SANDIA REPORT
SAND2010-4314
Unlimited Release
Printed June 2010

Economic Impacts of a Hypothetical H1N1 Pandemic:
A Cross-Sectional Analysis

Authors
Braeton J. Smith, Calvin R. Shaneyfelt

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

U.S. Department of Energy

Office of Scientific and Technical Information

P.O. Box 62

Oak Ridge, TN 37831

Telephone: (865) 576-8401

Facsimile: (865) 576-5728

E-Mail: reports@adonis.osti.gov

Online ordering: http://www.osti.gov/bridge

Available to the public from

U.S. Department of Commerce

National Technical Information Service

5285 Port Royal Rd.

Springfield, VA 22161

Telephone: (800) 553-6847

Facsimile: (703) 605-6900

E-Mail: orders@ntis.fedworld.gov

Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2010-4314
Unlimited Release
Printed June 2010

Economic Impacts of a Hypothetical H1N1 Pandemic

A Cross-Sectional Analysis

Braeton J. Smith, Calvin R. Shaneyfelt
Computational Economics Group
Org. 06371, Infrastructure and Economic Systems Analysis
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-MS1137

Abstract

A NISAC study on the economic effects of a hypothetical H1N1 pandemic was done in order to assess the differential impacts at the state and industry levels given changes in absenteeism, mortality, and consumer spending rates. Part of the analysis was to determine if there were any direct relationships between pandemic impacts and gross domestic product (GDP) losses. Multiple regression analysis was used because it shows very clearly which predictors are significant in their impact on GDP. GDP impact data taken from the REMI PI+ (Regional Economic Models, Inc., Policy Insight +) model was used to serve as the response variable. NISAC economists selected the average absenteeism rate, mortality rate, and consumer spending categories as the predictor variables. Two outliers were found in the data: Nevada and Washington, DC. The analysis was done twice, with the outliers removed for the second analysis. The second set of regressions yielded a cleaner model, but for the purposes of this study, the analysts deemed it not as useful because particular interest was placed on determining the differential impacts to states. Hospitals and accommodation were found to be the most important predictors of percentage change in GDP among the consumer spending variables.

ACKNOWLEDGMENTS

The authors are grateful to fellow economists Mark Ehlen, Drake Warren, Verne Loose, and Vanessa Vargas of the National Infrastructure Simulation and Analysis Center (NISAC) for their invaluable commentary and support in the research and writing of this report. We are also grateful to Theresa Brown, NISAC Program Manager, under whose leadership this report was prepared. Finally, editors at SNL were invaluable in ensuring that descriptions of the analysis were clear, crisp, and accurate. We are indebted to them but do not hold them responsible for any errors, misconceptions, or omissions. Those are the responsibility solely of the authors.

TABLE OF CONTENTS

1.1 National Infrastructure Simulations and Analysis Center (NISAC) H1N1 Economic Impact Analysis 7
1.1 Purpose and Scope of this Study 7
2.1 Regression Analysis 9
2.2 Variables Used 9
2.2.1 Percentage Change in 2009 State Gross Domestic Product (GDP) Level 10
2.2.2 Absenteeism 10
2.2.3 Mortality 10
2.2.4 Consumer Spending 10
2.3 Models and Their Use of Data 12
2.3.1 Models 12
2.3.2 Data Statistics and Relationships 13

3 Regressions 22
3.1. Initial findings 23
3.1.1 Preliminary Regressions 23
3.1.2 Full Model 25
3.2 Excluding Data on Nevada and District of Colombia 26
3.2.1 Preliminary Regressions 26
3.2.2 Full Model 28

4 Summary and Conclusions 30
4.1 Summary 31
4.2 Conclusions 31

FIGURES

Figure 2-1: Gross Domestic Product (GDP) versus Predictors: Including Nevada and Washington, DC ..... **Error! Bookmark not defined.**
Figure 2-2: : Gross Domestic Product (GDP) versus Predictors: Excluding Nevada and Washington, DC ..... **Error! Bookmark not defined.**
Figure 2-3: Matrix Plot of all Variables, Nevada and Washington, DC, Included 18
Figure 2-4: Matrix Plot of all Variables, Nevada and Washington DC Removed 19
Figure 2-5: Scatterplot of Gross Domestic Product (GDP) versus Important Consumer Spending Variables, Nevada and Washington, DC, Included 20

TABLES

Table 2-1: Selected Consumer Spending Categories and Corresponding REMI Sectors 12
Table 2-2: Descriptive Statistics of Variables, Nevada and Washington, DC, Included 16
Table 2-3: Descriptive Statistics of Variables, Nevada and Washington, DC, Removed 17
Table 3-1: Individual Regressions of Three Main Predictor Variables, Initial Findings 24
Table 3-2: Regression of Main Predictor Variables, Initial Findings 24

Table 3-3: Individual Regressions of Two Consumer Spending Variables, Initial Findings 24
Table 3-4: Regression of Important Consumer Spending Variables, Initial Findings.............................. 25
Table 3-5: Regression of Consumer Spending Variables, Initial Findings 25
Table 3-6: Full Model, Initial Findings.. 26
Table 3-7: Individual Regressions of Three Main Predictor Variables, Nevada and Washington, DC, Removed 27
Table 3-8: Regression of Main Predictor Variables, Nevada and Washington, DC, Removed 27
Table 3-9: Regression without Consumer Spending, Nevada and Washington, DC, Removed 27
Table 3-10: Individual Regressions of Two Consumer Spending Variables, Nevada and Washington, DC, Removed 28
Table 3-11: Full Model, Nevada and Washington, DC, Removed 28
Table 3-12: Comparison of Full Models.. 29

Acronyms and Abbreviations

Acronym	Definition
DHS	U.S. Department of Homeland Security
EpiSimS	Epidemiological Simulation System
GDP	Gross Domestic Product
LANL	Los Alamos National Laboratory
NAICS	North American Industry Classification System
NBIC	National Biosurveillance Integration Center
NISAC	National Infrastructure Simulation and Analysis Center
REMI	Regional Economic Models, Incorporated
REMI PI+	Regional Economic Models, Inc., Policy Insight +
SNL	Sandia National Laboratories

1 Introduction

1.1 National Infrastructure Simulation and Analysis Center (NISAC) H1N1 Economic Impact Analysis

The National Biosurveillance Integration Center (NBIC) of the U.S. Department of Homeland Security (DHS) tasked the economists at the National Infrastructure Simulation and Analysis Center (NISAC) to answer questions regarding the effects on the economy in the event of an A (H1N1), or *swine flu*, pandemic influenza[1].

The REMI PI+[2] model (Regional Economic Models Inc., Policy Insight +) is used by Sandia National Laboratories (SNL) to carry out some of its economic analyses. Part of the SNL REMI analysis projected changes in the level of gross domestic product (GDP), given certain levels of absenteeism, mortality, and consumer spending—all three of which served as input to the REMI PI+ model. The SNL analysts manipulated the absenteeism and mortality inputs to REMI based on results produced by NISAC analysts at Los Alamos National Laboratory (LANL) via their Epidemic Simulation System (EpiSimS)[3].

A regression analysis was requested to form a more comprehensive picture of the pandemic's effect on GDP. This report describes the results of the regression analysis performed on the REMI results to give a clearer idea of the different impacts experienced by each state, while also showing which economic sectors are most important. Specifically, regression analysis aids in determining which factors are most important in predicting and quantifying GDP losses for each state as a result of the pandemic. It also gives insight as to the differential impacts at the state level, making it far easier to see how different factors and industries affect individual states. Policy makers may find this information useful in conducting more effective legislation with a better understanding of the individual effects to each state. Breaking down the analysis to this level allows policy makers to more accurately and efficiently direct their decisions and legislation.

1.1 Purpose and Scope of this Study

The purposes of this study are to gain an understanding of the different impacts to specific states and to show which economic sectors are most affected by a hypothetical H1N1 pandemic influenza. Section 2 discusses basic theory behind multiple regression analysis and describes the methodology used to calculate inputs to the regression models. It also provides some discussion of the data. Section 3 describes the regressions and presents their analysis and results. Finally, Section 4 summarizes and presents conclusions from the process of research and analysis for the study.

---

1 For the complete study, see: National Infrastructure Simulation and Analysis Center, "2009 H1N1 Influenza Study: Impacts on the United States Population and Economy," Department of Homeland Security, Office of Health Affairs, Office of Infrastructure Protection, August 2009.

2 Regional Economic Models, Inc. "REMI PI+", v. 1.0.114, March 24, 2009 build, 51 region, 70 sector model, Amherst, MA.

3 Del Valle, SY, PD Stroud, JP Smith, SM Mniszewski, JM Riese, SJ Sydoriak, DA Kubicek. 2006. EpiSimS: Epidemic Simulation System. Los Alamos Unlimited Release (LAUR) 06-06714.

This page intentionally left blank.

2 Methodology

This section presents the basic theory for multiple linear regression, which is used in order to determine statistical relationships that predict basic outcomes, given changes to certain significant variables. This section also discusses the selection and calculation of variables included in the analysis and includes descriptions of the software models used (REMI and EpiSimS) to generate them. It also includes a basic discussion of the data.

2.1 Regression Analysis

Linear regression is a statistical modeling procedure that examines linear relationships between two or more variables by building an equation that represents the relationship. In *multiple linear regression*, a *response* variable (the dependent, "left-hand," or *y* variable) is held as dependent upon several *predictor* variables (the independent, "right-hand," or *x* variables). The equation is linear because each of the coefficients depicts a linear relationship when regressed singularly (and all are to the first power[4]). For this reason, they are sometimes called the *partial* regression coefficients to emphasize the fact that when all but one factor is held constant, the remaining factor measures all of the expected change in the response variable. The basic equation is of the linear form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

where:

- y is the response variable,
- β_1 through β_k are the coefficients associated with each factor (or predictor variable), and
- x_1 through x_k are vectors which contain the data pertaining to each factor, and
- ε is the error term.

There are a few assumptions associated with linear regression, namely

- Every variable is independent;
- The residuals, or the difference between the expected outcome and the observed outcome, must all sum to zero and be approximately normally distributed;
- The predictor variables must also be independent of one another, or un-correlated.

When regressing over time, there is also the common problem of *autocorrelation*, or the correlation of residuals, which has been eliminated in this analysis because the variables were regressed under a single period (called "cross-sectional" analysis). For this analysis, the percentage change in state GDP for 2009 as a result of a pandemic was regressed as a function of expected absenteeism and mortality rates as well as predicted changes in consumer spending for certain sectors.

2.2 Variables Used

The data used for this analysis were from output of both the EpiSimS and REMI models. They comprise information on expected changes in the 48 continental states and Washington, DC. Two outliers are immediately noticeable in the data (**Error! Reference source not found.** and **Error! Reference source**

---

[4] A "power" is the exponent to which a number is raised. If a number is raised to the first power, i.e. x^1, then it is equal to itself. In this case, if all variables x_1, x_2, …, x_k are to the first power, then the function is linear (as opposed to exponential).

not found.), representing Nevada and Washington, DC. These outliers are included in the first part of the regression analysis and then removed.

2.2.1 Percentage Change in 2009 State Gross Domestic Product (GDP) Level

The percentage change in state GDP is the REMI output for the changes in absenteeism, mortality, and consumer spending that were input to the model. This acts as the *response* or *dependent* variable in the model.

2.2.2 Absenteeism

The absenteeism rate is defined as the percentage of laborers absent from work. In the event of a pandemic, it is expected that the absenteeism rate would increase. The absenteeism rates were taken from the EpiSimS output that acted as one of the inputs to REMI. Because each industry has a different rate, each rate was multiplied by the total workers in each sector to incorporate the relative size of each industry. The values were then summed for each state and divided by the amount of total workers in each state. These values, the "average absenteeism" rates for each state, were used in this analysis.

2.2.3 Mortality

The mortality rate is defined as the percentage of people dead in each state as a result of the pandemic. The mortality rates for this analysis were calculated by taking the EpiSimS output for total dead in each state and dividing them by the total state populations found in REMI's base model. These values also served as inputs to REMI.

2.2.4 Consumer Spending

Change in consumer spending captures people's change in purchasing behavior. In the event of a pandemic, analysts expect that consumers will decrease spending in several areas of the economy while increasing spending in sectors like health care. Because H1N1 has a low mortality rate (and high morbidity rate), changes in consumer spending are not thought to be as great. For that reason, the initial inputs to REMI only had consumer spending changing by 1 percent for several selected REMI-specific industries. The initial list follows:

- Purchased meals and beverages
- Women's and children's clothing and accessories except shoes
- Men's and boys' clothing and accessories except shoes
- Drug preparations and sundries
- Other housing (hotels and other lodging places)
- Airline
- Physicians
- Nonprofit hospitals
- Proprietary hospitals
- Government hospitals
- Motion picture admissions

- Legitimate theater admissions
- Spectator sports admissions
- Commercial participant amusements
- Foreign travel by U.S. residents
- Foreign travel in the United States and other expenditures in the United States by nonresidents

Using REMI's National Input-Output (IO) Matrix to determine which parts of each sector would actually be affected, analysts shortened the list to the following (each listing corresponds directly to one of the above industries):

- Food services and drinking places
- Chemical manufacturing
- Accommodation
- Air transportation
- Ambulatory health care
- Hospitals
- Motion picture and sound recording
- Performing arts and spectator sports
- Amusement, gambling, and recreation
- Miscellaneous manufacturing

Analysts used these 10 variables to describe consumer spending in the model. Because the REMI inputs for each state and industry were the same (1 percent change in consumer spending), the analysts divided the value-added (from REMI) for each of the industries in each state by the GDP level for that state. Dividing by the GDP level incorporates the relative size (percentage) of each industry in each state. Analysts then aggregated the ten variables into the variable "total," standing for total consumer spending, which was also used in the analysis.

Table 2-1 shows how each consumer spending category corresponds to one of the industry sectors selected for this analysis.

Table 2-1: Selected Consumer Spending Categories and Corresponding REMI Sectors

REMI Consumer Spending Category	Primary REMI Industry Sector
Purchased meals and beverages	Food services and drinking places
Women's and children's clothing and accessories except shoes	--
Men's and boys' clothing and accessories except shoes	--
Drug preparations and sundries	Chemical manufacturing
Other housing (hotels and other lodging places)	Accommodation
Airline	Air transportation
Physicians	Ambulatory health care
Nonprofit hospitals	Hospitals
Proprietary hospitals	Hospitals
Government hospitals	Hospitals
Motion picture admissions	Motion picture and sound recording
Legitimate theater admissions	Performing arts and spectator sports
Spectator sports admissions	Performing arts and spectator sports
Commercial participant amusements	Amusement, gambling, and recreation
Foreign travel by U.S. residents	Miscellaneous manufacturing
Foreign travel in the U.S. and other expenditures in U.S. by nonresidents	--

2.3 Models and Their Use of Data

2.3.1 Models

EpiSimS is an epidemiological tool used to simulate and forecast the spread of disease in urban areas based on the demographic and geographic distribution of disease. The model works via a social network representing details of contacts between individuals based on their activity patterns. It provides estimates for how rapidly the disease spreads and shows the vulnerability of people, depending on the statistics of a specific disease. For this study, EpiSimS was used to estimate absenteeism and mortality data to be input to REMI.

The REMI model uses a built-in baseline scenario that models the current U.S. economy and may be altered based on modeled changes that reflect direct impacts to the economy. The model divides the economy into 51 regions (the 50 states and Washington, DC) and 70 sectors based on codes from the North American Industry Classification System (NAICS), though the industry classifications themselves are specific to REMI. Being organized this way allows REMI to give detailed output on specific effects to the economies of each state and to each industry (both as a whole and within each state), for any given impact. REMI also projects this change into the future (scenarios are run over a specified period) and notes effects in several other areas of the economy.

For this analysis, NISAC analysts ran the model based on an H1N1 pandemic impact reflecting changes in the absenteeism and mortality rates (both of which were calculated from the EpiSimS output) and in the selected consumer spending categories (listed in section 2.2.4. and in the left column of Table 2-1). As stated previously, these consumer spending categories correspond directly to specific REMI industries and changes in those consumer spending categories reflect changes in their corresponding industries. For this analysis, the REMI output used were the percentage change in GDP for each state in 2009 (current year) and the percentage change in each industry in each state as a result of the consumer spending. The REMI output for the consumer spending changes was used in the regression analysis, despite its being produced by the same algorithm (the reasoning for this is discussed in the Conclusions section of this report).

2.3.2 Data Statistics and Relationships

As stated in the introduction paragraph for section 2.2, there are two notable outliers in the data, representing Nevada and Washington, DC. For that reason, there were two separate sets of analysis done: one including the two outliers and the other with them removed. The descriptive statistics and matrix plots for both scenarios are shown below in Tables 2-2 and 2-3 and in Figures 2-3 and 2-4.

Perhaps the most noticeable difference between **Error! Reference source not found.** and **Error! Reference source not found.** is the scatterplot of mortality (which may actually be more accurately viewed here as a *survival* rate, as the rates are negative). The data are split into definite columns because the states were grouped into regions that had common mortality rates (Nevada, the point in the lower right corner of Figure 2-1, was clearly given its own mortality rate). The obvious changes between the two figures (by the removal of the outliers) are the relationships with percent GDP change of both mortality and total change in consumer spending. The presence of Nevada in Figure 2-1 seems to change the general correlation altogether (from positive to negative for mortality and from no relationship to negative for consumer spending). Figure 2-2 clarifies the actual relationships after the removal of the outliers.

Error! Reference source not found. and **Error! Reference source not found.** show the relationships of all of the variables to each other. As seen in the column labeled "Accommodation" in Figure 2-3, the outliers have a significant effect on its relationships with the other variables. It is not surprising, then, that it was found to be the most important consumer spending variable in the initial analysis (discussed in further detail in Chapter 0, the other being hospitals). Also, in looking at **Table 2-2** and Table 2-3, the descriptive statistics of the variables, note that there are no significant differences between the two tables except for the statistics on accommodation. Table 2-2 shows that accommodation has a mean of .0101, while in Table 2-3, it has a mean of only .0073. In Figures 2-4 and 2-6, accommodation no longer shows a relationship with anything, which is confirmed by the regression in section 3.2.1, Table 3-10.

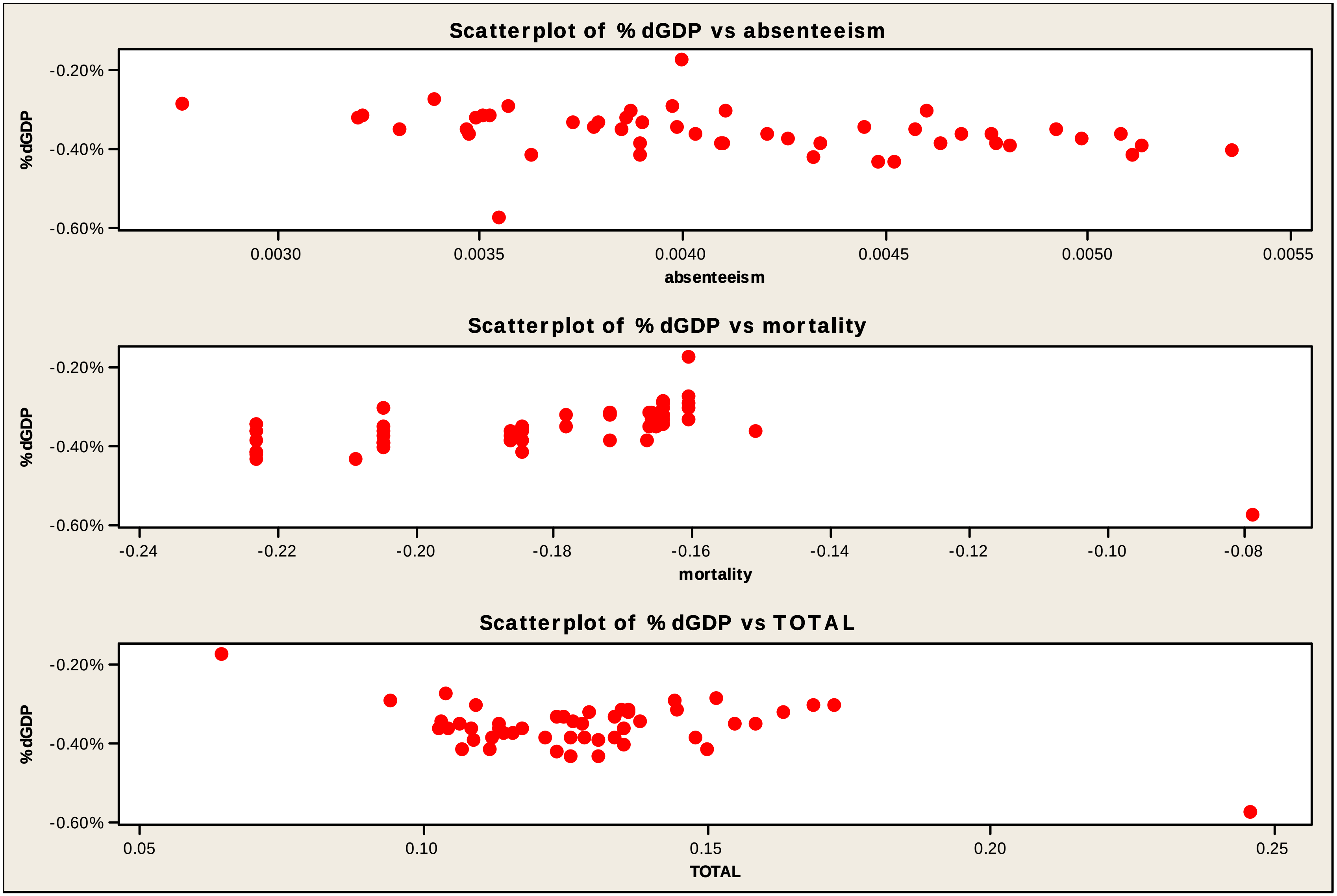

Figure 3-1: Gross Domestic Product (GDP) versus Predictors: Including Nevada and Washington, DC

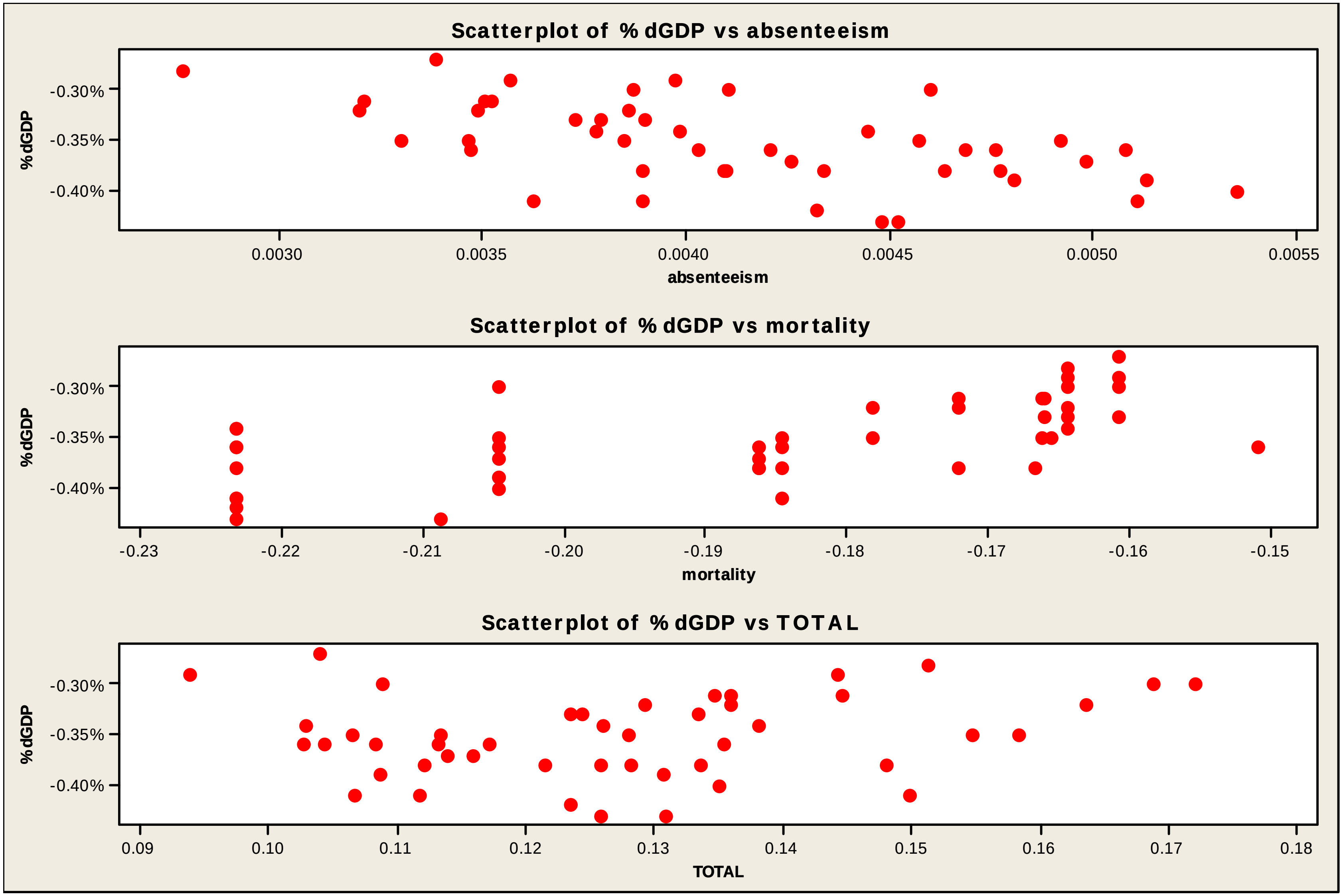

Figure 3-2: Gross Domestic Product (GDP) versus Predictors: Excluding Nevada and Washington, DC

Table 2-2: Descriptive Statistics of Variables, Nevada and Washington, DC, Included

Variable	Mean	StDev	Variance	Minimum	Median	Maximum
% Change in GDP	-0.003529	0.000571	0.000000	-0.005700	-0.003500	-0.001700
Absenteeism	0.004099	0.000595	0.000000	0.002765	0.003996	0.005355
Mortality	-0.18419	0.02776	0.00077	-0.22327	-0.18463	-0.07910
Food services and drinking places	0.017425	0.002473	0.000006	0.012054	0.017535	0.022406
Chemical manufacturing	0.01531	0.01188	0.00014	0.00090	0.01308	0.05162
Accommodation	0.01008	0.01912	0.00037	0.00293	0.00590	0.13667
Air transportation	0.005272	0.004647	0.000003	0.000375	0.003981	0.018170
Ambulatory health care	0.04185	0.00744	0.00006	0.01384	0.04253	0.05370
Hospitals	0.020480	0.006759	0.000046	0.006382	0.018982	0.035824
Motion picture and sound recording	0.001240	0.002132	0.000005	0.000275	0.000684	0.014110
Performing arts and spectator sports	0.002434	0.001598	0.000003	0.000416	0.002217	0.007169
Miscellaneous manufacturing	0.007755	0.005616	0.000032	0.000102	0.005779	0.027664
Amusement, gambling, and recreation	0.006881	0.003060	0.000009	0.001430	0.005844	0.015144
TOTAL	0.12873	0.02659	0.00071	0.06434	0.12603	0.24569

Table 2-3: Descriptive Statistics of Variables, Nevada and Washington, DC, Removed

Variable	Mean	StDev	Variance	Minimum	Median	Maximum
% Change in GDP	-0.003521	0.000406	0.000000	-0.004300	-0.003500	-0.002700
Absenteeism	0.004113	0.000602	0.000000	0.002765	0.004027	0.005355
Mortality	-0.18692	0.02334	0.00054	-0.22327	-0.18463	-0.15102
Food services and drinking	0.017433	0.002283	0.000005	0.012519	0.017535	0.022304
Chemical manufacturing	0.01591	0.01176	0.00014	0.00109	0.01310	0.05162
Accommodation	0.007314	0.005022	0.000025	0.002926	0.005762	0.027049
Air transportation	0.005364	0.004690	0.000022	0.000408	0.003981	0.018170
Ambulatory health care	0.042621	0.006219	0.000039	0.030435	0.042563	0.053704
Hospitals	0.020768	0.006734	0.000045	0.006382	0.019257	0.035824
Motion picture and sound recording	0.001238	0.002178	0.000005	0.000275	0.000650	0.014110
Performing arts and spectator sports	0.002345	0.001570	0.000002	0.000416	0.002027	0.007169
Miscellaneous manufacturing	0.007791	0.005555	0.000031	0.000557	0.005779	0.027664
Amusement, gambling, and recreation	0.006830	0.002790	0.000008	0.003849	0.005844	0.015144
TOTAL	0.12761	0.01868	0.00035	0.09383	0.12603	0.17208

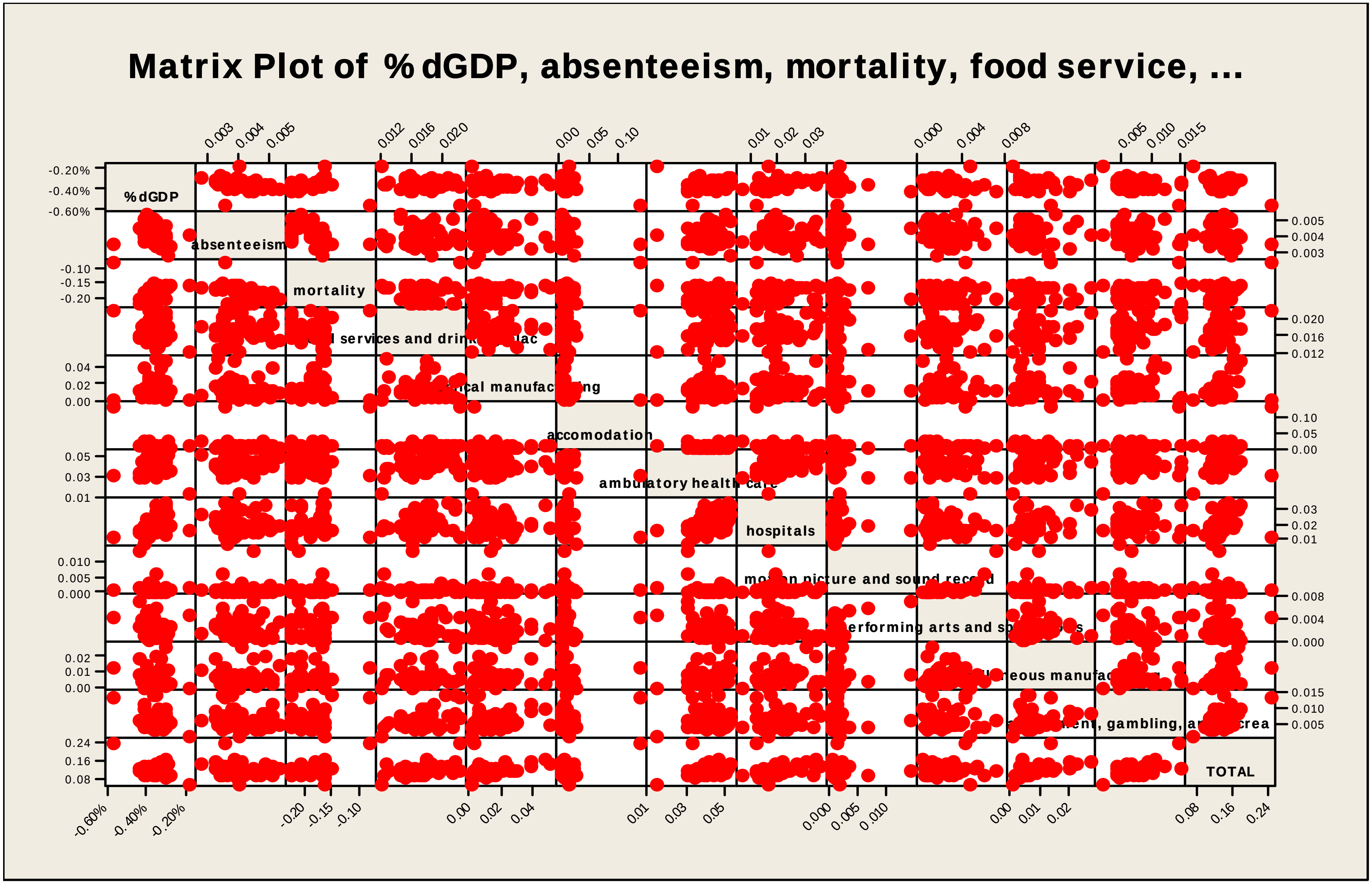

Figure 3-3: Matrix Plot of all Variables, Nevada and Washington, DC, Included

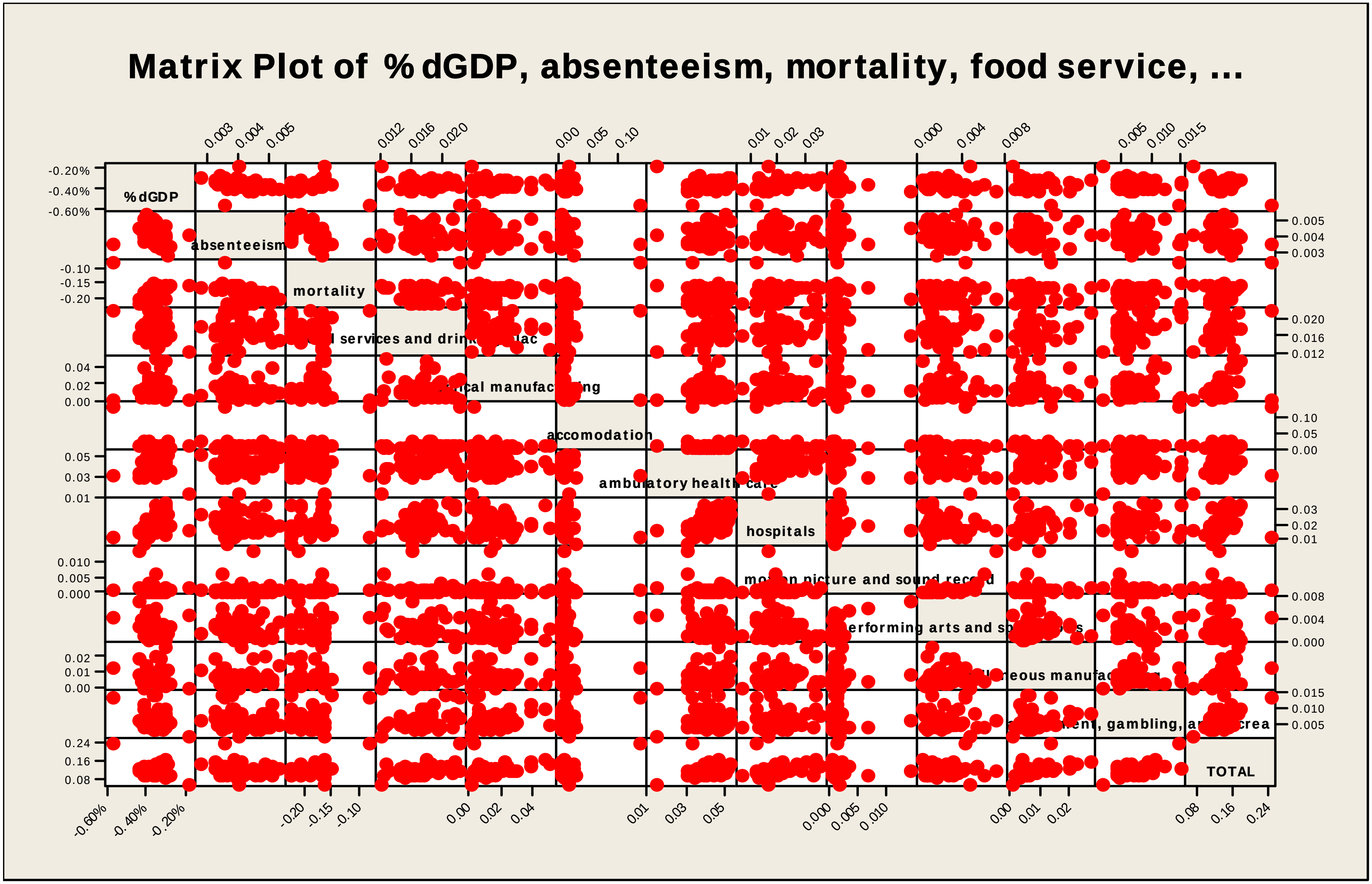

Figure 3-4: Matrix Plot of all Variables, Nevada and Washington, DC, Removed

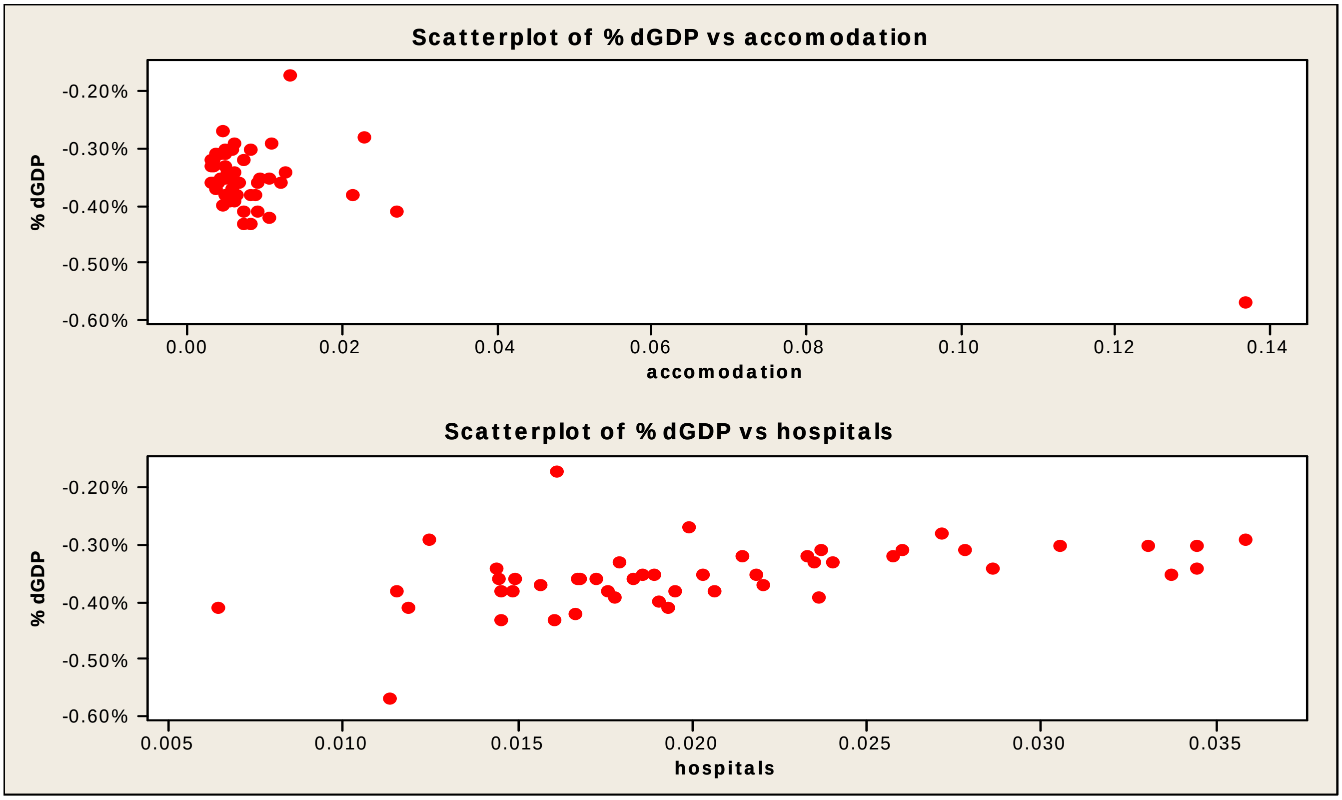

Figure 3-5: Scatterplot of Gross Domestic Product (GDP) versus Important Consumer Spending Variables, Nevada and Washington, DC, Included

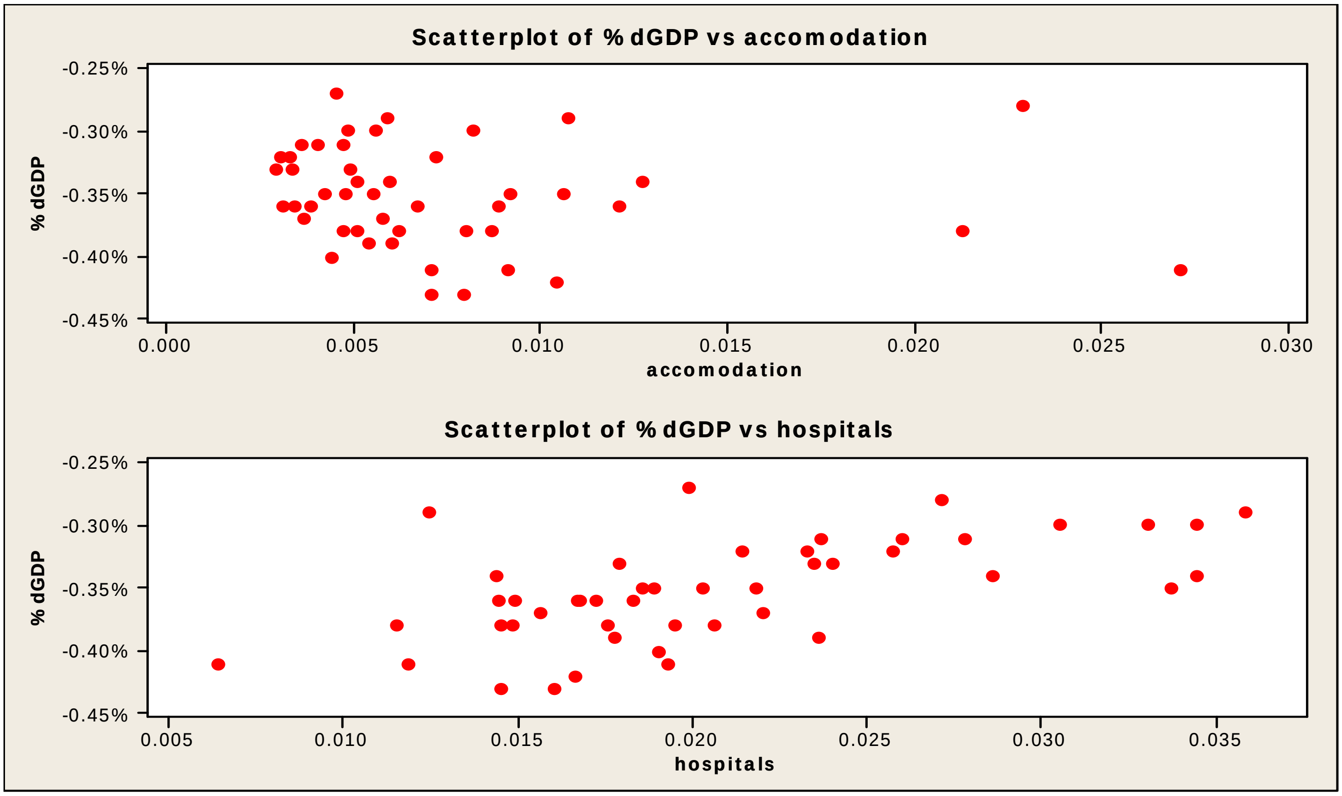

Figure 3-6: Scatterplot of Gross Domestic Product (GDP) versus Important Consumer Spending Variables, Nevada and Washington, DC, Removed

This page intentionally left blank.

3 Regressions

3.1. Initial findings

As stated previously, there are two outliers contained in the data, Nevada and Washington, DC. These outliers, which greatly influence the results, are included in the initial analysis. Instead of nullifying the initial results, the outliers can be viewed as showing explicitly the effects regulation will have on the two areas. When included in the regressions, the areas also show their impacts to the rest of the United States.

3.1.1 Preliminary Regressions

To begin, the NISAC analysts performed singular regressions on the three variables absenteeism, mortality, and total consumer spending, to determine which had the greatest effect. Those with the higher *R*-squared value, or those that explain the higher percentage of variation in percentage change GDP, are typically found to be the more important predictor variables. The *R*-squared value is a measure of the fit of the model to the data. Of even more importance is the *p*-value, which must be low (lower than the significance level α, generally 0.05) in order for the variable to be of significance in any given model. In this first series of regressions (there are three, one for each variable), only mortality is insignificant and total consumer spending is the most significant, alone explaining 17 percent of the variation in the percentage change in state GDP for 2009 (Table 3-1). Next, the three variables were run together in the same model (Table 3-2). Though the model has an *R*-squared value of 40.2 percent, mortality is only significant at the 0.1 level.

The next stage of the analysis was to evaluate the important consumer spending variables instead of using the all-encompassing “total” variable. Analysts ran all 10 variables together, of which only 6 were significant. Removing the other four yielded a model explaining 66.4 percent of the variation in percentage change in GDP (Table 3-5). The two most important factors were “accommodation” and “hospitals,” with *R*-squared values of 29.4 percent and 22 percent, respectively, when regressed on their own (Table 3-3). For “accommodation,” the strong negative relationship found by the model can be attributed to the outlier of Nevada. The two variables were then incorporated into the same model (Table 3-4), which explains 42.2 percent of the variation in percentage change GDP, better than the first model above (Table 3-2) as it contains all highly significant components.

Table 3-1: Individual Regressions of Three Main Predictor Variables, Initial Findings

Predictor	Coefficient	p-value	
Constant	-0.0022476	0.000	
Absenteeism	-0.3125	0.022	
		R-sq	*10.6%*
Constant	-0.0031033	0.000	
Mortality	0.002309	0.442	
		R-sq	*1.3%*
Constant	-0.0023876	0.000	
Total consumer spending	-0.008863	0.003	
		R-sq	*17.0%*

Table 3-2: Regression of Main Predictor Variables, Initial Findings

Predictor	Coefficient	*p*-value	
Constant	0.0005870	0.476	
Absenteeism	-0.3201	0.016	
Mortality	0.005705	0.074	
Total consumer spending	-0.013616	0.000	
		R-sq	*40.2%*

Table 3-3: Individual Regressions of Two Consumer Spending Variables, Initial Findings

Predictor	Coefficient	*p*-value	
Constant	-0.00336555	0.000	
Accommodation	-0.016179	0.000	
		R-sq	*29.4%*
Constant	-0.0043403	0.000	
Hospitals	0.03964	0.001	
		R-sq	*22.0%*

Table 3-4: Regression of Important Consumer Spending Variables, Initial Findings

Predictor	Coefficient	*p*-value	
Constant	-0.0040254	0.000	
Accommodation	-0.013752	0.000	
Hospitals	0.031025	0.003	
		R-sq	*42.2%*

Table 3-5: Regression of Consumer Spending Variables, Initial Findings

Predictor	Coefficient	*p*-value	
Constant	-0.0024683	0.000	
Accommodation	-0.011651	0.001	
Air transportation	-0.02900	0.017	
Ambulatory health care	-0.030240	0.001	
Hospitals	0.045022	0.000	
Motion picture and sound recording	-0.06652	0.011	
Amusement, gambling, and recreation	-0.05289	0.009	
		R-sq	*66.4%*

3.1.2 Full Model

To complete the initial analysis, absenteeism and mortality were added back into the model with the significant consumer spending variables (of which now there are seven). The full model explains 85.4 percent of the variation in the percentage change in 2009 state-level GDP with all highly significant factors (except for the constant). Although this model is less *parsimonious*[5] than the initial model that used the "total consumer spending" variable, it is much stronger. Mortality is now a highly significant factor (with a *p*-value of 0.001) and the model has a higher *R*-squared value by over 45 percent (Table 3-6). This model, however, improves greatly when the outliers are dropped.

[5] In statistics terminology, the word *parsimonious* refers to the simplicity of the model, or the number of factors. Typically, the most parsimonious model is used as it explains very nearly the same amount of variation explained by the chosen factors (R^2), but with less factors (only including those to be found significant with a *p*-value below a certain significance level α).

Table 3-6: Full Model, Initial Findings

Predictor	Coefficient	*p*-value	
Constant	0.0001386	0.761	
Absenteeism	-0.27826	0.000	
Mortality	0.006744	0.001	
Chemical manufacturing	-0.011467	0.002	
Accommodation	-0.021147	0.000	
Air transportation	-0.033403	0.006	
Ambulatory health care	-0.023273	0.000	
Hospitals	0.029583	0.000	
Motion picture and sound recording	-0.06047	0.001	
Amusement, gambling, and recreation	-0.04015	0.000	
		R-sq	*85.4%*

3.2 Excluding Data on Nevada and District of Colombia

The analysis was repeated a second time after the outliers from the initial analysis, Nevada and Washington, DC, were removed. The new regressions show some striking differences.

3.2.1 Preliminary Regressions

Having removed the outliers from the data, analysts ran the same series of regressions again. Analysts regressed the three variables (absenteeism, mortality, and total consumer spending) singularly, and then regressed them together, yielding notably different results from the previous analysis. Mortality becomes highly significant on its own, with a *p*-value of 0 to 3 decimal places and alone explaining about 39 percent of the variation in the percentage change in GDP (which is almost as high as the initial model in the previous analysis). Absenteeism is also highly significant, much more than in the previous analysis. Total consumer spending, which was previously the most important predictor, is now completely insignificant with a high *p*-value and explains only 4.2 percentage of the variation. It also has a slightly positive slope, as opposed to its negative slope in the previous analysis (Table 3-7). Because consumer spending is still insignificant when the three variables are run together (Table 3-8), it was not included in the regression. Regressing only absenteeism and mortality explains about 47 percent of the variation in percentage change in state GDP for 2009 (Table 3-9).

Breaking up the consumer spending variables also shows different results (as shown by the change in results for the “total consumer spending” variable). In a regression including all consumer spending variables, only “hospitals” is significant; it, alone, has an *R*-squared of almost 36 percent, compared with accommodation, which is no longer significant and explains just over 2 percent of the variation (Table 3-10). “Accommodation” is, however, included in the full model.

Table 3-7: Individual Regressions of Three Main Predictor Variables, Nevada and Washington, DC, Removed

Predictor	Coefficient	*p*-value	
Constant	-0.0019596	0.000	
Absenteeism	-0.37970	0.000	
		R-sq	*31.7%*
Constant	-0.0014943	0.000	
Mortality	0.010844	0.000	
		R-sq	*38.9%*
Constant	-0.0040889	0.000	
Total Consumer Spending	0.004448	0.168	
		R-sq	*4.2%*

Table 3-8: Regression of Main Predictor Variables, Nevada and Washington, DC, Removed

Predictor	Coefficient	*p*-value	
Constant	-0.0009381	0.139	
Absenteeism	-0.22420	0.014	
Mortality	0.008188	0.001	
Total consumer spending	-0.001023	0.696	
		R-sq	*47.1%*

Table 3-9: Regression without Consumer Spending, Nevada and Washington, DC, Removed

Predictor	Coefficient	*p*-value	
Constant	-0.0011284	0.006	
Absenteeism	-0.22290	0.013	
Mortality	0.007897	0.001	
		R-sq	*46.9%*

Table 3-10: Individual Regressions of Two Consumer Spending Variables, Nevada and Washington, DC, Removed

Predictor	Coefficient	*p*-value	
Constant	-0.0034328	0.000	
Accommodation	-0.01210	0.315	
		R-sq	*2.2%*
Constant	-0.0042681	0.000	
Hospitals	0.035960	0.000	
		R-sq	*35.6%*

3.2.2 Full Model

The full model, as in the previous analysis, includes absenteeism, mortality, and the significant consumer spending variables (Table 3-11). This model is more parsimonious than the initial full model as it has one less consumer spending variable, ambulatory health care. All of the components are highly significant, including the constant. Although the model accounts for less of the variation in the percentage change in state GDP level for 2009 as the model from the initial analysis (82.2 percent as opposed to 85.4 percent), this is the better, more accurate model due to the consistency in the data. The residual plots show they are distributed closer to normal than in the previous analysis and the plot of residuals versus the fitted values is more random. The fact this model is the best, however, does not imply the other is useless. In terms of this analysis, the first model is probably more useful, as it shows the differing effects across the states. This is discussed in greater detail in Section 0. Table 3-12 shows a comparison of the two full models.

Table 3-11: Full Model, Nevada and Washington, DC, Removed

Predictor	Coefficient	*p*-value	
Constant	-0.0012026	0.002	
Absenteeism	-0.30651	0.000	
Mortality	0.005238	0.003	
Chemical manufacturing	-0.007441	0.009	
Accommodation	-0.014937	0.023	
Air transportation	-0.019915	0.005	
Hospitals	0.023362	0.022	
Motion picture and sound recording	-0.04056	0.003	
Amusement, gambling, and recreation	-0.02624	0.000	
		R-sq	*82.2%*

Table 3-12: Comparison of Full Models

Initial Model				Model with Nevada and Washington, DC Removed			
Predictor	**Coefficient**	***p*-value**		**Predictor**	**Coefficient**	***p*-value**	
Constant	0.0001386	0.761		Constant	-0.0012026	0.002	
Absenteeism	-0.27826	0.000		*Absenteeism*	-0.30651	0.000	
Mortality	0.006744	0.001		*Mortality*	0.005238	0.003	
Chemical manufacturing	-0.011467	0.002		*Chemical manufacturing*	-0.007441	0.009	
Accommodation	-0.021147	0.000		*Accommodation*	-0.014937	0.023	
Air transportation	-0.033403	0.000		*Air transportation*	-0.019915	0.005	
Ambulatory health care	-0.023273	0.000		--	--	--	
Hospitals	0.029583	0.000		*Hospitals*	0.023362	0.000	
Motion picture & sound recording	-0.06047	0.001		*Motion picture & sound recording*	-0.04056	0.003	
Amusement, gambling, & recreation	-0.04015	0.006		*Amusement, gambling, & recreation*	-0.02624	0.022	
		R-sq:	*85.4%*			*R-sq:*	*82.2%*

This page intentionally left blank.

4 Summary and Conclusions

4.1 Summary

Analysts carried out a study on the economic effects of a hypothetical H1N1 pandemic to assess the differential impacts at the state and industry levels, given changes in absenteeism, mortality, and consumer spending rates. Part of the analysis, which is included in this report, was to determine if there are any direct relationships between pandemic impacts and GDP losses. To do this, analysts used multiple regression analysis because it shows very clearly which predictors are significant in their impact on GDP, and it provides clear graphics that make it easy to see relationships and determine outliers in the data.

GDP impact data were taken from the REMI PI+ model to serve as the response (or dependent) variable. Analysts used the average absenteeism rate, mortality rate, and several consumer spending categories selected by NISAC economists as the predictor variables. Analysts derived absenteeism and mortality rates from EpiSimS output and transformed the consumer spending categories into their industry sector counterparts via REMI's National IO Matrix. All of the predictor variables were used as REMI inputs in the primary part of this study (the GDP data were part of the output).

Because there were two outliers in the initial data, Nevada and Washington, DC, analysts ran the analysis two times and removed the outliers for the second run. The second set of regressions yielded a much cleaner model; however, for the purposes of this study, it was not deemed as useful because particular interest was placed on determining the differential impacts to states. Hospitals and accommodation were found to be the most important predictors of percentage change in GDP among the consumer spending variables.

4.2 Conclusions

Although the final model as described in Section 3.2.2 (Table 3-11) is the best, it is not the most useful in terms of this analysis. Typically, regressions are used to analyze relationships and to form predictions using certain variables, but in this case, the regressions may not be used explicitly for that purpose. Because the inputs to REMI are the predictor (right-hand) variables and the output from REMI is the response variable, there is an inherent bias among them. The variables are not independent of one another (inputs to REMI used to predict output from REMI, clearly a relationship will exist), making the regressions *biased* and somewhat invalid. They may, however, be used differently.

That being said, the first part of the analysis (in Section 3.1), which includes the outliers Nevada and Washington, DC, is the most useful. Because the purpose of this analysis is to assess the differential impacts at the state level, the regressions are meant to serve this purpose. The scatter plots for the three major variables (absenteeism, mortality, and total consumer spending) against the percentage change in GDP show (in the initial analysis) Nevada as having the highest and Washington, DC, as having the lowest percentage GDP loss of all the states. This relationship is generally true of each consumer spending variable (Nevada typically with the highest value of change and Washington, DC with the least), particularly for variables such as accommodation and amusement, gambling, and recreation. Nevada, whose economy is largely based on tourism-related activities, is expected to suffer in terms of GDP loss where other states are generally not as adversely affected. (Note, however, that Nevada is not the state with the greatest GDP loss, but with the greatest *percentage* GDP loss.) Washington, DC, on the other hand, is only minimally affected. An explanation for this could be that the economy of Washington, DC is not as dependent on consumption and the other variables because most of its spending is related to government.

Insofar as the purpose of this report is to illustrate the different impacts across the states and industry sectors, it succeeds to a large degree. Nevada is by far the most affected by a hypothetical H1N1 pandemic, and the effect to Washington, DC is significantly smaller than to the rest of the states. The two industries most affected by change in consumer spending are hospitals and accommodation (the latter of which is no longer significant with the removal of Nevada and Washington, DC from the data). It is important to note, however, that the negative effects of H1N1 to GDP are minimal across the board.

REFERENCES AND BIBLIOGRAPHY

1. American Hospital Association. April 2008. Beyond Health Care: The Economic Contribution of Hospitals. *Trendwatch*, 6:1, *www.aha.org/aha/**trendwatch**/2008/twapr2008econcontrib.pdf.*

2. Bell, Clive, and Maureen Lewis. October 2005. The Economic Implications of Epidemics Old and New. *World Economics* Vol. 5, No. 4: 137-174.

3. Brahmbhatt, Milian. November 7-9, 2005. Avian and Human Pandemic Influenza—Economic and Social Impacts *World Bank*, http://go.worldbank.org/QIB54LB5L0. (accessed September 23, 2005).

4. Bumiputra, Commerce. 2005. Avian Influenza: The Risk of a Pandemic and Its Economic Implications. Planning and Research, Vol. 9, www.bcb.com.my/corporate_info/economic_report.htm.

5. Congressional Budget Office. December 8, 2005; revised July 2006. A Potential Influenza Pandemic: Possible Macroeconomic Effects and Policy Issues. A Congressional Budget Office Assessment, U.S. Congress, www.cbo.gov/ftpdocs/69xx/doc6946/12-08-BirdFlu.pdf.

6. Del Valle, S.Y., P.D. Stroud, J.P. Smith, S.M. Mniszewski, J.M. Riese, S.J. Sydoriak, and D.A. Kubicek, 2006. EpiSimS: Epidemic Simulation System. Los Alamos Unlimited Release (LAUR) 06-06714.

7. Flaim, Silvio J. and Brian K. Edwards. October 11, 2006. Workforce Participation during an Influenza Pandemic. Los Alamos National Laboratory.

8. The Avian Flu Working Group. February 28, 2006. The Global Economic and Financial Impacts of an Avian Flu Pandemic and the Role of the IMF. *International Monetary Fund,* http://www.imf.org/external/pubs/ft/afp/2006/eng/022806.pdf.

9. James, Steven and timothy Sargent. June 2006. The Economic Impact of an Influenza Pandemic. Economic Analysis and Forecasting Division, Department of Finance, Government of Canada.

10. Marron, Donald B. July 27, 2006. A Potential Influenza Pandemic: An Update on Possible Macroeconomic Effects and Policy Issues. A Congressional Budget Office Assessment, 1-18, Congressional Budget Office, www.cbo.gov/ftpdocs/72xx/doc7214/05-22-Avian%20Flu.pdf (accessed June 26, 2009).

11. National Infrastructure Simulation and Analysis Center. August 2009. 2009 H1N1 Influenza Study: Impacts on the United States Population and Economy. Department of Homeland Security, Office of Health Affairs, Office of Infrastructure Protection.

12. Regional Economic Models, Inc. “REMI PI+”, v. 1.0.114, March 24, 2009 build, 51 region, 70 sector model, Amherst, MA.Regional Economic Models, Inc. “REMI PI+”, v. 1.0.114, March 24, 2009 build, 51 region, 70 sector model, Amherst, MA.

13. Timmer, Hans, Andrew Burns, and Dominique van der Mensbrugghe. June 24, 2009. Evaluating the Economic Consequences of Avian Influenza. *Global Development Finance*, 2008: 1-5. World Bank Development Economics Directorate.

Distribution

Copies	Mail Stop	Employee	Org.
1	MS1138	Theresa J. Brown	06371
1	MS1138	Lillian A. Snyder	06371
1	MS1138	Mark A. Ehlen	06371
1	MS1138	Braeton J. Smith	06371
1	MS1138	Calvin R. Shaneyfelt	06371
1	MS0899	Technical Library (electronic copy)	9536

Sandia National Laboratories