from work.	WC data and medical records were linked by probabilistic model.
35	Workers' Compensation Management In Construction	Donn Hancher and Jesus Garza	1995	USA	EMR, RIR, LTIR and WCCIR all is greater for small companies. Cost of WC can be reduced by the employers' active participation in the WC programs, operations and decisions.	Literature review and interview of contractor's owners and insurance companies. Info was collected from WC Research Institute, National Council on WCI and US Chamber of Commerce.

3. METHODOLOGY AND EXPLORATORY DATA ANALYSIS

3.1. Introduction

The path towards following the problem statement included extensive literature review. Research papers from various databases were downloaded which not only include engineering and construction databases but the business administration database too since this topic is closely related to business administration and finance. The research papers selected were further categorized into four groups namely: 1. Safety 2. Introduction/Importance of Workers' Compensation 3. Cost of Accidents and 4. Study on Workers' Compensation.

The study is mainly based on the statistical analysis of the data acquired by government authorities. There are two types of data which includes employment data from Bureau of Labor Statistics (BLS) and total benefits paid in dollars against workers' compensation from National Academy of Social Insurance (NASI).

3.2. Research Tasks

The overall research is divided into different tasks. These tasks are reflected in Figure 3.1.

- Task One: Formation of problem statement.
- Task Two: Identifying literature databases and previous studies through literature review.
- Task Three: Data Acquisition and preparation consisted of identifying the data sources, downloading the employment data from the Bureau of Labor Statistics, downloading the annual reports of the National Academy of Social Insurance and extracting the data from the reports. A digital file (excel format) was created by entering the values manually from the reports.

- Task Four: Exploratory data analysis included the creation of descriptive statistics table for each state, finding the trend analysis and correlation analysis. Analysis methods were then identified with the help of statistical consultants at NDSU campus.
- Task Five: A prediction model was developed with SAS software using multiple regression. The predicted values of total benefits paid were calculated for fifteen years and compared with the actual values. The residuals were calculated and plotted against predicted values and the results were concluded.
- Task Six: States were compared on the basis of predicted values and it was identified that which states were most fit for the model.
- Task Fifteen: Discussions were provided.
- Task Sixteen: Conclusions and future recommendations were given.

The data from NASI and BLS is present from every state from the year 2001 to 2015. The number of employment acquired by U.S. Department of Labor cannot be directly used for the statistical analysis because that number does not represent the number of covered workers. There are certain exclusions for small firms, agricultural workers and in Texas there are only about 80% of employees are covered since workers' compensation is elective in Texas. The number of covered workers is represented after the exclusions of non-covered workers.

SAS Integration Technologies Configuration was the software used for the statistical analysis. First of all the correlation between the data was checked. When we see the correlation between number of workers and the benefits paid for each state in a certain year that data is highly; correlated. There is a strong positive correlation. But when we see the correlation between number of workers and the benefits paid for each year in a certain state, that data is not very much correlated. It shows positive as well as negative correlation. Correlation by state is not significant because each state has different laws of workers' compensation insurance. The states are independent hence they have their own trends. The number of average employees have too much variation among the states as there are some states with a very large number of employees and there is very few employees working in some states.

The following figure shows the step wise advancement of the thesis. It tells about the agencies approached to acquire data regarding workers' compensation. There were several other agencies contacted; especially the state workers' compensation administration but there was no success. This data is highly sensitive as no one likes to share their accident records with public and the amount they are paying against each accident. In past there have been a lot of suing for underpaying the workers in case of accident or illness. The insurance companies are usually open to share their manual rates and the nature of claims but they are reluctant to share the specific details of the severity of the accident and the amount which was paid against the insurance claim. Now the companies are very careful with whom they share the data. The source of help in statistical analysis is also mentioned in the figure below. North Dakota State University has a wonderful basis of help for the students doing their research and need help with the statistical analysis. It saves a lot of time for a student to go through all the analyzing possibilities by themselves.

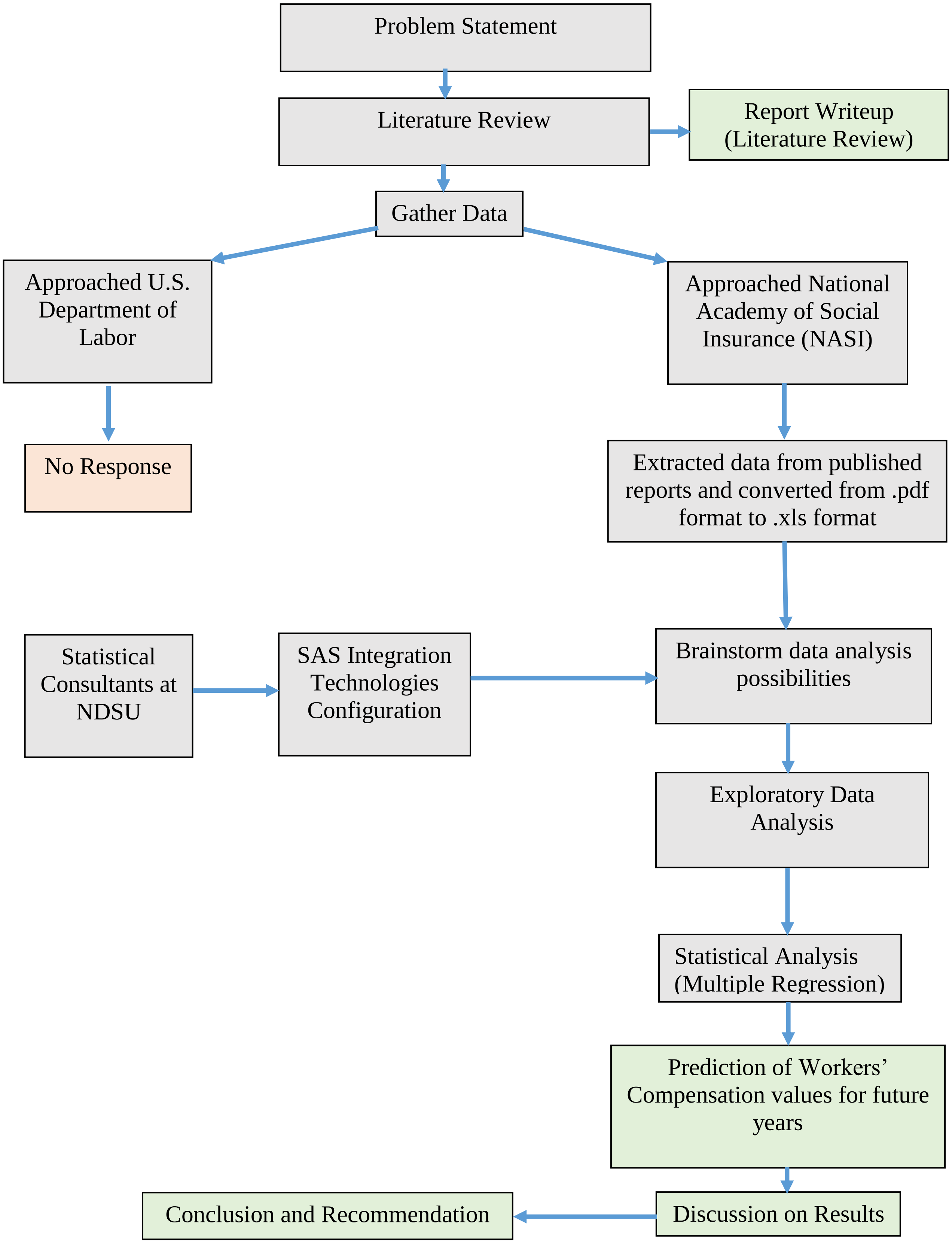


Figure 3.1. Research Methodology Flowchart

3.3. Data Acquisition

Accident reporting and the amount paid against the accidents is a sensitive topic. Companies usually do not want others to know the details of the accidents occurred and their expenses. Many insurance companies were contacted to acquire the data of number of employees and the amount of benefits paid against each accident. There was not a positive response from any one of them. U.S. Department of Labor was contacted and asked to provide the insurance and workers' compensation data but they also denied. The Freedom of Information Act (FOIA) was then used to ask the U.S. Department of Labor. The request was again denied saying:

> *"The information you are seeking appears to fall under the jurisdiction of state workers' compensation law. OWCP has responsibility for workers' compensation programs established under federal law. Since OWCP has responsibility for federal workers' compensation programs, records only exist within OWCP for federal employees. Therefore, no records exist for individuals that are not federal employees."*

After searching for several other resources to get the desired data, National Academy of Social Insurance (NASI) was identified and contacted for the insurance data. The only data we were able to get from them was in the form of published annual reports. Since those reports were in .pdf format and the table were not present in the excel (.xls) format, it was impossible to perform any statistical analysis on the data. Values from past 15 years were manually typed in the excel sheet and then transported to SAS for statistical analysis.

On the other hand, employment data was relatively easy to get. The average number of employees in every state between the years 2001-2015 was acquired from the website of Bureau of Labor Statistics bls.gov.

3.4. Sample Data

Below is the sample of the data in tabular form which was acquired from multiple resources and agencies. There was two kinds of data acquired. Firstly, the employment data form US Bureau of Labor Statistics which is shown in Table 3.1. Here in this table, there are only 6 states shown rather than 51 because this is only the sample of the data. The six states are chosen in descending order of ranking. California, Georgia, Maryland, Kansas, Maine and Wyoming represent the first, eleventh, twenty-first, thirty-first, forty-first and fifty-first ranking according to the number of covered workers in each state, respectively.

Secondly, the workers' compensation data for all US states from 2001-2015 was acquired from National Academy of Social Insurance. Table 3.2 represents the workers' compensation data. There are only six states represented in the table since it is just a sample. The six states are chosen in descending order of ranking. California, Georgia, Connecticut, West Virginia, Montana and South Dakota represent the first, eleventh, twenty-first, thirty-first, forty-first and fifty-first ranking according to the dollar amount of total benefits paid in each state, respectively.

Table 3.1. Number of Covered Workers (in Thousands) of Selected States 2001-2015

States	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
California	14728	14588	14553	14706	14992	15256	15395	15248	14377	14171	14310	14674	15139	15567	16051
Georgia	3682	3624	3597	3663	3751	3838	3891	3831	3592	3543	3594	3644	3722	3834	3954
Maryland	2295	2299	2306	2332	2372	2405	2422	2407	2326	2310	2330	2363	2384	2406	2443
Kansas	1286	1270	1251	1263	1272	1293	1324	1342	1283	1261	1268	1285	1303	1322	1332
Maine	579	577	577	583	581	584	588	585	564	559	562	565	569	573	578
Wyoming	228	230	232	240	247	260	270	279	267	263	267	271	272	277	275

The table above shows the number of covered workers in all the US states from the year 2001 to 2015. The figures in the table represents the number of covered workers in thousands. For example, the average number of covered California workers in 2001 is 14,728,000

Table 3.2. Amount of Total Benefits Paid (in Thousands) of Selected States 2001-2015

States	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
California	10082580	10974355	12409808	12459589	10938475	10017099	9608884	9529739	9392835	9396443	10850879	11535904	12113656	12097277	12065579
Georgia	1006721	917266	981142	1114154	1197521	1397771	1499306	1596051	1527428	1410753	1383560	1431794	1381721	1386071	1362480
Connecticut	641341	675895	677088	711237	713275	719758	734425	785133	842840	788701	892920	914723	955329	909138	908069
West Virginia	686808	791762	823300	796680	695771	433258	356717	319877	341717	362372	523130	476927	435709	419656	414958
Montana	181770	196197	216715	223048	239498	228347	236993	244114	246233	266850	251981	250090	248039	245858	253017
South Dakota	70736	73478	74241	77409	85889	109030	119567	111184	93578	100348	95373	92251	99084	97595	106594

The table above shows the amount in US dollars which was paid as the compensation to injured workers in all the US states from the year 2001 to 2015. The figures in the table represents the dollar amount in thousands. For example, the compensation paid to California workers in 2001 is $10,082,580,000.

3.5. Exploratory Data Analysis

Exploratory data analysis is performed to see the behavior of the data. Graphics play an important role in doing the exploratory analysis of the data. Histograms, box plots, trend lines, dot plots, density plots, normality plots are the basic graphical representation of the exploratory data analysis. The value of the data is maximized because of exploratory data analysis (Jebb et al. 2017). Mainly this analysis is to design a prediction model for the dollar amount of total benefits paid in future years. But first the exploratory data analysis is presented to see the overall picture of the data. Descriptive statistics, histograms, normality plot, trend analysis and correlation are suitable for our data and they can be found ahead in this chapter.

First of all, histogram charts are plotted to get the picture of average number of covered workers in each state and the average total benefits paid to the covered workers in each state. This is the average of 2001 – 2015. California is leading the chart with a huge margin followed by New York and Texas. Wyoming has the least average number of covered workers but the state which has the least amount of benefits paid is South Dakota.

Secondly, descriptive analysis is done to see the mean, median, standard deviation, minimum and maximum values etc. descriptive analysis is conducted for each state separately. The states are arranged in the alphabetical order. If someone want to know about the ranking of the states with respect to the number of workers and the total benefits paid, they can see the descriptive analysis tables which are present in appendix A.

Thirdly, normality tests are done on both kinds of the data to see whether the data for each state is normally distributed or not.

Fourthly, the hypothesis testing is done to see the goodness of fit. This hypothesis testing is also necessary for the development of prediction model.

Fifthly, box plots are created to see the ranges of the quartiles and the outliers.

Then, trend analysis is performed and the graphs are created to represent the increasing and decreasing trend in number of covered workers and the total benefits paid in a given year.

Lastly, correlation analysis is done to check how much the data is correlated.

3.5.1. Overview through Histograms

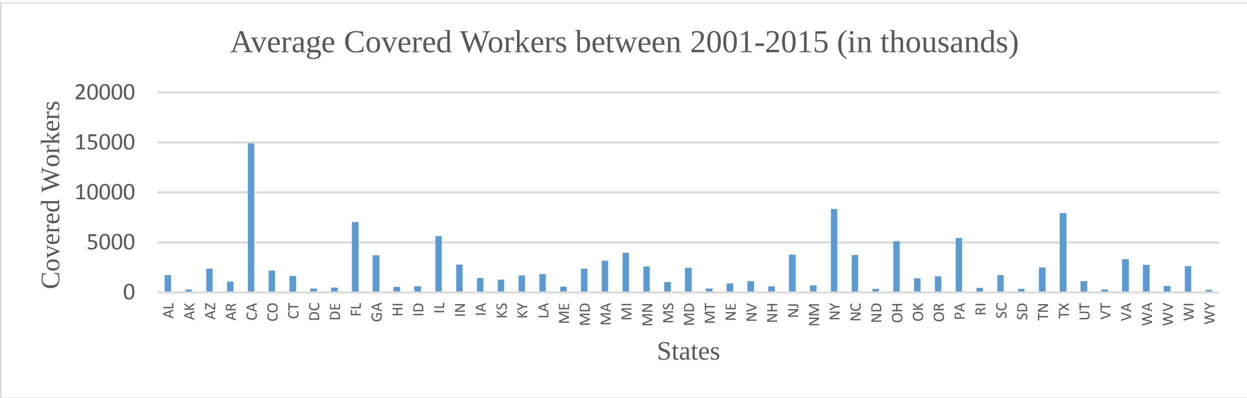


Figure 3.2. Average Number of Covered Workers Between 2001-2015 in each State (in thousands)

The figure above shows the average number of covered workers in all 51 states. The states are arranged in the alphabetical order.

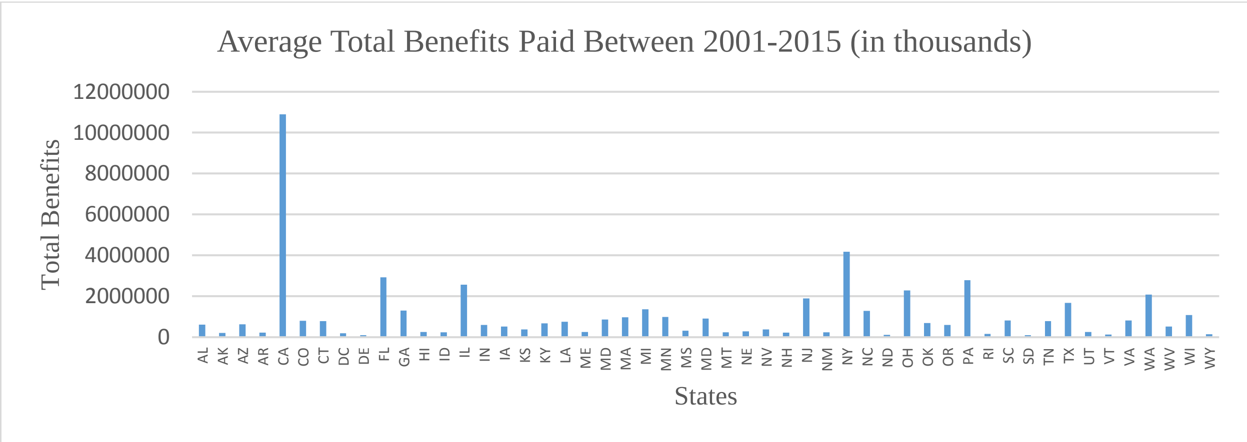


Figure 3.3. Average Number of Total Benefits Paid Between 2001-2015 in each State (in thousands)

The figure above shows the average amount paid against workers' compensation in all 51 states. The states are arranged in the alphabetical order.

Looking at both the graphs of number of covered workers and total benefits paid, it can be seen that the states with large number of covered workers are paying more against the claims of workers' compensation. But this is not necessarily true for every state. Especially the states with the small number of workers have varying ranking with respect to the total benefits paid.

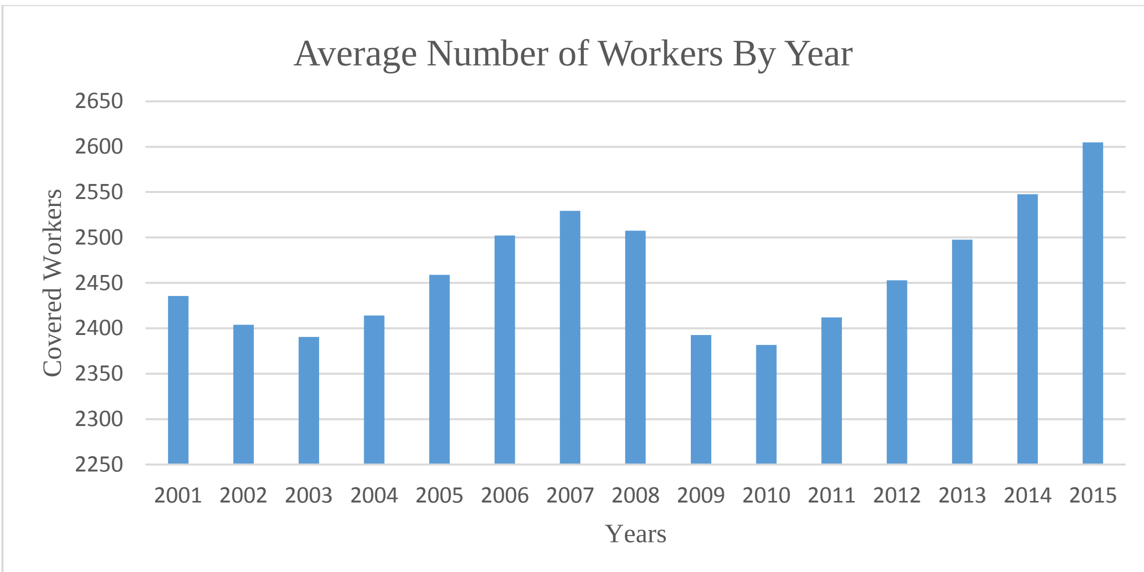


Figure 3.4. Average Number of Covered Workers by Year (in thousands)

The figure above represents the average number of workers in the US from year 2001-2015. It can be noted that there has been a drastic drop in the employment in the year 2009. United States was struck by a massive financial crisis in 2008 which resulted in a large number of lost jobs in the years following 2008 (Kotz 2009).

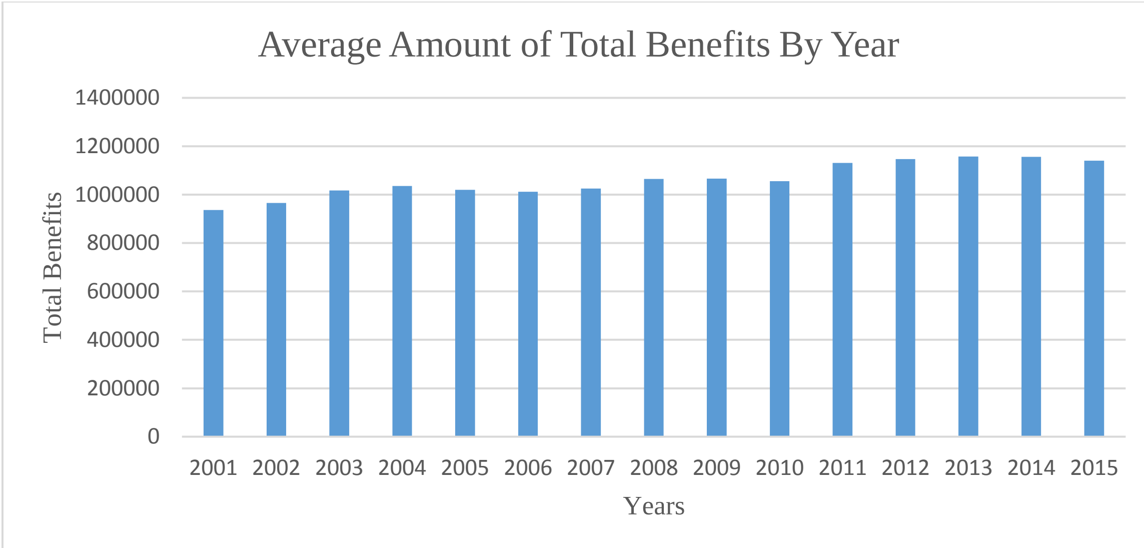


Figure 3.5. Average Number of Total Benefits Paid by Year (in thousands)

The figure above represents the average amount of benefits paid to the injured workers as compensation from the year 2001-2015.

3.5.2. Descriptive Statistics

Descriptive statistics or explorative statistics is a technique to get the overview of the data. It is used to summarize the data in order to make it easy for the reader to grasp. Trends, insights and characteristics are presented in the descriptive statistics (Marshall and Jonker 2010). Appendix A contains the descriptive analysis of 50 U.S. states and the D.C. As mentioned above, the data consists of number of covered workers and the total benefits paid. The descriptive analysis shows the basic characteristics of the data such as mean, median, standard deviation, range, minimum and maximum value etc.

3.5.3. Normality Plots on the Basis of Number of Covered Workers

Figure 3.6 below is showing the normality test performed on each and every state with respect to the number of covered workers. Among 50 states and the D.C., most of them are normally distributed with a p-value greater than 0.05.

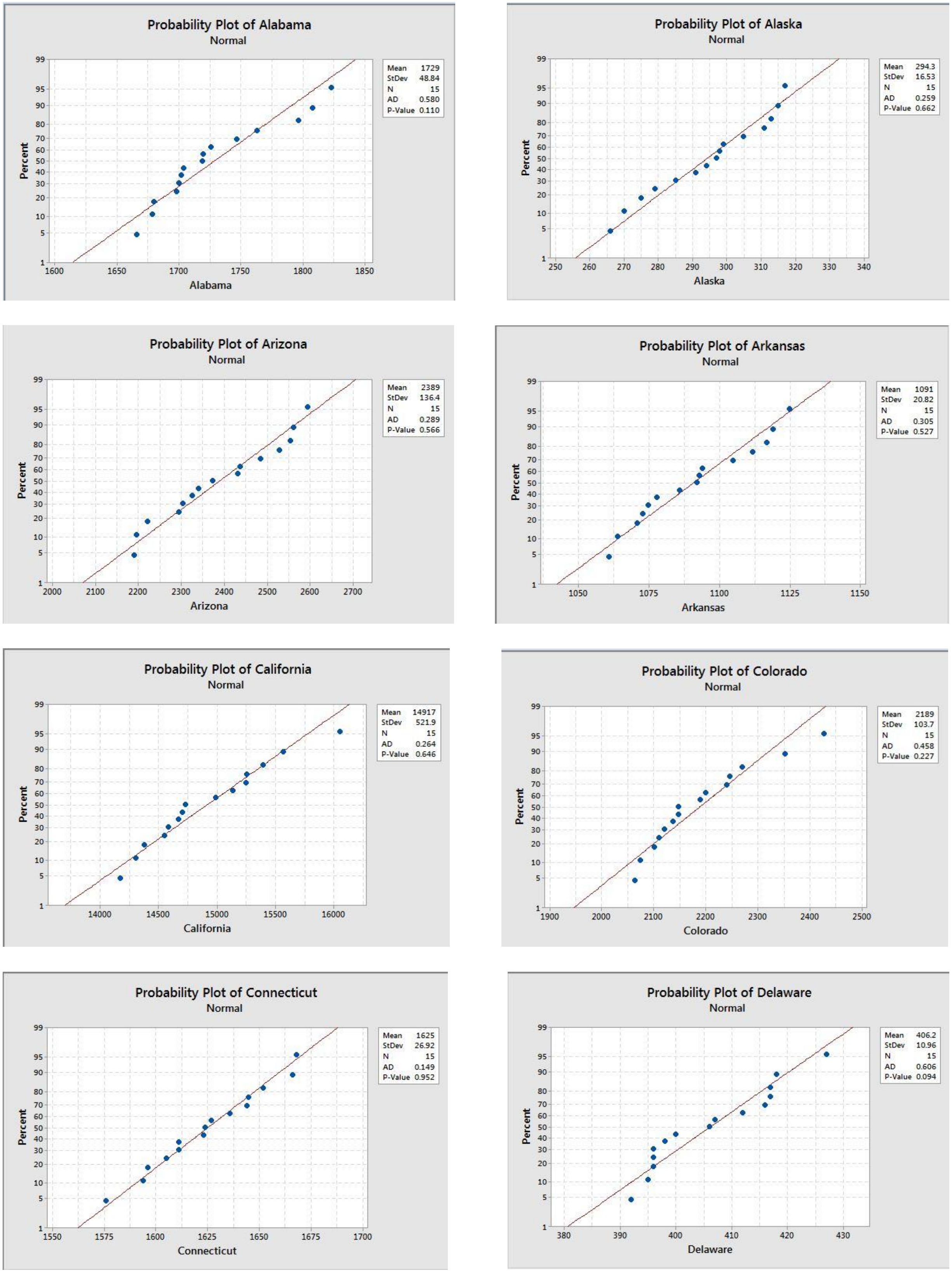


Figure 3.6. Normality Plots of Number of Covered Workers

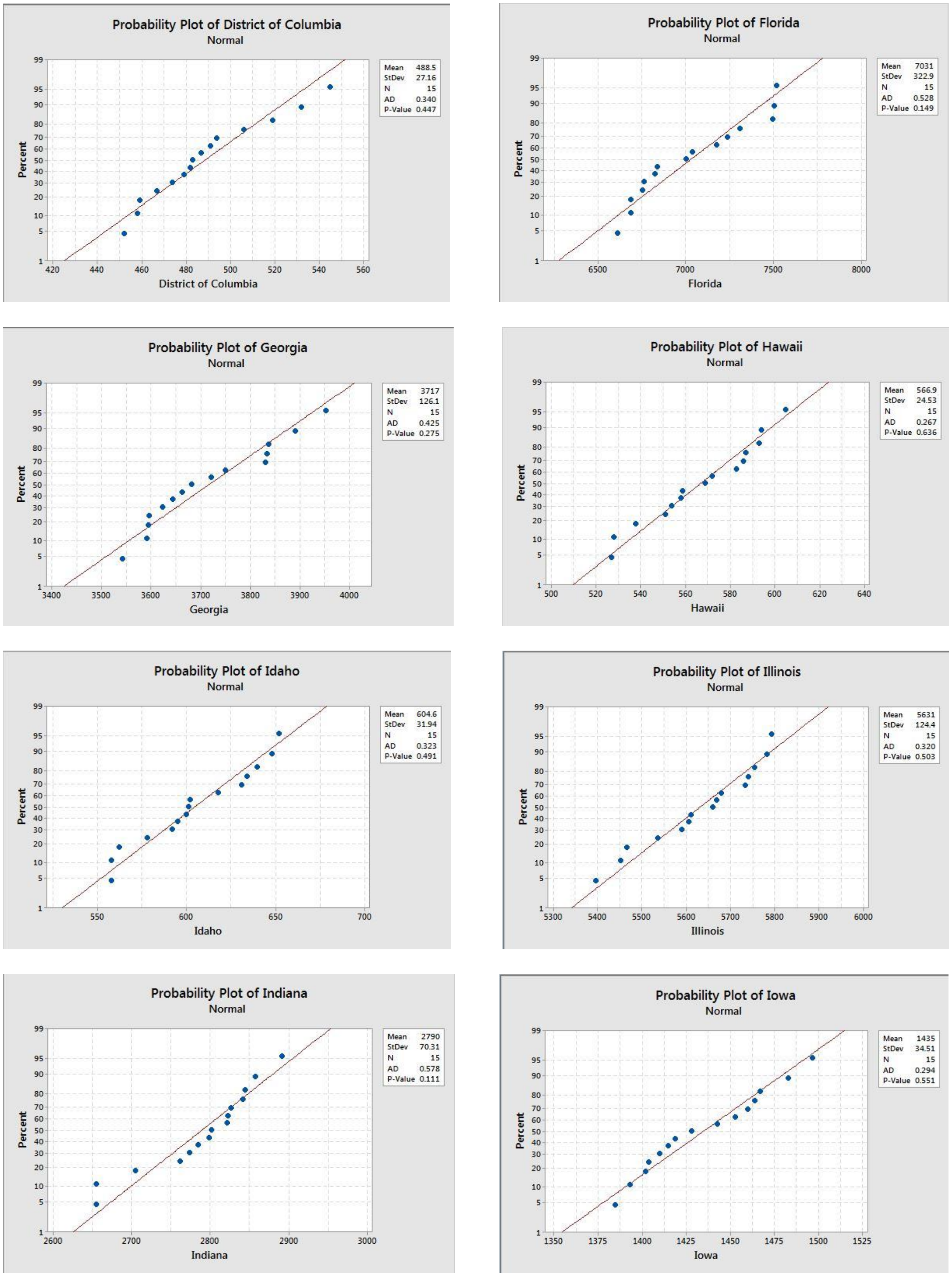


Figure 3.6. Normality Plots of Number of Covered Workers (Continued)

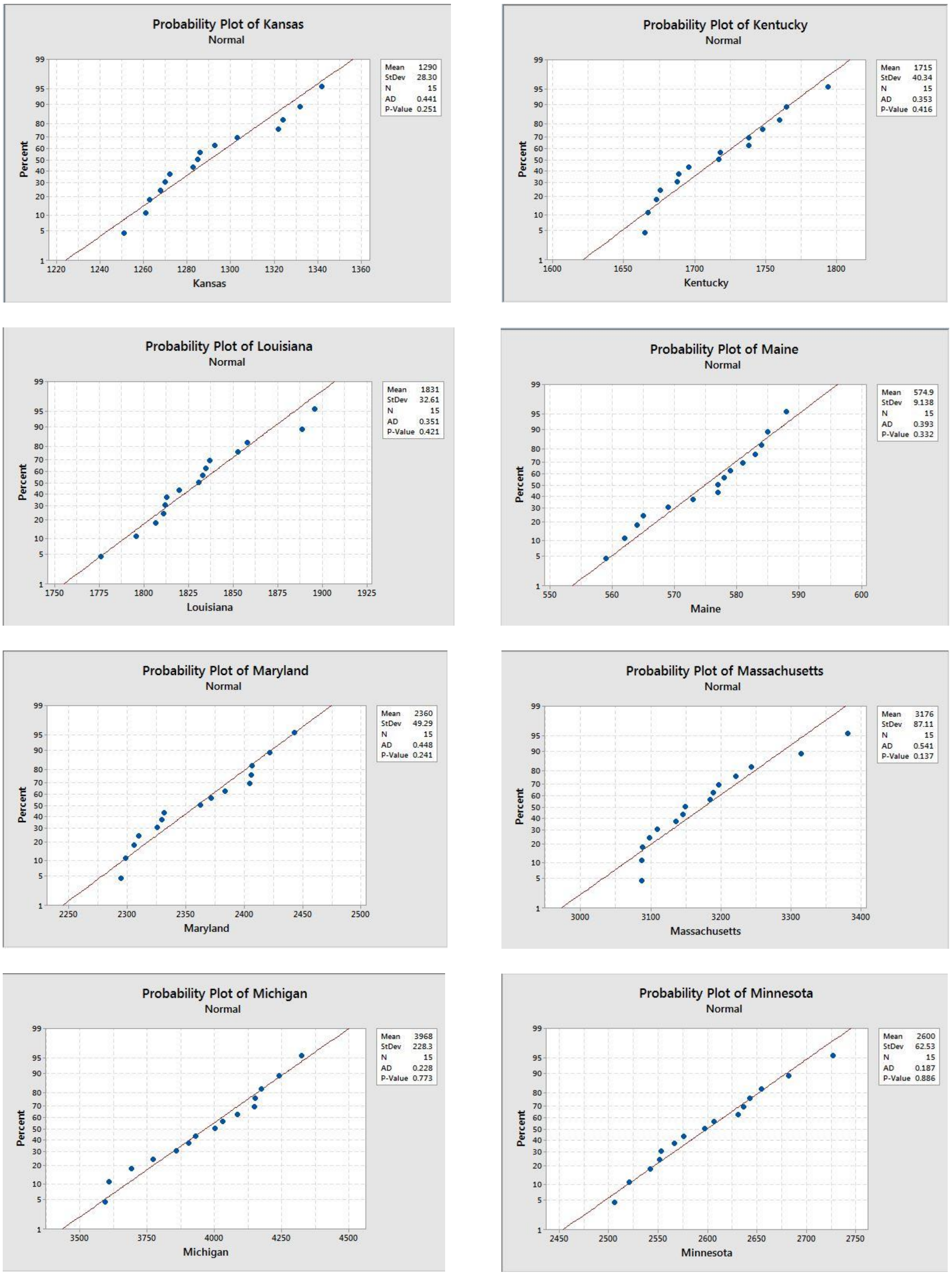


Figure 3.6. Normality Plots of Number of Covered Workers (Continued)

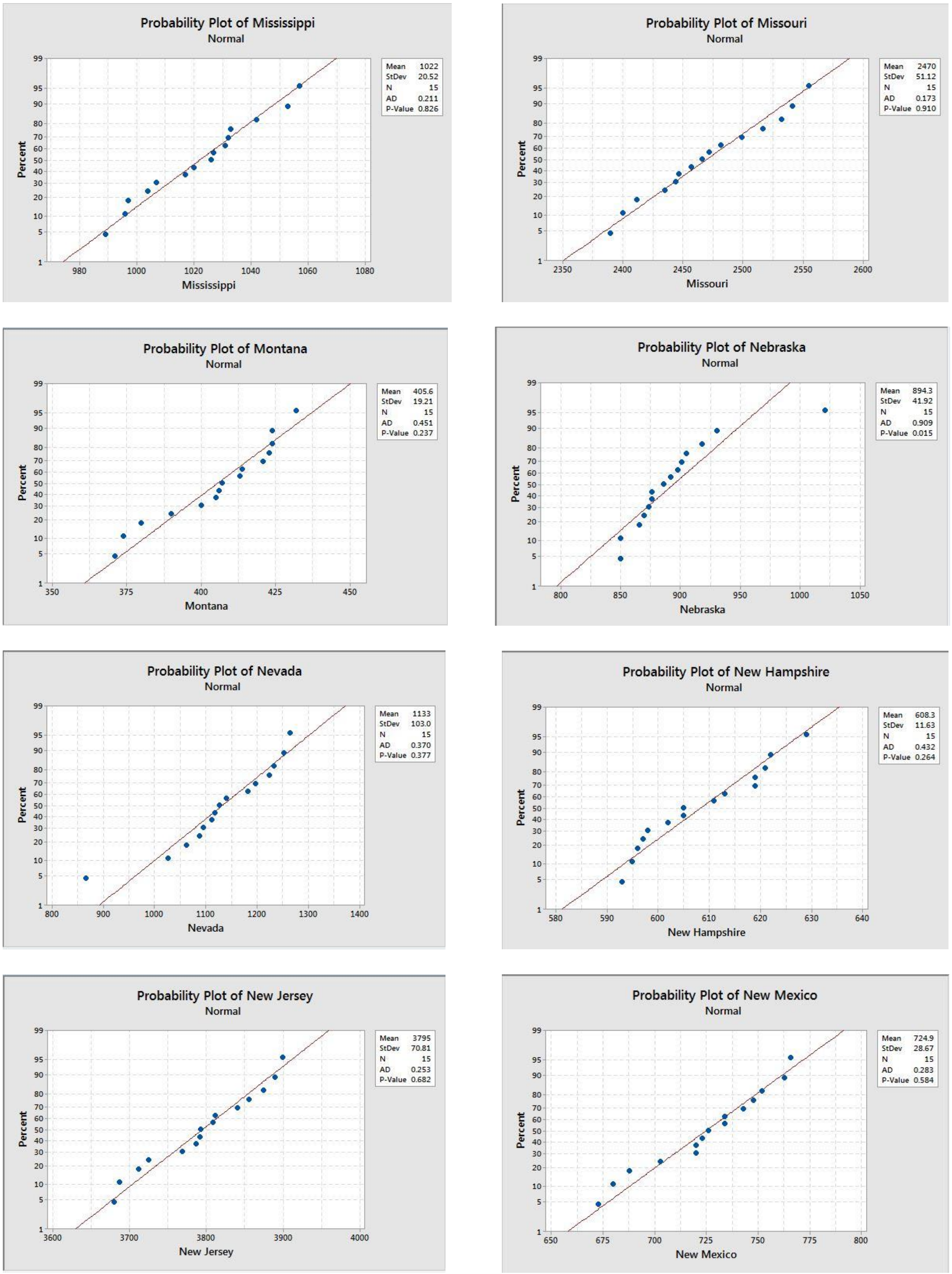


Figure 3.6. Normality Plots of Number of Covered Workers (Continued)

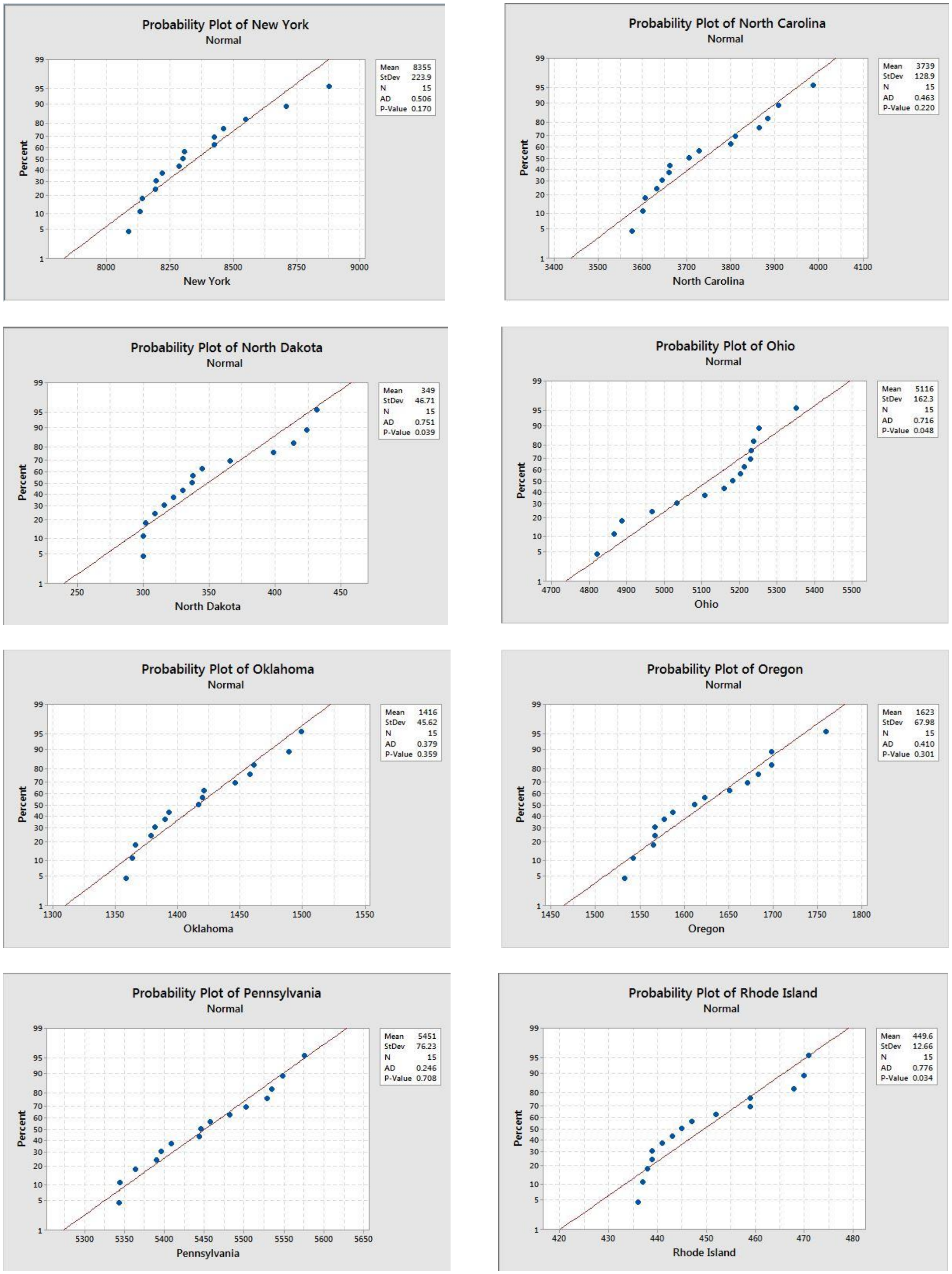


Figure 3.6. Normality Plots of Number of Covered Workers (Continued)

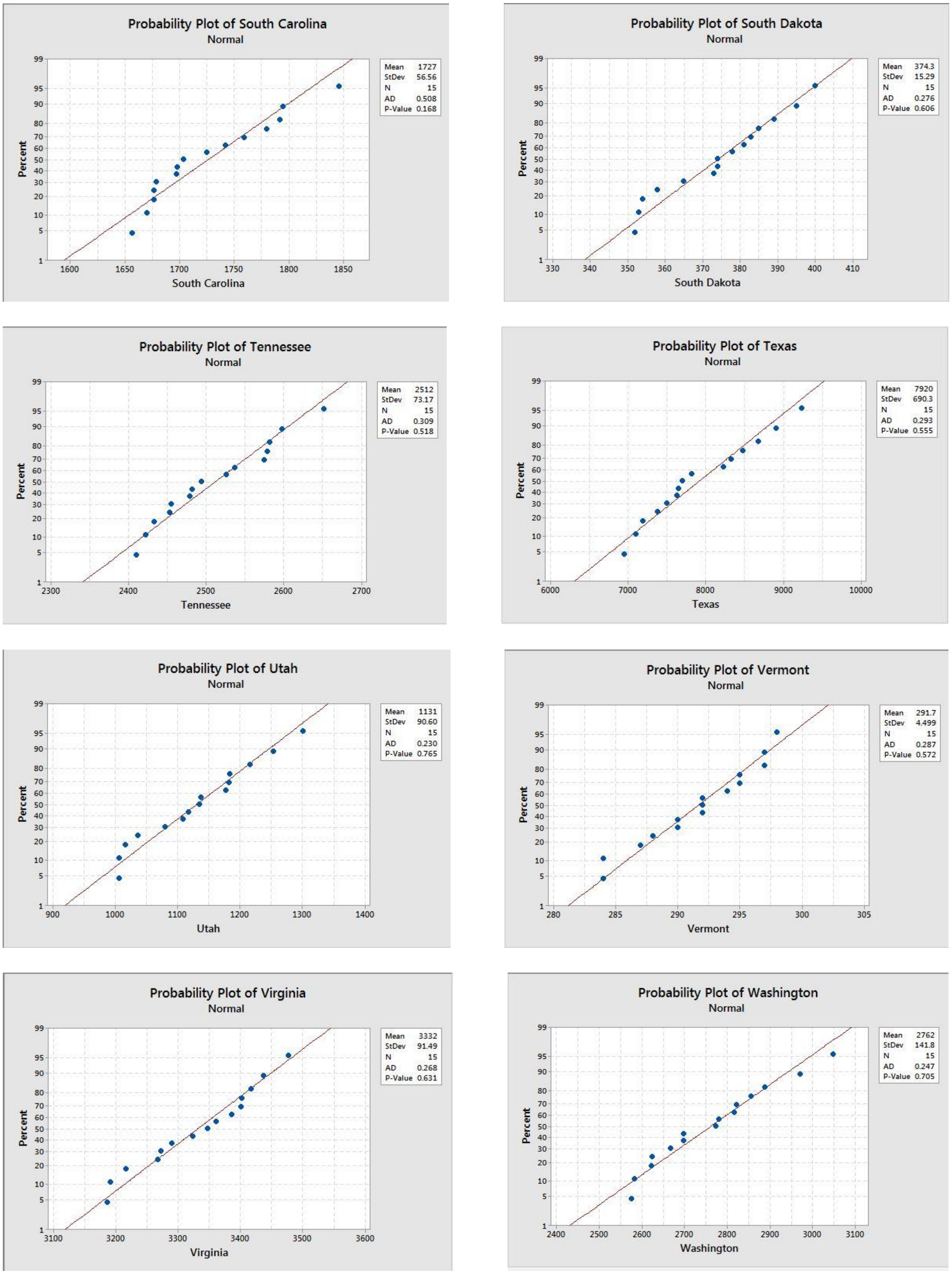
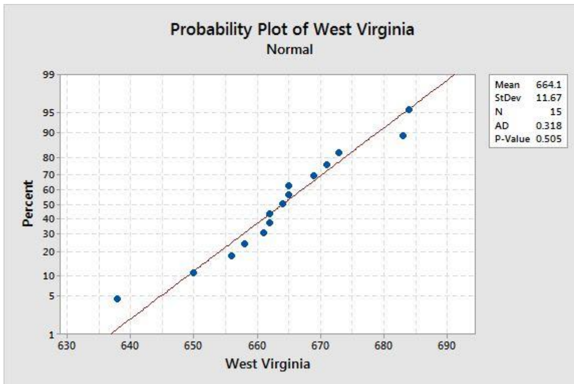
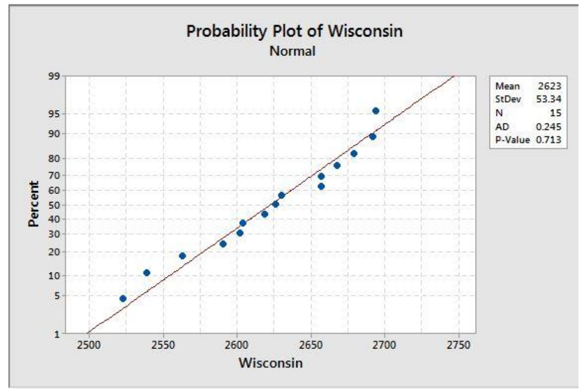


Figure 3.6. Normality Plots of Number of Covered Workers (Continued)





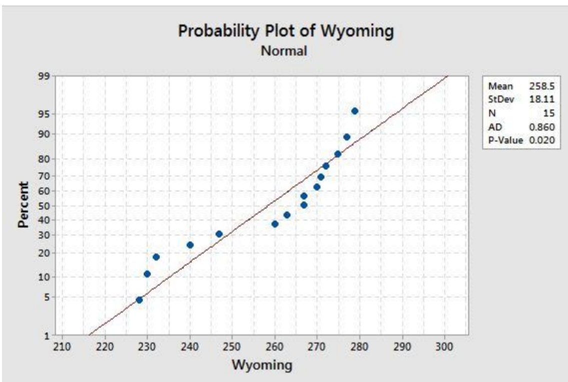


Figure 3.6. Normality Plots of Number of Covered Workers (Continued)

Table 3.3 below shows the list of p-values associated with the normality test on the number of covered workers. It can be seen that most U.S. states have the p-value greater than 0.05 which means that the data of number of covered workers for most of the U.S. states is normally distributed. Nebraska, North Dakota, Ohio, Rhode Island and Wyoming are the states whose data is not normally distributed.

Table 3.3. List of Normality Test p-values of Covered Workers

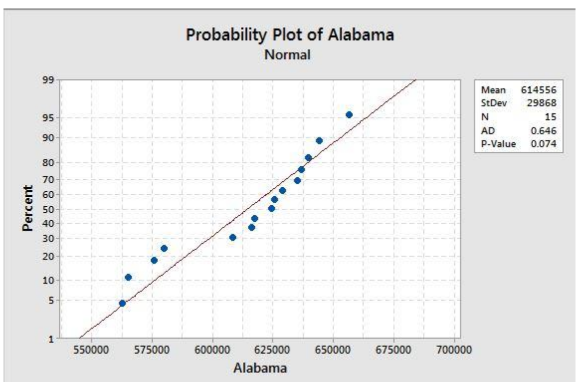
S. No.	States	p-value
1	**Alabama**	0.11
2	**Alaska**	0.26
3	**Arizona**	0.56
4	**Arkansas**	0.52
5	**California**	0.64
6	**Colorado**	0.22
7	**Connecticut**	0.95
8	**Delaware**	0.09
9	**District of Columbia**	0.44
10	**Florida**	0.15
11	**Georgia**	0.27
12	**Hawaii**	0.63
13	**Idaho**	0.49
14	**Illinois**	0.5
15	**Indiana**	0.11
16	**Iowa**	0.55
17	**Kansas**	0.25
18	**Kentucky**	0.41
19	**Louisiana**	0.42
20	**Maine**	0.33
21	**Maryland**	0.24
22	**Massachusetts**	0.13
23	**Michigan**	0.77
24	**Minnesota**	0.88
25	**Mississippi**	0.82
26	**Missouri**	0.91
27	**Montana**	0.23
28	**Nebraska**	0.01
29	**Nevada**	0.37
30	**New Hampshire**	0.26
31	**New Jersey**	0.68
32	**New Mexico**	0.58
33	**New York**	0.17
34	**North Carolina**	0.22
35	**North Dakota**	0.03
36	**Ohio**	0.04
37	**Oklahoma**	0.35
38	**Oregon**	0.3
39	**Pennsylvania**	0.7

Table 3.3. List of p-values of Normality Test of Covered Workers (Continued)

S. No.	States	p-value
40	**Rhode Island**	0.03
41	**South Carolina**	0.16
42	**South Dakota**	0.06
43	**Tennessee**	0.51
44	**Texas**	0.55
45	**Utah**	0.76
46	**Vermont**	0.57
47	**Virginia**	0.63
48	**Washington**	0.7
49	**West Virginia**	0.5
50	**Wisconsin**	0.71
51	**Wyoming**	0.02

3.5.4. Normality Plots on the Basis of Total Benefits Paid

Figure 3.7 below is showing the normality test performed on each and every state with respect to amount of total benefits. Among 50 states and the D.C., most of them are normally distributed with a p-value greater than 0.05.



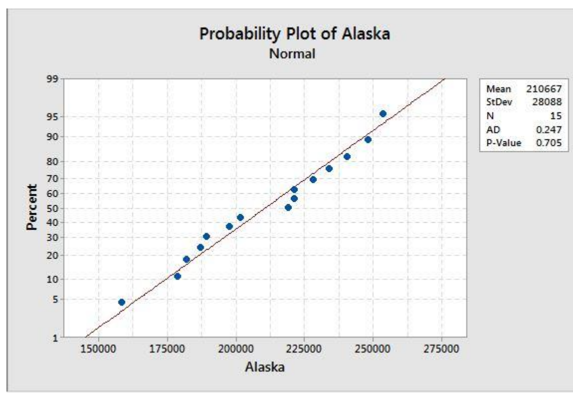


Figure 3.7. Normality Plots of Total Benefits

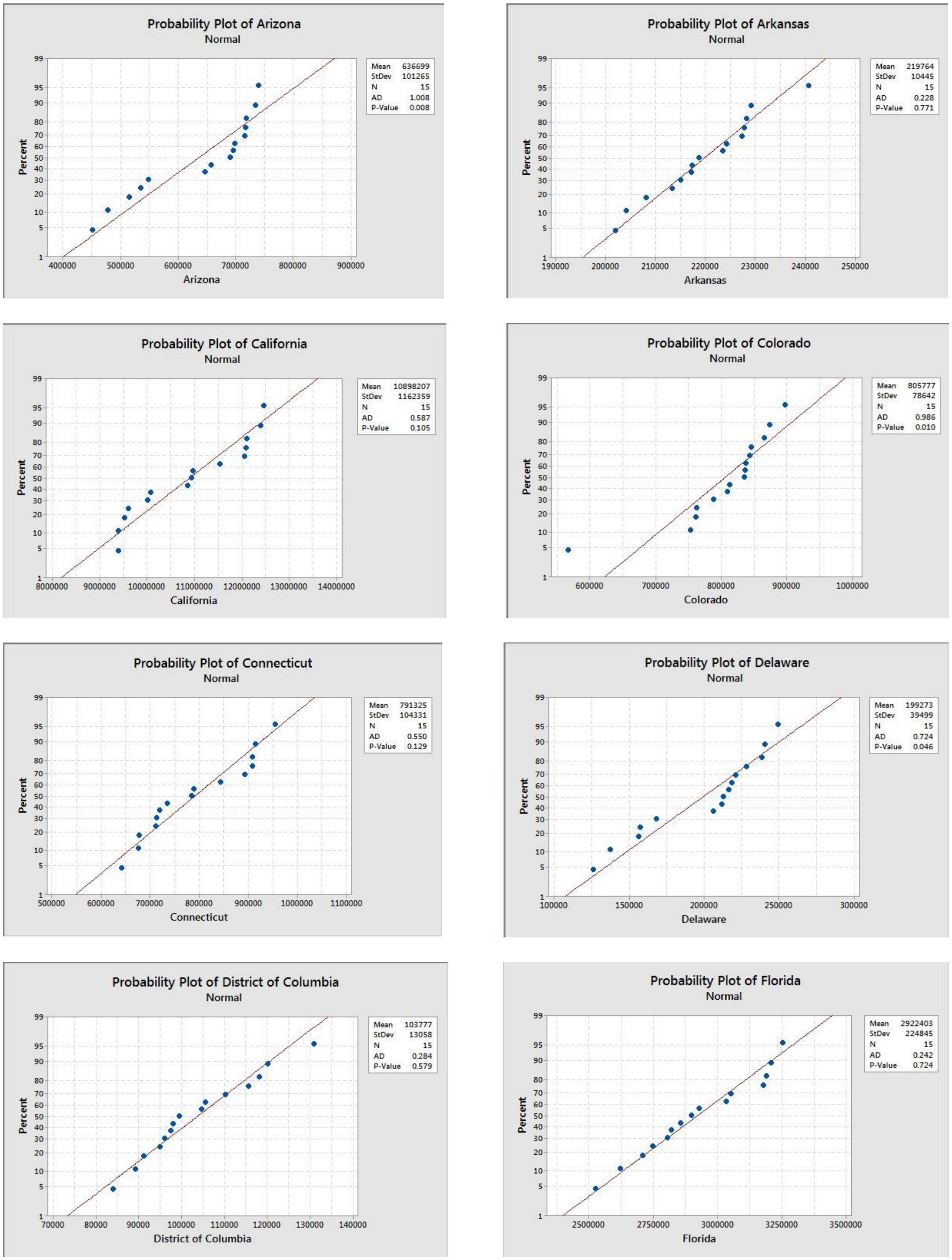


Figure 3.7. Normality Plots of Total Benefits (Continued)

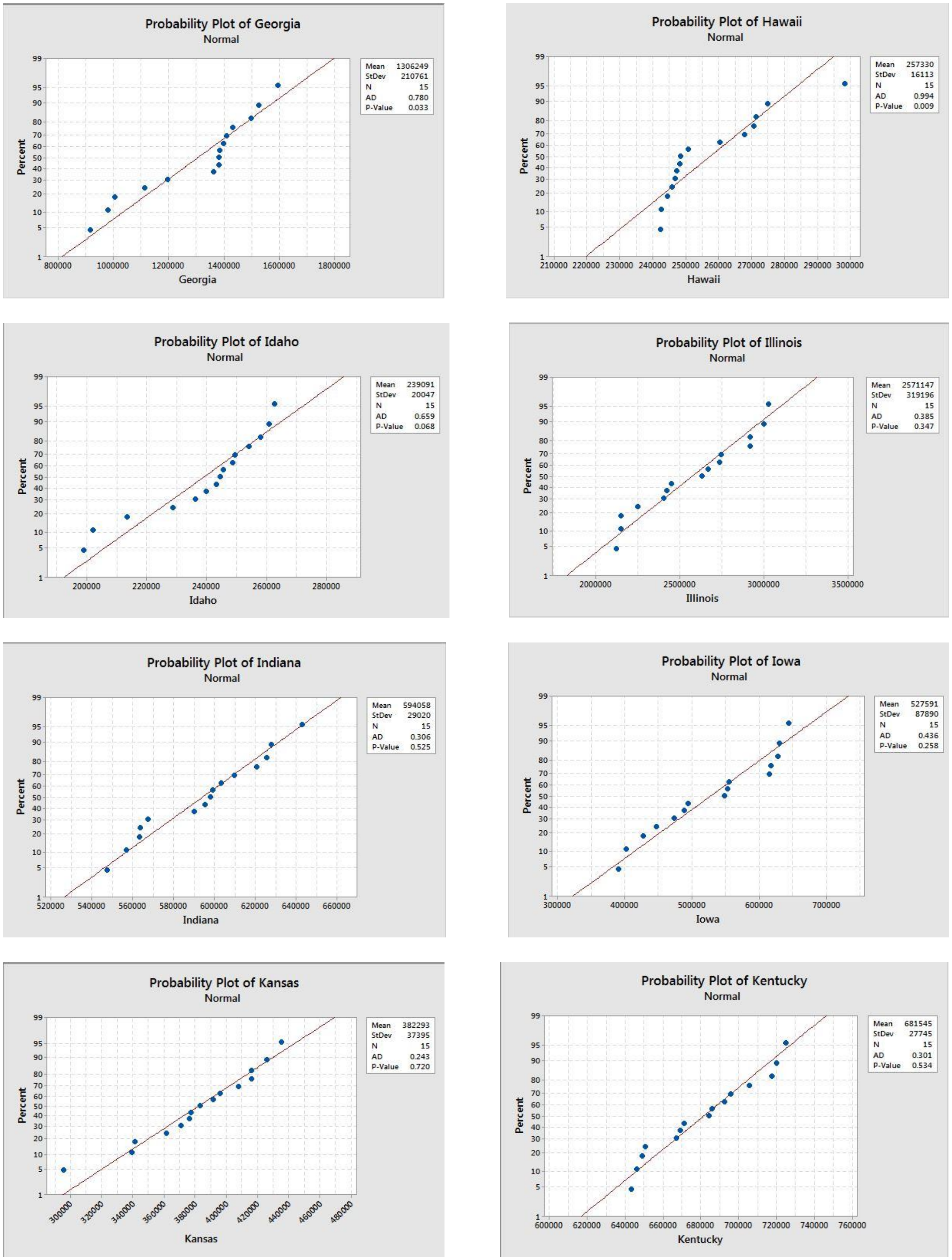


Figure 3.7. Normality Plots of Total Benefits (Continued)

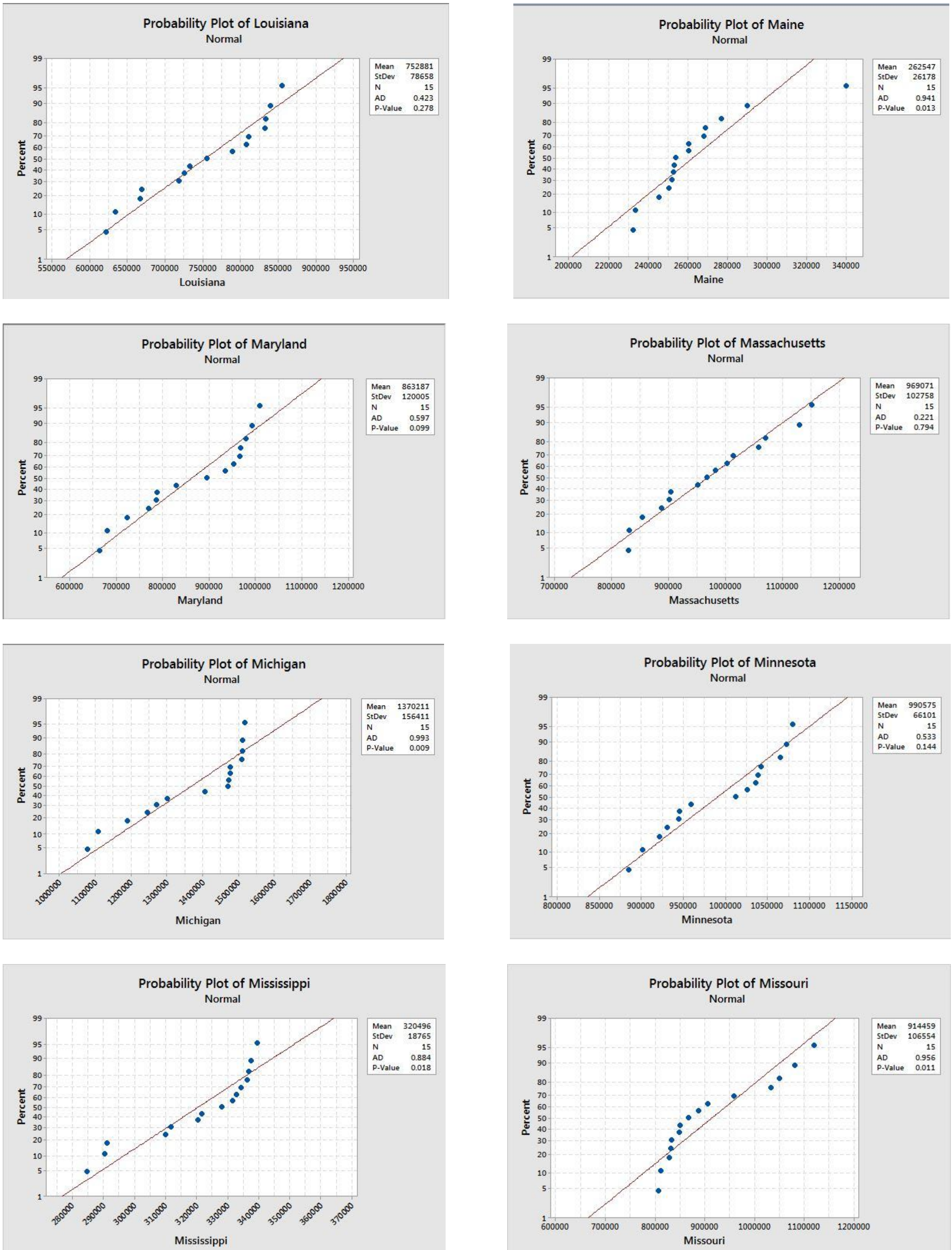


Figure 3.7. Normality Plots of Total Benefits (Continued)

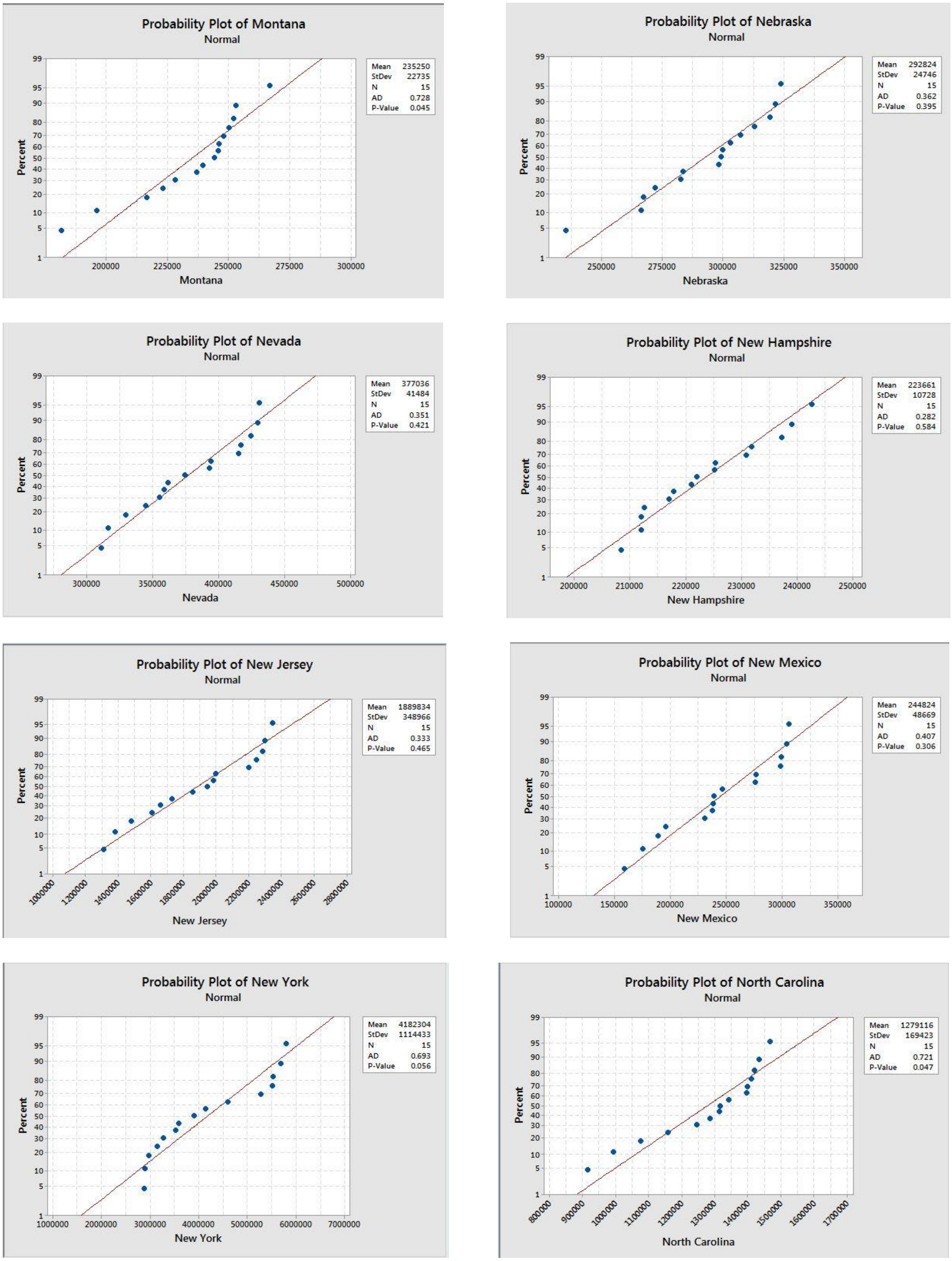


Figure 3.7. Normality Plots of Total Benefits (Continued)

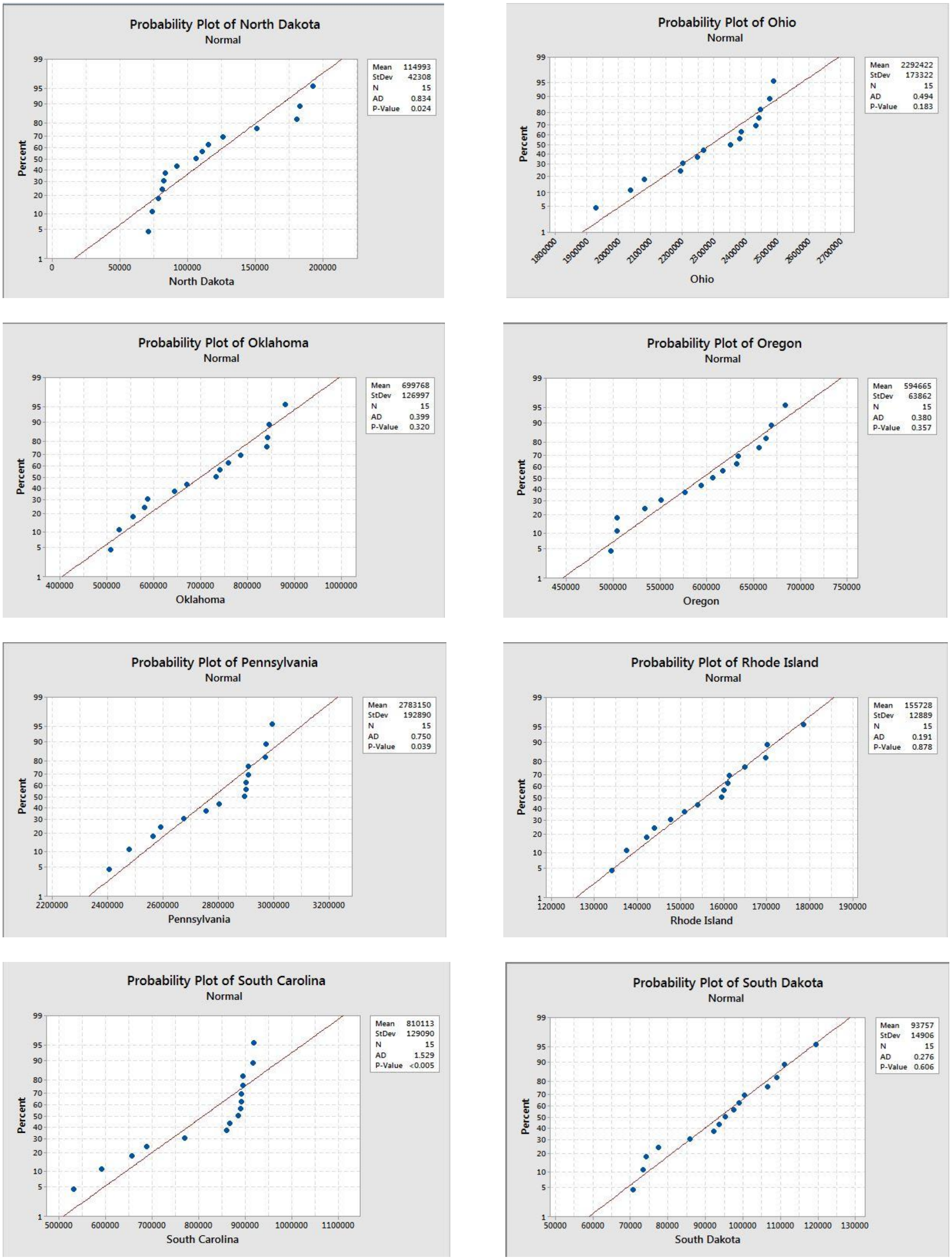


Figure 3.7. Normality Plots of Total Benefits (Continued)

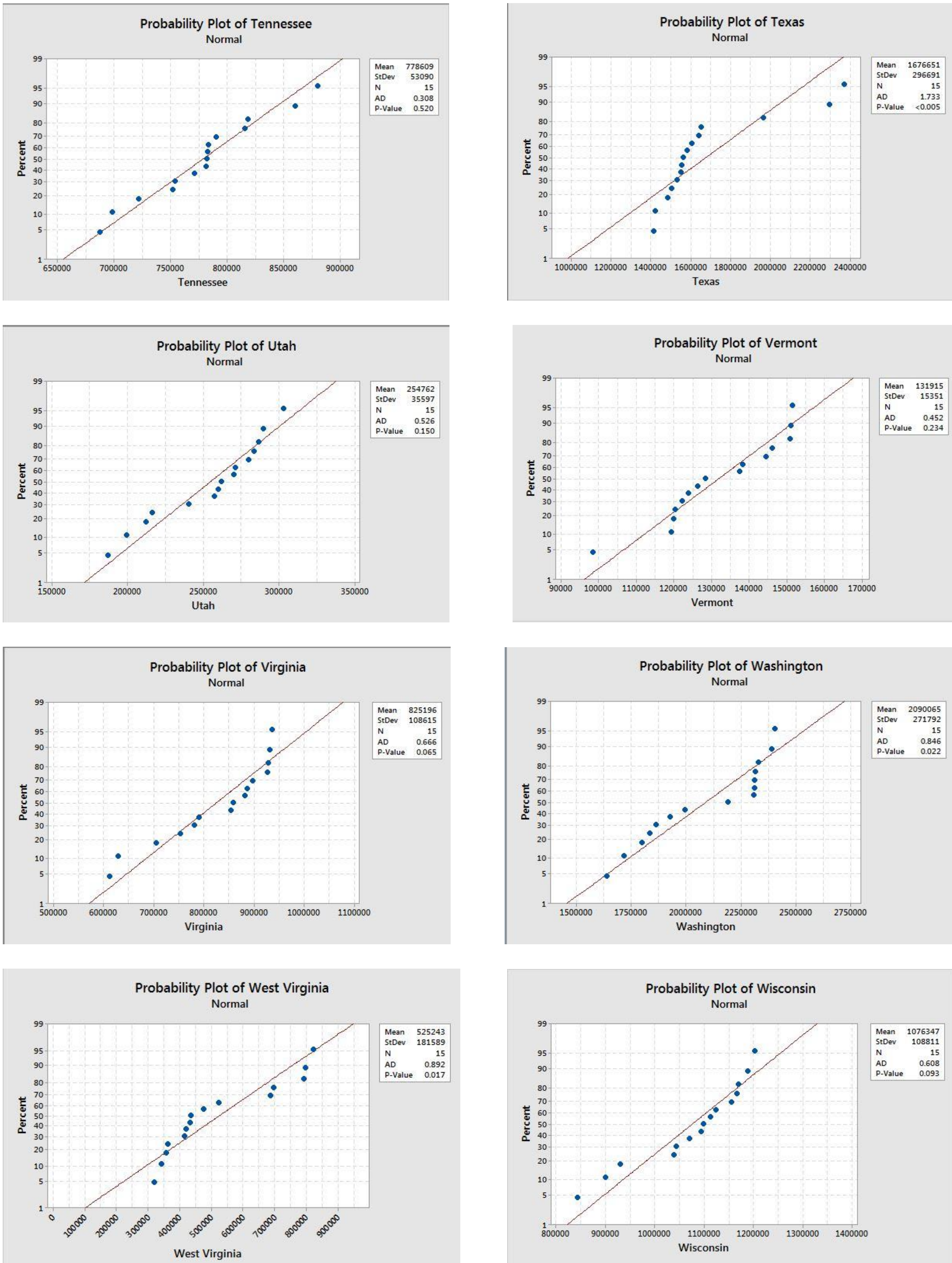


Figure 3.7. Normality Plots of Total Benefits (Continued)

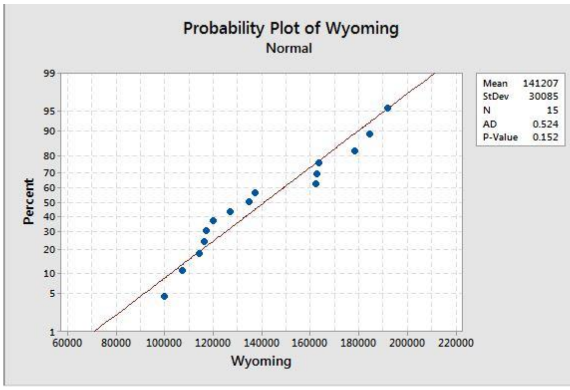


Figure 3.7. Normality Plots of Total Benefits (Continued)

Table 3.4 below shows the list of p-values associated with the normality test on the total benefits paid. It can be seen that most U.S. states have the p-value greater than 0.05 which means that the employment data for most of the U.S. states is normally distributed. However, there is a large number of state whose data is not normally distributed. There are a total of sixteen states whose data for total benefits paid is not normally distributed.

Table 3.4. List of Normality Test p-values of Total Benefits

S. No.	States	p-value
1	**Alabama**	0.07
2	**Alaska**	0.7
3	**Arizona**	0.008
4	**Arkansas**	0.77
5	**California**	0.1
6	**Colorado**	0.01
7	**Connecticut**	0.12
8	**Delaware**	0.04
9	**District of Columbia**	0.57
10	**Florida**	0.72
11	**Georgia**	0.03
12	**Hawaii**	0.009
13	**Idaho**	0.06
14	**Illinois**	0.34
15	**Indiana**	0.52
16	**Iowa**	0.25
17	**Kansas**	0.72
18	**Kentucky**	0.53
19	**Louisiana**	0.27

Table 3.4. List of Normality Test p-values of Total Benefits (Continued)

S. No.	States	p-value
20	**Maine**	0.01
21	**Maryland**	0.09
22	**Massachusetts**	0.79
23	**Michigan**	0.009
24	**Minnesota**	0.14
25	**Mississippi**	0.01
26	**Missouri**	0.01
27	**Montana**	0.04
28	**Nebraska**	0.39
29	**Nevada**	0.42
30	**New Hampshire**	0.58
31	**New Jersey**	0.46
32	**New Mexico**	0.3
33	**New York**	0.05
34	**North Carolina**	0.04
35	**North Dakota**	0.02
36	**Ohio**	0.18
37	**Oklahoma**	0.32
38	**Oregon**	0.35
39	**Pennsylvania**	0.03
40	**Rhode Island**	0.87
41	**South Carolina**	0.005
42	**South Dakota**	0.6
43	**Tennessee**	0.52
44	**Texas**	0.005
45	**Utah**	0.15
46	**Vermont**	0.23
47	**Virginia**	0.06
48	**Washington**	0.02
49	**West Virginia**	0.01
50	**Wisconsin**	0.09
51	**Wyoming**	0.15

3.5.5. Hypothesis Test

Table 3.5. SAS Output of Hypothesis Test (I)

Source	F Value	Pr > F
Model	386.18	<.0001

This is the p-value for the overall model. The null hypothesis associated with the overall F-test is that the fit of the intercept-only model (i.e. model with no predictors) and this model (i.e. model with the predictors) are equal. Since the p-value is less than the alpha reference value of 0.05, we reject the null hypothesis and conclude the model with the predictors is a better fit.

Table 3.6. SAS Output of Hypothesis Test (II)

R-Square	Total Paid Mean
0.983286	1062247

The R-Squared indicates approximately 98% of the variation of the Total Benefits paid can be explained by the model.

Table 3.7. SAS Output of Hypothesis Test (III)

Source	F Value	Pr > F
AvgEmployed	1.04	0.3081
State	3.68	<.0001
AvgEmployed*State	4.24	<.0001

The last column provides the p-values for each of the predictors in the model. State and the interaction between average employed and state are both statistically significant. Average employed on its own is not statistically significant, however, since the interaction is significant we need to leave it in the model. If the average employed and the interaction of average employed with state i.e. AvgEmployed*State were not significant then it can be taken out of the model and the equation would run without the coefficient (β values) of average employed and avgemployed*state.

The significant p-value associated with the interaction between the variables 'AvgEmployed' and 'State' indicates the slopes are not the same, and thus, the interaction variable needs to be included in the model. If this interaction was not significant, we could assume the total amount paid for each state increases and decreases as the same rate, and thus the regression lines for each state would be parallel. Looking at the covariance plot in figure 4.4, it can be seen that is not the case.

3.5.6. Box Plots on the Basis of Number of Workers

There are 710 observation within the box limit as shown in figure 3.8. Most of the values lies between the boundaries of first and third quartile. Only 55 observations are outliers starting from 6,840,000 workers in Florida in 2003 to 16,051,000 workers in California in 2015. The outlier states are California, New York, Texas and Florida.

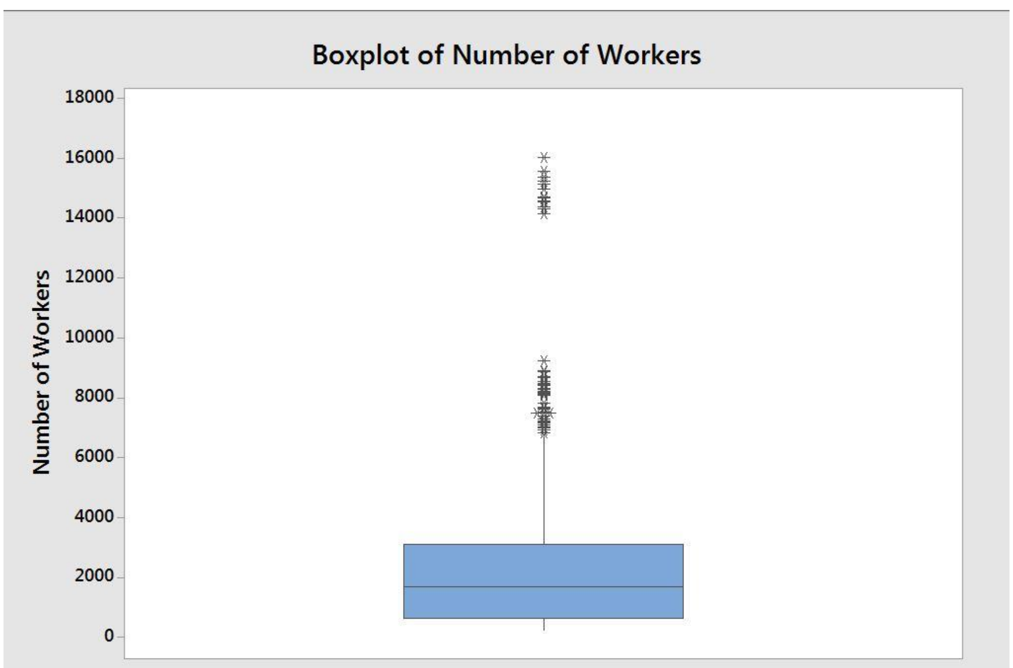


Figure 3.8. Box Plot of Number of Covered Workers

3.5.7. Box Plots on the Basis of Total Benefits

There are 690 observation within the box limit. Most of the values lies between the boundaries of first and third quartile as shown in figure 3.9. Only 75 observations are outliers starting from $2,406,272,000 paid in Pennsylvania in 2001 to the maximum of $12,459,589,000 paid to California workers in 2004. The outlier states are California, New York, Florida, Pennsylvania and Illinois.

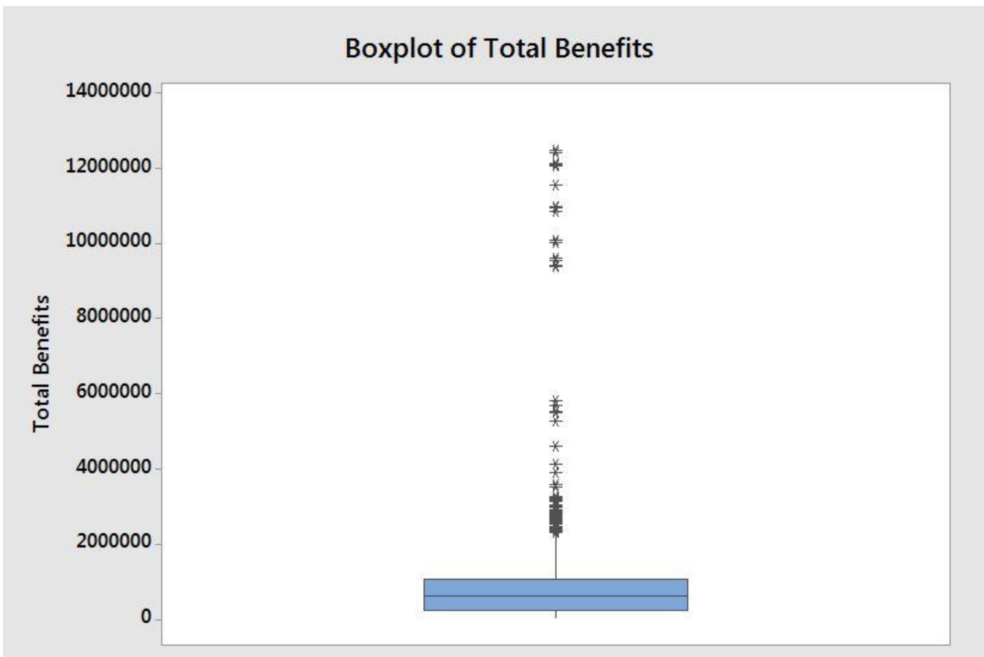


Figure 3.9. Box Plot of Total Benefits Paid

3.5.8. Trend Analysis

The series of figures below are representing the time series plots average number of employees and the total benefits paid for 15 years in each state. As it can be seen in the graphs below that almost all the states had a depression during the financial crisis in number of employees between the years 2009 – 2011. However the depression in the amount of benefits paid is not proportional to the average employment. The blue solid line in the figure represents the amount of total benefits paid from the year 2001 – 2015. There is a graph present for all 50 states of the United States and D. C. The red dotted line represents the amount of total benefits paid from the year 2001 – 2015. Here the term total benefits is used because benefits can be cash benefits and medical benefits. There is a graph present for all 50 states of the United States and D.C. In the graphs below both the red and the blue line are shown for every state in the same graph. Since there are 51 states and if we were to develop a graph for both benefits payment and employment, it would become 102 figures. Total benefits paid and number of employment are merged together to be able to comprehend easily.

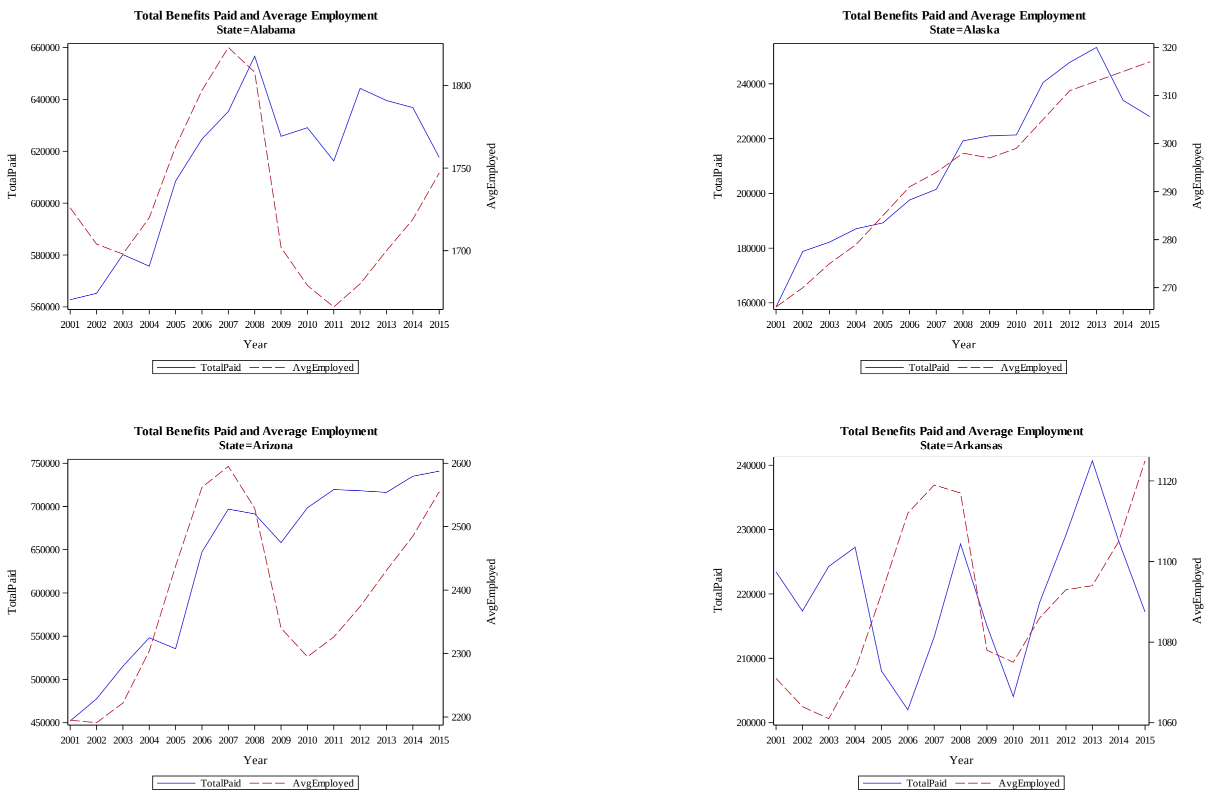


Figure 3.10. Trend Analysis of Covered Workers and Total Benefits by Year

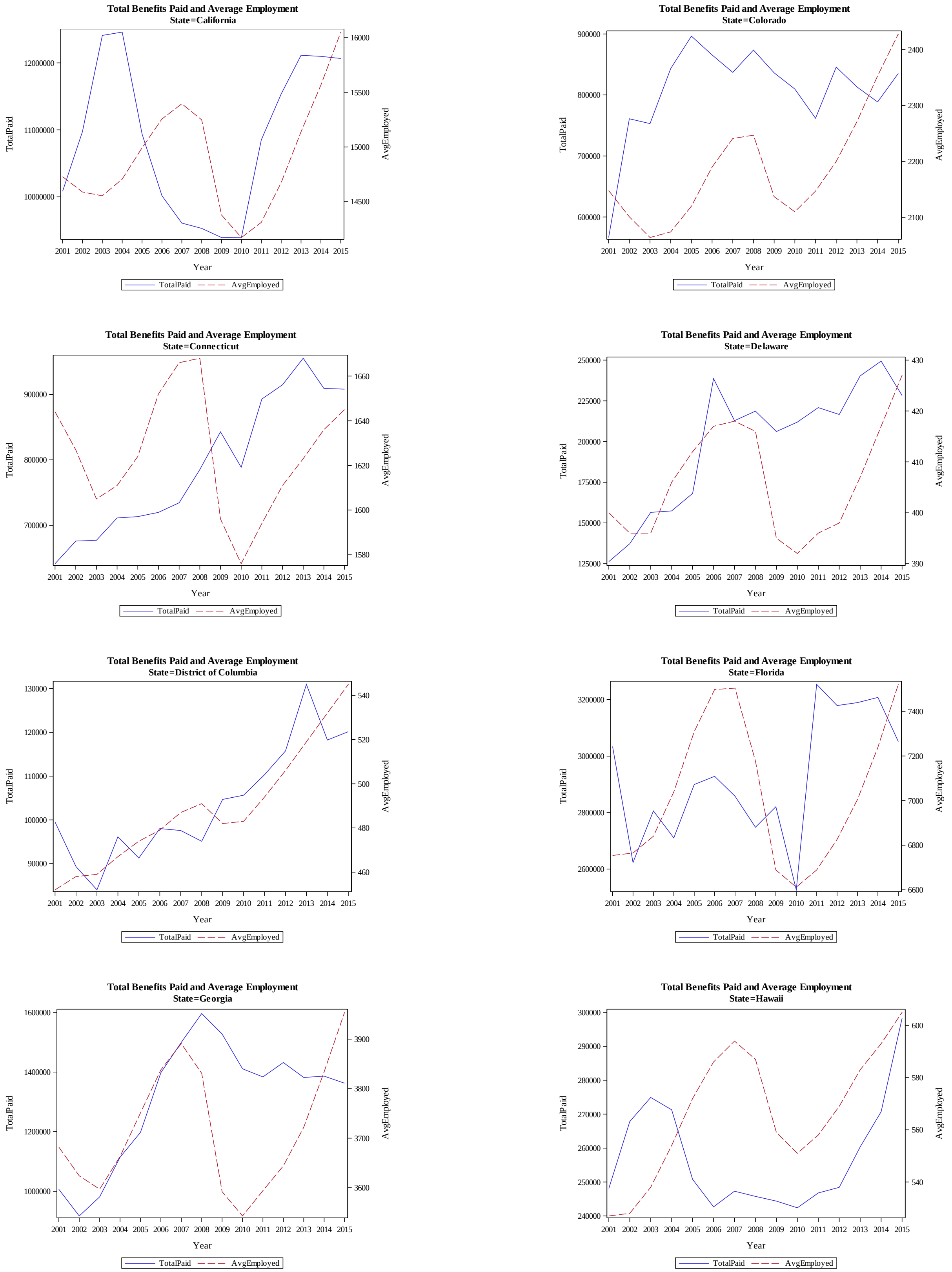


Figure 3.10. Trend Analysis of Covered Workers and Total Benefits by Year (Continued)

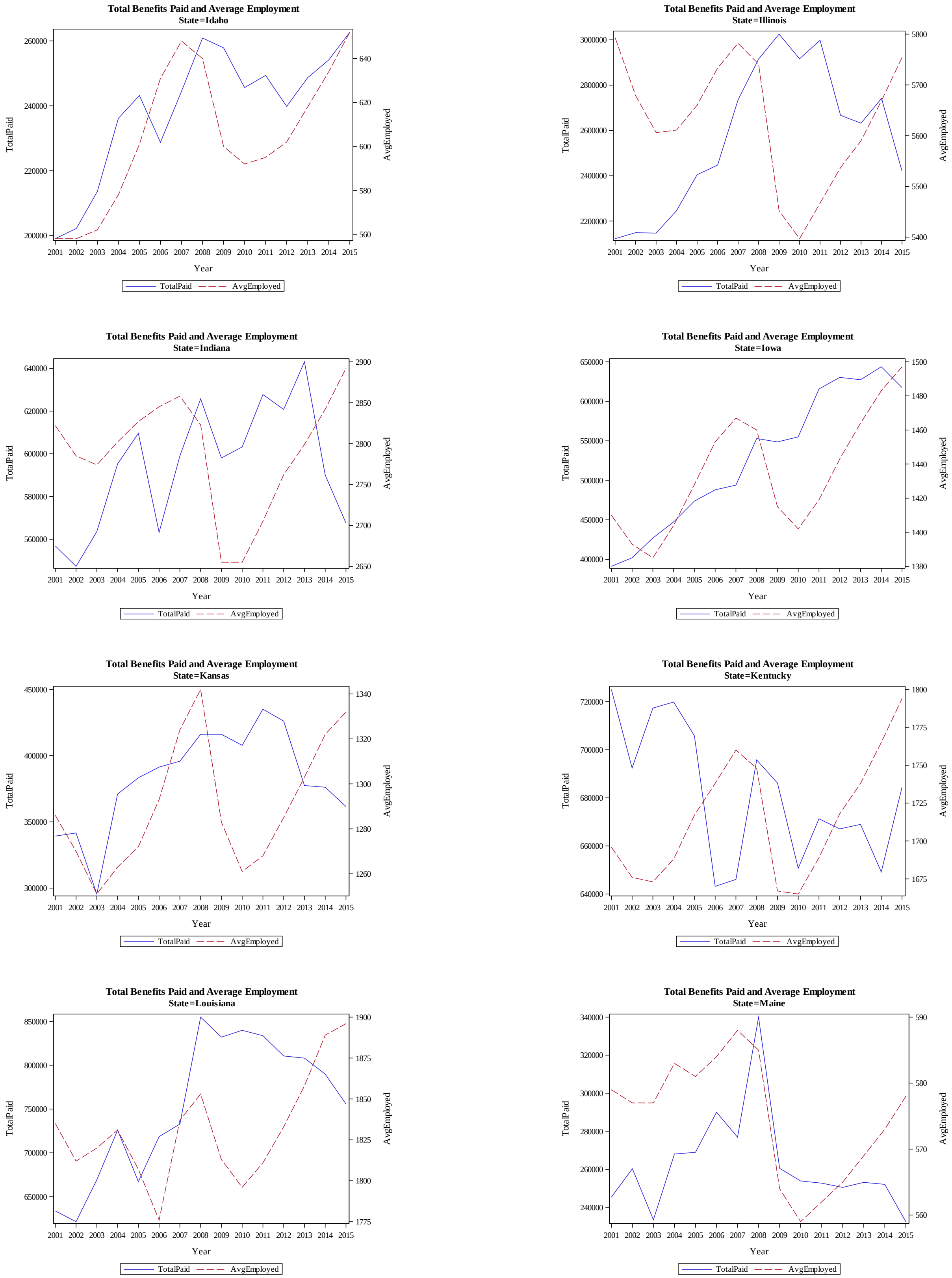


Figure 3.10. Trend Analysis of Covered Workers and Total Benefits by Year (Continued)

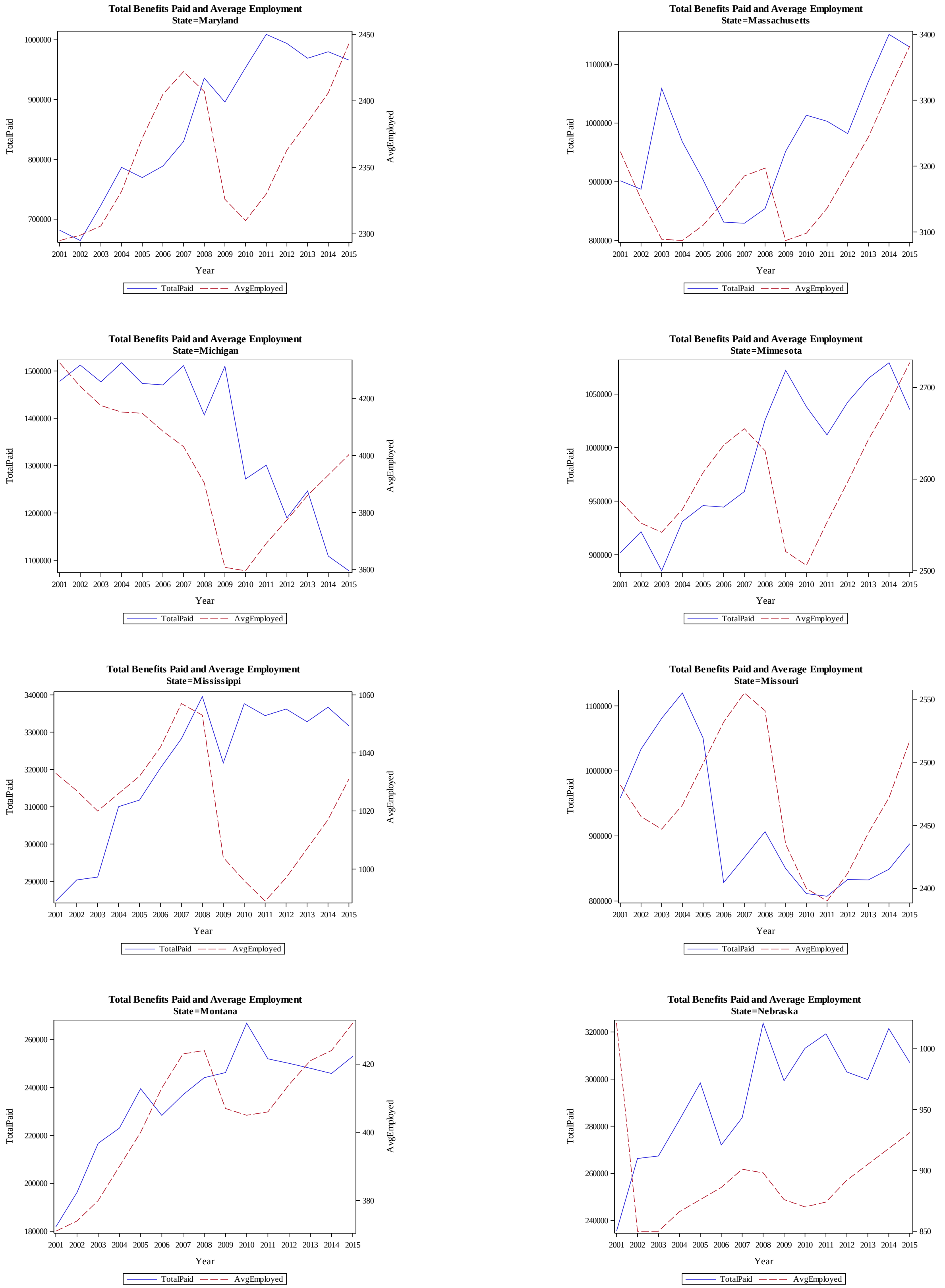


Figure 3.10. Trend Analysis of Covered Workers and Total Benefits by Year (Continued)

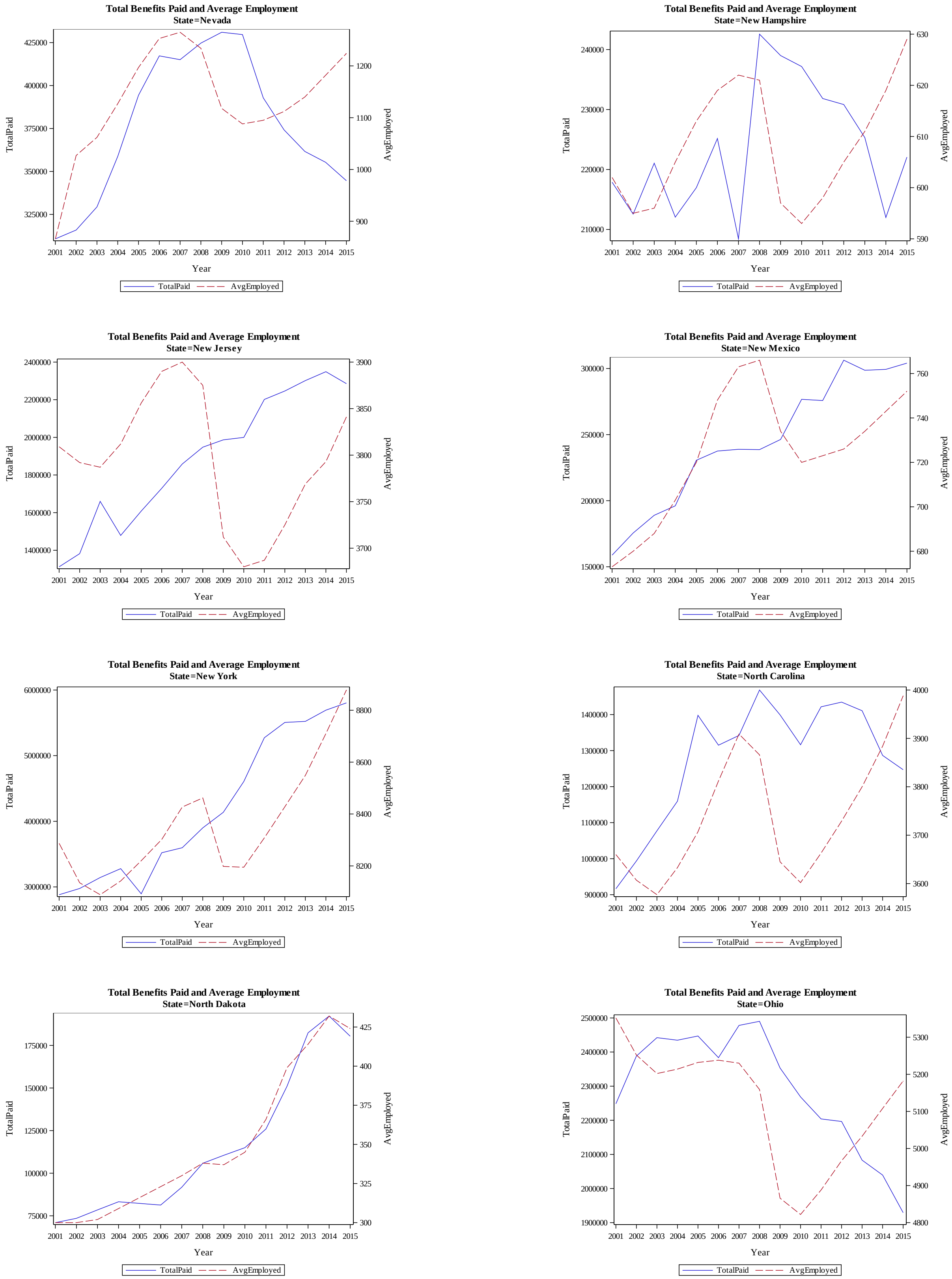


Figure 3.10. Trend Analysis of Covered Workers and Total Benefits by Year (Continued)

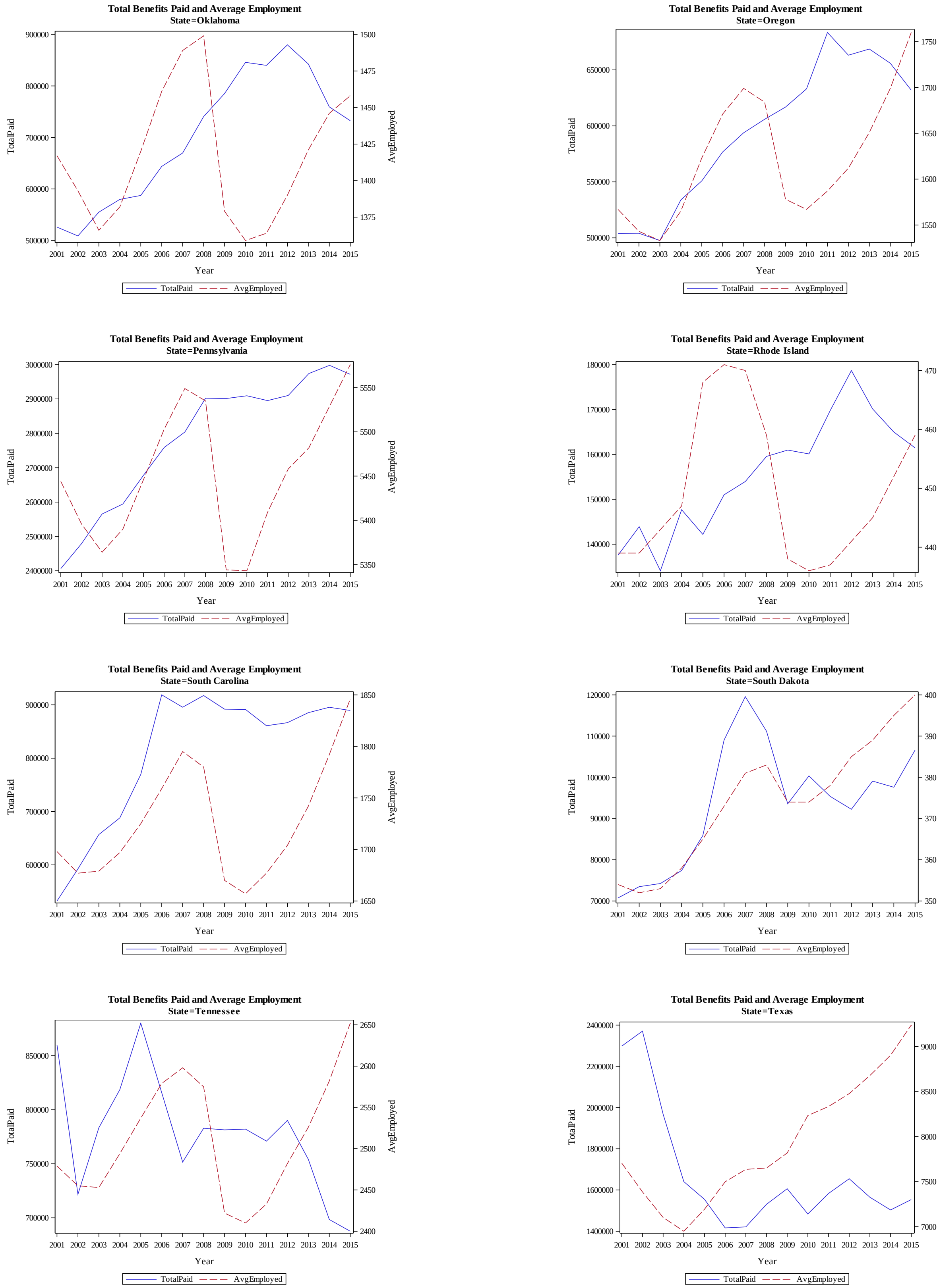
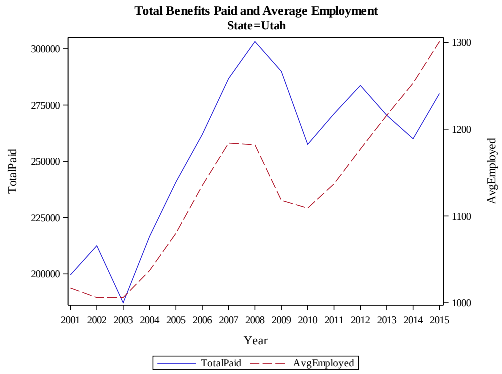
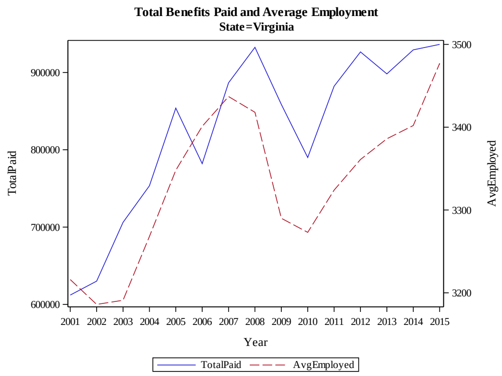
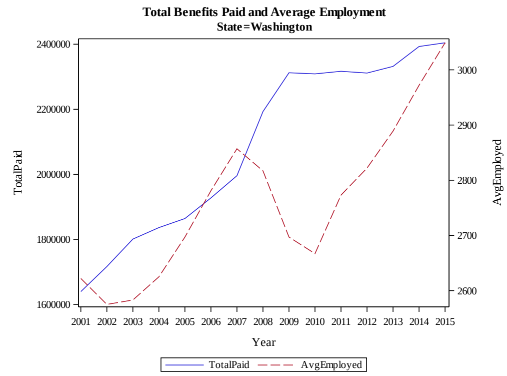
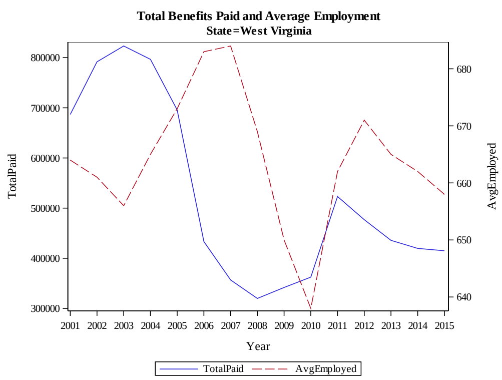


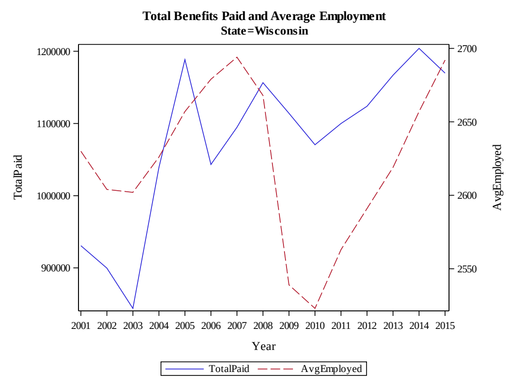
Figure 3.10. Trend Analysis of Covered Workers and Total Benefits by Year (Continued)











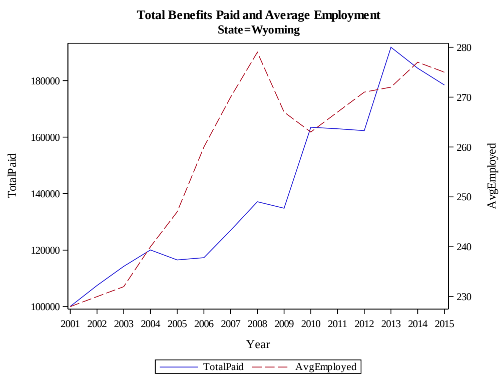


Figure 3.10. Trend Analysis of Covered Workers and Total Benefits by Year (Continued)

3.5.9. Correlation Analysis

In the exploratory data analysis phase, correlation between the states and correlation between years has been done. Following table shows the correlation by state. Here it can be noted that the correlation coefficient or R^2 values of correlation by state are not significant. They are positive as well as negative and they vary from a high of 0.99 (which shows strong positive correlation) to the low of -0.01(which shows no correlation at all). The reason for these varying figures will be discussed in the next chapter.

Table 3.8. List of Pearson's Correlation Coefficient as a Result of Correlation by State

States	Pearson's Correlation Coefficient
Alabama	0.25
Alaska	0.94
Arizona	0.67
Arkansas	-0.07
California	0.24
Colorado	0.18
Connecticut	-0.18
Delaware	0.44
District of Columbia	0.84
Florida	0.16
Georgia	0.35
Hawaii	0.15
Idaho	0.77
Illinois	-0.52
Indiana	-0.28
Iowa	0.65
Kansas	0.23
Kentucky	-0.38
Louisiana	0.19
Maine	0.46
Maryland	0.45
Massachusetts	0.46
Michigan	0.43
Minnesota	0.26
Mississippi	-0.23
Missouri	0.16
Montana	0.79

Table 3.8. List of Pearson's Correlation Coefficient as a Result of Correlation by State (Continued)

States	Pearson's Correlation Coefficient
Nebraska	-0.31
Nevada	0.60
New Hampshire	-0.21
New Jersey	-0.33
New Mexico	0.67
New York	0.74
North Carolina	0.38
North Dakota	0.99
Ohio	0.28
Oklahoma	-0.07
Oregon	0.48
Pennsylvania	0.38
Rhode Island	-0.15
South Carolina	0.46
South Dakota	0.77
Tennessee	-0.30
Texas	-0.32
Utah	0.79
Vermont	-0.09
Virginia	0.85
Washington	0.73
West Virginia	-0.01
Wisconsin	0.23
Wyoming	0.78

The next table represents the correlation by year for a particular state. The value of correlation coefficient in this table are very significant which shows the data is very much correlated by year.

Table 3.9. List of Pearson's Correlation Coefficient as a Result of Correlation by Year

Year	Pearson's Correlation Coefficient
2001	0.93
2002	0.92
2003	0.91
2004	0.90
2005	0.92
2006	0.93
2007	0.94
2008	0.94
2009	0.94
2010	0.92
2011	0.92
2012	0.92
2013	0.91
2014	0.91
2015	0.91

3.6. Statistical Analysis Approach

After the development of research question, gathering of data and performing the exploratory data analysis, the next step is to analyze the data for predictions of future values of workers' compensation. This is mainly done in SAS software. Multiple regression is done but the approach used in SAS is ANCOVA. The reason ANCOVA is used in the model development rather than conventional regression model is because if we make the model using conventional regression methods, there would be a lot of dummy variables incorporated in each calculation. There are 51 observations of states and 15 observations of years.

They combine to form a lot of dummy variable which in real do not have any impact on the calculation since their value would be zero. For example if we calculate the amount of total benefits paid in 2001 in North Dakota state, we only have to put the β value of North Dakota rather than including each and every state in the equation and eliminating them in the next step. The equation developed for the model is as follows:

$$Total\ Paid = Intercept + \ \beta_1 AvgEmployed + \beta_2 State + \ \beta_3 AvgEmployed * State$$

3.6.1. Analysis of Covariance (ANCOVA)

Analysis of covariance (ANCOVA) is one of the mainly used statistical method which is used to analyze the quantitative data acquired from experimental studies. It is mainly used in the field of psychology and education. The basis of ANCOVA is on assumptions, however this is a linear model therefore assumptions are straight forward (Leppink 2018).

The reason analysis of covariance is performed on this data is because it has an adding numerical value in the data. Average number of employees in every state represents a numerical value. If the data did not contain any numerical value we would perform analysis of variance (ANOVA) on the data. There is a numerical value designated to the states and the interaction of average employment and state (AvgEmployed*State). These coefficients are further used in the model equation to give the predicting values for every U.S. state from the year 2001-2015.

3.6.2. ANCOVA Assumptions

Leppink (2018) has given an elaborate overview of ANCOVA and the assumptions used in the process. The assumptions are as follows which are quoted from the journal article of Leppink (2018).

1. “The residuals are assumed to be independent
2. The residuals are assumed to have a mean of zero regardless of the grouping variable or the level of the covariate.
3. The residuals are assumed to be normally distributed.
4. The variance of the residual is the same regardless of the grouping variable or covariate.
5. The grouping variable and covariate are assumed to be fixed and measured without error.

6. The slope of the linear relation between response variable and covariate is assumed to be the same across groups."

4. WORKERS' COMPENSATION PREDICTION MODEL DEVELOPMENT

4.1. Introduction

Multiple regression can either be used for the explanation of the data or prediction. Here in this chapter multiple regression is used as a prediction tool. When multiple regression is used as predictor, there is an equation created using the sample. The equation then detects the phenomenon of the particular data and make the predictions. The predictors, which have to be entered into the equation, have to rely on their statistical properties (Osborne 2000). In this case the predictors are average number of employees, states, and the interaction of number of employees and states.

4.2. Model Development

The following table 4.1 provides the parameter estimates for the 50 states and D.C. Here we have a different model for every state. An example of how to calculate the predicted mean for North Dakota is presented.

The model is as follows:

$$Total\ Paid = Intercept + \ \beta_1 AvgEmployed + \beta_2 + \ \beta_3 * AvgEmployed$$

The intercept and parameter estimate for Avg Employed will be the same for all states. The parameter estimates for State and AvgEmployed*State will vary state to state. Using the above model, the predicted average amount paid in benefits in 2015 for North Dakota is listed below. The average number of employed in ND in 2015 was 424.

Total Paid = 1124596.42 – 258.27(424) – 1322832.80 + 1155.77(424)

Predicted Total Paid = $182,303.62

The actual total paid for ND for 2015 was $180,401.

Therefore, the residual for North Dakota for the year 2015 = 182,303 – 180,401 = $1,902.

There must be a state nominated as a reference state while developing the model in SAS. Kentucky was chosen as the reference state; therefore, there are no coefficients associated with Kentucky in table 4.1. The reason behind choosing Kentucky as a reference state is because the mean number of employees for Kentucky is the median value of all the states combined. It doesn't affect the calculation much. If another state was chosen as the reference, the output predicted values would be the same but the estimate values would be different.

Table 4.1 contains four columns. The first column is parameter, second is the symbol associated with that parameter. The third column is the estimate values. These are the values that go into the model equation for calculating the total benefit for a certain year. The fourth column is the t-values. These values are of not much use in the model itself but they tell how closely related a certain state is with our reference state, Kentucky.

Table 4.1. Calculation Matrix for Prediction Model

Parameter		Coefficient	Estimate	t-value
Intercept		β	1124596.42	0.6658
AvgEmployed		β_1	-258.27	0.8648
State	**Alabama**	β_2	-772940.46	0.8195
State	**Alaska**	β_2	-1385542.83	0.6236
State	**Arizona**	β_2	-1670895.28	0.5531
State	**Arkansas**	β_2	-868093.06	0.8336
State	**California**	β_2	1696826.13	0.5886
State	**Colorado**	β_2	-618509.68	0.8315
State	**Connecticut**	β_2	798998.16	0.8597
State	**Delaware**	β_2	-1573705.23	0.6487
State	**District of Columbia**	β_2	-1218626.32	0.6665
State	**Florida**	β_2	1027682.63	0.7254
State	**Georgia**	β_2	-1966718.24	0.5348
State	**Hawaii**	β_2	-922801.34	0.7555

Table 4.1. Calculation Matrix for Prediction Model (Continued)

	Parameter	Coefficient	Estimate	t-value
State	**Idaho**	β_2	-1177005.74	0.6797
State	**Illinois**	β_2	8896594.31	0.0196
State	**Indiana**	β_2	-205841.24	0.9539
State	**Iowa**	β_2	-2967919.34	0.4151
State	**Kansas**	β_2	-1135564.22	0.7661
State	**Louisiana**	β_2	-1215265.36	0.7781
State	**Maine**	β_2	-1626764.53	0.7264
State	**Maryland**	β_2	-2853126.55	0.4668
State	**Massachusetts**	β_2	-1894556.13	0.5807
State	**Michigan**	β_2	-933360.16	0.7401
State	**Minnesota**	β_2	-845028.08	0.8165
State	**Mississippi**	β_2	-588981.27	0.8832
State	**Missouri**	β_2	-1013314.26	0.7971
State	**Montana**	β_2	-1270554.85	0.6621
State	**Nebraska**	β_2	-669968.18	0.8181
State	**Nevada**	β_2	-1019024.43	0.7048
State	**New Hampshire**	β_2	-781830.13	0.8497
State	**New Jersey**	β_2	7010689.04	0.0945
State	**New Mexico**	β_2	-1703379.06	0.5740
State	**New York**	β_2	-27753704.46	<.0001
State	**North Carolina**	β_2	-1734450.90	0.5821
State	**North Dakota**	β_2	-1322832.80	0.6169
State	**Ohio**	β_2	-356036.52	0.9125
State	**Oklahoma**	β_2	-142035.08	0.9649
State	**Oregon**	β_2	-1259511.92	0.6732
State	**Pennsylvania**	β_2	-3520129.30	0.4895
State	**Rhode Island**	β_2	-898097.02	0.7912
State	**South Carolina**	β_2	-2123486.99	0.5077
State	**South Dakota**	β_2	-1311765.44	0.6624

Table 4.1. Calculation Matrix for Prediction Model (Continued)

Parameter		Coefficient	Estimate	t-value
State	**Tennessee**	β_2	207711.64	0.9505
State	**Texas**	β_2	1642921.89	0.5425
State	**Utah**	β_2	-1221163.71	0.6528
State	**Vermont**	β_2	-906767.62	0.8485
State	**Virginia**	β_2	-3661502.13	0.2857
State	**Washington**	β_2	-2887689.66	0.3135
State	**West Virginia**	β_2	-505498.33	0.9075
State	**Wisconsin**	β_2	-1290991.76	0.7457
State	**Wyoming**	β_2	-1316367.61	0.6318
State	**Kentucky**	β_2	0.00	.
AvgEmployed*State	**Alabama**	β_3	410.34	0.8348
AvgEmployed*State	**Alaska**	β_3	1860.58	0.6420
AvgEmployed*State	**Arizona**	β_3	753.36	0.6340
AvgEmployed*State	**Arkansas**	β_3	224.59	0.9459
AvgEmployed*State	**California**	β_3	799.72	0.5993
AvgEmployed*State	**Colorado**	β_3	395.19	0.8082
AvgEmployed*State	**Connecticut**	β_3	-438.43	0.8726
AvgEmployed*State	**Delaware**	β_3	1854.48	0.7487
AvgEmployed*State	**District of Columbia**	β_3	663.17	0.8071
AvgEmployed*State	**Florida**	β_3	367.80	0.8099
AvgEmployed*State	**Georgia**	β_3	836.20	0.5997
AvgEmployed*State	**Hawaii**	β_3	356.23	0.9029
AvgEmployed*State	**Idaho**	β_3	740.41	0.7620
AvgEmployed*State	**Illinois**	β_3	-1064.68	0.5046
AvgEmployed*State	**Indiana**	β_3	141.88	0.9354
AvgEmployed*State	**Iowa**	β_3	1910.63	0.4132
AvgEmployed*State	**Kansas**	β_3	563.04	0.8312
AvgEmployed*State	**Louisiana**	β_3	718.94	0.7658
AvgEmployed*State	**Maine**	β_3	1588.36	0.8171

Table 4.1. Calculation Matrix for Prediction Model (Continued)

Parameter	Coefficient	Estimate	t-value
AvgEmployed*State Maryland	β_3	1356.45	0.4891
AvgEmployed*State Massachusetts	β_3	805.85	0.6299
AvgEmployed*State Michigan	β_3	555.36	0.7185
AvgEmployed*State Minnesota	β_3	531.76	0.7684
AvgEmployed*State Mississippi	β_3	47.79	0.9886
AvgEmployed*State Missouri	β_3	583.45	0.7628
AvgEmployed*State Montana	β_3	1198.13	0.7342
AvgEmployed*State Nebraska	β_3	77.33	0.9707
AvgEmployed*State Nevada	β_3	497.94	0.7599
AvgEmployed*State New Hampshire	β_3	62.48	0.9909
AvgEmployed*State New Jersey	β_3	-1387.35	0.4270
AvgEmployed*State New Mexico	β_3	1394.49	0.5945
AvgEmployed*State New York	β_3	3945.90	0.0107
AvgEmployed*State North Carolina	β_3	763.50	0.6311
AvgEmployed*State North Dakota	β_3	1155.77	0.5643
AvgEmployed*State Ohio	β_3	556.11	0.7221
AvgEmployed*State Oklahoma	β_3	58.59	0.9769
AvgEmployed*State Oregon	β_3	707.91	0.6883
AvgEmployed*State Pennsylvania	β_3	1208.28	0.4816
AvgEmployed*State Rhode Island	β_3	100.86	0.9841
AvgEmployed*State South Carolina	β_3	1306.04	0.4835
AvgEmployed*State South Dakota	β_3	1008.87	0.8137
AvgEmployed*State Tennessee	β_3	37.83	0.9826
AvgEmployed*State Texas	β_3	120.54	0.9368
AvgEmployed*State Utah	β_3	569.03	0.7319
AvgEmployed*State Vermont	β_3	-36.29	0.9979
AvgEmployed*State Virginia	β_3	1267.34	0.4448
AvgEmployed*State Washington	β_3	1653.46	0.2948
AvgEmployed*State West Virginia	β_3	116.93	0.9829

Table 4.1. Calculation Matrix for Prediction Model (Continued)

Parameter	Coefficient	Estimate	t-value
AvgEmployed*State Wisconsin	β_3	732.07	0.7004
AvgEmployed*State Wyoming	β_3	1546.22	0.6764
AvgEmployed*State Kentucky	β_3	0.00	.

4.2. Discussions

The calculation shown in section 4.1 was carried out for each year and each state, so a total of 765 observations were generated. The residual was also calculated for each observation and values of residual are plotted on the graph shown in figure 4.1.

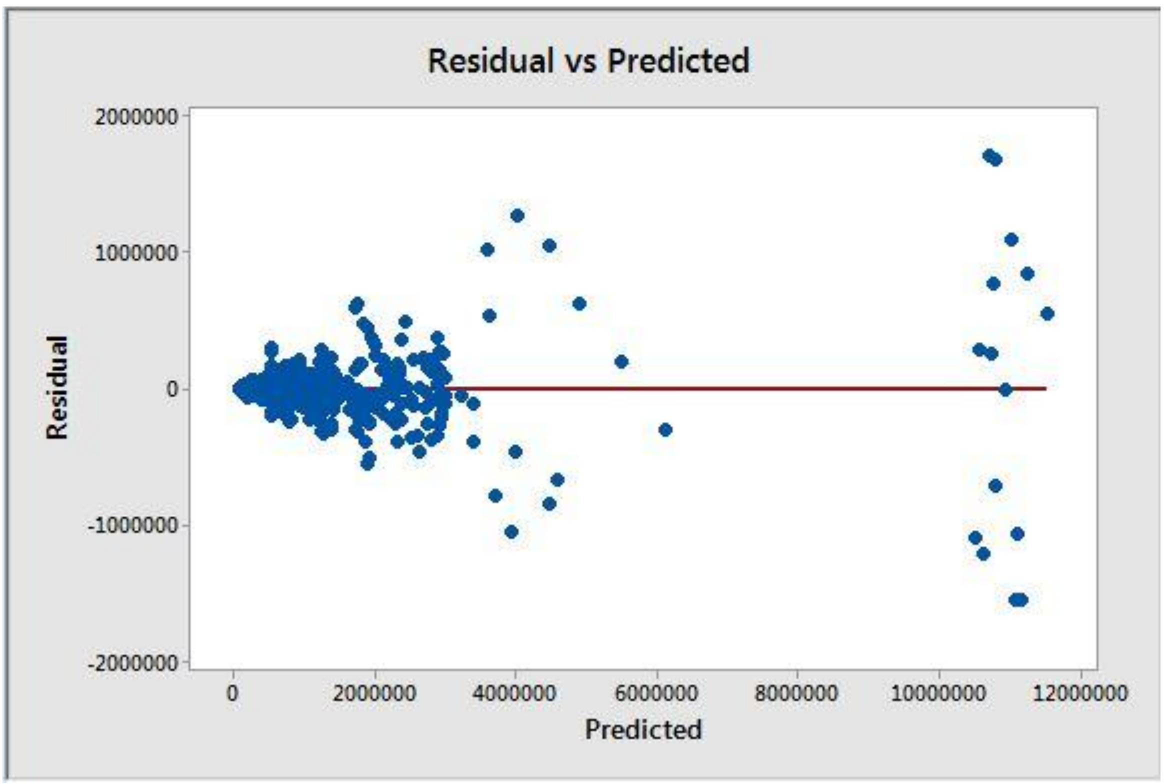


Figure 4.1. Residual vs Predicted Values

Figure 4.1 shows the diagnostic plots that SAS produces to assess the model fit. We want to see the residuals randomly scattered about zero. It can be seen in the top plot that the spread of the residuals increases as the total paid increases, so this is not necessarily a random scattered plot. There are some values toward the right hand side which are away from zero. These residuals have a very large value hence we can say that states which have a larger amount of workers do not exactly fit in this model. In this particular case all the values towards right belong to California. As it can be seen in figure 3.2, California has an extraordinarily large number of workers compared to other states. It has double the number of workers than its immediate successor, New York.

Below, in figures 4.2 and figure 4.3 the histogram of residuals and the normality test on the residuals is shown, respectively.

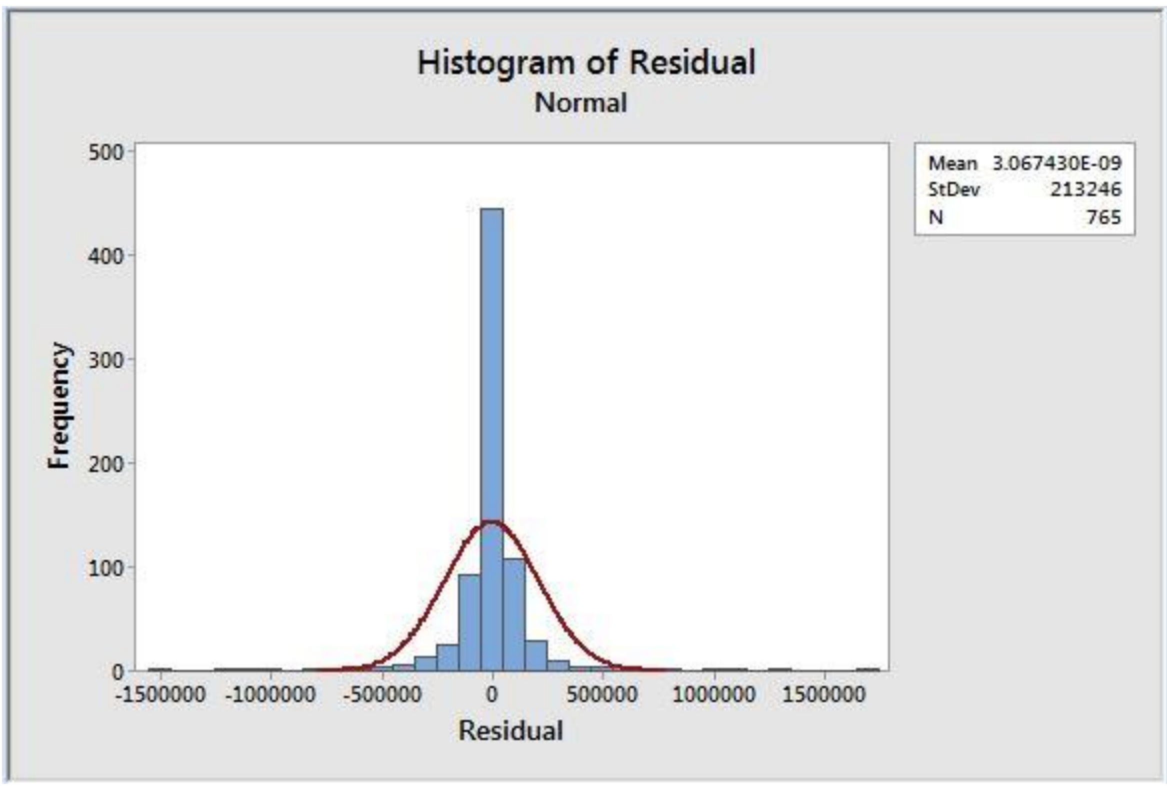


Figure 4.2. Histogram of Residuals

Figure 4.2 shows the histogram of the residual values. Most of the residuals are closer to zero meaning the predicted values are not deviating from the actual values too much. However, as discussed before, the states with an extraordinarily large number of covered workers can affect the calculation and the predicted values.

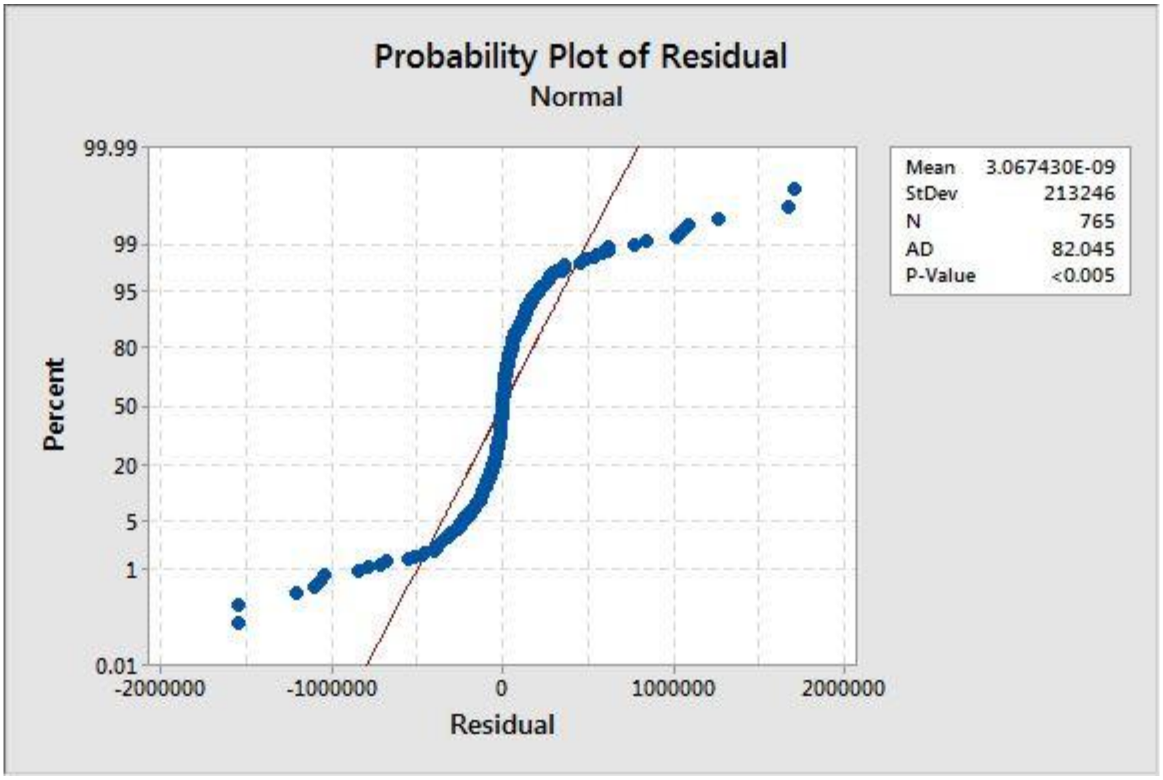


Figure 4.3. Residual Normality Test

In the plot of figure 4.3, we can assess the normality of the residuals. We want to see the residuals fall as closely on the diagonal line as possible. It can be noted that the residuals have somewhat of an 'S' pattern. The normality and equal variance assumptions are not met with this model.

Some assumptions are not met but since our model is robust i.e. we have large number of observations, we can override the fact that assumptions are not met. One way to improve the fit of the model is to transform the dependent variable (i.e. Total Paid), usually a log transformation is the best fit. The problem with this is that the interpretation of the results aren't exactly intuitive. Most often, the model will be developed on the log (Total Paid), and then after the model is created, the model will exponentiate the result so it is easier to interpret. This method was tried to see how it compares to the original model, and it was found that there was more variation using this method. It is recommended we leave the model as is, and not transform the total benefits paid.

The complete breakdown of the residual values is present in the Appendix B. The table shows the difference in the predicted value and the actual value. The predicted value was generated using the model discussed above. The predicted value, actual value, and the residual is shown for each state for all fifteen years.

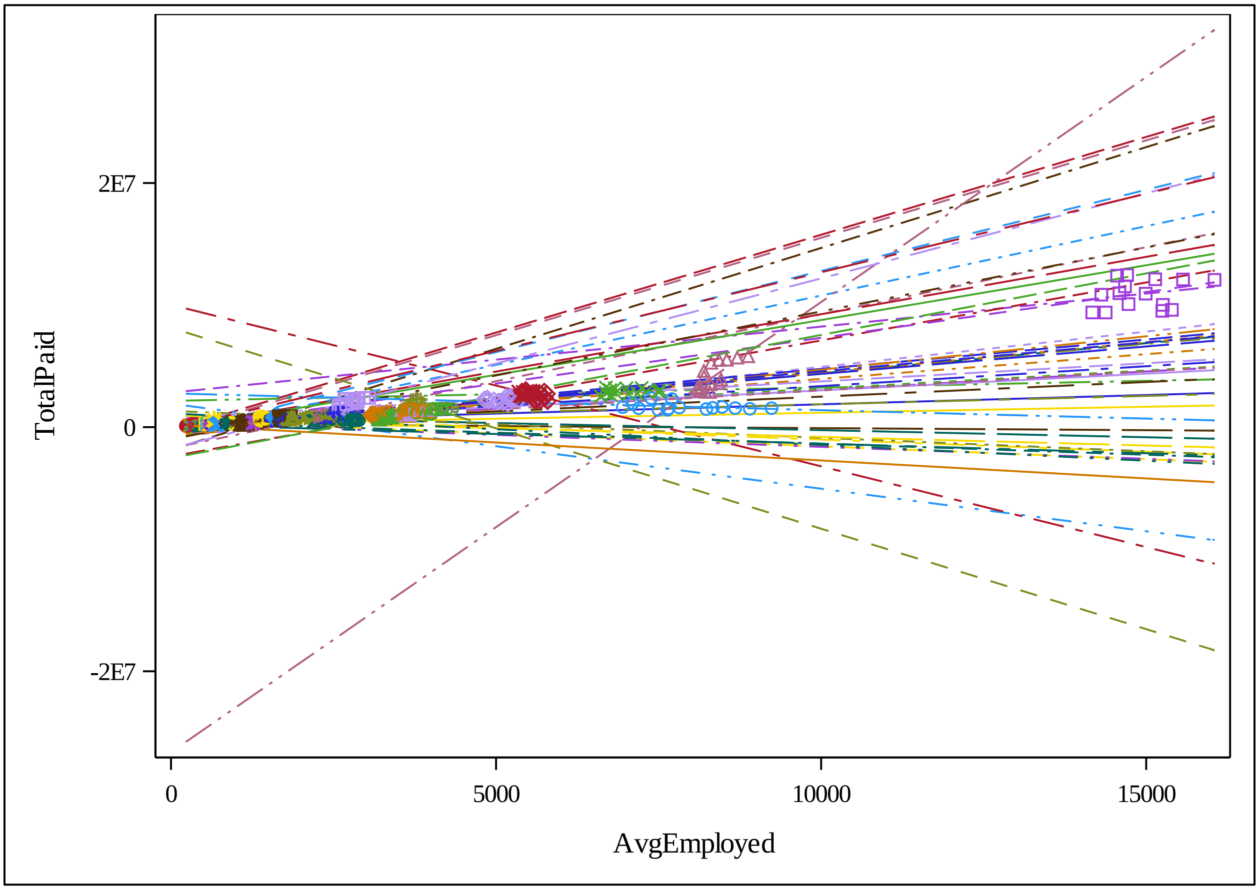


Figure 4.4. Analysis of Covariance for Total Paid

Here in figure 4.4, we can see that the slopes are not the same, and thus, the interaction variable needs to be included in the model. By interaction variable, AvgEmployed*State is meant as shown in table 4.1. If this interaction was not significant, we would assume the total amount paid for each state increases and decreases as the same rate, and thus the regression lines for each state would be parallel. Looking at the covariance plot, it can be seen that is not the case.

4.3. Comparison of States

The model generates a prediction value of total benefits paid for every state in all 15 years. So a total of 765 observations. These observations are then used to calculate the deviation of predicted values form the actual values. The states with higher number of workers have higher amount of error (residual) and the states with lower number of workers have lower error. After calculating the percentage difference of each and every observation, it is seen that performance of West Virginia is not very good in this model. Some of the prediction values of West Virginia have more than 30% difference from the actual value.

Most of the observations lie between the error range of -10% to +10%. Out of 765 observation, 522 observations have error from -10% to +10%. Alabama, Alaska, Arkansas, Indiana, Kentucky and New Hampshire are the states with the prediction values of all 15 years lying within the above mentioned range. Then there are some states with 14 years lying within range and then some states with 13 years and so on. Couple years of West Virginia are also present in this group. Most of the values of West Virginia are deviating more than ±10% of the actual value. The reason behind this is, when we look at the West Virginia raw data of total benefits paid, we see a lot of ups and downs in the data. Therefore, the difference between the actual values and predicted values varies too much. Figure 4.5 is the visual representation of the comparison of states and their number of years lying within ±10% range of error. The table of regression predicted values is present in appendix B which can be used for the calculation reference.

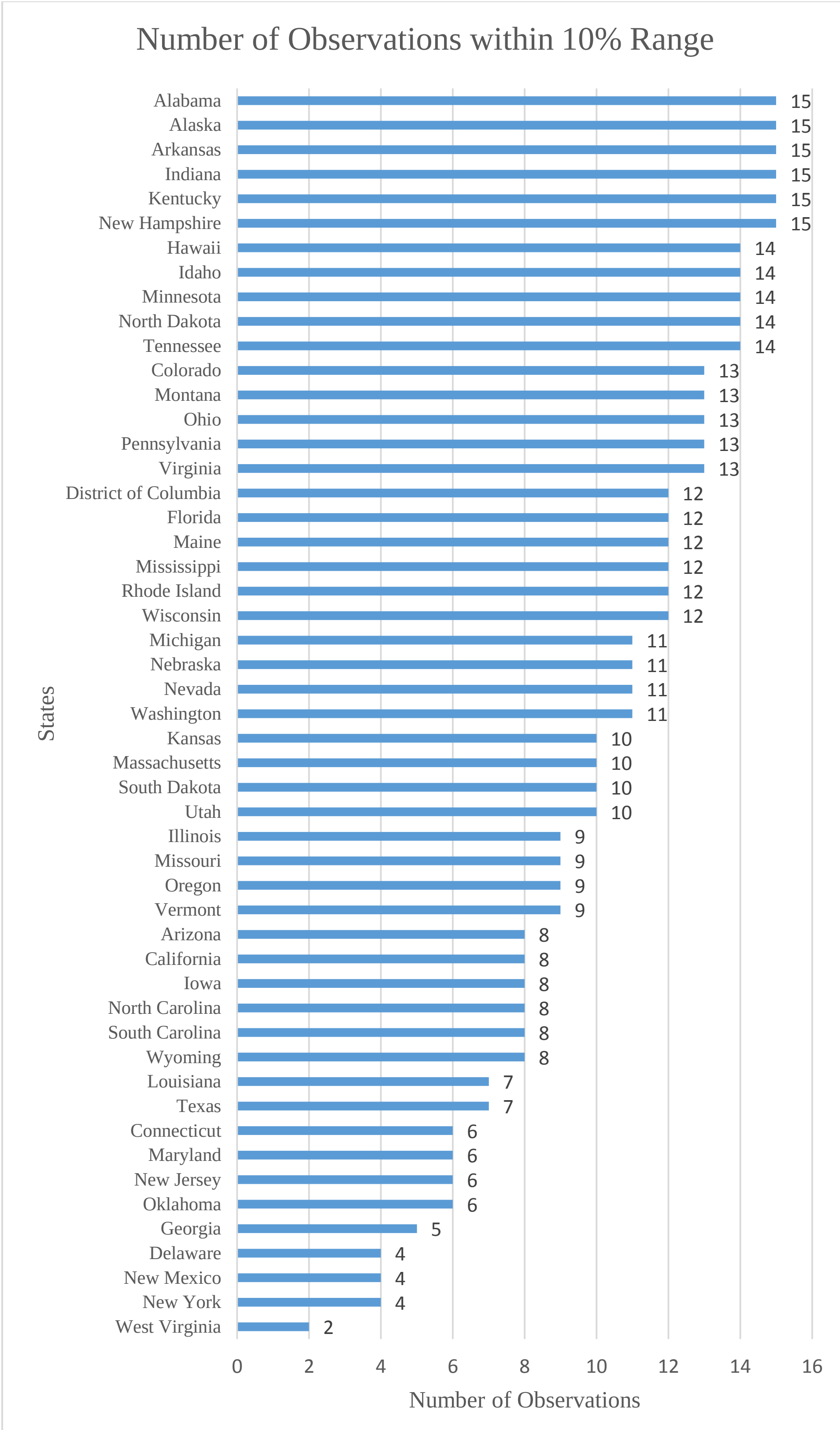


Figure 4.5. Number of Observation within ±10% Range

4.4. SAS Coding

SAS was used in the statistical analysis. Below is the coding which was used in SAS to achieve our final goal of the workers' compensation prediction values.

```
proc import datafile='Covered workers and Benefits paid.xlsx'
 out=Workers_Initial (rename=(B=State
Number_of_Covered_Workers__in_Th=Year2001  D=Year2002  E=Year2003
                                F=Year2004  G=Year2005  H=Year2006
I=Year2007  J=Year2008  K=Year2009
                                L=Year2010 M=Year2011  N=Year2012  O=Year2013
P=Year2014  Q=Year2015)
                    drop=A)
 dbms=xlsx
 replace;
 datarow=4;
 getnames=yes;
 sheet='Employment';
 run;

data workers;
 set Workers_Initial;
 if state ne '';
run;


proc sort data=workers;
 by state;
       run;
```

First of all, the raw data was imported into the SAS software. As it can be seen in the code above, the file named "Covered workers and Benefits paid" is the name of our raw data file in excel (.xlsx) format. It had two tabs. One for the number of covered workers and the other for total benefits paid.

After importing the raw data into SAS, some data management is done so the data could run easily according to SAS programs.

```
*** Univariate Data Set for Yearly Average Employment ***;

proc transpose data=workers out=workers_uv (rename=(COL1=AvgEmployed)) ;
 by state;
 var Year2001 Year2002 Year2003 Year2004 Year2005 Year2006 Year2007 Year2008
Year2009 Year2010 Year2011 Year2012 Year2013 Year2014 Year2015 ;
 run;


data workers_uv2;
 retain State Year  AvgEmployed;
 set workers_uv;
 Year=substr(_name_,5,4);
 drop _name_ _label_;
 run;


proc print data=workers_uv2 (obs=20);
 title 'Verify Yearly Employment Data';
run;




proc import datafile='Covered workers and Benefits paid.xlsx'
 out=Comp_Initial  (rename=(B=State
Amount_of_Total_Benefits_Paid__i=Year2001  D=Year2002  E=Year2003
                                F=Year2004  G=Year2005  H=Year2006
I=Year2007  J=Year2008  K=Year2009
                                L=Year2010 M=Year2011  N=Year2012  O=Year2013
P=Year2014  Q=Year2015)
                    drop=A)
 dbms=xlsx
 replace;
 datarow=4;
 getnames=yes;
 sheet='Compensation';
 run;

data Compensation;
 set Comp_Initial;
 if state ne '';
 run;


proc sort data=Compensation;
 by state;
 run;


*** Univariate Data Set for Yearly Total Benefits Paid ***;

proc transpose data=Compensation out=Compensation_uv
(rename=(COL1=TotalPaid)) ;
 by state;
```

```
 var Year2001 Year2002 Year2003 Year2004 Year2005 Year2006 Year2007 Year2008
Year2009 Year2010 Year2011 Year2012 Year2013 Year2014 Year2015 ;
 run;

data Compensation_uv2;
 retain State Year TotalPaid;
 set Compensation_uv;
 Year=substr(_name_,5,4);
 drop _name_ _label_;
 run;


proc print data=Compensation_uv2 (obs=20);
 title 'Verify Yearly Benefits Paid Data';
run;




proc sort data=workers_uv2;
 by state year;
 run;

proc sort data=compensation_uv2;
 by state year;
 run;


 *** Merge Employment and benefits data by state and year ***;

data combined;
 merge workers_uv2 compensation_uv2;
 by state year;
 label State='State';
 run;


proc print data=combined (obs=20);
 title 'Preview of Combined Data Set';
       run;
```

The codes mentioned above are all associated with the data management. The raw data was in multivariate form i.e. there was one state and there were 15 values written in front of it for each year. So, a total of 16 columns. SAS requires data to be univariate to run our desired regression program. The SAS code above first transposes the raw data of average employment and then it transposes the data of total benefits. In the end, it combines the two transposed data together.

```
** Create a data set to calculate the 15 year average for each state **;
proc means data=combined  noprint;
 by state;
 var avgemployed ;
 output out=means(drop= _type_ _freq_) Mean=;
 run;

** Find the Median employment value **;
proc univariate data=means noprint  ;
 var avgemployed;
 output out=median median=Avgemployed;
 run;

proc print data=median;
 run;

proc print data=means;
 where 1715 <= avgemployed <= 1716;
 title 'State with Median Avg Employment';
 run;
```

This code serves as identifying the reference state. This step is not necessary. If we do not identify the state our self, the software will itself pick the last state in the list. Initially, it picked Wyoming as a reference state but then this code was introduced in the program to pick Kentucky. This code basically takes the mean of all the values for 15 years and then sorts it from largest to smallest. It then pick the median value and identify the state associated with that median value.

```
*ods rtf file='Output\Regression Model for Benefits Paid.rtf';


proc glm data=combined plots=(diagnostics ) ;
 class state(ref='Kentucky') ;
 model totalpaid = AvgEmployed state state*AvgEmployed /  ss3 solution;
 output out=TotPred (drop=logtotalpd) predicted=Predicted  r=Residual;
 title 'Multiple Linear Regression Analysis';
run;

*ods rtf close;


/*
proc export data=TotPred
 dbms=xlsx
 outfile='S:\VPIT\Stats_Consulting\Arsalan Azmi\Output\Regression Predicted values.xlsx'
 replace;
 run;
*/
```

In the end, regression is carried out. This code gives the following results after running.

- Number of observations.
- P-value for overall model.
- R-Squared value.
- Model fitness.
- Parameter estimates
- List of predicted values and residuals.
- Fit diagnosis for total paid.
- Analysis of covariance of total paid.

5. CONCLUSION AND FUTURE WORK

5.1. Conclusion

Workers' compensation insurance (WCI) is the greatest expense of an accident because the insurance premium rises with each claim and the employer ends up paying a very large amount for the accident and the payments continue for a very long time. The WCI expense is part of the direct cost of an accident. Direct costs are usually hidden in variable cost portions of financial reports. There is a great loss of money resulting from an accident, especially in the form of insurance payments as they rise drastically. WCI is a complex and expensive part for any business, most often in terms of direct cost, which cannot be neglected. There is a need for better understanding the workers' compensation system and how payments are made over the years. This can help the regulatory bodies develop rules and policies of WCI accordingly.

Through a better understanding of the WCI system and its payments, the regulatory bodies can develop rules and policies of WCI accordingly. The policies of WCI have changed almost every year and vary state by state. There are some studies which talk about reducing the cost of insurance premiums and making the claim process faster and more efficient. Furthermore, the literature review told about the new techniques used in the WCI field and introduced new methods to have employees insured. There was a comparison of the workers' compensation system of the United States with other countries. Most developing and under developed countries either don't have the WCI system or they pay insufficiently after the employee gets injured or dies.

The availability of a prediction tool will ease the process and policy makers can make policy more accurately by obtaining the predicted values of total benefits paid. The final output of this thesis is the prediction model. Sources are identified where the employment data and the workers' compensation data could be acquired. Employment data was acquired from the Bureau of Labor Statistics and workers' compensation data was acquired from the National Academy of Social Insurance. SAS was identified as an adequate software to develop a regression model through multiple regression. When calculating all 50 states and D.C. simultaneously multiple regression was overwhelmed by the equation. Therefore, ANCOVA was a great technique in combination with regression to ease the process. ANCOVA simplified the equation and omitted all the unnecessary elements from the equation. The reason ANCOVA was used is because there were numerical values associated with the variates. If these were not numerical values than the same procedure could be run using ANOVA.

This predictive model is fit for all of the U.S. states. The model best describes the states of Alabama, Alaska, Arkansas, Indiana, Kentucky, and New Hampshire. The model generated predicted values of all of these states within the error margin of ±10%. However, when states have a high average number of covered workers the residual in the end prediction value can become too large. This is previously shown in figure 4.1. The states include California, New York, and Texas. Predicted values can be generated for these states from the model but they can deviate a lot from the actual values. These states are paying a large amount for total benefits so greatness of error can be expected. When we look at the percentage change it is found that West Virginia is not doing so well with the predicted values with a deviation of up to 36%. Most of the observations fall within the error margin of ±10%. Out of 765 observations, 522 have an error of 10% or less.

5.2. Limitations of the Work

The workers' compensation prediction model is statistically proven to be significant though it has a few limitations. As shown in figure 4.5, there are some states with minimum error and some with huge error. West Virginia happens to be the one with the largest percent deviation in predicted values form the actual values of total benefits. The predicted values from the model are somewhat in a linear form but in actual the value of total benefits vary a lot from year to year. Mining is one of the biggest industries in West Virginia and it depends on the number of accidents that how much the claims are occurring for the accidents. When there is mining going on, there would be more claims and in a year when there is not a lot of mining, the number of claims would be reduced. Hence, it is highly recommended to run the model for each industry separately in order to obtain more accurate predicted value of total benefits. There is a significant difference in the number of accidents from industry to industry. Similarly, the amount of workers' compensation insurance purchase is also significantly different. For example in North Dakota in 2017-2018, professional athlete is insured for $49.75 and a banker is insured for $0.23 per $100 in payroll, respectively. It all depends on the nature of work and probability of accidents. West Virginia has a huge mining industry which is also responsible for most number of accidents.

5.3. Future Work

Following are the recommendations which can be considered for future work:

1. The raw data of total benefits was available for all industries combined. Therefore, the analysis and the prediction model is based for all industries. In the future, when agencies are able to provide data separately for each industry then the prediction model would be much more accurate. Currently, employment data is available for each industry but the total benefits data is not.

2. Cluster analysis is a technique which could be used to analyze this data set. Clustering is the grouping of the qualities of any observation and once the system recognizes the input it tells the range where the observation belongs.
3. Statistical analysis other than multiple regression can be explored and used if it better fits the requirements and assumptions this data set is offering.
4. Software other than SAS can be used to perform the same task and see whether there is any change in the prediction values and the residuals.
5. Only total benefits were used for the analysis. Total benefits is the combination of medical benefits and cash benefits which are available in the reports published by NASI. Medical benefits can be used to assess the prediction value of cost for future medical benefits and similarly for cash benefits.
6. Another layer of number of accidents can be added to the data of number of workers and total benefits.

REFERENCES

Agarwal, P., and Everett, J. G. (1997). "Strategies for construction contractors to reduce workers' compensation costs." *Journal of Management in Engineering*, 13(5), 70-75.

Berecki-Gisolf, J., Clay, F. J., Collie, A., and McClure, R. J. (2012). "Predictors of sustained return to work after work-related injury or disease: insights from workers' compensation claims records." *Journal of occupational rehabilitation*, 22(3), 283-291.

Berreth, C. A. (1994). "Workers' compensation laws: Significant changes in 1993." *Monthly Labor Review*, 117(1), 53.

Boyd, L. W. (1999). "Workers Compensation Reform Past and Present: An Analysis of Issues and Changes in Benefits." *Labor Studies Journal*, 24(2), 45-62.

Brahmasrene, T., and Smith, S. S. (2008). "Empirical Evidence of Factors Affecting Experience Modification Rate Used by the US Insurance Industry." *Journal of Transnational Management*, 13(3), 244-258.

Emerson, R. D., Minchin Jr, R. E., and Gruneberg, S. (2013). "Workers' Compensation in Construction: Workers' Benefits under Alternative Dispute Resolution Systems." *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, 5(3), 113-121.

Everett, J. G., and Frank Jr, P. B. (1996). "Costs of accidents and injuries to the construction industry." *Journal of Construction Engineering and Management*, 122(2), 158-164.

Everett, J. G., and Thompson, W. S. (1995). "Experience modification rating for workers' compensation insurance." *Journal of Construction Engineering and Management*, 121(1), 66-79.

Everett, J. G., and Yang, I.-T. (1997). "Workers' comp. premiums: Disparities in penalties for identical losses." *Journal of construction engineering and management*, 123(3), 312-317.

Fortunato III, B. R., Hallowell, M. R., Behm, M., and Dewlaney, K. (2011). "Identification of safety risks for high-performance sustainable construction projects." *Journal of Construction Engineering and Management*, 138(4), 499-508.

Friedman, L., Krupczak, C., Brandt-Rauf, S., and Forst, L. (2013). "Occupational amputations in Illinois 2000–2007: BLS vs. data linkage of trauma registry, hospital discharge, workers compensation databases and OSHA citations." *Injury*, 44(5), 667-673.

Friedman, L. S., Ruestow, P., and Forst, L. (2012). "Analysis of ethnic disparities in workers' compensation claims using data linkage." *Journal of occupational and environmental medicine*, 54(10), 1246-1252.

Guyton, G. P. (1999). "A brief history of workers' compensation." *The Iowa orthopaedic journal*, 19, 106.

Hallowell, M. R., and Gambatese, J. A. (2009). "Construction safety risk mitigation." *Journal of Construction Engineering and Management*, 135(12), 1316-1323.

Hancher, D. E., and de la Garza, J. M. (1995). "Workers' Compensation Management in Construction." *Proc., Construction Congress*, ASCE, 471-478.

Hancher, D. E., Garza, J. M. d. l., and Eckert, G. K. (1997). "Improving workers' compensation management in construction." *Journal of construction engineering and management*, 123(3), 285-291.

Haslam, R. A., Hide, S. A., Gibb, A. G., Gyi, D. E., Pavitt, T., Atkinson, S., and Duff, A. (2005). "Contributing factors in construction accidents." *Applied ergonomics*, 36(4), 401-415.

Hinze, J., and Appelgate, L. L. (1991). "Costs of construction injuries." *Journal of Construction Engineering and Management*, 117(3), 537-550.

Hinze, J., Bren, D. C., and Piepho, N. (1995). "Experience modification rating as measure of safety performance." *Journal of construction engineering and management*, 121(4), 455-458.

Hinze, J., and Russell, D. B. (1995). "Analysis of fatalities recorded by OSHA." *Journal of Construction Engineering and Management*, 121(2), 209-214.

Ikpe, E., Hammon, F., and Oloke, D. (2012). "Cost-benefit analysis for accident prevention in construction projects." *Journal of Construction Engineering and Management*, 138(8), 991-998.

Imriyas, K., Low, S., Teo, A., and Chan, S. (2008). "Premium-rating model for workers' compensation insurance in construction." *Journal of Construction Engineering and Management*, 134(8), 601-617.

Imriyas, K., Pheng, L. S., and Teo, E. A. L. (2007). "A framework for computing workers' compensation insurance premiums in construction." *Construction Management and Economics*, 25(6), 563-584.

Imriyas, K., Pheng, L. S., and Teo, E. A. L. (2007). "A fuzzy knowledge-based system for premium rating of workers' compensation insurance for building projects." *Construction Management and Economics*, 25(11), 1177-1195.

Jebb, A. T., Parrigon, S., and Woo, S. E. (2017). "Exploratory data analysis as a foundation of inductive research." *Human Resource Management Review*, 27(2), 265-276.

Kotz, D. M. (2009). "The financial and economic crisis of 2008: A systemic crisis of neoliberal capitalism." *Review of Radical Political Economics*, 41(3), 305-317.

Leppink, J. (2018). "Analysis of Covariance (ANCOVA) vs. Moderated Regression (MODREG): Why the Interaction Matters." *Health Professions Education*.

Lessard, J. L. (2000). "Temperature effects of dams on cold-water fish and macroinvertebrate communities in Michigan." Master of Science, Michigan State University, East Lansing.

Liao, C.-W., and Chiang, T.-L. (2015). "The examination of workers' compensation for occupational fatalities in the construction industry." *Safety Science*, 72, 363-370.

Lippel, K. (2012). "Preserving workers' dignity in workers' compensation systems: An international perspective." *American journal of industrial medicine*, 55(6), 519-536.

Marshall, G., and Jonker, L. (2010). "An introduction to descriptive statistics: A review and practical guide." *Radiography*, 16(4), e1-e7.

Moehrle, T. (2010). "Compensation of residential and nonresidential construction workers." *Monthly Labor Review*, 133(4).

Mroz, T. M., Carlini, A. R., Archer, K. R., Wegener, S. T., Hoolachan, J. I., Stiers, W., Shore, R. A., and Castillo, R. C. (2014). "Frequency and cost of claims by injury type from a state workers' compensation fund from 1998 through 2008." *Archives of physical medicine and rehabilitation*, 95(6), 1048-1054. e1046.

Murray, B. (2010). "The combined and differential roles of working memory mechanisms in academic achievement." Doctor of Philosophy, Indiana University of Pennsylvania, Indiana.

Recarte Suazo, G. A., and Jaselskis, E. J. (1993). "Comparison of construction safety codes in United States and Honduras." *Journal of construction Engineering and Management*, 119(3), 560-572.

Ruseckaite, R., and Collie, A. (2011). "Repeat workers' compensation claims: risk factors, costs and work disability." *BMC Public Health*, 11(1), 492.

Schexnayder, C. J., Weber, S. L., and David, S. A. (2004). "Transportation agency use of owner-controlled insurance programs." *Journal of construction engineering and management*, 130(4), 517-524.

Shannon, H. S., and Lowe, G. S. (2002). "How many injured workers do not file claims for workers' compensation benefits?" *American journal of industrial medicine*, 42(6), 467-473.

Spieler, E. A., and Burton, J. F. (2012). "The lack of correspondence between work-related disability and receipt of workers' compensation benefits." *American journal of industrial medicine*, 55(6), 487-505.

Strunin, L., and Boden, L. I. (2004). "The workers' compensation system: worker friend or foe?" *American Journal of Industrial Medicine*, 45(4), 338-345.

Veltri, A. (1990). "An accident cost impact model: The direct cost component." *Journal of Safety Research*, 21(2), 67-73.

Zhou, Z., Goh, Y. M., and Li, Q. (2015). "Overview and analysis of safety management studies in the construction industry." *Safety science*, 72, 337-350.

APPENDIX A. DESCRIPTIVE STATISTICS BY STATES

Alabama						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1729	614556		Skewness	1	-1
Median	1719	624685		Range	157	93834
Standard Deviation	49	29868		Minimum	1666	562773
Sample Variance	2385	892102011		Maximum	1823	656607
Kurtosis	0	-1		Count	15	15

Alaska						
	Employment	Total Benefits			Employment	Total Benefits
Mean	294	210667		Skewness	0	0
Median	297	219163		Range	51	94777
Standard Deviation	17	28088		Minimum	266	158520
Sample Variance	273	788932851		Maximum	317	253297
Kurtosis	-1	-1		Count	15	15

Arizona						
	Employment	Total Benefits			Employment	Total Benefits
Mean	2389	636699		Skewness	0	-1
Median	2374	691384		Range	404	288772
Standard Deviation	136	101265		Minimum	2191	452011
Sample Variance	18592	10254686659		Maximum	2595	740783
Kurtosis	-1	-1		Count	15	15

Arkansas						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1091	219764		Skewness	0	0
Median	1092	218670		Range	64	38670
Standard Deviation	21	10445		Minimum	1061	202006
Sample Variance	434	109104660		Maximum	1125	240676
Kurtosis	-1	0		Count	15	15

California						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	14917	10898207		Skewness	0.59	0
Median	14728	10938475		Range	1880	3066754
Standard Deviation	521.87	1162359		Minimum	14171	9392835
Sample Variance	272351.71	1351079173954		Maximum	16051	12459589
Kurtosis	-0.10	-2		Count	15	15

Colorado						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	2189	805777		Skewness	1	-2
Median	2148	835265		Range	364	330076
Standard Deviation	104	78642		Minimum	2064	566354
Sample Variance	10762	6184635960		Maximum	2428	896430
Kurtosis	1	6		Count	15	15

Connecticut						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	1625	791325		Skewness	0	0
Median	1624	785133		Range	92	313988
Standard Deviation	27	104331		Minimum	1576	641341
Sample Variance	725	10884908677		Maximum	1668	955329
Kurtosis	-1	-2		Count	15	15

District of Columbia						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	489	103777		Skewness	1	0.563440988
Median	483	99496		Range	93	46953
Standard Deviation	27	13057.62883		Minimum	452	84015
Sample Variance	738	170501670.7		Maximum	545	130968
Kurtosis	0	-0.300973434		Count	15	15

Delaware						
	Employment	Total Benefits			Employment	Total Benefits
Mean	406	199273		Skewness	0	-1
Median	406	212805		Range	35	123115
Standard Deviation	11	39499		Minimum	392	126270
Sample Variance	120	1560147256		Maximum	427	249385
Kurtosis	-1	-1		Count	15	15

Florida						
	Employment	Total Benefits			Employment	Total Benefits
Mean	7031	2922403		Skewness	0	0
Median	7005	2899301		Range	909	727422
Standard Deviation	323	224845		Minimum	6612	2526580
Sample Variance	104286	50555171791		Maximum	7521	3254002
Kurtosis	-1	-1		Count	15	15

Georgia						
	Employment	Total Benefits			Employment	Total Benefits
Mean	3717	1306249		Skewness	0	-1
Median	3682	1383560		Range	411	678785
Standard Deviation	126	210761		Minimum	3543	917266
Sample Variance	15893	44420241436		Maximum	3954	1596051
Kurtosis	-1	-1		Count	15	15

Hawaii						
	Employment	Total Benefits			Employment	Total Benefits
Mean	567	257330		Skewness	0	1
Median	569	248433		Range	78	55837
Standard Deviation	25	16113		Minimum	527	242400
Sample Variance	602	259617975		Maximum	605	298237
Kurtosis	-1	1		Count	15	15

Idaho						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	605	239091		Skewness	0	-1
Median	601	244451		Range	94	63630
Standard Deviation	32	20047		Minimum	558	199044
Sample Variance	1020	401874499		Maximum	652	262674
Kurtosis	-1	0		Count	15	15

Illinois						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	5631	2571147		Skewness	-1	0
Median	5660	2632204		Range	396	903051
Standard Deviation	124	319196		Minimum	5397	2122283
Sample Variance	15464	101886086061		Maximum	5793	3025334
Kurtosis	-1	-1		Count	15	15

Indiana						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	2790	594058		Skewness	-1	0
Median	2802	598048		Range	237	95763
Standard Deviation	70	29020		Minimum	2655	547305
Sample Variance	4943	842133988		Maximum	2892	643068
Kurtosis	0	-1		Count	15	15

Iowa						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	1435	527591		Skewness	0	0
Median	1428	548605		Range	112	252545
Standard Deviation	35	87890		Minimum	1385	391156
Sample Variance	1191	7724613251		Maximum	1497	643701
Kurtosis	-1	-1		Count	15	15

Kansas						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1290	382293		Skewness	1	-1
Median	1285	383283		Range	91	139625
Standard Deviation	28	37395		Minimum	1251	295520
Sample Variance	801	1398374262		Maximum	1342	435145
Kurtosis	-1	1		Count	15	15

Kentucky						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1715	681545		Skewness	0	0
Median	1717	684422		Range	129	81864
Standard Deviation	40	27745		Minimum	1665	643192
Sample Variance	1627	769776437		Maximum	1794	725056
Kurtosis	-1	-1		Count	15	15

Louisiana						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1831	752881		Skewness	1	0
Median	1831	755714		Range	120	233399
Standard Deviation	33	78658		Minimum	1776	621449
Sample Variance	1064	6187010968		Maximum	1896	854848
Kurtosis	0	-1		Count	15	15

Maine						
	Employment	Total Benefits			Employment	Total Benefits
Mean	575	262547		Skewness	0	2
Median	577	253872		Range	29	107495
Standard Deviation	9	26178		Minimum	559	232464
Sample Variance	83	685281069		Maximum	588	339959
Kurtosis	-1	5		Count	15	15

Maryland						
	Employment	Total Benefits			Employment	Total Benefits
Mean	2360	863187		Skewness	0	0
Median	2363	895905		Range	148	344744
Standard Deviation	49	120005		Minimum	2295	664282
Sample Variance	2430	14401272225		Maximum	2443	1009026
Kurtosis	-1	-1		Count	15	15

Massachusetts						
	Employment	Total Benefits			Employment	Total Benefits
Mean	3176	969071		Skewness	1	0
Median	3150	968085		Range	295	321509
Standard Deviation	87	102758		Minimum	3087	829449
Sample Variance	7589	10559166850		Maximum	3382	1150958
Kurtosis	1	-1		Count	15	15

Michigan						
	Employment	Total Benefits			Employment	Total Benefits
Mean	3968	1370211		Skewness	0	-1
Median	4003	1470574		Range	729	439439
Standard Deviation	228	156411		Minimum	3596	1077947
Sample Variance	52103	24464261245		Maximum	4325	1517386
Kurtosis	-1	-1		Count	15	15

Minnesota						
	Employment	Total Benefits			Employment	Total Benefits
Mean	2600	990575		Skewness	0	0
Median	2597	1011890		Range	221	194257
Standard Deviation	63	66101		Minimum	2506	885006
Sample Variance	3909	4369331806		Maximum	2727	1079263
Kurtosis	0	-2		Count	15	15

Mississippi						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1022	320496		Skewness	0	-1
Median	1026	328234		Range	68	54802
Standard Deviation	21	18765		Minimum	989	284729
Sample Variance	421	352126929		Maximum	1057	339531
Kurtosis	-1	-1		Count	15	15

Missouri						
	Employment	Total Benefits			Employment	Total Benefits
Mean	2470	914459		Skewness	0	1
Median	2466	867153		Range	165	312577
Standard Deviation	51	106554		Minimum	2390	807294
Sample Variance	2613	11353674419		Maximum	2555	1119871
Kurtosis	-1	-1		Count	15	15

Montana						
	Employment	Total Benefits			Employment	Total Benefits
Mean	406	235250		Skewness	-1	-1
Median	407	244114		Range	61	85080
Standard Deviation	19	22735		Minimum	371	181770
Sample Variance	369	516862887		Maximum	432	266850
Kurtosis	-1	1		Count	15	15

Nebraska						
	Employment	Total Benefits			Employment	Total Benefits
Mean	894	292824		Skewness	2	-1
Median	886	299292		Range	171	88292
Standard Deviation	42	24746		Minimum	850	235434
Sample Variance	1758	612365421		Maximum	1021	323726
Kurtosis	6	0		Count	15	15

Nevada						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1133	377036		Skewness	-1	0
Median	1127	374209		Range	399	120287
Standard Deviation	103	41484		Minimum	866	310750
Sample Variance	10613	1720945933		Maximum	1265	431037
Kurtosis	2	-1		Count	15	15

New Hampshire						
	Employment	Total Benefits			Employment	Total Benefits
Mean	608	223661		Skewness	0	0
Median	605	222064		Range	36	34130
Standard Deviation	12	10728		Minimum	593	208437
Sample Variance	135	115085298		Maximum	629	242567
Kurtosis	-1	-1		Count	15	15

New Jersey						
	Employment	Total Benefits			Employment	Total Benefits
Mean	14917	1889834		Skewness	0.59	0
Median	14728	1947752		Range	1880	1036568
Standard Deviation	522	348966		Minimum	14171	1312381
Sample Variance	272352	121777201070		Maximum	16051	2348949
Kurtosis	-0.10	-1		Count	15	15

New Mexico						
	Employment	Total Benefits			Employment	Total Benefits
Mean	725	244824		Skewness	0	0
Median	726	238881		Range	93	147489
Standard Deviation	29	48669		Minimum	673	158815
Sample Variance	822	2368630595		Maximum	766	306304
Kurtosis	-1	-1		Count	15	15

New York						
	Employment	Total Benefits			Employment	Total Benefits
Mean	8355	4182304		Skewness	1	0
Median	8302	3899911		Range	789	2922187
Standard Deviation	224	1114433		Minimum	8089	2881566
Sample Variance	50128	1241960199468		Maximum	8878	5803753
Kurtosis	1	-2		Count	15	15

North Carolina						
	Employment	Total Benefits			Employment	Total Benefits
Mean	3739	1279116		Skewness	1	-1
Median	3707	1316291		Range	411	551713
Standard Deviation	129	169423		Minimum	3577	916541
Sample Variance	16625	28704075430		Maximum	3988	1468254
Kurtosis	-1	0		Count	15	15

North Dakota						
	Employment	Total Benefits			Employment	Total Benefits
Mean	349	114993		Skewness	1	1
Median	337	105837		Range	132	121253
Standard Deviation	47	42308		Minimum	300	70984
Sample Variance	2182	1789927588		Maximum	432	192237
Kurtosis	-1	-1		Count	15	15

Ohio						
	Employment	Total Benefits			Employment	Total Benefits
Mean	5116	2292422		Skewness	-1	-1
Median	5182	2353384		Range	530	560818
Standard Deviation	162	173322		Minimum	4822	1929262
Sample Variance	26343	30040379131		Maximum	5352	2490080
Kurtosis	-1	0		Count	15	15

Oklahoma						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1416	699768		Skewness	0	0
Median	1417	732542		Range	140	370764
Standard Deviation	46	126997		Minimum	1359	508931
Sample Variance	2081	16128223762		Maximum	1499	879695
Kurtosis	-1	-1		Count	15	15

Oregon						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1623	594665		Skewness	1	0
Median	1612	605897		Range	227	185840
Standard Deviation	68	63862		Minimum	1533	497612
Sample Variance	4622	4078344977		Maximum	1760	683452
Kurtosis	-1	-1		Count	15	15

Pennsylvania						
	Employment	Total Benefits			Employment	Total Benefits
Mean	5451	2783150		Skewness	0	-1
Median	5446	2895338		Range	233	591658
Standard Deviation	76	192890		Minimum	5343	2406272
Sample Variance	5811	37206576654		Maximum	5576	2997930
Kurtosis	-1	-1		Count	15	15

Rhode Island						
	Employment	Total Benefits			Employment	Total Benefits
Mean	450	155728		Skewness	1	0
Median	445	159550		Range	35	44609
Standard Deviation	13	12889		Minimum	436	134072
Sample Variance	160	166119892		Maximum	471	178681
Kurtosis	-1	-1		Count	15	15

South Carolina						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1727	810113		Skewness	1	-1
Median	1704	885307		Range	189	386276
Standard Deviation	57	129090		Minimum	1657	532374
Sample Variance	3199	16664328846		Maximum	1846	918650
Kurtosis	0	0		Count	15	15

South Dakota						
	Employment	Total Benefits			Employment	Total Benefits
Mean	374	93757		Skewness	0	0
Median	374	95373		Range	48	48831
Standard Deviation	15	14906		Minimum	352	70736
Sample Variance	234	222186470		Maximum	400	119567
Kurtosis	-1	-1		Count	15	15

Tennessee						
	Employment	Total Benefits			Employment	Total Benefits
Mean	2512	93757		Skewness	0	0
Median	2494	95373		Range	242	48831
Standard Deviation	73	14906		Minimum	2410	70736
Sample Variance	5354	222186470		Maximum	2652	119567
Kurtosis	-1	-1		Count	15	15

Texas						
	Employment	Total Benefits			Employment	Total Benefits
Mean	7920	778609		Skewness	0	0
Median	7705	782091		Range	2289	192505
Standard Deviation	690	53090		Minimum	6949	687595
Sample Variance	476543	2818583344		Maximum	9238	880100
Kurtosis	-1	0		Count	15	15

Utah						
	Employment	Total Benefits			Employment	Total Benefits
Mean	1131	1676651		Skewness	0	2
Median	1135	1564956		Range	295	954510
Standard Deviation	91	296691		Minimum	1006	1416287
Sample Variance	8209	88025448974		Maximum	1301	2370797
Kurtosis	-1	2		Count	15	15

Vermont						
	Employment	Total Benefits			Employment	Total Benefits
Mean	292	131915		Skewness	0	0
Median	292	128305		Range	14	53026
Standard Deviation	4	15351		Minimum	284	98518
Sample Variance	20	235644209		Maximum	298	151544
Kurtosis	-1	0		Count	15	15

Virginia						
	Employment	Total Benefits			Employment	Total Benefits
Mean	3332	825196		Skewness	0	-1
Median	3348	858665		Range	291	324239
Standard Deviation	91	108615		Minimum	3186	612083
Sample Variance	8370	11797144273		Maximum	3477	936322
Kurtosis	-1	0		Count	15	15

Washington						
	Employment	Total Benefits			Employment	Total Benefits
Mean	2762	2090065		Skewness	1	0
Median	2773	2192885		Range	474	764929
Standard Deviation	142	271792		Minimum	2575	1639435
Sample Variance	20113	73871137305		Maximum	3049	2404364
Kurtosis	0	-2		Count	15	15

West Virginia						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	664	525243		Skewness	0	1
Median	664	435709		Range	46	503423
Standard Deviation	12	181589		Minimum	638	319877
Sample Variance	136	32974591524		Maximum	684	823300
Kurtosis	1	-1		Count	15	15

Wisconsin						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	2623	1076347		Skewness	0	-1
Median	2626	1099950		Range	171	360114
Standard Deviation	53	108811		Minimum	2523	843888
Sample Variance	2845	11839732732		Maximum	2694	1204002
Kurtosis	-1	0		Count	15	15

Wyoming						
	Employment	**Total Benefits**			**Employment**	**Total Benefits**
Mean	259	141207		Skewness	-1	0
Median	267	134835		Range	51	91749
Standard Deviation	18	30085		Minimum	228	100076
Sample Variance	328	905089445		Maximum	279	191825
Kurtosis	-1	-1		Count	15	15

APPENDIX B. REGRESSION PREDICTED VALUES

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Alabama	2001	1726	562773	614130	-51357	-9.1
Alabama	2002	1704	565264	610785	-45521	-8.1
Alabama	2003	1698	580184	609872	-29688	-5.1
Alabama	2004	1720	575697	613218	-37521	-6.5
Alabama	2005	1763	608522	619757	-11235	-1.8
Alabama	2006	1797	624685	624927	-242	0.0
Alabama	2007	1823	635315	628881	6434	1.0
Alabama	2008	1808	656607	626600	30007	4.6
Alabama	2009	1702	625755	610480	15275	2.4
Alabama	2010	1679	629069	606983	22086	3.5
Alabama	2011	1666	616260	605006	11254	1.8
Alabama	2012	1680	644224	607135	37089	5.8
Alabama	2013	1700	639549	610176	29373	4.6
Alabama	2014	1719	636813	613066	23747	3.7
Alabama	2015	1747	617622	617324	298	0.0
Alaska	2001	266	158520	165269	-6749	-4.3
Alaska	2002	270	178789	171678	7111	4.0
Alaska	2003	275	182204	179689	2515	1.4
Alaska	2004	279	187080	186099	981	0.5
Alaska	2005	285	189212	195712	-6500	-3.4
Alaska	2006	291	197580	205326	-7746	-3.9
Alaska	2007	294	201477	210133	-8656	-4.3
Alaska	2008	298	219163	216542	2621	1.2
Alaska	2009	297	221021	214940	6081	2.8
Alaska	2010	299	221327	218145	3182	1.4
Alaska	2011	305	240482	227759	12723	5.3
Alaska	2012	311	247862	237373	10489	4.2
Alaska	2013	313	253297	240577	12720	5.0
Alaska	2014	315	233962	243782	-9820	-4.2
Alaska	2015	317	228034	246986	-18952	-8.3
Arizona	2001	2195	452011	540421	-88410	-19.6
Arizona	2002	2191	477568	538440	-60872	-12.7
Arizona	2003	2222	515231	553788	-38557	-7.5
Arizona	2004	2304	548172	594385	-46213	-8.4
Arizona	2005	2438	535539	660727	-125188	-23.4
Arizona	2006	2562	647463	722118	-74655	-11.5
Arizona	2007	2595	696908	738456	-41548	-6.0
Arizona	2008	2529	691384	705780	-14396	-2.1

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Arizona	2009	2340	658115	612208	45907	7.0
Arizona	2010	2295	698459	589929	108530	15.5
Arizona	2011	2326	719537	605277	114260	15.9
Arizona	2012	2374	718152	629041	89111	12.4
Arizona	2013	2431	716253	657262	58991	8.2
Arizona	2014	2485	734908	683996	50912	6.9
Arizona	2015	2555	740783	718653	22130	3.0
Arkansas	2001	1071	223416	220437	2979	1.3
Arkansas	2002	1064	217346	220673	-3327	-1.5
Arkansas	2003	1061	224275	220774	3501	1.6
Arkansas	2004	1073	227243	220370	6873	3.0
Arkansas	2005	1092	208021	219730	-11709	-5.6
Arkansas	2006	1112	202006	219057	-17051	-8.4
Arkansas	2007	1119	213337	218821	-5484	-2.6
Arkansas	2008	1117	227769	218888	8881	3.9
Arkansas	2009	1078	215067	220202	-5135	-2.4
Arkansas	2010	1075	204066	220303	-16237	-8.0
Arkansas	2011	1086	218670	219932	-1262	-0.6
Arkansas	2012	1093	229180	219696	9484	4.1
Arkansas	2013	1094	240676	219663	21013	8.7
Arkansas	2014	1105	228195	219292	8903	3.9
Arkansas	2015	1125	217190	218619	-1429	-0.7
California	2001	14728	10082580	10795873	-713293	-7.1
California	2002	14588	10974355	10720070	254285	2.3
California	2003	14553	12409808	10701120	1708688	13.8
California	2004	14706	12459589	10783961	1675628	13.4
California	2005	14992	10938475	10938815	-340	0.0
California	2006	15256	10017099	11081758	-1064659	-10.6
California	2007	15395	9608884	11157019	-1548135	-16.1
California	2008	15248	9529739	11077426	-1547687	-16.2
California	2009	14377	9392835	10605825	-1212990	-12.9
California	2010	14171	9396443	10494286	-1097843	-11.7
California	2011	14310	10850879	10569548	281331	2.6
California	2012	14674	11535904	10766635	769269	6.7
California	2013	15139	12113656	11018408	1095248	9.0

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
California	2014	15567	12097277	11250148	847129	7.0
California	2015	16051	12065579	11512209	553370	4.6
Colorado	2001	2148	566354	800200	-233846	-41.3
Colorado	2002	2101	760958	793764	-32806	-4.3
Colorado	2003	2064	753049	788698	-35649	-4.7
Colorado	2004	2074	843256	790067	53189	6.3
Colorado	2005	2120	896430	796366	100064	11.2
Colorado	2006	2190	865585	805950	59635	6.9
Colorado	2007	2241	837004	812934	24070	2.9
Colorado	2008	2247	873643	813755	59888	6.9
Colorado	2009	2137	836238	798693	37545	4.5
Colorado	2010	2110	809707	794997	14710	1.8
Colorado	2011	2147	761760	800063	-38303	-5.0
Colorado	2012	2200	845654	807320	38334	4.5
Colorado	2013	2271	813193	817041	-3848	-0.5
Colorado	2014	2353	788559	828269	-39710	-5.0
Colorado	2015	2428	835265	838538	-3273	-0.4
Connecticut	2001	1644	641341	778227	-136886	-21.3
Connecticut	2002	1627	675895	790071	-114176	-16.9
Connecticut	2003	1605	677088	805398	-128310	-19.0
Connecticut	2004	1611	711237	801218	-89981	-12.7
Connecticut	2005	1624	713275	792161	-78886	-11.1
Connecticut	2006	1652	719758	772653	-52895	-7.3
Connecticut	2007	1666	734425	762900	-28475	-3.9
Connecticut	2008	1668	785133	761506	23627	3.0
Connecticut	2009	1596	842840	811668	31172	3.7
Connecticut	2010	1576	788701	825602	-36901	-4.7
Connecticut	2011	1594	892920	813062	79858	8.9
Connecticut	2012	1611	914723	801218	113505	12.4
Connecticut	2013	1623	955329	792858	162471	17.0
Connecticut	2014	1636	909138	783800	125338	13.8
Connecticut	2015	1645	908069	777530	130539	14.4
Delaware	2001	400	126270	189377	-63107	-50.0
Delaware	2002	396	137264	182992	-45728	-33.3
Delaware	2003	396	156494	182992	-26498	-16.9
Delaware	2004	406	157398	198954	-41556	-26.4
Delaware	2005	412	168146	208532	-40386	-24.0
Delaware	2006	417	238638	216513	22125	9.3
Delaware	2007	418	212805	218109	-5304	-2.5

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Delaware	2008	416	218665	214916	3749	1.7
Delaware	2009	395	206145	181396	24749	12.0
Delaware	2010	392	211921	176607	35314	16.7
Delaware	2011	396	220830	182992	37838	17.1
Delaware	2012	398	216588	186185	30403	14.0
Delaware	2013	407	240313	200550	39763	16.5
Delaware	2014	417	249385	216513	32872	13.2
Delaware	2015	427	228240	232475	-4235	-1.9
District of Columbia	2001	452	99496	88985	10511	10.6
District of Columbia	2002	458	89315	91414	-2099	-2.4
District of Columbia	2003	459	84015	91819	-7804	-9.3
District of Columbia	2004	467	96141	95058	1083	1.1
District of Columbia	2005	474	91270	97892	-6622	-7.3
District of Columbia	2006	479	98016	99917	-1901	-1.9
District of Columbia	2007	487	97564	103156	-5592	-5.7
District of Columbia	2008	491	95100	104776	-9676	-10.2
District of Columbia	2009	482	104672	101132	3540	3.4
District of Columbia	2010	483	105636	101537	4099	3.9
District of Columbia	2011	494	110316	105990	4326	3.9
District of Columbia	2012	506	115743	110849	4894	4.2
District of Columbia	2013	519	130968	116113	14855	11.3
District of Columbia	2014	532	118249	121377	-3128	-2.6
District of Columbia	2015	545	120154	126640	-6486	-5.4
Florida	2001	6754	3033955	2892049	141906	4.7
Florida	2002	6765	2623239	2893254	-270015	-10.3
Florida	2003	6840	2805941	2901469	-95528	-3.4
Florida	2004	7039	2710272	2923265	-212993	-7.9

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Florida	2005	7309	2899301	2952838	-53537	-1.8
Florida	2006	7498	2928460	2973540	-45080	-1.5
Florida	2007	7504	2857930	2974197	-116267	-4.1
Florida	2008	7177	2748092	2938380	-190288	-6.9
Florida	2009	6689	2820747	2884929	-64182	-2.3
Florida	2010	6612	2526580	2876496	-349916	-13.8
Florida	2011	6689	3254002	2884929	369073	11.3
Florida	2012	6826	3178981	2899935	279046	8.8
Florida	2013	7005	3189393	2919541	269852	8.5
Florida	2014	7239	3207769	2945171	262598	8.2
Florida	2015	7521	3051390	2976059	75331	2.5
Georgia	2001	3682	1006721	1285829	-279108	-27.7
Georgia	2002	3624	917266	1252309	-335043	-36.5
Georgia	2003	3597	981142	1236705	-255563	-26.0
Georgia	2004	3663	1114154	1274848	-160694	-14.4
Georgia	2005	3751	1197521	1325706	-128185	-10.7
Georgia	2006	3838	1397771	1375987	21784	1.6
Georgia	2007	3891	1499306	1406617	92689	6.2
Georgia	2008	3831	1596051	1371941	224110	14.0
Georgia	2009	3592	1527428	1233815	293613	19.2
Georgia	2010	3543	1410753	1205496	205257	14.5
Georgia	2011	3594	1383560	1234971	148589	10.7
Georgia	2012	3644	1431794	1263867	167927	11.7
Georgia	2013	3722	1381721	1308946	72775	5.3
Georgia	2014	3834	1386071	1373675	12396	0.9
Georgia	2015	3954	1362480	1443027	-80547	-5.9
Hawaii	2001	527	248100	253419	-5319	-2.1
Hawaii	2002	528	267827	253517	14310	5.3
Hawaii	2003	538	274922	254496	20426	7.4
Hawaii	2004	554	271290	256064	15226	5.6
Hawaii	2005	572	250779	257827	-7048	-2.8
Hawaii	2006	586	242685	259198	-16513	-6.8
Hawaii	2007	594	247294	259982	-12688	-5.1
Hawaii	2008	587	245763	259296	-13533	-5.5
Hawaii	2009	559	244375	256553	-12178	-5.0
Hawaii	2010	551	242400	255770	-13370	-5.5
Hawaii	2011	558	246780	256455	-9675	-3.9
Hawaii	2012	569	248433	257533	-9100	-3.7
Hawaii	2013	583	260352	258904	1448	0.6

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Hawaii	2014	593	270720	259884	10836	4.0
Hawaii	2015	605	298237	261059	37178	12.5
Idaho	2001	558	199044	216624	-17580	-8.8
Idaho	2002	558	202181	216624	-14443	-7.1
Idaho	2003	562	213604	218552	-4948	-2.3
Idaho	2004	578	236149	226266	9883	4.2
Idaho	2005	601	243168	237356	5812	2.4
Idaho	2006	631	228764	251820	-23056	-10.1
Idaho	2007	648	244451	260016	-15565	-6.4
Idaho	2008	640	260881	256159	4722	1.8
Idaho	2009	600	257868	236873	20995	8.1
Idaho	2010	592	245622	233016	12606	5.1
Idaho	2011	595	249368	234463	14905	6.0
Idaho	2012	602	239807	237838	1969	0.8
Idaho	2013	618	248667	245552	3115	1.3
Idaho	2014	634	254120	253266	854	0.3
Idaho	2015	652	262674	261945	729	0.3
Illinois	2001	5793	2122283	2357358	-235075	-11.1
Illinois	2002	5679	2148757	2508174	-359417	-16.7
Illinois	2003	5606	2146926	2604749	-457823	-21.3
Illinois	2004	5611	2246186	2598135	-351949	-15.7
Illinois	2005	5660	2404456	2533310	-128854	-5.4
Illinois	2006	5733	2447104	2436735	10369	0.4
Illinois	2007	5782	2735393	2371911	363482	13.3
Illinois	2008	5741	2915102	2426152	488950	16.8
Illinois	2009	5452	3025334	2808483	216851	7.2
Illinois	2010	5397	2916379	2881245	35134	1.2
Illinois	2011	5467	2998181	2788639	209542	7.0
Illinois	2012	5537	2666873	2696033	-29160	-1.1
Illinois	2013	5590	2632204	2625917	6287	0.2
Illinois	2014	5669	2741604	2521404	220200	8.0
Illinois	2015	5754	2420417	2408953	11464	0.5
Indiana	2001	2822	556866	590302	-33436	-6.0
Indiana	2002	2785	547305	594609	-47304	-8.6
Indiana	2003	2774	563577	595889	-32312	-5.7
Indiana	2004	2802	595245	592630	2615	0.4
Indiana	2005	2827	609596	589720	19876	3.3
Indiana	2006	2845	563190	587625	-24435	-4.3
Indiana	2007	2858	598973	586112	12861	2.1

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Indiana	2008	2823	625721	590186	35535	5.7
Indiana	2009	2655	598048	609739	-11691	-2.0
Indiana	2010	2655	603193	609739	-6546	-1.1
Indiana	2011	2705	627737	603920	23817	3.8
Indiana	2012	2762	620780	597286	23494	3.8
Indiana	2013	2799	643068	592979	50089	7.8
Indiana	2014	2842	590031	587974	2057	0.3
Indiana	2015	2892	567536	582155	-14619	-2.6
Iowa	2001	1410	391156	486503	-95347	-24.4
Iowa	2002	1393	401983	458412	-56429	-14.0
Iowa	2003	1385	427030	445194	-18164	-4.3
Iowa	2004	1404	447343	476588	-29245	-6.5
Iowa	2005	1428	473724	516245	-42521	-9.0
Iowa	2006	1453	487985	557554	-69569	-14.3
Iowa	2007	1467	493953	580687	-86734	-17.6
Iowa	2008	1460	552913	569120	-16207	-2.9
Iowa	2009	1415	548605	494764	53841	9.8
Iowa	2010	1402	554973	473284	81689	14.7
Iowa	2011	1419	615544	501374	114170	18.5
Iowa	2012	1443	630303	541030	89273	14.2
Iowa	2013	1464	627280	575730	51550	8.2
Iowa	2014	1483	643701	607125	36576	5.7
Iowa	2015	1497	617375	630258	-12883	-2.1
Kansas	2001	1286	339258	380972	-41714	-12.3
Kansas	2002	1270	341606	376096	-34490	-10.1
Kansas	2003	1251	295520	370305	-74785	-25.3
Kansas	2004	1263	371011	373962	-2951	-0.8
Kansas	2005	1272	383283	376705	6578	1.7
Kansas	2006	1293	391381	383106	8275	2.1
Kansas	2007	1324	395836	392554	3282	0.8
Kansas	2008	1342	416157	398040	18117	4.4
Kansas	2009	1283	416157	380058	36099	8.7
Kansas	2010	1261	407776	373353	34423	8.4
Kansas	2011	1268	435145	375486	59659	13.7
Kansas	2012	1285	426096	380667	45429	10.7
Kansas	2013	1303	377452	386153	-8701	-2.3
Kansas	2014	1322	376158	391944	-15786	-4.2
Kansas	2015	1332	361558	394992	-33434	-9.2
Kentucky	2001	1696	725056	686572	38484	5.3

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Kentucky	2002	1676	692398	691738	660	0.1
Kentucky	2003	1673	717309	692513	24796	3.5
Kentucky	2004	1688	719833	688639	31194	4.3
Kentucky	2005	1717	705802	681149	24653	3.5
Kentucky	2006	1738	643192	675725	-32533	-5.1
Kentucky	2007	1760	646066	670043	-23977	-3.7
Kentucky	2008	1748	695746	673142	22604	3.2
Kentucky	2009	1667	686142	694062	-7920	-1.2
Kentucky	2010	1665	650701	694579	-43878	-6.7
Kentucky	2011	1689	671282	688380	-17098	-2.5
Kentucky	2012	1718	667084	680890	-13806	-2.1
Kentucky	2013	1738	668956	675725	-6769	-1.0
Kentucky	2014	1765	649182	668752	-19570	-3.0
Kentucky	2015	1794	684422	661262	23160	3.4
Louisiana	2001	1835	633703	754662	-120959	-19.1
Louisiana	2002	1812	621449	744067	-122618	-19.7
Louisiana	2003	1820	669218	747752	-78534	-11.7
Louisiana	2004	1831	726004	752820	-26816	-3.7
Louisiana	2005	1807	667097	741763	-74666	-11.2
Louisiana	2006	1776	718542	727483	-8941	-1.2
Louisiana	2007	1837	732788	755584	-22796	-3.1
Louisiana	2008	1853	854848	762954	91894	10.7
Louisiana	2009	1813	831997	744527	87470	10.5
Louisiana	2010	1796	839821	736696	103125	12.3
Louisiana	2011	1811	833632	743606	90026	10.8
Louisiana	2012	1833	810539	753741	56798	7.0
Louisiana	2013	1858	808073	765258	42815	5.3
Louisiana	2014	1889	789789	779538	10251	1.3
Louisiana	2015	1896	755714	782763	-27049	-3.6
Maine	2001	579	245343	267956	-22613	-9.2
Maine	2002	577	260310	265296	-4986	-1.9
Maine	2003	577	233458	265296	-31838	-13.6
Maine	2004	583	268040	273277	-5237	-2.0
Maine	2005	581	268936	270617	-1681	-0.6
Maine	2006	584	289994	274607	15387	5.3
Maine	2007	588	276880	279927	-3047	-1.1
Maine	2008	585	339959	275937	64022	18.8
Maine	2009	564	260526	248005	12521	4.8
Maine	2010	559	253872	241355	12517	4.9

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Maine	2011	562	252726	245345	7381	2.9
Maine	2012	565	250479	249335	1144	0.5
Maine	2013	569	253139	254655	-1516	-0.6
Maine	2014	573	252084	259976	-7892	-3.1
Maine	2015	578	232464	266626	-34162	-14.7
Maryland	2001	2295	681633	791805	-110172	-16.2
Maryland	2002	2299	664282	796198	-131916	-19.9
Maryland	2003	2306	723475	803885	-80410	-11.1
Maryland	2004	2332	786631	832438	-45807	-5.8
Maryland	2005	2372	769563	876365	-106802	-13.9
Maryland	2006	2405	788874	912606	-123732	-15.7
Maryland	2007	2422	829914	931275	-101361	-12.2
Maryland	2008	2407	935948	914802	21146	2.3
Maryland	2009	2326	895905	825849	70056	7.8
Maryland	2010	2310	953533	808278	145255	15.2
Maryland	2011	2330	1009026	830242	178784	17.7
Maryland	2012	2363	993842	866482	127360	12.8
Maryland	2013	2384	969103	889544	79559	8.2
Maryland	2014	2406	980011	913704	66307	6.8
Maryland	2015	2443	966069	954337	11732	1.2
Massachusetts	2001	3222	901729	994333	-92604	-10.3
Massachusetts	2002	3150	887313	954907	-67594	-7.6
Massachusetts	2003	3089	1058838	921505	137333	13.0
Massachusetts	2004	3087	968085	920410	47675	4.9
Massachusetts	2005	3110	903555	933004	-29449	-3.3
Massachusetts	2006	3146	831373	952717	-121344	-14.6
Massachusetts	2007	3185	829449	974072	-144623	-17.4
Massachusetts	2008	3197	854351	980643	-126292	-14.8
Massachusetts	2009	3087	952081	920410	31671	3.3
Massachusetts	2010	3098	1013343	926433	86910	8.6
Massachusetts	2011	3136	1003138	947241	55897	5.6
Massachusetts	2012	3190	982005	976810	5195	0.5
Massachusetts	2013	3244	1070458	1006380	64078	6.0
Massachusetts	2014	3315	1150958	1045257	105701	9.2
Massachusetts	2015	3382	1129393	1081945	47448	4.2
Michigan	2001	4325	1477986	1476154	1832	0.1
Michigan	2002	4242	1512457	1451495	60962	4.0
Michigan	2003	4175	1476850	1431590	45260	3.1
Michigan	2004	4152	1517386	1424757	92629	6.1

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Michigan	2005	4148	1473598	1423569	50029	3.4
Michigan	2006	4085	1470574	1404852	65722	4.5
Michigan	2007	4031	1511282	1388809	122473	8.1
Michigan	2008	3904	1407282	1351079	56203	4.0
Michigan	2009	3608	1509881	1263140	246741	16.3
Michigan	2010	3596	1271892	1259575	12317	1.0
Michigan	2011	3692	1301061	1288095	12966	1.0
Michigan	2012	3774	1189483	1312457	-122974	-10.3
Michigan	2013	3860	1246512	1338007	-91495	-7.3
Michigan	2014	3931	1108978	1359100	-250122	-22.6
Michigan	2015	4003	1077947	1380491	-302544	-28.1
Minnesota	2001	2576	901780	984084	-82304	-9.1
Minnesota	2002	2552	921473	977520	-56047	-6.1
Minnesota	2003	2542	885006	974785	-89779	-10.1
Minnesota	2004	2567	931005	981622	-50617	-5.4
Minnesota	2005	2607	945888	992562	-46674	-4.9
Minnesota	2006	2637	944448	1000767	-56319	-6.0
Minnesota	2007	2655	958984	1005690	-46706	-4.9
Minnesota	2008	2631	1025671	999126	26545	2.6
Minnesota	2009	2521	1072122	969042	103080	9.6
Minnesota	2010	2506	1038272	964939	73333	7.1
Minnesota	2011	2553	1011890	977794	34096	3.4
Minnesota	2012	2597	1042478	989827	52651	5.1
Minnesota	2013	2643	1064684	1002408	62276	5.8
Minnesota	2014	2682	1079263	1013074	66189	6.1
Minnesota	2015	2727	1035657	1025381	10276	1.0
Mississippi	2001	1033	284729	318195	-33466	-11.8
Mississippi	2002	1027	290378	319458	-29080	-10.0
Mississippi	2003	1020	291151	320931	-29780	-10.2
Mississippi	2004	1026	310030	319669	-9639	-3.1
Mississippi	2005	1032	311796	318406	-6610	-2.1
Mississippi	2006	1042	320394	316301	4093	1.3
Mississippi	2007	1057	328234	313144	15090	4.6
Mississippi	2008	1053	339531	313986	25545	7.5
Mississippi	2009	1004	321771	324299	-2528	-0.8
Mississippi	2010	996	337633	325983	11650	3.5
Mississippi	2011	989	334430	327456	6974	2.1
Mississippi	2012	997	336208	325772	10436	3.1
Mississippi	2013	1007	332790	323668	9122	2.7

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Mississippi	2014	1017	336689	321563	15126	4.5
Mississippi	2015	1031	331683	318616	13067	3.9
Missouri	2001	2482	958708	918383	40325	4.2
Missouri	2002	2457	1033458	910253	123205	11.9
Missouri	2003	2447	1080870	907001	173869	16.1
Missouri	2004	2466	1119871	913180	206691	18.5
Missouri	2005	2499	1050889	923911	126978	12.1
Missouri	2006	2532	828370	934642	-106272	-12.8
Missouri	2007	2555	867153	942121	-74968	-8.6
Missouri	2008	2541	906587	937569	-30982	-3.4
Missouri	2009	2435	849798	903099	-53301	-6.3
Missouri	2010	2400	811427	891718	-80291	-9.9
Missouri	2011	2390	807294	888466	-81172	-10.1
Missouri	2012	2412	833119	895620	-62501	-7.5
Missouri	2013	2444	832469	906026	-73557	-8.8
Missouri	2014	2472	848867	915131	-66264	-7.8
Missouri	2015	2517	888004	929764	-41760	-4.7
Montana	2001	371	181770	202731	-20961	-11.5
Montana	2002	374	196197	205550	-9353	-4.8
Montana	2003	380	216715	211190	5525	2.5
Montana	2004	390	223048	220588	2460	1.1
Montana	2005	400	239498	229987	9511	4.0
Montana	2006	413	228347	242205	-13858	-6.1
Montana	2007	423	236993	251604	-14611	-6.2
Montana	2008	424	244114	252543	-8429	-3.5
Montana	2009	407	246233	236566	9667	3.9
Montana	2010	405	266850	234686	32164	12.1
Montana	2011	406	251981	235626	16355	6.5
Montana	2012	414	250090	243145	6945	2.8
Montana	2013	421	248039	249724	-1685	-0.7
Montana	2014	424	245858	252543	-6685	-2.7
Montana	2015	432	253017	260062	-7045	-2.8
Nebraska	2001	1021	235434	269893	-34459	-14.6
Nebraska	2002	850	266304	300833	-34529	-13.0
Nebraska	2003	850	267372	300833	-33461	-12.5
Nebraska	2004	866	282636	297938	-15302	-5.4
Nebraska	2005	876	298366	296129	2237	0.7
Nebraska	2006	886	272039	294319	-22280	-8.2
Nebraska	2007	901	283619	291605	-7986	-2.8

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Nebraska	2008	898	323726	292148	31578	9.8
Nebraska	2009	876	299292	296129	3163	1.1
Nebraska	2010	870	313066	297214	15852	5.1
Nebraska	2011	874	319228	296490	22738	7.1
Nebraska	2012	892	303014	293234	9780	3.2
Nebraska	2013	905	299774	290881	8893	3.0
Nebraska	2014	918	321449	288529	32920	10.2
Nebraska	2015	931	307034	286177	20857	6.8
Nevada	2001	866	310750	313125	-2375	-0.8
Nevada	2002	1027	315886	351711	-35825	-11.3
Nevada	2003	1062	329333	360100	-30767	-9.3
Nevada	2004	1127	358732	375678	-16946	-4.7
Nevada	2005	1197	394373	392455	1918	0.5
Nevada	2006	1253	417285	405876	11409	2.7
Nevada	2007	1265	415085	408752	6333	1.5
Nevada	2008	1234	424729	401323	23406	5.5
Nevada	2009	1118	431037	373521	57516	13.3
Nevada	2010	1088	429686	366331	63355	14.7
Nevada	2011	1095	392862	368009	24853	6.3
Nevada	2012	1112	374209	372083	2126	0.6
Nevada	2013	1140	361651	378794	-17143	-4.7
Nevada	2014	1182	355323	388860	-33537	-9.4
Nevada	2015	1224	344604	398926	-54322	-15.8
New Hampshire	2001	602	217879	224901	-7022	-3.2
New Hampshire	2002	595	212571	226271	-13700	-6.4
New Hampshire	2003	596	221050	226075	-5025	-2.3
New Hampshire	2004	605	212060	224313	-12253	-5.8
New Hampshire	2005	613	216968	222747	-5779	-2.7
New Hampshire	2006	619	225161	221572	3589	1.6
New Hampshire	2007	622	208437	220985	-12548	-6.0
New Hampshire	2008	621	242567	221181	21386	8.8
New Hampshire	2009	597	238998	225880	13118	5.5
New Hampshire	2010	593	237168	226663	10505	4.4
New Hampshire	2011	598	231835	225684	6151	2.7
New Hampshire	2012	605	230831	224313	6518	2.8
New Hampshire	2013	611	225320	223139	2181	1.0
New Hampshire	2014	619	212002	221572	-9570	-4.5
New Hampshire	2015	629	222064	219614	2450	1.1
New Jersey	2001	3809	1312381	1867124	-554743	-42.3

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
New Jersey	2002	3792	1382123	1895100	-512977	-37.1
New Jersey	2003	3787	1659898	1903328	-243430	-14.7
New Jersey	2004	3812	1478882	1862188	-383306	-25.9
New Jersey	2005	3856	1608345	1789780	-181435	-11.3
New Jersey	2006	3890	1729356	1733829	-4473	-0.3
New Jersey	2007	3900	1858396	1717373	141023	7.6
New Jersey	2008	3875	1947752	1758514	189238	9.7
New Jersey	2009	3712	1986725	2026749	-40024	-2.0
New Jersey	2010	3680	1999801	2079409	-79608	-4.0
New Jersey	2011	3687	2201474	2067890	133584	6.1
New Jersey	2012	3725	2246386	2005356	241030	10.7
New Jersey	2013	3769	2301663	1932949	368714	16.0
New Jersey	2014	3793	2348949	1893454	455495	19.4
New Jersey	2015	3841	2285378	1814465	470913	20.6
New Mexico	2001	673	158815	185892	-27077	-17.0
New Mexico	2002	680	175551	193845	-18294	-10.4
New Mexico	2003	688	188959	202935	-13976	-7.4
New Mexico	2004	703	196123	219978	-23855	-12.2
New Mexico	2005	720	230591	239294	-8703	-3.8
New Mexico	2006	748	237551	271108	-33557	-14.1
New Mexico	2007	763	238881	288151	-49270	-20.6
New Mexico	2008	766	238649	291560	-52911	-22.2
New Mexico	2009	734	246325	255201	-8876	-3.6
New Mexico	2010	720	276697	239294	37403	13.5
New Mexico	2011	723	275783	242703	33080	12.0
New Mexico	2012	726	306304	246111	60193	19.7
New Mexico	2013	734	298690	255201	43489	14.6
New Mexico	2014	743	299359	265427	33932	11.3
New Mexico	2015	752	304077	275653	28424	9.3
New York	2001	8287	2881566	3930316	-1048750	-36.4
New York	2002	8135	2976380	3369795	-393415	-13.2
New York	2003	8089	3143350	3200164	-56814	-1.8
New York	2004	8142	3278654	3395609	-116955	-3.6
New York	2005	8220	2895331	3683244	-787913	-27.2
New York	2006	8302	3520913	3985630	-464717	-13.2
New York	2007	8427	3597478	4446585	-849107	-23.6
New York	2008	8462	3899911	4575652	-675741	-17.3
New York	2009	8198	4136960	3602116	534844	12.9
New York	2010	8195	4606295	3591053	1015242	22.0

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
New York	2011	8308	5272629	4007756	1264873	24.0
New York	2012	8428	5506370	4450272	1056098	19.2
New York	2013	8549	5522078	4896476	625602	11.3
New York	2014	8710	5692894	5490185	202709	3.6
New York	2015	8878	5803753	6109708	-305955	-5.3
North Carolina	2001	3660	916541	1239303	-322762	-35.2
North Carolina	2002	3607	993658	1212526	-218868	-22.0
North Carolina	2003	3577	1077322	1197369	-120047	-11.1
North Carolina	2004	3633	1159566	1225662	-66096	-5.7
North Carolina	2005	3707	1398001	1263049	134952	9.7
North Carolina	2006	3812	1315059	1316099	-1040	-0.1
North Carolina	2007	3909	1342188	1365107	-22919	-1.7
North Carolina	2008	3866	1468254	1343381	124873	8.5
North Carolina	2009	3645	1399275	1231725	167550	12.0
North Carolina	2010	3602	1316291	1210000	106291	8.1
North Carolina	2011	3663	1421576	1240819	180757	12.7
North Carolina	2012	3729	1434643	1274164	160479	11.2
North Carolina	2013	3800	1410746	1310036	100710	7.1
North Carolina	2014	3884	1286647	1352476	-65829	-5.1
North Carolina	2015	3988	1246968	1405020	-158052	-12.7
North Dakota	2001	300	70984	71015	-31	0.0
North Dakota	2002	300	73517	71015	2502	3.4
North Dakota	2003	302	78453	72810	5643	7.2
North Dakota	2004	309	83237	79093	4144	5.0
North Dakota	2005	316	82282	85375	-3093	-3.8
North Dakota	2006	323	81297	91658	-10361	-12.7
North Dakota	2007	330	91741	97940	-6199	-6.8
North Dakota	2008	338	105837	105121	716	0.7
North Dakota	2009	337	110526	104223	6303	5.7
North Dakota	2010	345	114985	111403	3582	3.1
North Dakota	2011	366	125960	130251	-4291	-3.4
North Dakota	2012	399	151034	159868	-8834	-5.8
North Dakota	2013	414	182405	173331	9074	5.0
North Dakota	2014	432	192237	189486	2751	1.4
North Dakota	2015	424	180401	182306	-1905	-1.1
Ohio	2001	5352	2248369	2362614	-114245	-5.1
Ohio	2002	5252	2388186	2332830	55356	2.3
Ohio	2003	5202	2442187	2317938	124249	5.1
Ohio	2004	5214	2434715	2321512	113203	4.6

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Ohio	2005	5232	2447038	2326873	120165	4.9
Ohio	2006	5238	2383544	2328660	54884	2.3
Ohio	2007	5230	2478080	2326277	151803	6.1
Ohio	2008	5159	2490080	2305130	184950	7.4
Ohio	2009	4866	2353384	2217863	135521	5.8
Ohio	2010	4822	2268515	2204757	63758	2.8
Ohio	2011	4888	2203962	2224415	-20453	-0.9
Ohio	2012	4967	2196508	2247945	-51437	-2.3
Ohio	2013	5033	2083101	2267602	-184501	-8.9
Ohio	2014	5108	2039406	2289940	-250534	-12.3
Ohio	2015	5182	1929262	2311981	-382719	-19.8
Oklahoma	2001	1417	526070	699621	-173551	-33.0
Oklahoma	2002	1393	508931	704413	-195482	-38.4
Oklahoma	2003	1366	555127	709805	-154678	-27.9
Oklahoma	2004	1382	579795	706610	-126815	-21.9
Oklahoma	2005	1420	587523	699022	-111499	-19.0
Oklahoma	2006	1461	643817	690835	-47018	-7.3
Oklahoma	2007	1489	669863	685245	-15382	-2.3
Oklahoma	2008	1499	740434	683248	57186	7.7
Oklahoma	2009	1379	785218	707209	78009	9.9
Oklahoma	2010	1359	845726	711202	134524	15.9
Oklahoma	2011	1364	839922	710204	129718	15.4
Oklahoma	2012	1390	879695	705012	174683	19.9
Oklahoma	2013	1421	842466	698822	143644	17.1
Oklahoma	2014	1446	759385	693831	65554	8.6
Oklahoma	2015	1458	732542	691434	41108	5.6
Oregon	2001	1567	503895	569665	-65770	-13.1
Oregon	2002	1543	504085	558873	-54788	-10.9
Oregon	2003	1533	497612	554377	-56765	-11.4
Oregon	2004	1565	533831	568765	-34934	-6.5
Oregon	2005	1623	550878	594844	-43966	-8.0
Oregon	2006	1671	576778	616427	-39649	-6.9
Oregon	2007	1699	593872	629017	-35145	-5.9
Oregon	2008	1684	605897	622272	-16375	-2.7
Oregon	2009	1578	616869	574611	42258	6.9
Oregon	2010	1567	633054	569665	63389	10.0
Oregon	2011	1587	683452	578657	104795	15.3
Oregon	2012	1612	663181	589898	73283	11.1
Oregon	2013	1651	668686	607434	61252	9.2

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Oregon	2014	1699	655971	629017	26954	4.1
Oregon	2015	1760	631907	656445	-24538	-3.9
Pennsylvania	2001	5444	2406272	2776310	-370038	-15.4
Pennsylvania	2002	5396	2478709	2730710	-252001	-10.2
Pennsylvania	2003	5364	2565344	2700310	-134966	-5.3
Pennsylvania	2004	5390	2594238	2725010	-130772	-5.0
Pennsylvania	2005	5446	2677899	2778210	-100311	-3.7
Pennsylvania	2006	5503	2758784	2832361	-73577	-2.7
Pennsylvania	2007	5549	2803819	2876061	-72242	-2.6
Pennsylvania	2008	5535	2902243	2862761	39482	1.4
Pennsylvania	2009	5344	2901339	2681310	220029	7.6
Pennsylvania	2010	5343	2909341	2680360	228981	7.9
Pennsylvania	2011	5409	2895338	2743060	152278	5.3
Pennsylvania	2012	5458	2910221	2789610	120611	4.1
Pennsylvania	2013	5482	2974135	2812411	161724	5.4
Pennsylvania	2014	5529	2997930	2857061	140869	4.7
Pennsylvania	2015	5576	2971644	2901711	69933	2.4
Rhode Island	2001	439	137518	157396	-19878	-14.5
Rhode Island	2002	439	143894	157396	-13502	-9.4
Rhode Island	2003	443	134072	156767	-22695	-16.9
Rhode Island	2004	447	147674	156137	-8463	-5.7
Rhode Island	2005	468	142170	152831	-10661	-7.5
Rhode Island	2006	471	150999	152359	-1360	-0.9
Rhode Island	2007	470	153954	152516	1438	0.9
Rhode Island	2008	459	159550	154248	5302	3.3
Rhode Island	2009	438	160964	157554	3410	2.1
Rhode Island	2010	436	160105	157868	2237	1.4
Rhode Island	2011	437	169754	157711	12043	7.1
Rhode Island	2012	441	178681	157081	21600	12.1
Rhode Island	2013	445	170136	156452	13684	8.0
Rhode Island	2014	452	164983	155350	9633	5.8
Rhode Island	2015	459	161460	154248	7212	4.5
South Carolina	2001	1698	532374	780216	-247842	-46.6
South Carolina	2002	1677	592530	758213	-165683	-28.0
South Carolina	2003	1679	656935	760309	-103374	-15.7
South Carolina	2004	1697	688115	779169	-91054	-13.2
South Carolina	2005	1725	769553	808506	-38953	-5.1
South Carolina	2006	1759	918650	844130	74520	8.1
South Carolina	2007	1795	895503	881850	13653	1.5

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
South Carolina	2008	1780	917419	866133	51286	5.6
South Carolina	2009	1670	891830	750879	140951	15.8
South Carolina	2010	1657	891283	737258	154025	17.3
South Carolina	2011	1677	860818	758213	102605	11.9
South Carolina	2012	1704	866545	786503	80042	9.2
South Carolina	2013	1742	885307	826318	58989	6.7
South Carolina	2014	1792	895401	878706	16695	1.9
South Carolina	2015	1846	889428	935286	-45858	-5.2
South Dakota	2001	354	70736	78545	-7809	-11.0
South Dakota	2002	352	73478	77044	-3566	-4.9
South Dakota	2003	353	74241	77794	-3553	-4.8
South Dakota	2004	358	77409	81547	-4138	-5.3
South Dakota	2005	365	85889	86802	-913	-1.1
South Dakota	2006	373	109030	92806	16224	14.9
South Dakota	2007	381	119567	98811	20756	17.4
South Dakota	2008	383	111184	100312	10872	9.8
South Dakota	2009	374	93578	93557	21	0.0
South Dakota	2010	374	100348	93557	6791	6.8
South Dakota	2011	378	95373	96559	-1186	-1.2
South Dakota	2012	385	92251	101814	-9563	-10.4
South Dakota	2013	389	99084	104816	-5732	-5.8
South Dakota	2014	395	97595	109320	-11725	-12.0
South Dakota	2015	400	106594	113073	-6479	-6.1
Tennessee	2001	2479	860144	785839	74305	8.6
Tennessee	2002	2455	721733	791130	-69397	-9.6
Tennessee	2003	2453	783400	791571	-8171	-1.0
Tennessee	2004	2494	818627	782533	36094	4.4
Tennessee	2005	2537	880100	773054	107046	12.2
Tennessee	2006	2579	815808	763796	52012	6.4
Tennessee	2007	2598	751615	759607	-7992	-1.1
Tennessee	2008	2575	782894	764677	18217	2.3
Tennessee	2009	2422	781426	798405	-16979	-2.2
Tennessee	2010	2410	782091	801050	-18959	-2.4
Tennessee	2011	2433	771006	795980	-24974	-3.2
Tennessee	2012	2482	790158	785178	4980	0.6
Tennessee	2013	2526	754091	775479	-21388	-2.8
Tennessee	2014	2582	698448	763134	-64686	-9.3
Tennessee	2015	2652	687595	747704	-60109	-8.7
Texas	2001	7705	2298129	1706282	591847	25.8

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Texas	2002	7386	2370797	1750219	620578	26.2
Texas	2003	7102	1967609	1789336	178273	9.1
Texas	2004	6949	1640765	1810409	-169644	-10.3
Texas	2005	7193	1554796	1776802	-222006	-14.3
Texas	2006	7498	1416287	1734793	-318506	-22.5
Texas	2007	7636	1421056	1715786	-294730	-20.7
Texas	2008	7651	1530772	1713720	-182948	-12.0
Texas	2009	7818	1606267	1690719	-84452	-5.3
Texas	2010	8234	1483708	1633421	-149713	-10.1
Texas	2011	8334	1583205	1619648	-36443	-2.3
Texas	2012	8477	1654624	1599952	54672	3.3
Texas	2013	8678	1564956	1572268	-7312	-0.5
Texas	2014	8903	1503302	1541278	-37976	-2.5
Texas	2015	9238	1553497	1495137	58360	3.8
Utah	2001	1017	199567	219480	-19913	-10.0
Utah	2002	1006	212537	216062	-3525	-1.7
Utah	2003	1006	187182	216062	-28880	-15.4
Utah	2004	1037	216599	225695	-9096	-4.2
Utah	2005	1080	240767	239058	1709	0.7
Utah	2006	1135	261896	256150	5746	2.2
Utah	2007	1184	286757	271378	15379	5.4
Utah	2008	1182	303223	270756	32467	10.7
Utah	2009	1118	289952	250867	39085	13.5
Utah	2010	1109	257522	248070	9452	3.7
Utah	2011	1137	271124	256772	14352	5.3
Utah	2012	1177	283714	269202	14512	5.1
Utah	2013	1216	270444	281322	-10878	-4.0
Utah	2014	1253	260024	292820	-32796	-12.6
Utah	2015	1301	280124	307737	-27613	-9.9
Vermont	2001	292	98518	131817	-33299	-33.8
Vermont	2002	290	119329	132406	-13077	-11.0
Vermont	2003	288	120009	132996	-12987	-10.8
Vermont	2004	292	123823	131817	-7994	-6.5
Vermont	2005	295	122160	130934	-8774	-7.2
Vermont	2006	297	126287	130344	-4057	-3.2
Vermont	2007	297	120382	130344	-9962	-8.3
Vermont	2008	294	128305	131228	-2923	-2.3
Vermont	2009	284	144565	134174	10391	7.2
Vermont	2010	284	138370	134174	4196	3.0

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Vermont	2011	287	137359	133290	4069	3.0
Vermont	2012	290	146149	132406	13743	9.4
Vermont	2013	292	151088	131817	19271	12.8
Vermont	2014	295	150844	130934	19910	13.2
Vermont	2015	298	151544	130050	21494	14.2
Virginia	2001	3216	612083	708278	-96195	-15.7
Virginia	2002	3186	630107	678006	-47899	-7.6
Virginia	2003	3191	706110	683051	23059	3.3
Virginia	2004	3268	753409	760750	-7341	-1.0
Virginia	2005	3348	853877	841476	12401	1.5
Virginia	2006	3401	782062	894957	-112895	-14.4
Virginia	2007	3437	886657	931284	-44627	-5.0
Virginia	2008	3418	932492	912111	20381	2.2
Virginia	2009	3290	858665	782950	75715	8.8
Virginia	2010	3273	790025	765795	24230	3.1
Virginia	2011	3324	882193	817258	64935	7.4
Virginia	2012	3361	926568	854594	71974	7.8
Virginia	2013	3386	898149	879821	18328	2.0
Virginia	2014	3402	929225	895966	33259	3.6
Virginia	2015	3477	936322	971647	-35325	-3.8
Washington	2001	2622	1639435	1895110	-255675	-15.6
Washington	2002	2575	1716435	1829536	-113101	-6.6
Washington	2003	2583	1800849	1840697	-39848	-2.2
Washington	2004	2625	1836174	1899296	-63122	-3.4
Washington	2005	2697	1864015	1999750	-135735	-7.3
Washington	2006	2781	1927431	2116946	-189515	-9.8
Washington	2007	2857	1995744	2222981	-227237	-11.4
Washington	2008	2817	2192885	2167173	25712	1.2
Washington	2009	2697	2312186	1999750	312436	13.5
Washington	2010	2667	2308748	1957894	350854	15.2
Washington	2011	2773	2316713	2105785	210928	9.1
Washington	2012	2822	2311299	2174149	137150	5.9
Washington	2013	2889	2331783	2267627	64156	2.8
Washington	2014	2972	2392919	2383428	9491	0.4
Washington	2015	3049	2404364	2490859	-86495	-3.6
West Virginia	2001	664	686808	525252	161556	23.5
West Virginia	2002	661	791762	525676	266086	33.6
West Virginia	2003	656	823300	526383	296917	36.1
West Virginia	2004	665	796680	525111	271569	34.1

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
West Virginia	2005	673	695771	523980	171791	24.7
West Virginia	2006	683	433258	522567	-89309	-20.6
West Virginia	2007	684	356717	522426	-165709	-46.5
West Virginia	2008	669	319877	524546	-204669	-64.0
West Virginia	2009	650	341717	527231	-185514	-54.3
West Virginia	2010	638	362372	528927	-166555	-46.0
West Virginia	2011	662	523130	525535	-2405	-0.5
West Virginia	2012	671	476927	524263	-47336	-9.9
West Virginia	2013	665	435709	525111	-89402	-20.5
West Virginia	2014	662	419656	525535	-105879	-25.2
West Virginia	2015	658	414958	526100	-111142	-26.8
Wisconsin	2001	2630	930762	1079696	-148934	-16.0
Wisconsin	2002	2604	899700	1067377	-167677	-18.6
Wisconsin	2003	2602	843888	1066429	-222541	-26.4
Wisconsin	2004	2626	1038893	1077800	-38907	-3.7
Wisconsin	2005	2657	1188459	1092488	95971	8.1
Wisconsin	2006	2679	1043244	1102912	-59668	-5.7
Wisconsin	2007	2694	1094685	1110019	-15334	-1.4
Wisconsin	2008	2668	1156519	1097700	58819	5.1
Wisconsin	2009	2539	1114089	1036580	77509	7.0
Wisconsin	2010	2523	1070534	1028999	41535	3.9
Wisconsin	2011	2563	1099950	1047951	51999	4.7
Wisconsin	2012	2591	1123861	1061217	62644	5.6
Wisconsin	2013	2619	1166872	1074484	92388	7.9
Wisconsin	2014	2657	1204002	1092488	111514	9.3
Wisconsin	2015	2692	1169754	1109071	60683	5.2
Wyoming	2001	228	100076	101882	-1806	-1.8
Wyoming	2002	230	107475	104458	3017	2.8
Wyoming	2003	232	114252	107034	7218	6.3
Wyoming	2004	240	120062	117337	2725	2.3
Wyoming	2005	247	116528	126353	-9825	-8.4
Wyoming	2006	260	117324	143096	-25772	-22.0
Wyoming	2007	270	126996	155976	-28980	-22.8
Wyoming	2008	279	137133	167567	-30434	-22.2
Wyoming	2009	267	134835	152112	-17277	-12.8
Wyoming	2010	263	163497	146960	16537	10.1
Wyoming	2011	267	162960	152112	10848	6.7
Wyoming	2012	271	162304	157264	5040	3.1
Wyoming	2013	272	191825	158552	33273	17.3

State	Year	Average Employed	Total Paid	Predicted	Residual	% Diff
Wyoming	2014	277	184398	164991	19407	10.5
Wyoming	2015	275	178444	162416	16028	9.0