STATEWISE CORRELATES OF CIVIL NUCLEAR ENERGY

A Thesis

by

NISCHAL KAFLE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Paul Nelson Jr.
Committee Members,	Michael Sherman
	Pavel Tsvetkov
Head of Department,	Yassin Hassan

August 2014

Major Subject: Nuclear Engineering

ABSTRACT

Quantitative empirical analysis has been used in several works, over the past decade or so, to identify correlates of states motivation for pursuing military nuclear technology. Nelson and Sprecher used such methodology to identify various national attributes that correlate to states peaceful use of nuclear power for electricity generation, which was termed as "Nuclear Reliance." The major initial objective for the present work was to replace a dichotomous subjective independent variable used by Nelson and Sprecher to represent engagement in international commerce in civil nuclear technology with more objectively defined variables carrying a similar representation. Ordinary least squares stepwise regression was applied to a dataset consisting of 27 independent variables that was created for this study. Data for 13 of 27 independent variables were added to the dataset from previous study, and 9 of 14 previous attributes data were updated. Supervised stepwise regression was used to create a linear regression model with five predictors having acceptable confidence level ($p < 0.01$) and coefficient of determination ($R^2 \approx 0.51$). Results from stepwise linear regression showed that states that trade knowledge and material for nuclear power technology are heavily involved in civil nuclear power that states that are not involved in international trade of such technology and material. Analyses of the individual steps at several different levels of aggregation showed that some predictors were included as a consequence of improvements to residuals only for a few states. Preliminary results show that an analysis based on change from some prior year (1980 was used, for illustrative purposes) has considerable promise.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Professor Paul Nelson for letting me work as his graduate research assistant. It has been a great opportunity for me to learn from such an experienced professor. I would also like to thank my committee members, Professor Micheal Sherman, Professor Pavel Tsvetkov, for their guidance and support throughout the course of this research. Special thanks to Professor Michael Sherman for providing guidance with backward elimination regression method.

I am grateful to my parents (Navaraj and Neeta Kafle), my girlfriend (Shikha), and my brother (Nimesh) who have sacrificed countless hours of their life, and their support in order for me to reach this level. Thanks also go to my friends and colleagues, and the department faculty and staff for making my time at Texas A&M University a great experience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE Page

1. INTRODUCTION

In recent years several authors have applied the statistical empirical analysis methodology native to the social sciences to questions involving nuclear technology, particularly to issues related to determining correlates of nuclear armament; see Section 1.1 (Literature Review) below for citations and discussion. More recently, Nelson and Sprecher [1] and Yim and McNelis [2] have applied similar methodologies, except with the state as unit of analysis rather than the more traditional state-year, to the determination of correlates of civil nuclear power. The Nelson-Sprecher methodology rather simplistically employed ordinary least squares, but the resulting model of Nuclear Reliance (NR $\equiv$ fraction of the electrical power generated in a state that derives from nuclear energy) nonetheless had several attributes that were aesthetically pleasing:

i) The resulting model contained enough independent variables to give a decent statistical fit to actual NR ($R^2 \approx .53$), but few enough to permit some reasonable conjecturing regarding plausible causal relations.

ii) The coefficients of the independent variables in the model were generally stable over the several stages of the underlying stepwise regression process that led to the final model, which evidences minimal multicollinerity [3] among the independent variables selected for the model.

iii) If one groups the candidate independent variables in the underlying database into the five natural categories "energy," "international (nuclear) commerce," "nuclear material processing," "people and economy," and "political," then four of the five categories are represented among the five independent variables selected for the final model in the stepwise regression process. Nonetheless, there are troubling aspects of the Nelson-Sprecher model, namely:

a) The selected variables that represent the "international commerce" and "nuclear material processing" categories are "dummy" (bivariate) variables, and that representing commerce is considerably subjective.

b) Very surprisingly, the category not represented in the selected variables is "people and economy," which includes such variables as GDP.

The objective of this thesis is to replace the more subjective dichotomous independent variables in the Nelson-Sprecher data, especially international commerce, by more objective data, and to do so in a way that results in a linear regression model that retains desirable features i)-iii) of the Nelson-Sprecher model. Somewhat surprisingly the resulting model essentially lost (see Section 3) desirable properties i) and iii) above. This led to a search for (Sections 4 and 5)alternate approaches to restore those properties. One of the approaches was to introduce additional variables in the dataset that would represent a wide range of the nuclear materials and technology trade, i.e. from frontend to backend of civil nuclear power production. A new statistical regression methodology,such as backward elimination, was used to capture additional features, if any, lost within the stepwise regression methodology.

The five natural categories listed in iii) are hypothesized to be most important factors that would influence the nuclear reliance of different states. The basic linear model that would identify attributes that are both reasonably representative of the extent to which a given state currently relies on is expected to select at least one or more independent variable from each of the five natural categories.

1.1 Literature Review

References [1] and [4] are foundational to the present work. Nelson and Sprecher [1, 4] studied the extent to which a given state depends on nuclear energy to generate electricity. Two major focuses of Reference [1] were understanding different national attributes that would "associate with the existing degree of reliance of various states upon nuclear energy to meet their electrical needs." Further analysis was conducted to corroborate the predicted model using two different near-term measurements of nuclear intent [5].

In [5] Nelson and Sprecher used two near-term measures of nuclear intent, "reactors under construction" and "reactors planned plus reactor under construction" as addition dependent variables. Nuclear deficit (defined as the difference between nuclear reliance predicted from the model developed in [1, 4] and then current (2008) nuclear reliance of the state) was studied for near-term nuclear intent using so-called composite error as the measure of accuracy. The Nelson-Sprecher model [1] was able to predict a long-term measure of planned reactors better than nuclear plants under construction.

Li, Yim and McNelis [2] explored the "relationship between nuclear proliferation and civilian nuclear power development." Using historical records from 1945 to 2000 and tools such as mutinomial logit regression and Weibull and Cox event modeling, estimation of a state's proliferation motives was conducted. Results from the study indicated the usefulness of quantitative model to predict or warn against attempts of nuclear proliferation. Another research article by Yim and Li [6] collects information on various national capabilities and conducts correlational analysis. Findings from the study showed that states committed to nuclear proliferation and presence in the nuclear weapons had an adverse effect toward civil nuclear power program. The study also indicated that "level of democracy and nuclear technological capabilities" were important for progress towards civilian nuclear power program.

Furthermore, in [7] Singh and Way used mutlinomial logit to create a data set on nuclear proliferation and identify pathways to obtain nuclear weapons. Results from their study showed that "level of economic development, the external threat of great-power security guarantees, and a low level of integration in the world economy" have a strong impact of nuclear weapons proliferation. Jo and Gartzke [8] also used quantitative techniques to estimate states probability of establishing nuclear weapons program. Most important contributing factors from the findings of Jo and Gartzkes' finding was that states tendency towards their nuclear weapons program depends on the states "security concerns and technological, economic capabilities, and domestic politics."

A large body of knowledge is available that applies quantitative methodology to identify states motivation for military nuclear technology. Rauchhaus, Kroenig, and Gartzke [9] argue that acquiring nuclear weapons "enhance the security and diplomatic influence of their possessors." Reference 9 provides numerous collection of scholarly journal articles that used similar methods to study nuclear proliferation by political and social scientists.

1.2 Limitation of Study

The data collection was one of the most challenging aspects during this research. Information collection on sensitive technology (e.g. covertly operated enrichment and reprocessing facilities) for countries such as Iran are not available. Moreover, independent variables chosen in the database had to come from unclassified source. Furthermore, in some cases most recent data for all independent variables for all states were not available. For example, nuclear technology export data for Mexico was from 2011, while data for Malaysia for the variable was from 2012 [10]. Some states have no data available for various variables; therefore a value of zero was given.

1.3 Organization of the Thesis

The text addresses three distinct features: design and data, model, and analysis of the model. Each feature is addressed distinctly in different sections. Section 2 focuses on describing the data and design. Section 2.1 explains selection criteria of the states in the database, and provides description of stepwise ordinary linear regression (OLS). The section also includes discussion of the Nelson-Sprecher model [1] which used the MatLab code STEPWISEFIT (explained in Section 2) to implement the ordinary least-squares linear regression technique. Section 2 concludes with introduction of candidate correlates in five natural categories in Section 2.2. Section 3.1 includes description of independent variables present in the database. Section 3.2 shows the basic linear model with updated database as employed by Nelson and Sprecher [1, 4]. Section 3.3 implements a new graphical interface from MatLab called STEPWISE to better understand the obtained model.

The poor statistical characteristics of the newly obtained basic linear model compared to model by Nelson and Sprecher [1, 4], prompted for the move towards expansion of database with new independent variables. Section 4 of this text provides steps involved in the building of the alternative model with additional independent variable. In Section 4.1 the stepwise evolution of the model is shown. Next in Section 4.2 the model obtained by removing and inserting reprocessing and enrichment is discussed. Then in Section 4.3 different stepwise regression tool (Backward Elimination) is introduced. The final section in Section 4 discusses the improved stepwise regression model, which was obtained by slight change in the tolerance of p-value. In Section 5 analyses of the newly created model were conducted using confusion matrices, scatter plots, and state by state residual analysis respectively, obtained in Section 4.4. Section 6 looks further into extending the model obtained in Section 4. Section 6.1 looks into the variables that could be added from categories not represented in the improved model in Section 4. In Section 6.2 a new variable called “persistence” variable (states that had civil nuclear power reactors in 1980) to minimized the problem from “overpredicted” cluster of states discussed in Section 5.1. Section 6.3 conducts residual analysis to study any improvement brought by the inclusion of “persistence” variable on aggregated square residual in the model Sections 4 and 5.

2. DESIGN AND DATA

This section is devoted to a description of the database underlying the study of correlates of civil nuclear energy to be presented in the future sections. The section will focus on criteria for creating the list of appropriate states to the study, discuss the technique of ordinary least squares, present the nuclear reliance data for each state, and categorize candidate variables. Section 2.1 introduces data and design, ordinary linear regression, and Nelson-Sprecher model, and the five natural categories of independent variables are discussed in Section 2.2.

2.1 Database and Design

To provide comparability same criteria for inclusion of states in the relevant database as in Nelson and Sprecher [4] has been maintained. The underlying idea was to focus on states that had the potential to actually build nuclear power plants (NPPs) for energy production. The States deemed to have capacity to construct 1000 MWe nuclear power plant, which tend to have a large capital cost (close to $3-5 billion) [1, 11], were typically considered. The cost of building NPPs could be significantly larger than their GDP for many states [12], therefore; potential list of the candidate states would be significantly smaller than the existing 193 member states of the U.N. [13]. Candidate states under the selection criteria were named as nuclear candidates (NC). These are the 89 states that, according to the most recently available data circa 2007, had either a population of at least 20 million or a gross domestic product of at least $20 billion, less three such states (Afghanistan, Puerto Rico, and Uganda) for which some of the required data were unavailable as of that time. These 86 nuclear candidate states (86NC) are listed alphabetically in Table 2.1, along with their respective values of the nuclear reliance [14] as of 2011;

the candidate independent variables were evaluated ("coded," in the terminology of political science) as described in the following subsection.

An ordinary linear regression model is created as follows:

a) Dependent variables y that comprise a measure of the extent to which the states in some ensemble of states currently employ, or intend to employ, nuclear energy to meet their civil needs for energy; and

b) values $X = \{x_i : i = 1, \ldots, N\}$ of some collection of state attributes (candidate independent variables), across the selected ensemble, that might be expected to correlate to such independent variables. Index 'i' represent different attributes chosen as candidate independent variables.

Further, given these it is wished to employ stepwise ordinary least-squares regression to fit $\hat{y}$ to a linear model, as shown in Eq. (2.1):

$$\widehat{y} = \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n + c. \tag{2.1}$$

Using ordinary least squares (OLS) stepwise regression the resulting basic linear model (BLM), in Eq. (2.2), by Nelson and Sprecher was obtained:

$$\begin{aligned}\widehat{NR} = {} & (.3 \pm .04) IC? + (.097 \pm .032) ALGN? - (.33 \pm .09) COAL \\ & - (.13 \pm .05) FCS? + (.11 \pm .04) PLTY - .034.\end{aligned} \tag{2.2}$$

The variables appearing here are defined in Appendix A. Here the italicized symbol following the brief description of an attribute is the symbol used to represent that attribute in mathematical formulas, especially in a linear regression. A name ending in the symbol "?" indicates a bivariate variable (sometimes termed as a "dummy" variable), which is to say a variable that takes on only the values zero

Table 2.1: 86NC states, and their respective nuclear reliances, circa 2011 [14].

State	Nuclear Reliance [%]	State	Nuclear Reliance [%]
Algeria	0.0	Mexico	3.6
Argentina	5.0	Morocco	0.0
Australia	0.0	Myanmar	0.0
Austria	0.0	Nepal	0.0
Bangladesh	0.0	Netherlands	3.6
Belarus	0.0	New Zealand	0.0
Belgium	54.0	Nigeria	0.0
Brazil	3.2	North Korea	0.0
Bulgaria	32.6	Norway	0.0
Canada	15.3	Pakistan	3.8
Chile	0.0	Peru	0.0
China	1.9	Philippines	0.0
Colombia	0.0	Poland	0.0
Congo-Kinshasa	0.0	Portugal	0.0
Croatia	0.0	Qatar	0.0
Cuba	0.0	Romania	19.0
Czech Republic	33.0	Russia	17.6
Denmark	0.0	Saudi Arabia	0.0
Egypt	0.0	Serbia	0.0
Ethiopia	0.0	Singapore	0.0
Finland	31.6	Slovakia	54.0
France	77.7	Slovenia	41.7
Germany	17.8	South Africa	5.2
Ghana	0.0	South Korea	34.6
Greece	0.0	Spain	19.5
Guatemala	0.0	Sri Lanka	0.0
Hong Kong	0.0	Sudan	0.0
Hungary	43.3	Sweden	39.6
India	3.7	Switzerland	40.9
Indonesia	0.0	Syria	0.0
Iran	0.0	Taiwan	19.02
Iraq	0.0	Tanzania	0.0
Ireland	0.0	Thailand	0.0
Israel	0.0	Tunisia	0.0
Italy	0.0	Turkey	0.0
Japan	18.1	United Arab Emirates	0.0
Kazakhstan	0.0	Ukraine	47.2
Kenya	0.0	United Kingdom	17.8
Kuwait	0.0	United States	19.3
Lebanon	0.0	Uzbekistan	0.0
Libya	0.0	Venezuela	0.0
Lithuania	0.0	Vietnam	0.0
Malaysia	0.0	Yemen	0.0

("no," or "false") and one ("yes,", or "true"). The values obtained in the above equation (coefficient and standard errors) were obtained from the STEPWISEFIT. Code in MatLab STEPWISEFIT performs "multilinear regression of the response values in the n-by-1 vector on the *p* predictive terms in the n-by-p matrix" [15]. It is a statistical tool in MatLab that uses an unsupervised stepwise regression algorithm to determine most significant independent variable for a given model [16]. Statistical test is done by using *p*-statistics (also known as *p*-value) [17] to determine the significance of an independent variable. Smaller *p*-value tells the probability of obtaining the test statistic is as unlikely as the observed value when the null hypothesis is true [17]. Default *p*-statistics for inclusion and expulsion for a independent variable into a model in STEPWISEFIT is 0.05 (95% confidence level) and 0.1 (90% confidence level) respectively. The various independent variables are listed in the order in which they are added to the model by the stepwise regression process, and not by their increasing *p*-values.

2.2 Candidate Correlates

The following 24 national attributes, grouped by the previously mentioned five broad categories, were considered as candidate correlates (independent variables) for the linear regression model Eq. (2.1).

1. ENERGY

   Coal reserves ($COAL$); Electricity Generation ($EGEN$); Energy Insecurity (EI); National Gas Reserves (GAS)

2. INTERNATIONAL (NUCLEAR) COMMERCE

   Members of Nuclear Suppliers Group (NSG?) ; International Commercialization (nIC); Nuclear Technology Export ($NUEXP$); International Commercialization of Conversion to UF_6 (IC_{CUF_6}); International Commercialization of Enrichment (IC_{ENRCH}); International Commercialization of Fuel Fabrication (IC_{FuFab}); International Commercialization of Mining and Milling (IC_{MING})

3. NUCLEAR MATERIAL PROCESSING

   Conversion to UF_6 (C_{UF_6}); Enrichment ($ENRCH$); Enrichment and Reprocessing (ENR); Fuel Fabrication ($FuFab$); Mining and Milling (MING); Reprocessing (RePro)

4. PEOPLE AND ECONOMY

   Gross Domestic Product (GDP); Per Capita Gross Domestic Product (Purchasing Power Parity) ($pcGDPppp$); Mid-Level Economic State ($MLES$); Population (POP)

5. POLITICAL

   De Jure Nuclear Weapon State (NWS?); De Facto Nuclear Weapon State (NWs?); Historic Alignment ($ALGN$?); Polity IV ($PLTY$); Primary Fuel Production State ($PFPS$?).

Precise measures of these various attributes are described subsequently in Appendix A. As used here to develop various linear models of the form Eq. (2.2), some of the attributes were unit normalized by dividing by the the maximum value that occurs in a given attribute from candidate states, while others by dividing by the number of respective state's population and maximum value within the given attribute. Unit normalization was also implemented by Nelson and Sprecher [1, 4]. The key idea is to scale the measures of every attributes so that its minimum value is near zero and maximum value is near one. This allow the sensitivity of the dependent variable to a particular correlate to be measured by the magnitude of the coefficient of that variable in the linear regression model.

3. MODEL

The initial explorations for this thesis were based on a dataset that was minimally revised, relative to that employed by Nelson and Sprecher [1, 4], as described in Section 3.1. In Section 3.1 below the application of the MatLab automated stepwise regression code STEPWISEFIT to this dataset, and the results of that application, are described. Somewhat surprisingly the resulting model is substantially different from, and statistically inferior to, that found in the earlier work of Nelson and Sprecher [1, 4]. Section 3.2 below is devoted to a description of an application of the interactive stepwise regression MATLAB code STEPWISE to obtain a better understanding of the reasons underlying that unexpected result.

3.1 Initial Basic Linear Model

As one of the main objectives of the research was to objectively define international commercialization, since it was the most significant variable in the Nelson-Sprecher model in reference [1], other bivariates were not changed, except all objectively defined variables in the dataset were updated to the latest available data. A new model was then created using STEPWISEFIT.

This resulted in the basic linear model (BLM) as shown in Eq. (3.1). A statistical summary shown of this BLM follows in Table 3.1. Numbers in square brackets underneath each selected independent variable are the respective p-values for each selected variable respectively. Looking at the p-value of each variable shows that 'new' international commercialization is the most significant variable with a p-value of $6.78e-5$, compared to polity with a p-value of 0.0024.

$$\widehat{NR} = (.42 \pm .11)nIC? + (.143 \pm .046)PLTY - .037 \qquad (3.1)$$
$$[6.78e-5] \qquad\qquad [.0024]$$

Table 3.1: Statistical summary of the model shown in Eq. (3.1)

RSME	R^2	**Adj** R^2	**f-stat**
0.1386	0.2853	0.2680	16.56

Appendix A provides detailed description of all metrics for the national attributes selected to create the dataset. Section A.1 contains description of the variables appearing in Eq. (3.1). The model shown above in Eq. (3.1) had successfully satisfied the research objective to replace dichotomous variable "old international commercialization" with nIC. The model, however, is unsatisfactory as it was unable to meet all of the three properties discussed in the preface of Section 1. One of three properties was to have a decent statistical fit, which in this case was very low ($R^2 \approx .28$) when compared to statistics obtained in previous study by Nelson and Sprecher ($R^2 \approx .53$) in Reference [1]. Moreover, model's constant term was about 3.7%, which is very high compared to the nuclear reliance, as seen in Table 2.1 for states, such as Brazil, China, India, Mexico, Netherlands, and Pakistan. With only two variables selected in the model, it was unable to represent the five natural categories, as shown in Section 2.2 previously, which was another attribute desired to be obtained from the new basic linear model (BLM). Reiterating the fact that there were four independent variables from four different categories in the Nelson-Sprecher model, which makes the model of Eq. (3.1) inferior to the BLM obtained in reference [1], as restates in the thesis as Eq. (2.2).

3.2 Graphical Interactive Tool STEPWISE

The MatLab tool STEPWISE was implemented for analysis of model obtained in Eq. (3.1), and also used later in Section 4 and Section 6. STEPWISE is also a stepwise regression statistical tool in MATLAB, but it is more interactive then STEPSWISEFIT [15].

STEPWISE serves as a graphical user interface that can be used to show each variable added to the model one at a time with "Next Step" button, or that can predict variables all at once with the "All Steps" button, as shown in Figure 3.1 below. The interface also displays the coefficient, t-statistics, and p-values candidate independent variables in a box listed as $X_1, \ldots, X_n$, along with 90% (red horizontal line), and 95% (grey horizontal line) confidence interval of the predictors. The interface also displays statistical summary and model history. This interface is also useful because it can be used to see how the model changes when some of the predictors are forced into the model.

After running the program with STEPWISE no predictor is selected at first. The user can select all steps to compute the model at once or select next step to see the model evolution that adds a new independent variable each time until no more variables can be added at the specified level of significance. Figure 3.1 below is as an example, which shows the first step before any variables are added in the model. By clicking on the *All Steps* in the interface shown in Figure 3.1, the model in Eq. (3.1) can be obtained, which was created by updating data-set and replacing the old dichotomous data for international commercialization by nIC.

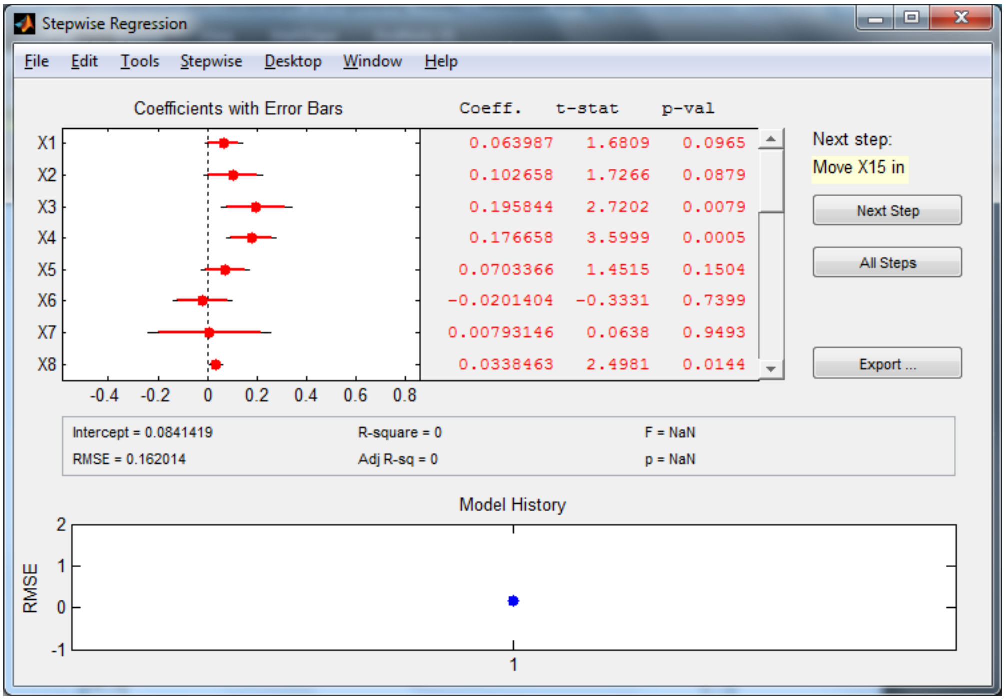

Figure 3.1: STEPWISE regression interface

Using STEPWISE it can be seen that the coefficients of different variables with 95% confidence interval crossing the zero line are eliminated from being selected for the model. Next step in this thesis could have been to create a model using variables that have coefficient with 90% or less confidence interval, i.e. narrower bound; however, addition of other independent variables into the dataset was chosen. The interactive nature of STEPWISE interface that is more supervised than the STEPWISEFIT, becomes useful in obtaining a acceptable model, which is later discussed in Section 4.4.

Original bivariate international commercialization was coded on the basis of states involved in international trade of different material and technologies, such as frontend and backend fuel supplies, knowledge and reactor components. The newly coded nIC only accounts for suppliers of the reactors and its peripheral components; therefore, it is necessary to add independent variables representing nuclear fuel cycle and nuclear trade in order to fully compensate the transition from subjectively defined IC to objectively defined nIC. Section 4 discusses these additional variables, the new basic linear model and its analysis.

4. MODEL WITH ADDITIONAL VARIABLES

After the unsatisfactory model obtained in Section 3 from the minimal revision in the database, additional variables were added to the model. This section focuses on listing new variables, creating a new and improved model with step-by-step evolution, and defining additional independent variables that appear in the new model.

4.1 STEPWISEFIT Model with Additional Variables

There were several variables added to the database, and following lists the acronym of variables added in different broad categories:

1. International Commerce: NSG, $NUEXP$, UF_6, IC_{FuFab}, IC_{MING}
2. Nuclear Material Processing: C_{UF_6}, $ENRCH$, $FuFab$, $MING$, $RePro$, ENR
3. People and Economy: $MLES$.

Above listed variables included in the database are described in detail in Appendix A. The MatLab code STEPWISEFIT, as described in Section 2, was applied using all variables listed above and variables from Ref. [1] dataset. Five evolutionary steps of the model are as shown below, as Eqs. (4.1)-(4.5). The resulting statistical data from each step are listed in Table 4.1.

Step by step evolution of the linear model is as follows:

$$\widehat{NR} = .08 \tag{4.1}$$

$$\widehat{NR} = (.16 \pm .03)NSG? + .006 \tag{4.2}$$

$$\widehat{NR} = (.14 \pm .028)NSG? + (.51 \pm .11)RePro + .006 \tag{4.3}$$

$$\begin{aligned}\widehat{NR} = &(.12 \pm .027)NSG? + (.51 \pm .11)RePro + (.33 \pm .11)NUEXP \\ &+ .006\end{aligned} \tag{4.4}$$

$$\begin{aligned}\widehat{NR} = &(.11 \pm .027)NSG? + (.46 \pm .11)RePro + (.26 \pm .11)NUEXP \\ &+ (.21 \pm .103)nIC + .006.\end{aligned} \tag{4.5}$$

Table 4.1: Statistical data for model obtained from STEPWISEFIT regression

Step	**p-value(s)**	R^2	**Adj** R^2	**f-stat**	**RSME**
1	–	–	–	NaN	.1620
2	$[5.6228e-7]$	.2669	.2582	30.58	.1395
3	$[1.958e-6]$ $[2.867e-5]$	.4058	.3915	28.34	.1264
4	$[1.4907e-5]$ $[1.312e-5]$ $[.0031]$	.4652	.4457	23.79	.1206
5	$[1.018e-4]$ $[7.72e-5]$ $[.0255]$, $[.0466]$	.4876	.4623	19.27	.1188

The first step in the model gives the estimated nuclear reliance of the actual nuclear reliance. Independent variables added in the model are primarily based on the significance level determined by the p-value for each independent variable. Default p-values set in STEPWISEFIT while selecting the independent variable are 0.05 (95%

confidence level) for inclusion and 0.10 (90% confidence level) for exclusion. During the selection process the variable with largest significance level, or equivalently the variable having the lowest p-value, is added. Depending on the selected variable the significance level of other variables could increase or decrease. A variable previously added in the model with smaller *p-value* might have lower significance than the independent variable added in subsequent steps, and can also be removed when the *p-value* become greater than 0.10. In above steps in Eqs. (4.1)- (4.5) all previously added variables remain in the model throughout all evolution steps. However, the significance of previously added variables has changed with addition of each variable. The model is not arranged in the order of increasing *p-value*, but in order of variables appearing first in the model. The final model created by the STEPWISEFIT using default p-values is shown above in Eq. (4.5). The model obtained will always have non-zero nuclear reliance because of the constant value, but the constant term is small and very close to zero (about 0.59%).

The stepwise forward regression model evolution in Eqs. (4.1) - (4.5) were obtained using the default inclusion and exclusion p-values. The model obtained in this case is, however, not satisfactory. The major for dissatisfaction was because only two of the five natural category was selected into the model. One of the three attributes for Nelson-Sprecher model in Eq. (2.2) was to contain variables from different categories, and it was one of the important reasons for seeking an improved model beyond Eq. (3.1) (which also had undesirable statistical data). Additional techniques seem necessary in order to include variables that could try to search for better model. Following section discusses on variables that were once selected into the BLM during the course of the research, but later were omitted. Section 4.2 also concentrates on selecting subset of candidate variables from the database that could be added to Eq. (4.5).

4.2 Search for Better Model

Alternative approach was to look at reducing the number of independent variables from the dataset by eliminating variables showing similar objective in the dataset. One of those variables was ENR; however it is subjective in nature, which defines a physical presence of enrichment or reprocessing plant or technological capability of building such plants for a given state. Two variables tried to be relegated from the dataset were $ENRCH$ and $RePro$, and four different models were generated without including either $ENRCH$ or $RePro$ or both and including both of them. Four models created were mainly focused on, as shown in Eq. 4.6- Eq. 4.9. The first model and second models are labeled as asymmetric model because reprocessing was excluded from model (4.6), and $ENRCH$ is left out from model (4.7). Next model is labeled symmetric, which is identical to Eq. (4.11), where both ENRCH and RePro were included in the database, and finally the fourth model without either of the variable was included in the database to create model is shown in Eq. (4.9). Following list presents the three model described in this section.

1. Asymmetric Model (includes only enrichment but not reprocessing)

$$\begin{aligned} \widehat{NR} =& (.14 \pm .03)NSG? + (.32 \pm .10)nIC + (.46 \pm .12)ENRCH \\ & - (.32 \pm .14)GDP - (.16 \pm .081)MING + .011 \end{aligned} \tag{4.6}$$

2. Asymmetric Model (includes only reprocessing but not enrichment)

$$\begin{aligned} \widehat{NR} =& (.11 \pm .03)NSG? + (.46 \pm .11)RePro + (.25 \pm .11)NUEXP \\ & + (.21 \pm .10)nIC + .006 \end{aligned} \tag{4.7}$$

3. Symmetric Model (includes both enrichment and reprocessing)

$$\widehat{NR} = (.11 \pm .027)NSG? + (.46 \pm .11)RePro + (.26 \pm .11)NUEXP + (.21 \pm .10)nIC + .006 \tag{4.8}$$

4. Model without enrichment and reprocessing

$$\widehat{NR} = (.13 \pm .028)NSG? + (.21 \pm .11)nIC - (.26 \pm .086)MING + (.32 \pm .097)CONVR + (.24 \pm .11)NUEXP + .007 \tag{4.9}$$

The nine variables that appear in the models Eqs. (4.6) - (4.9): $NSG?$, nIC, $ENRCH$, GDP, $MING$, *Repro*, $NUEXP$, Conversion ($CONVR$) and constant term were then selected for inclusion in a backward model selection technique. Backward elimination is a variation of quantitative statistical model analysis; it is further discussed in the following section.

4.3 Backward Elimination General Linear Model

Backward linear regression method uses opposite regression methodology when compared to stepwise regression. In backward regression, all variables are initially selected in the model. One variable is then removed in each subsequent step until the stopping criteria is reached. The variable which is the least contributor in the model is removed. For a backward elimination independent variable that has the least significant F statistics is dropped from the model [18]. SAS defines the F statistics for removal to be as shown in Eq. 4.10

$$F = (RSS_{p-k} - RSS_p)/k/(RSS_p/(n-p-k)), \tag{4.10}$$

where RSS is residual sum of residual sum of squares, p is the number of variables in the current step, and k is the degrees of freedom, and n is the total number of variables.

The subset of the database with the nine variables mentioned in Section 4.2 was used to perform the backward elimination analysis in order to find the most significant variables among those nine selected in Eqs. (4.6)-(4.9). The model (4.11) and the resulting statistics in Table 4.2 were thus obtained using 0.05 for inclusion and 0.1 for expulsion p-value criterion.

$$\begin{aligned} \widehat{NR} =& (.12 \pm .028)NSG? + (.44 \pm .11)RePro + (.22 \pm .11)NUEXP \\ &+ (.26 \pm .11)nIC - (.13 \pm .08)MING + .006 \end{aligned} \tag{4.11}$$

Table 4.2: Statistical data for model obtained from STEPWISEFIT regression

Step	**p-value(s)**	R^2	**Adj** R^2	**f-stat**	**RSME**
6	$[3.086e-5]$ $[1.498e-4]$ [.0554], [.0182], [.0969]	.5051	.4742	16.33	.117

Section A.2 contains description of variables appearing in Eqs. (4.1) - (4.11). Comparing models (4.5) and (4.11), MING is the only new variable added to the model (4.5) which gives (4.11). Coefficients of the variables included in both models are very similar, suggesting no multicollinearity between coefficient estimates after adding MING to model (4.5). Additionally inclusion of MING, with a negative coefficient, is intuitively satisfying, in that it can be viewed as an instance oft-discussed as "resource curse," or paradox of plenty [19], which implies that states having abundant natural resources tend to have less economic growth then states deprived of it. Further in the following section STEPWISE interactive interface is used to analyze the addition of MING into the forward stepwise regression model in Eq. (4.5).

4.4 Improved Stepwise Model

In order to improve the stepwise regression model obtained in Eq. (4.5), STEPWISE interactive interface (discussed in Section 3.2) was useful. Figure 4.1 shows a screenshot of STEPWISE graphical interface for model (4.5). The figure shows $MING$ (variable '*X21*' in Figure 4.1) was not included into the model because it's p-value was above 0.05 inclusion limit. All variables not selected into the model have their coefficient 95% confidence intervals crossing the "null" line (represented by the dotted line in the box labeled *Coefficient with Error Bars*). This section only utilizes STEPWISE graphical interface to add any viable independent variables into the BLM in Eq. (4.5). Of the the variables there were in Nelson-Spercher model and left out from model (4.11) $PLTY$ with p-value of 0.34 was the mostly likely candidate.

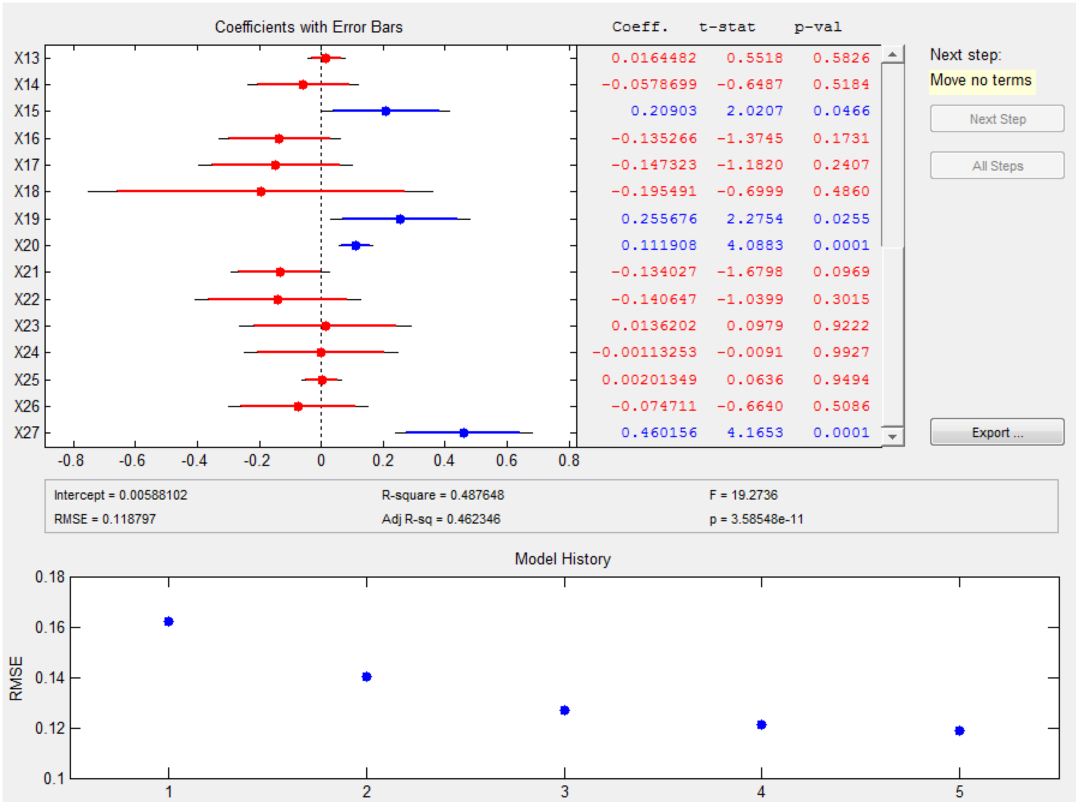

Figure 4.1: Stepwise graphical representation after default run at p-values of 0.05 inclusion and 0.10 exclusion

The next variable that would be chosen in the variable is $MING$ (variable '*X21*' in Figure 4.1) which has a coefficient confidence interval of 90% not crossing the "null" line. The forward stepwise regression with a weaker threshold (p inclusion of 0.1 and p expulsion of 0.15) in fact leads to the five-variable model as presented in Eq. (4.11). Selecting MING gives a better statistics, and also the negative coefficient for this variable improves the estimation of nuclear reliance for states such as Australia and Kazakhstan.

The graphical tool, STEPWISE, was further used after the addition of MING to see additional variables from five natural category could be added to model (4.11).

Figure 4.2 presents the screenshot of STEPWISE tool after the addition of MING into to the model (4.5). Next variable that could be added to the model is $NWS?$ (variable '*X3*'), which is not shown in Figure 4.1, with p-value of more than 0.15. Number of natural categories representation would be on par for Nelson-Sprecher model (2.2) and model (4.11) with addition of $NWS?$. However, the low significance (i.e. high p-value) of $NWS?$, and instability in coefficient of $RePro$, which jumps up 13 %, after the addition of $NWS?$ is not desirable.

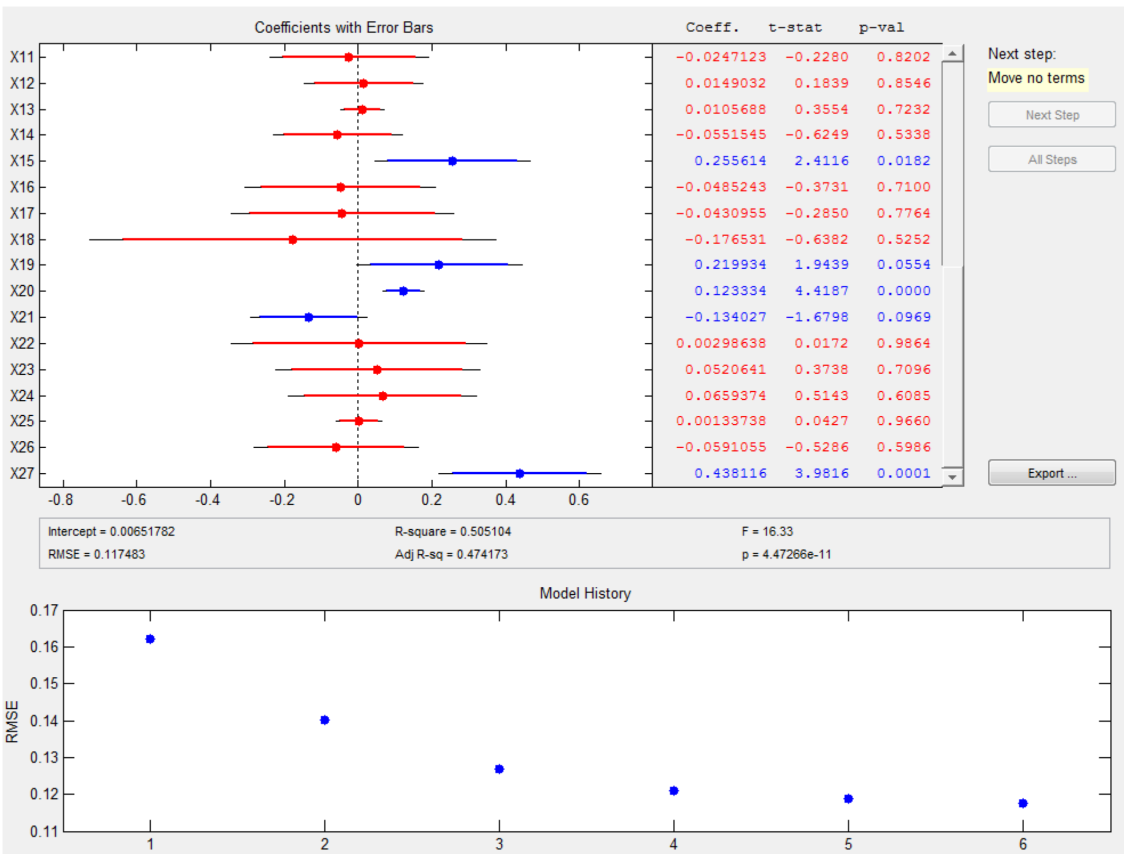

Figure 4.2: Stepwise graphical representation after addition of mining and milling variable at p-values of 0.10 inclusion and 0.15 exclusion

Addition of remaining variables into the model is dubious because of very low confidence interval of their coefficients. Eq. (4.11) therefore appears to be the best possible model that can be obtained from the stepwise regression method, and the present database. Basic Linear Model in Eq. (4.11), obtained in this section has successfully included more objectively defined variables when compared to N-S model. section 5 will ruminate obtained stepwise regression models using confusion matrices, scatter plots and residual table; three analyses methodology, in order to understand reduction of square residual with an addition of each variable.

5. MODEL ANALYSIS

In the following section, three levels of analyses for errors are done on the model (4.1)-(4.5) and (4.11), as obtained using STEPWISEFIT. From the macro to the micro-level, the selection of the dependent variables is analyzed, with a view toward better understanding how the selection of particular variables occurs during the model evolution described by Eqs. (4.1)-(4.5) and (4.11), and which states primarily contribute to residual model errors at the various steps. Analyses are conducted using confusion matrices in Section 5.1, in Section 5.2 scatter plots are generated to visualize the change in the prior and posterior residual after addition of each independent variable, Section 5.3 analyses square residual scatter plots before and after each new independent variable is added into the evolutionary model, and finally, in Section 5.4 individual state by state residual (residual is the difference between the estimated and actual nuclear reliance) analysis is conducted. Section 5 also discusses usefulness of addition of various fuel cycle variable as discussed in the objective of Section 1.

5.1 Confusion Matrices

Confusion matrices provide a convenient comparison of actual and estimated values of the independent variable, in a matrix format [20]. A confusion matrix can contain two classes or more (i.e. two by two or bigger matrix) and must be a square matrix. The row of the matrix contains corresponds to ranges of the estimated values of the dependent variable and the column to the corresponding ranges of the actual variable. A perfect model or a perfect classifier would have all instances fall in the diagonal elements and all the non-diagonal elements would be zero. Nonzero contents in the non-diagonal elements shows the model cannot estimate the actual data accurately. From the confusion matrix one can deduce the models effectiveness and where the model is not performing well. For this research project, the clustering for the confusion matrix was done to evaluate the generated model to distinguish four levels of nuclear reliance: states without any nuclear reliance, states with some/emerging nuclear reliance, and states that are heavily dependent on civil nuclear power. A sizable gap in the nuclear reliances among all states is observed at 0%[1], 0-10%, 10-25%, and above 25%; hence distinguishing states nuclear reliances into four categories. For convenience these four categories of the nuclear reliance in this thesis will be termed as "*de minimis*," "threshold," "moderate," and "aggressive," respectively. Below are the construction of confusion matrix for addition of each new variable, in (4.1)-(4.5) and (4.11), as shown in Tables 5.1-5.6.

First step if the confusion matrix is the addition of constant term into the model. The constant term is just the average nuclear reliance for the chosen 86 states. Below in Table 5.1 all the states fall in "threshold" category of estimated NR category as the average of NR is 8.41%. The confusion matrix at this stage predicts with about 9.3% accuracy, which was measured by taking the ratio of number of states in the diagonal cells of the confusion matrix to all states under in the dataset. The accuracy of the confusion matrices is expected to increase with each evolutionary step.

The new confusion matrix in Table 5.2, after the addition of NSG? in Eq. (4.2), has a large number of reshuffling of states, which shows a drastic increase in the accuracy of the confusion matrix from 9.3% to 58.1%. In comparison to the prior confusion matrix none of the states remain in "threshold estimated" category. Most of the states with 0%, and 8 out of 9 of the "moderate" NR states have been estimated correctly. However, 12 "aggressive" and 15 "*de minimis*" states are largest off-

[1]Estimated Nuclear Reliance which is less than 1% is considered to be zero because of the constant term (0.0065) present in models (4.1)-(4.5) and (4.11).

Table 5.1: Confusion matrix after constant term is added to the model in (4.1)

		Actual				
	NR	0%	0-10%	10-25%	> 25%
Estimated	0%	0	0	0	0
	$0-10\%$	57	8	9	12
	$10-25\%$	0	0	0	0
	> 25%	0	0	0	0

diagonal terms. The 'fifteen' states are a bit more of a concern, because it is distance two from the diagonal, so there is no possibility of those states being misestimated, but only closely, because either their actual or estimated values are near the boundary between clusters. Two clusters of misestimated states are named "overestimated" for the group of 15 states, and "underestimated" for the group of 12 states with actual NR of zero and larger than 25% respectively.

Table 5.2: Confusion matrix after *NSG*? is added to the model in (4.2)

		Actual			
	NR	0%	0-10%	10-25%	> 25%
Estimated	0%	42	2	1	0
	$0-10\%$	0	0	0	0
	$10-25\%$	15	6	8	12
	> 25%	0	0	0	0

Next Table 5.3 presents confusion matrix after addition of RePro for model (4.3). There are only minimal changes in this stage compared to previous step when NSG? was added. The variable *RePro* makes three changes by moving two states from actual and estimated-moderate cell to actual-moderate estimated-aggressive cell,

Table 5.3: Confusion matrix after *RePro* is added to the model in (4.3)

		Actual				
		NR	0%	0-10%	10-25%	> 25%
Estimated	0%	42	2	1	0	
	$0-10\%$	0	0	0	0	
	$10-25\%$	15	6	6	11	
	$>25\%$	0	0	2	1	

and moving one from actual-aggressive estimated-moderate cell to actual-estimated-aggressive cell. In this step, one state, France with presumable improved estimation moves from actual-aggressive estimated-moderate cell to actual and estimated aggressive cell; however, two states from actual and estimated-moderate cell move to cell actual-moderate estimated-aggressive cell, which decreases the confusion matrix accuracy from 58.1% to 57%. Next two states UK and Russia are respectively second and third largest producers of reprocessed nuclear material, but the inclusion of *RePro* overestimates their nuclear reliance and shifts them into to estimated-aggressive row of the confusion matrix. The selection of variables thus can be influenced by a very small number of states or maybe even only one state. In this step no changes in the "overestimated" states in cell actual-*de minimus* estimated-moderate cell is observed.

The next variable selected in the model is NUEXP (4.4), confusion matrix in Table 5.4. This again brings minimal changes when compared to the previous confusion matrix in Table 5.3. The accuracy of the confusion matrix moves up from 57% to 59.3%. Only change observed was two states moving from actual aggressive-estimated moderate cell to actual and estimated-aggressive cell. As in the previous step, the variable selection is influenced by a small number of states in the database. In this case, Sweden and Belgium, which have the highest *NUEXP* per capita value,

Table 5.4: Confusion matrix after $NUEXP$ is added to the model in (4.4)

	Actual				
Estimated	NR	0%	0-10%	10-25%	> 25%
0%	42	2	1	0	
$0-10\%$	0	0	0	0	
$10-25\%$	15	6	6	9	
$> 25\%$	0	0	2	3	

Table 5.5: Confusion matrix after nIC is added to the model in (4.5)

	Actual				
Estimated	NR	0%	0-10%	10-25%	> 25%
0%	42	2	1	0	
$0-10\%$	0	0	0	0	
$10-25\%$	15	6	6	9	
$> 25\%$	0	0	2	3	

are the two states moving from actual-aggressive estimated-moderate cell to the diagonal actual and estimated aggressive cell.

Table 5.5 shows the confusion matrix after the addition of nIC into the model (4.5). At this step, no changes in the confusion matrix are observed. Further analysis with scatter plot and residual table analysis in this section will shed some light for the selection of this variable.

The confusion matrix in Table 5.6, as created after the addition of $MING$ (4.11) shows some changes in actual-*de minimis* estimated-threshold and actual-threshold estimated-moderate cells, which were left unchanged in all previous steps. The confusion matrix accuracy with the inclusion of this variable increases from 59.3% to 60.4%. Most substantial changes can be seen on the overestimated states of actual-*de minimus* estimated-moderate cell. Negative coefficient for this variable shows that

Table 5.6: Confusion matrix after $MING$ is added to the model in (4.11)

		Actual				
		NR	0%	0-10%	10-25%	> 25%
Estimated	0%	42	2	1	0	
	$0-10\%$	2	1	0	0	
	$10-25\%$	13	5	6	9	
	$>25\%$	0	0	2	3	

mining and milling decreases civil nuclear power production for electricity. Australia and Kazakhstan are two states among the 15 "overestimated" states that move towards the diagonal of the confusion matrix. Likewise South Africa moves to the diagonal of the matrix from actual-threshold estimated-moderate cell to actual and estimated threshold cell.

With an exception of NSG? in the first step, the various steps have contributed small changes or no changes to the confusion matrices. Minimal improvement in the severely "overestimated" and "underestimated" clusters created by the selection of NSG remains a major concern. Twelve states that are in the "overestimated" cell of the confusion matrix presumably are either dependent on nuclear energy from other states, or they oppose nuclear energy for reasons such as problems nuclear waste management, nuclear proliferation, scarce fissile material resource. On the other-hand, variables added after *NSG?* show no significant improvement to the "underestimated" cluster of states. The states belonging to the "underestimated" cluster do not make major contribution in terms of international commerce and nuclear material trade, at this level of analysis. Therefore, influence of small number of states for the variable selection into the model seem dominant. Subsequent analyses at a less aggregated level could shed more light on the degree of improvement after each variable addition.

5.2 Nuclear Reliances Scatter Plots

In the scatter plots of both this and the following section, one data point is plotted, on a two-dimensional scatter plot, for each of the 86NC states in our database. In both this and the following section the horizontal coordinate of a data point is the actual nuclear reliance of the corresponding state. In this section the vertical coordinate is the estimated value of nuclear reliance provided by the particular model from one of (4.3)-(4.5) or (4.11) that is under consideration. Each actual vs. estimated nuclear reliance scatter plot, in this section, is intended to provide a slightly more detailed picture, compared to confusion matrices. To facilitate comparison with Section 5.1, the plots of this section have horizontal and vertical lines representing boundaries between the four categories of nuclear reliance, represented in the confusion matrices of the previous section. Graphical representation of actual nuclear reliance vs. estimated nuclear reliance can show migration of the nuclear reliances associated to selection of an additional variable to the model. The use of "Estimated Nuclear Reliance Scatter Plot" is to graphically compare changes in estimated nuclear reliance before and after each variable is added; i.e. each section step in the model evolution represented in (4.1)-(4.5) and (4.11). The estimated nuclear reliance scatter plot for the initial steps (i.e.(4.1) and (4.2)), in which only the constant term and the NSG? term appear are not shown. That is because initially all states will have estimated nuclear reliance equal to the mean value of the nuclear reliance of states; in the subsequent step, when NSG? is added, all 41 member states of Nuclear Supplier's Group will have one estimated nuclear reliance value and rest of the states will have zero estimated nuclear reliance.

Figure 5.1 is a scatter plot of estimated nuclear reliances, both before (red crosses) and after (green circles) the step (4.3) in which $RePro$ is selected as the independent variable to be added to the model. Most states fall in a cluster very near the origin, corresponding to *de minimis* values for both actual and estimated nuclear reliances, or in a band across the moderate range of estimated nuclear reliances. The latter represent states that are members of the Nuclear Suppliers Group. After addition of RePro to the model the estimated value of NR for those states not engaged in reprocessing decreases by about 0.02, corresponding to the reduction of approximately .02 in the coefficient of NSG that accompanies this step. These data points still lie in the moderate range of estimated values of Nuclear Reliance, just slightly lower (approximately 0.15, vs. the previous value of approximately 0.17). Because the four clusters used for the confusion matrices are well separated, this migration is not sufficient to cross cell boundaries in the confusion matrix during the transition when NSG? was added to when $RePro$ was added.

In the discussion of the preceding section it was determined that three states moved across confusion-matrix (Table 5.3) cell boundaries in the transition depicted in Figure 5.1. In more detail, these inter-cell transfers took the form of a transfer of one state from actual aggressive and estimated moderate nuclear reliance to actual and estimated aggressive (an improved estimation) and of two states from actual and estimated moderate nuclear reliance to actual moderate and estimated aggressive nuclear reliance (a worse estimation). Residual value listed in detail later in Section 5.4 below reveals the identity of the former state as France, and the latter two as the UK and Russia when NSG was added to the model. This conclusively demonstrates that improved estimation for a single state can influence selection of a variable to the model. Even more remarkably, it demonstrates that this can happen, even when selecting that variable imposes a worse estimation (albeit by a smaller amount) on more than one state. A better understanding of this phenomenon will be achieved in the following two sections.

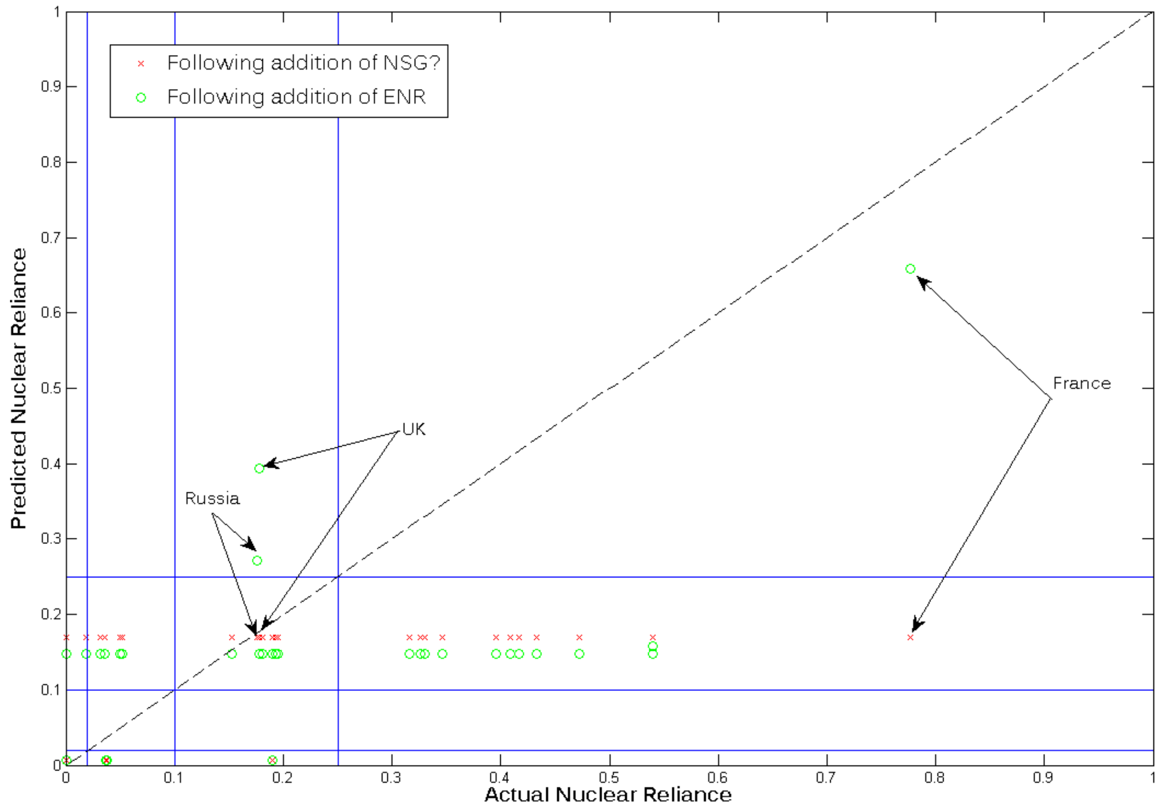

Figure 5.1: Actual versus estimated Nuclear Reliances before and after adding *RePro* (4.3)

Figure 5.1 also reveals some information regarding the severely overestimated and the underestimated states. Specifically it at least strongly suggests that all are members of the NSG, but members that have diametrically opposed (respectively, *de minimis* vs. aggressive) positions *vis a vis* reliance on nuclear energy. The identities of the members of these groups of NSG members having contrasting positions on nuclear energy will be further discussed in the following two sections. Yet another interesting identity that will be revealed is the identity of the single state that the confusion matrices of the preceding section suggest persistently lies in the cell corresponding to actual NR is equal to moderate and estimated NR equal to *de minimis*. Already Figure 5.1 suggests that particular seeming outlier state is not a member of the NSG, but nonetheless has an actual Nuclear Reliance of just under 20%.

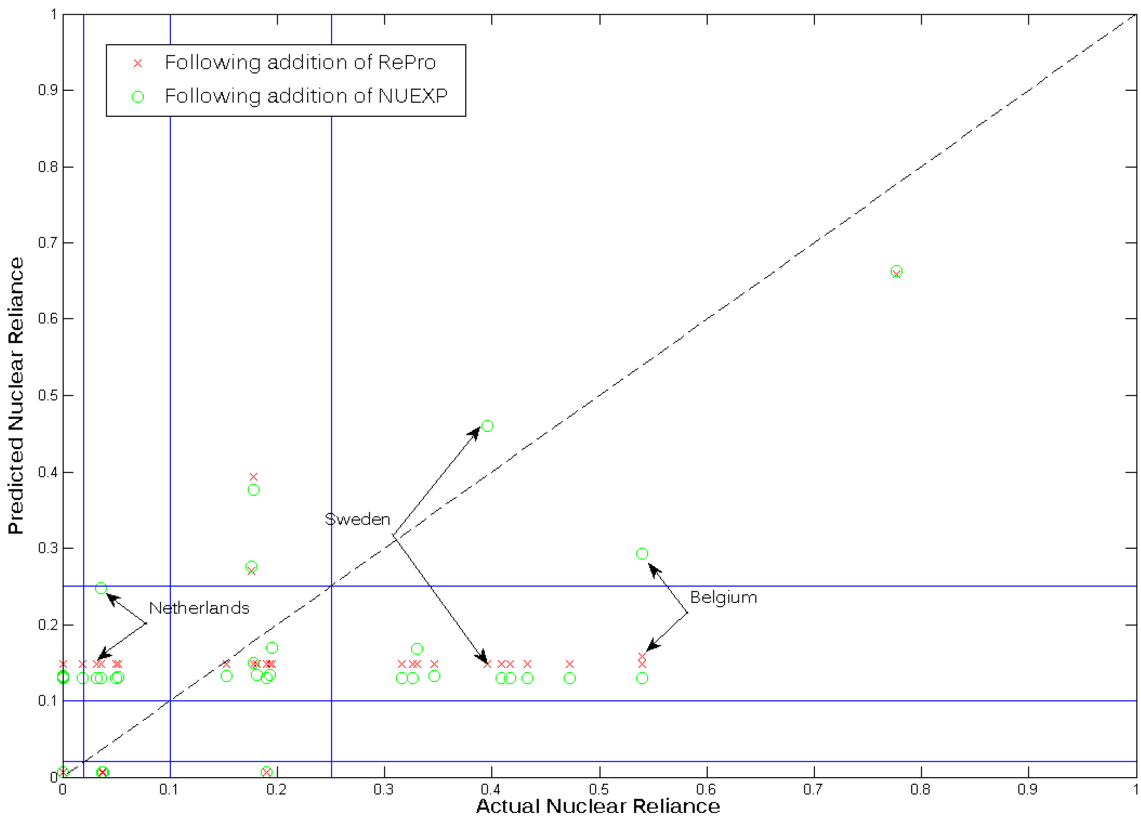

Figure 5.2: Actual Nuclear Reliance versus estimated Nuclear Reliance before and after and addition of *NUEXP* (4.4)

Figure 5.2 is a similar estimated nuclear reliance scatter plot for the stepwise regression model in which the independent variable $NUEXP$ is selected for addition to the model. Again the bulk of the posterior estimated nuclear reliance data points (the circles) lie just below the corresponding prior data (the x's). This corresponds to the slight decrease in the coefficients of the prior independent variables, particularly that for NSG, that accompanies this step. But there are two states for which this step introduces a significant increase in the estimated NR. Belgium and Sweden, as so labeled in the subject figure was found from the square residual value shown in detail in Section 5.4 when $NUEXP$ was added. These are the two states that, in the transition from the confusion matrix of Table 5.3 to that of Table 5.4, migrated from actual NR is equal to aggressive, estimated NR is equal to moderate to the improved estimation represented by membership in the cell actual NR is equal to estimated NR is equal to aggressive.

Figure 5.2 also shows that a couple of states, one in the cell actual and estimated moderate nuclear reliance and one in the cell corresponding to actual aggressive and estimated moderate reliance, achieve an increased estimated NR, and thereby one of improved accuracy, from this step. But it also shows there is a state having an actual NR of .036 that in this step transitions from a estimated NR of nearly 0.15 to one of not quite 0.25. Thus again we see that addition of a new variable significantly worsens the accuracy of the estimation for at least one state, albeit in this case not by (quite) enough to occasion migration of the corresponding data point across a cell boundary in the confusion matrix. The Netherlands is the latter state that moves towards to the edge of the cell boundary but not quite crosses it shown in detail in Section 5.4 in the table when $NUEXP$ was added.

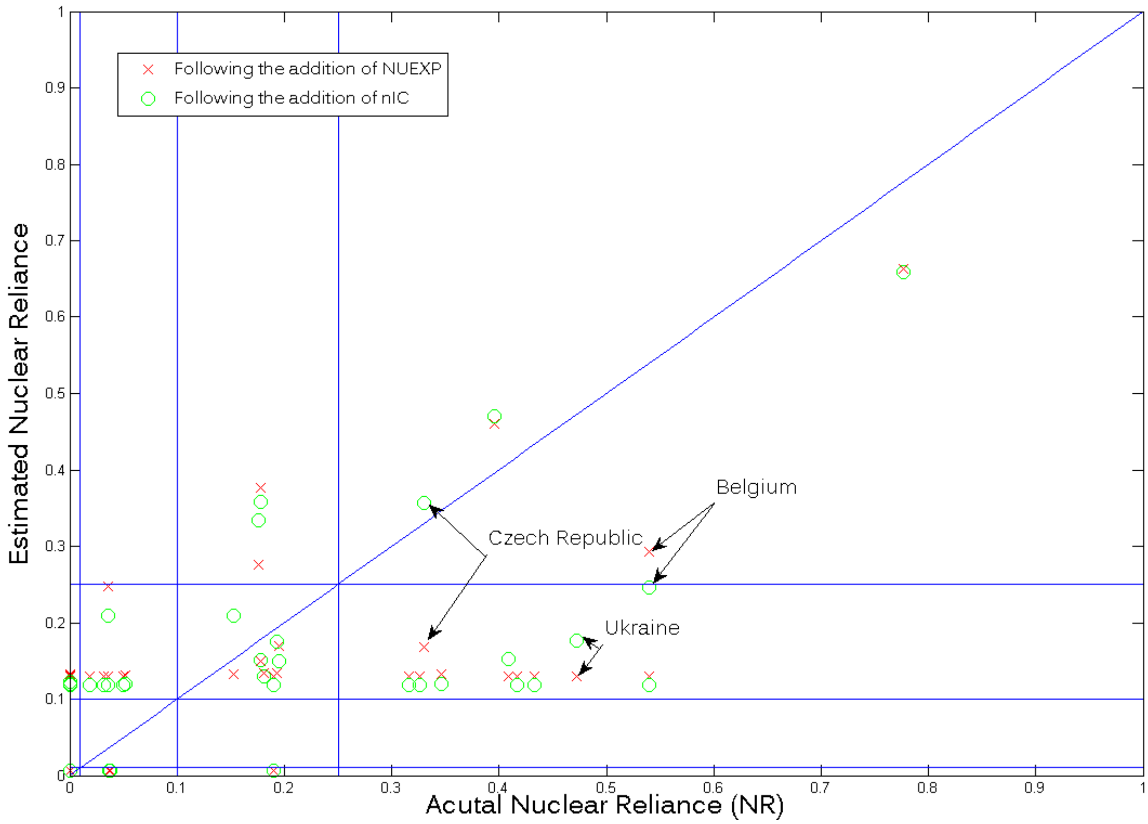

Figure 5.3: Actual Nuclear Reliance versus estimated Nuclear Reliance before and after and addition of nIC (4.5)

Figure 5.3 is the estimated nuclear reliance scatter plot for the stepwise regression model (4.5) in which the independent variable nIC is added to the model. Again the bulk of the posterior estimated nuclear reliance data points (the circles) lie just below the corresponding prior data (the x's). As before, this corresponds to a slight reduction in the coefficients of (all) prior independent variables that accompanies this step. In this step there is no *net* transfer between cells of the confusion matrix; however, there are transfers, one in each direction between the cells corresponding to actual-NR is equal to aggressive, estimated-NR is equal to moderate and actual NR

is equal to estimated NR is equal to aggressive. See the data points labeled "Czech Republic" and "Belgium" in Figure 5.3 where Czech Republic NR estimation is improved considerably, but the NR estimation of Belgium worsens quite a bit; as a results no *net* change of states moving across the cell boundary (i.e. $NR > 25\%$) has occurred.

Overall it is, from the vantage point of the associated estimated nuclear reliance scatter plot, difficult to see any consistent pattern underlying the selection of nIC as a model variable. Only major advantage of adding nIC, from Figure 5.3, is that NR for Czech Republic is the only state that appears to reduce the overall residual of the model significantly.

With the addition of $MING$ in passing to the model (4.11), there are some changes in the nuclear reliances of "*de minimis*" states shown in Figure 5.4. Likewise to the previous step the bulk of the posterior estimated nuclear reliance data points (the circles) lie below the corresponding prior data (the x's). There has been three *net* transfer between cells of the confusion matrix. Largest movement observed from the addition was in Canada, Australia, Kazakhstan, and South Africa as labeled in Figure 5.4, of which all except Canada cross a cell boundary.

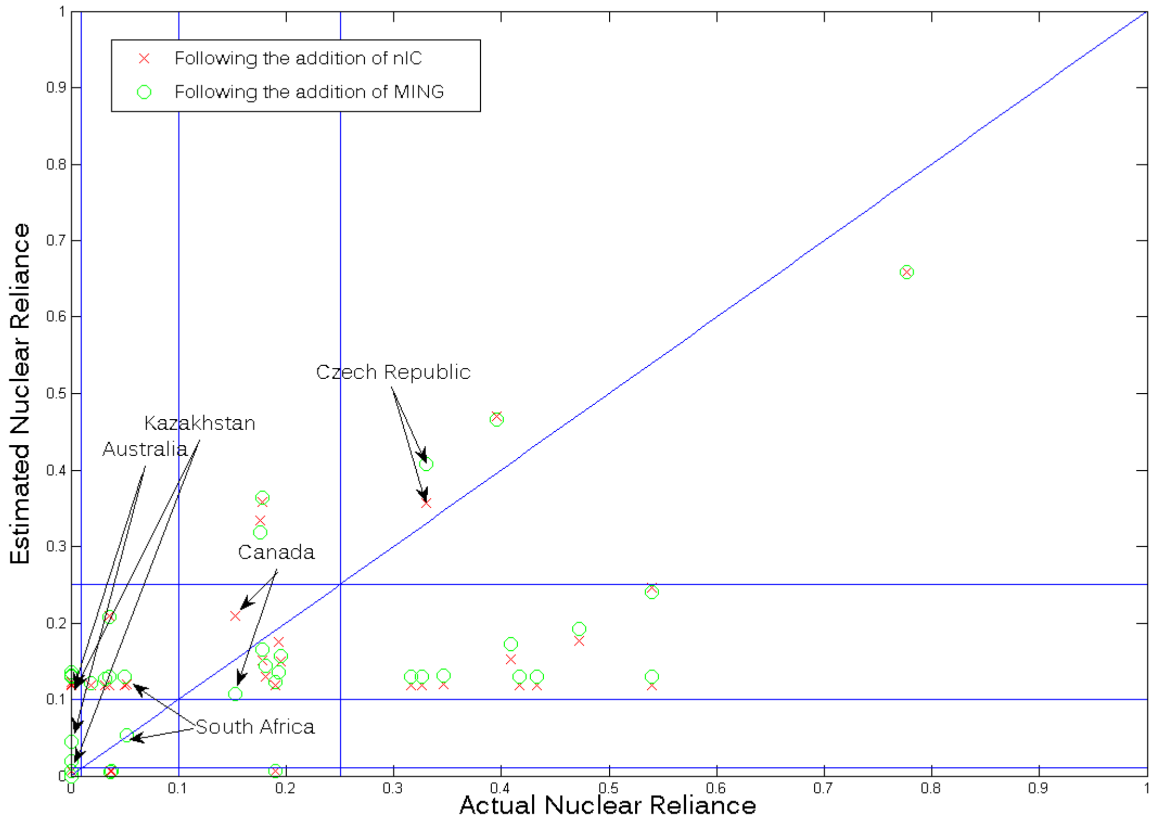

Figure 5.4: Actual Nuclear Reliance versus estimated Nuclear Reliance before and after and addition of *MING* (4.11)

5.3 Residual Scatter Plots

In this section the evolution of the model through the prism of plots of the residuals, both before and after each step, vs. the actual nuclear reliance is viewed, where "residual" is defined as actual nuclear reliance minus estimated nuclear reliance. In this and the following section a question of particular interest is how it can happen that an improved fit for a single state can effectively lead to selection of a variable for the model. This objective leads to the choice of residual for the ordinate in the scatter plots of this section, because large positive and negative residuals in each step can be observed. Residual tables in Section 5.4 were used to obtain and discuss square residual data because the ordinary least-squares regression underlying this work selects the linear fit that, at each step, minimizes the sum of the square residuals.

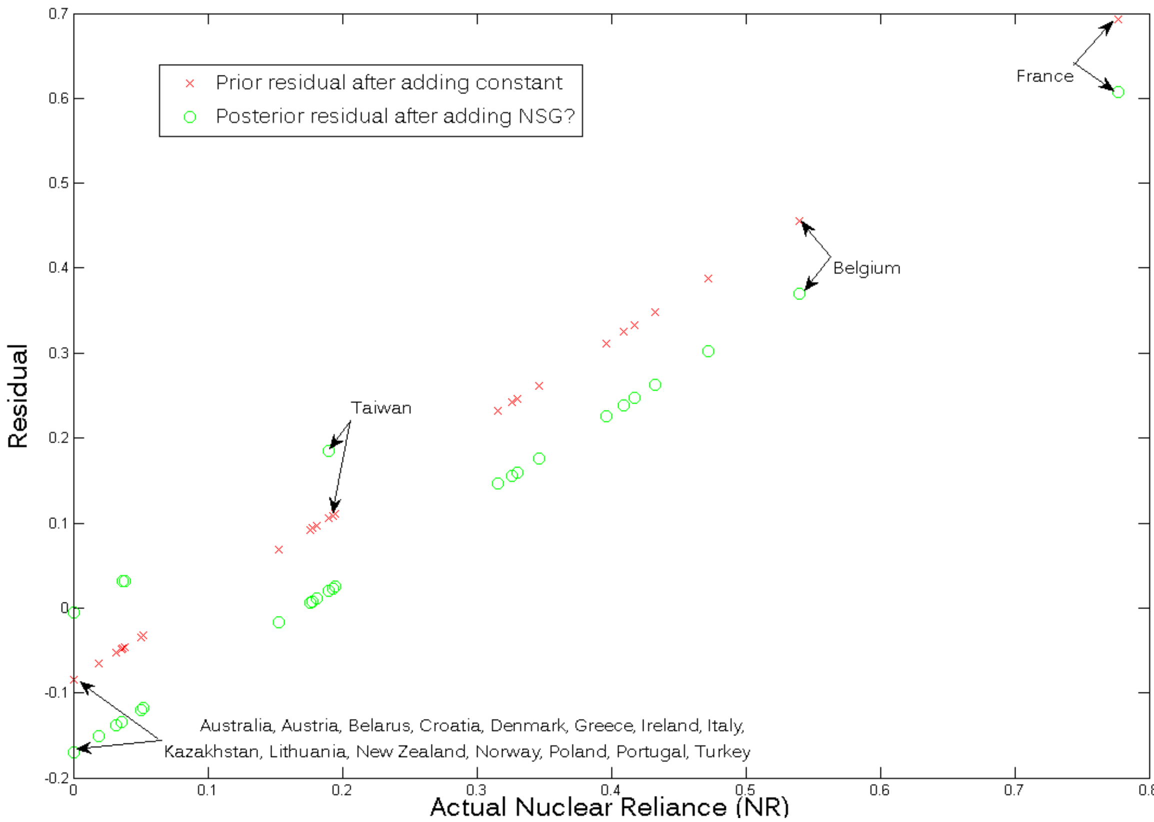

Figure 5.5: Prior and posterior residual after the addition of $NSG?$ versus Nuclear Reliance

Figure 5.5 shows a residual scatter plot, for the step in which $NSG?$ is the variable added (i.e., the step leading to the model (4.2). Here the x's are the residuals prior to selection of NSG, and the circles reflect the residual after the selection. The most striking feature of these data is that selection of $NSG?$ significantly reduces the residual of all (12) states having nuclear reliance greater than 0.25; i.e., having aggressive actual NR. Thus in this case the variable selection seems associated with improving the estimation (decreasing the magnitude of the residual) for a sizable number of states, presumably the bulk of the states that in Section 5.1 were termed as "underestimated." (Note: Only 11 different data points are perceptible in this graph, but that at value 0.34 for the actual NR represents *two* states that have exactly the same actual and estimated values of NR, to the accuracy retained in the data from which this figure was plotted.) But addition of NSG? to the model does not occasion a uniform reduction of the square residual, across all states. The data point at actual NR equal to zero, square residual ≈ 0.03, reflects the block of 15 states in the *de minimis*-actual NR, moderate-estimated NR cell of the confusion matrix 5.2; the square residuals for those states, termed as severely overestimated in Section 5.1, actually increase from this step. Likewise so does that of the state having $NR \approx .19$, which is the single occupant (Taiwan) of the cell (moderate, *de minimis*) in the confusion matrix of Table 5.2 in Section 5.1.

Figure 5.6 is the residual scatter plot for the step leading to the model (4.3), in which $RePro$ is the independent variable selected for addition to the model. By contrast to the preceding such scatter plot, the striking feature of it is the huge reduction of residual in the outlying state (France) having (by far) the largest residual prior to addition of $RePro$ to the model. The residuals of the other states having aggressive nuclear reliances seem slightly to increase as a result of this selection, as do those of the two states (Russia and the UK) having actual moderate nuclear reliance that undergo any significant change. But these deteriorated estimations are in some part compensated by improved estimates, in the sense of residual having a smaller magnitude for the states having moderate or *de minimis* actual nuclear reliances. But none of these latter changes are comparable in magnitude, for an individual state, to the reduction in magnitude square residual achieved for France. It is, however, still possible that in aggregate these effects could be comparable. That possibility will be further considered in the following section.

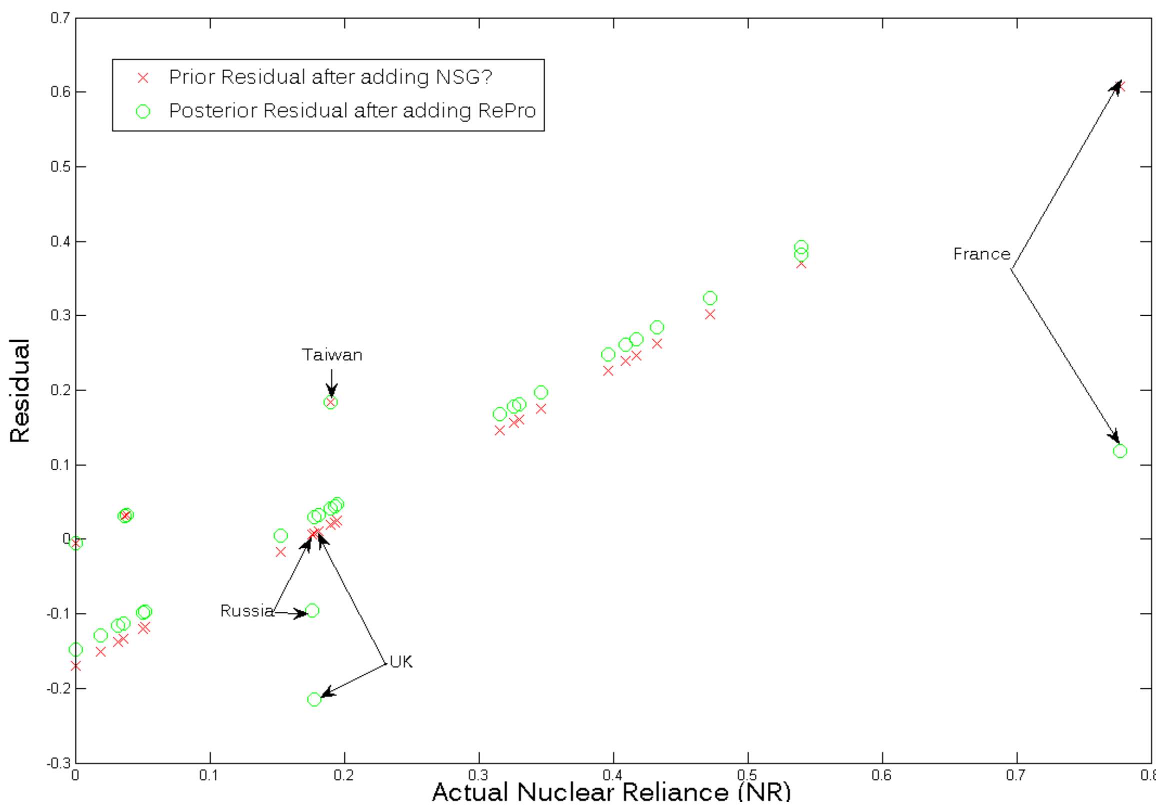

Figure 5.6: Prior and posterior residual after the addition of RePro versus Nuclear Reliance

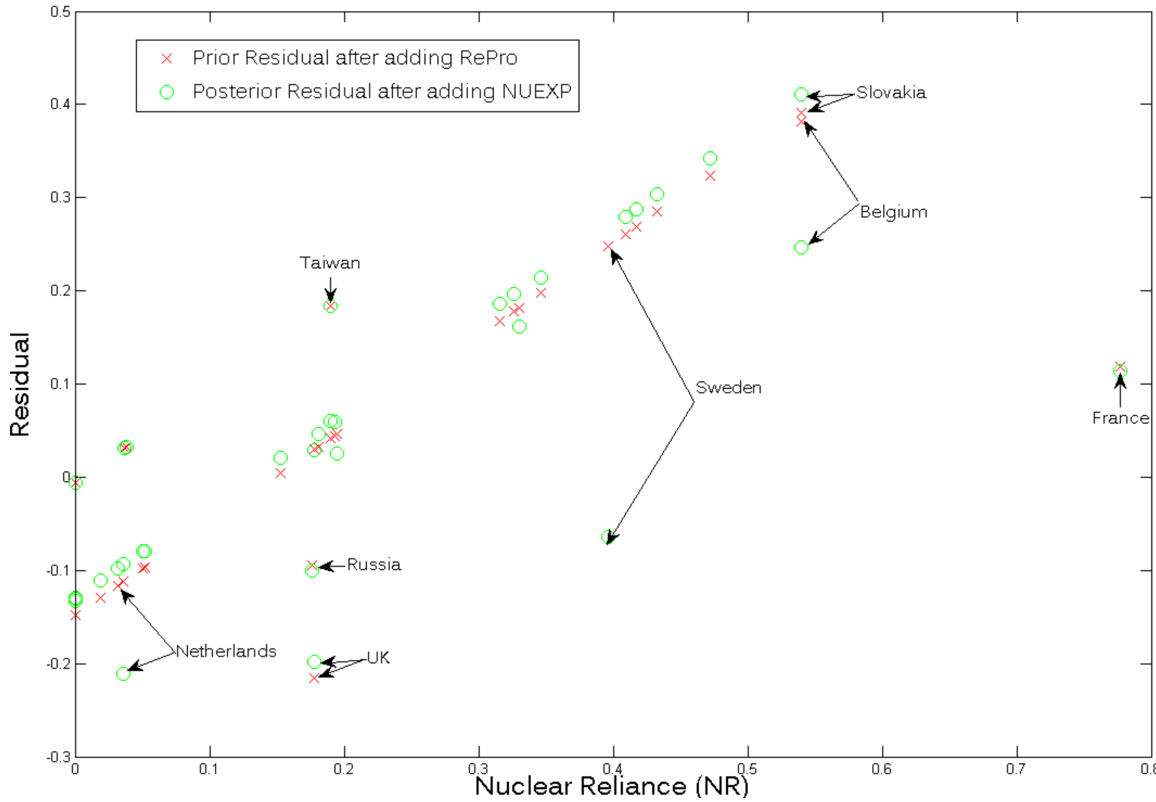

Figure 5.7: Prior and posterior residual after the addition of $NUEXP$ versus Nuclear Reliance

Sources of the reduction in residual from adding $NUEXP$, as a result in the model (4.4) are even more difficult to discern from the corresponding residual scatter plot, displayed as Figure 5.7. Two things that do stand out are that Belgium (as identified from the residual table when $NUEXP$ was added shown in Section 5.4) undergoes the largest reduction in magnitude of square residual from this step, and Sweden (as similarly identified) undergoes by far the second largest such reduction. These are two of the states in the previously identified band of underestimated states, in fact the two previously identified as moving in this step from the confusion-matrix cell actual aggressive estimated moderate nuclear reliance to the improved estimation represented by membership in the cell actual and estimated aggressive nuclear reliance. On the other hand, these significant reductions in square residual are at least partially offset by the fact that the remaining states in the underestimated band suffer slightly enlarged magnitude of square residuals from this step, as do the UK and Netherlands.

Past these more-or-less obviously visible feature, other changes in residual as seen from the residual scatter plot of Figure 5.7 appear rather small. However, it turns out that another significant contributing factor is the relatively smaller reduction in square residual displayed by the multiple states corresponding to the data points plotted at or near the lower left-hand corner of Figure 5.7; this group of states include the 15 states classified above, in the discussion of the corresponding confusion matrices of Tables 5.2 and 5.4, as severely overestimated. Because of the number of these states, the aggregate impact of their small reductions in square residuals is sizable, even though the individual impact of each is rather negligible. This effect is difficult to discern from the residual scatter plot of Figure 5.7, because many of the corresponding data points lie exactly atop each other. For that reason we defer further discussion of this matter until the more detailed picture provided by the residual tables of the following section.

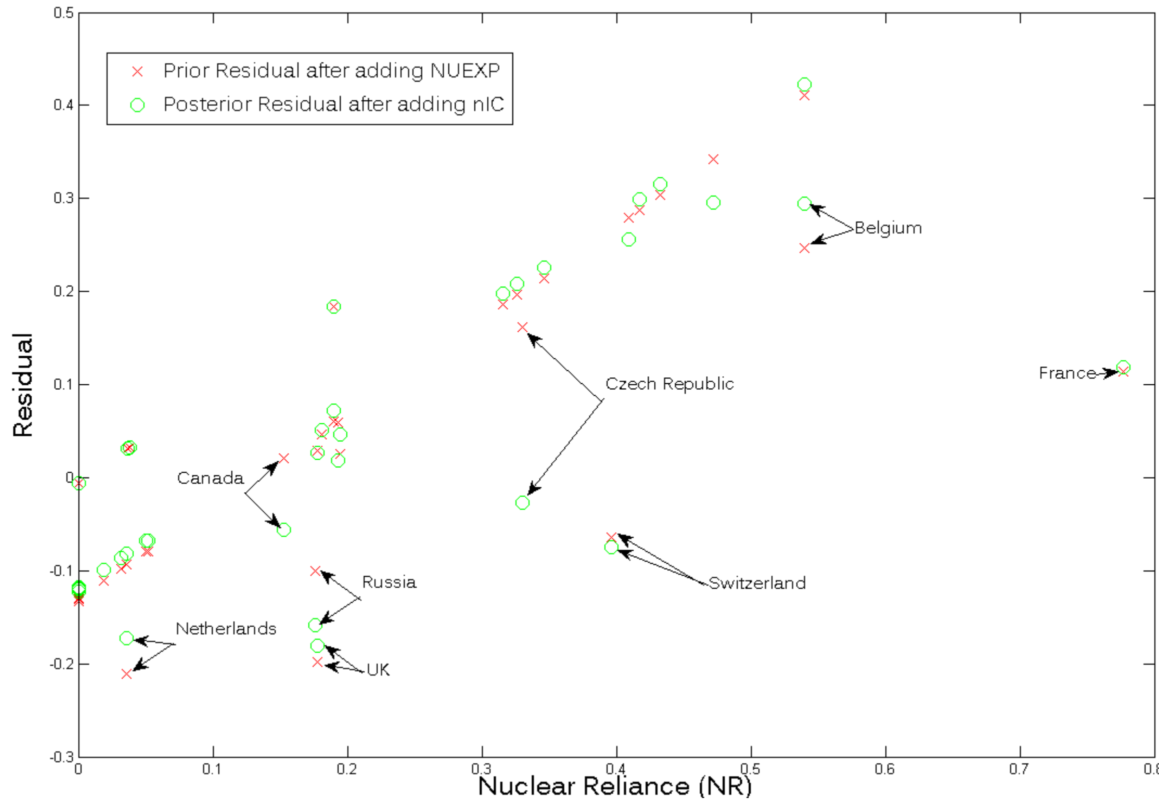

Figure 5.8: Prior and posterior residuals after the addition of nIC versus Nuclear Reliance

Figure 5.8 displays the residual scatter plot for the step in which nIC is selected as the independent variable to be added to the model, which results in the model of Eq. (4.5). What it basically displays is some modest amount of fine tuning of estimated reliances, according as actual nuclear reliances correlate with the variable nIC. This fine tuning occurs primarily among those states having moderate to aggressive actual nuclear reliances (Actual $NR \geq 10\%$). For example, the magnitudes of the square residuals for Ukraine, the Czech Republic, Netherlands, Switzerland and the UK all notably decrease, by decreasing amounts, while those for Belgium, Russia and Slovakia all increase, by decreasing amounts respectively as seen in the residual table when nIC was added in Section 5.4.

The variable nIC, as described in detail in Appendix A, is a measure of a given states trade of nuclear reactors and major nuclear components for an operating commercial civil nuclear reactor unit. It was the first variable that was added to the dataset, which was used as a replacement for the old "subjective" international commercialization. Values used for this variable stemmed from is subjected to data collected from nuclear reactor vendors and major construction and component suppliers; therefore is different than $NUEXP$, which in addition to nuclear reactor supplies considers additional nuclear technology and material used in research, hospitals and so forth (See Appendix A for details).

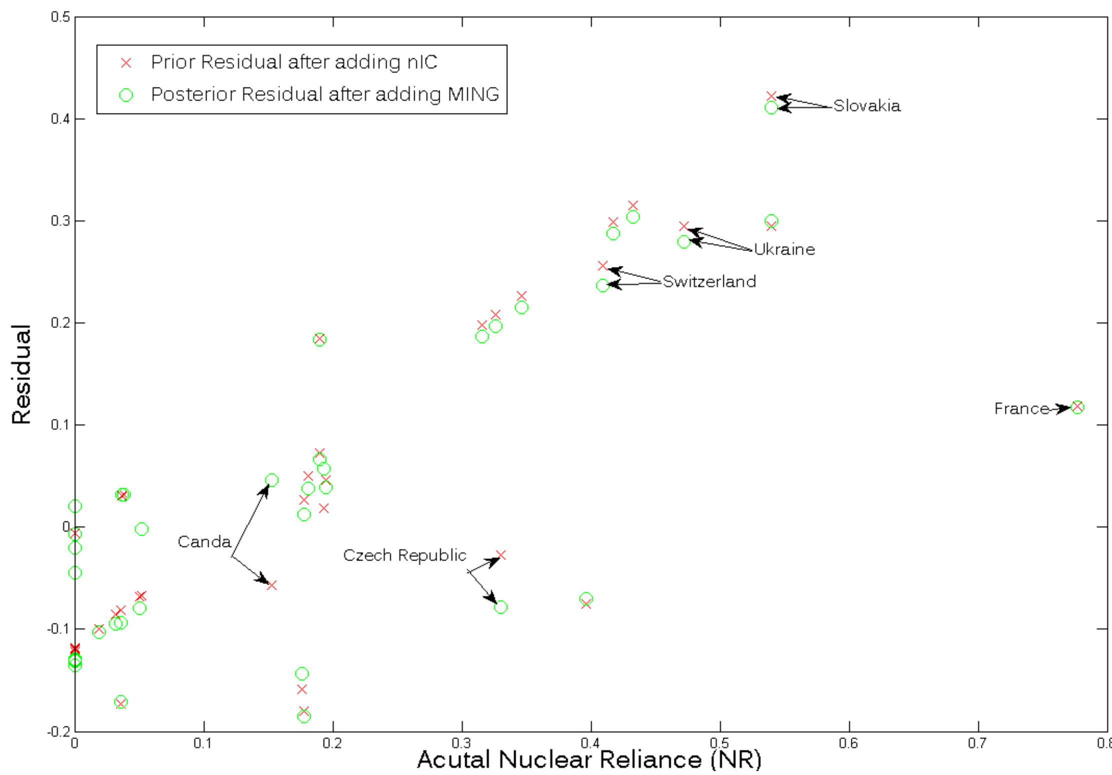

Figure 5.9: Prior and posterior residual after the addition of $MING$ versus Nuclear Reliance

Figure 5.9 displays residual scatter plot for the step in which $MING$ is selected as the independent variable to be added to the model, which results in the model of Eq. (4.11). In Section 5.4, the square residual values shows the reductions of the magnitudes of the square residuals for Kazakhstan and Australia are strong because they are two of three world's largest uranium mining nations when $MING$ was added to the model. They are two states that are members of the NSG, and in previous steps have been included in the "severely overestimated" group; the MING variable effectively identifies their interest in Nuclear Suppliers Group membership as deriving from possession of large reserves of uranium ore, not nuclear reactors.

States that were termed as "underestimated" states in Section 5.1, and primarily Slovakia, Switzerland, Ukraine, undergo relatively large reduction in their respective square residual. Total increase in about 10% nuclear reliance of three of the four coefficients has contributed to the decrease in square residual of above mentioned three states. All of the states that remain in that band undergo model reductions in square residual at this step, although their square residuals remain among the larger, even after this step. At the end of all steps Slovakia contributes with largest residual compared to other states.

5.4 Tabular Residual Analyses

Table 5.7 is the residual table for the step in which the null hypothesis (all states have zero nuclear reliance) is replaced by the mean of the actual nuclear reliances as the uniform estimator. These data can be used to refer back to the confusion matrix in Table 5.1. This section provides residual data after the addition of each variables into the basic linear model (BLM). Also as seen in the corresponding confusion matrix (Table 5.1), estimated NR for all states is the average nuclear reliance for 86 states. In this step, prior squared residual is the square of the actual nuclear reliances.

Table 5.7: Residual table after the addition of the constant term

State	NR	Estimated NR	Posterior Residual	Prior Residual Squared	Post Residual Squared
Algeria	0	0.0841	-0.0841	0.00000	0.00707
Argentina	0.05	0.0841	-0.0341	0.00250	0.00116
Australia	0	0.0841	-0.0841	0.00000	0.00707
Austria	0	0.0841	-0.0841	0.00000	0.00707
Bangladesh	0	0.0841	-0.0841	0.00000	0.00707
Belarus	0	0.0841	-0.0841	0.00000	0.00707
Belgium	0.54	0.0841	0.4559	0.29160	0.20784
Brazil	0.032	0.0841	-0.0521	0.00102	0.00271
Bulgaria	0.326	0.0841	0.2419	0.10628	0.05852
Canada	0.153	0.0841	0.0689	0.02341	0.00475
Chile	0	0.0841	-0.0841	0.00000	0.00707
China	0.019	0.0841	-0.0651	0.00036	0.00424
Colombia	0	0.0841	-0.0841	0.00000	0.00707
Congo-Kinshasa	0	0.0841	-0.0841	0.00000	0.00707
Croatia	0	0.0841	-0.0841	0.00000	0.00707
Cuba	0	0.0841	-0.0841	0.00000	0.00707
Czech Republic	0.33	0.0841	0.2459	0.10890	0.06047
Denmark	0	0.0841	-0.0841	0.00000	0.00707
Egypt	0	0.0841	-0.0841	0.00000	0.00707
Ethiopia	0	0.0841	-0.0841	0.00000	0.00707
Finland	0.316	0.0841	0.2319	0.09986	0.05378
France	0.777	0.0841	0.6929	0.60373	0.48011
Germany	0.178	0.0841	0.0939	0.03168	0.00882
Ghana	0	0.0841	-0.0841	0.00000	0.00707
Greece	0	0.0841	-0.0841	0.00000	0.00707
Guatemala	0	0.0841	-0.0841	0.00000	0.00707
Hong Kong	0	0.0841	-0.0841	0.00000	0.00707
Hungary	0.433	0.0841	0.3489	0.18749	0.12173
India	0.037	0.0841	-0.0471	0.00137	0.00222
Indonesia	0	0.0841	-0.0841	0.00000	0.00707
Iran	0	0.0841	-0.0841	0.00000	0.00707
Iraq	0	0.0841	-0.0841	0.00000	0.00707
Ireland	0	0.0841	-0.0841	0.00000	0.00707
Israel	0	0.0841	-0.0841	0.00000	0.00707
Italy	0	0.0841	-0.0841	0.00000	0.00707
Japan	0.181	0.0841	0.0969	0.03276	0.00939
Kazakhstan	0	0.0841	-0.0841	0.00000	0.00707
Kenya	0	0.0841	-0.0841	0.00000	0.00707
Kuwait	0	0.0841	-0.0841	0.00000	0.00707
Lebanon	0	0.0841	-0.0841	0.00000	0.00707
Libya	0	0.0841	-0.0841	0.00000	0.00707
Lithuania	0	0.0841	-0.0841	0.00000	0.00707
Malaysia	0	0.0841	-0.0841	0.00000	0.00707
Mexico	0.036	0.0841	-0.0481	0.00130	0.00231
Morocco	0	0.0841	-0.0841	0.00000	0.00707
Myanmar	0	0.0841	-0.0841	0.00000	0.00707
Nepal	0	0.0841	-0.0841	0.00000	0.00707
Netherlands	0.036	0.0841	-0.0481	0.00130	0.00231
New Zealand	0	0.0841	-0.0841	0.00000	0.00707
Nigeria	0	0.0841	-0.0841	0.00000	0.00707
North Korea	0	0.0841	-0.0841	0.00000	0.00707
Norway	0	0.0841	-0.0841	0.00000	0.00707
Pakistan	0.038	0.0841	-0.0461	0.00144	0.00213
Peru	0	0.0841	-0.0841	0.00000	0.00707
Philippines	0	0.0841	-0.0841	0.00000	0.00707
Poland	0	0.0841	-0.0841	0.00000	0.00707
Portugal	0	0.0841	-0.0841	0.00000	0.00707
Qatar	0	0.0841	-0.0841	0.00000	0.00707
Romania	0.19	0.0841	0.1059	0.03610	0.01121
Russia	0.176	0.0841	0.0919	0.03098	0.00845

Table 5.7: Continued

State	NR	Estimated NR	Posterior Residual	Prior Residual Squared	Post Residual Squared
Saudi Arabia	0	0.0841	-0.0841	0.00000	0.00707
Serbia	0	0.0841	-0.0841	0.00000	0.00707
Singapore	0	0.0841	-0.0841	0.00000	0.00707
Slovakia	0.54	0.0841	0.4559	0.29160	0.20784
Slovenia	0.417	0.0841	0.3329	0.17389	0.11082
South Africa	0.052	0.0841	-0.0321	0.00270	0.00103
South Korea	0.346	0.0841	0.2619	0.11972	0.06859
Spain	0.195	0.0841	0.1109	0.03803	0.01230
Sri Lanka	0	0.0841	-0.0841	0.00000	0.00707
Sudan	0	0.0841	-0.0841	0.00000	0.00707
Sweden	0.396	0.0841	0.3119	0.15682	0.09728
Switzerland	0.409	0.0841	0.3249	0.16728	0.10556
Syria	0	0.0841	-0.0841	0.00000	0.00707
Taiwan	0.1902	0.0841	0.1061	0.03618	0.01126
Tanzania	0	0.0841	-0.0841	0.00000	0.00707
Thailand	0	0.0841	-0.0841	0.00000	0.00707
Tunisia	0	0.0841	-0.0841	0.00000	0.00707
Turkey	0	0.0841	-0.0841	0.00000	0.00707
UAE	0	0.0841	-0.0841	0.00000	0.00707
UK	0.178	0.0841	0.0939	0.03168	0.00882
Ukraine	0.472	0.0841	0.3879	0.22278	0.15047
USA	0.193	0.0841	0.1089	0.03725	0.01186
Uzbekistan	0	0.0841	-0.0841	0.00000	0.00707
Venezuela	0	0.0841	-0.0841	0.00000	0.00707
Vietnam	0	0.0841	-0.0841	0.00000	0.00707
Yemen	0	0.0841	-0.0841	0.00000	0.00707

Table 5.8 is the residual table for the step in which NSG? is the variable selected for addition to the model. If these data are sorted on actual Nuclear Reliances, then the 15 severely overestimated states in the corresponding confusion matrix (Table 5.2) are revealed as: Australia, Austria, Belarus, Croatia, Denmark, Greece, Ireland, Italy, Kazakhstan, Lithuania, New Zealand, Norway, Poland, Portugal and Turkey. These are all members of the Nuclear Suppliers Group that had no operating nuclear power plant within their boundaries, as of 2011. A few of them, notably Lithuania, Poland and Turkey, currently are considering building NPPs [21]. Similarly, 12 underestimated states of that confusion matrix are identifiable, in order of decreasing (actual) Nuclear Reliances as: France, Belgium, Slovakia, Ukraine, Hungary, Slovenia, Switzerland, Sweden, South Korea, Czech Republic, Bulgaria and Finland. It

seems unclear what those states have in common, beyond being members of the Nuclear Suppliers Group and having an aggressive Nuclear Reliance. The only other cell in this corresponding confusion matrix (Table 5.2) having nonzero entry and distance two from the diagonal is that corresponding to moderate actual Nuclear Reliance and *de minimis* estimated Nuclear Reliance. The one entry in that cell represents Taiwan, which is not a member of the Nuclear Suppliers Group, but nonetheless has a Nuclear Reliance just short of 20%.

Table 5.8: Residual table after the addition of NSG?

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Algeria	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Argentina	0.05	0.170	-0.0341	-0.12002	0.00116	0.01441
Australia	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Austria	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Bangladesh	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Belarus	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Belgium	0.54	0.170	0.4559	0.36998	0.20784	0.13688
Brazil	0.032	0.170	-0.0521	-0.13802	0.00271	0.01905
Bulgaria	0.326	0.170	0.2419	0.15598	0.05852	0.02433
Canada	0.153	0.170	0.0689	-0.01702	0.00475	0.00029
Chile	0	0.006	-0.0841	-0.00589	0.00707	0.00003
China	0.019	0.170	-0.0651	-0.15102	0.00424	0.02281
Colombia	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Congo-Kinshasa	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Croatia	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Cuba	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Czech Republic	0.33	0.170	0.2459	0.15998	0.06047	0.02559
Denmark	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Egypt	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Ethiopia	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Finland	0.316	0.170	0.2319	0.14598	0.05378	0.02131
France	0.777	0.170	0.6929	0.60698	0.48011	0.36842
Germany	0.178	0.170	0.0939	0.00798	0.00882	0.00006
Ghana	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Greece	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Guatemala	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Hong Kong	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Hungary	0.433	0.170	0.3489	0.26298	0.12173	0.06916
India	0.037	0.006	-0.0471	0.03111	0.00222	0.00097
Indonesia	0	0.006	-0.0841	-0.00589	0.00707	0.00003

Table 5.8: Continued

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Iran	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Iraq	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Ireland	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Israel	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Italy	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Japan	0.181	0.170	0.0969	0.01098	0.00939	0.00012
Kazakhstan	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Kenya	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Kuwait	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Lebanon	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Libya	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Lithuania	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Malaysia	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Mexico	0.036	0.170	-0.0481	-0.13402	0.00231	0.01796
Morocco	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Myanmar	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Nepal	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Netherlands	0.036	0.170	-0.0481	-0.13402	0.00231	0.01796
New Zealand	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Nigeria	0	0.006	-0.0841	-0.00589	0.00707	0.00003
North Korea	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Norway	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Pakistan	0.038	0.006	-0.0461	0.03211	0.00213	0.00103
Peru	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Philippines	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Poland	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Portugal	0	0.170	-0.0841	-0.17002	0.00707	0.02891
Qatar	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Romania	0.19	0.170	0.1059	0.01998	0.01121	0.00040
Russia	0.176	0.170	0.0919	0.00598	0.00845	0.00004
Saudi Arabia	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Serbia	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Singapore	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Slovakia	0.54	0.170	0.4559	0.36998	0.20784	0.13688
Slovenia	0.417	0.170	0.3329	0.24698	0.11082	0.06100
South Africa	0.052	0.170	-0.0321	-0.11802	0.00103	0.01393
South Korea	0.346	0.170	0.2619	0.17598	0.06859	0.03097
Spain	0.195	0.170	0.1109	0.02498	0.01230	0.00062
Sri Lanka	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Sudan	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Sweden	0.396	0.170	0.3119	0.22598	0.09728	0.05107
Switzerland	0.409	0.170	0.3249	0.23898	0.10556	0.05711
Syria	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Taiwan	0.1902	0.006	0.1061	0.18431	0.01126	0.03397
Tanzania	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Thailand	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Tunisia	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Turkey	0	0.170	-0.0841	-0.17002	0.00707	0.02891
UAE	0	0.006	-0.0841	-0.00589	0.00707	0.00003
UK	0.178	0.170	0.0939	0.00798	0.00882	0.00006
Ukraine	0.472	0.170	0.3879	0.30198	0.15047	0.09119
USA	0.193	0.170	0.1089	0.02298	0.01186	0.00053
Uzbekistan	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Venezuela	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Vietnam	0	0.006	-0.0841	-0.00589	0.00707	0.00003
Yemen	0	0.006	-0.0841	-0.00589	0.00707	0.00003

Table 5.9 is the residual table for the step in which the variable $REPRO$ is selected for inclusion, which results in the model (4.3). The only states having nonzero values for this variable are, in order of decreasing value, France, UK, Russia and Belgium. These correspond respectively to the following *civil* reprocessing facilities [22]:

1. Areva NC La Hague UP2-800, France
2. Areva NC La Hague UP23, France
3. Areva NC Melox, France
4. NDA B205 Magnox Reprocessing, Sellafield, UK
5. RT-1, Combined Mayak, Chelyabinsk, Ozersk, Russia
6. FBFC International - MOX, Belgium

The changes in estimated NR and residuals occasioned by addition of $RePro$ to the model have already been illustrated in Figures 5.1 and 5.7, with callout for the effect on the above-listed states (except for Belgium, for which this impact is rather small) that actually engage in civil reprocessing.

The net impact, of addition of $RePro$ to the model, on the sum of the squares of the residuals is also of interest. That discussion is deferred to the summary of this section given in the following section in p. 59, in order to consider more-or-less simultaneously the impact of all variables added to the model. Table 5.9 was obtained when $RePro$ was added to the model (4.3)

Table 5.9: Residual table after the addition of *RePro*

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Algeria	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Argentina	0.05	0.12957	-0.09839	-0.07957	0.00968	0.00633
Australia	0	0.12959	-0.14839	-0.12959	0.02202	0.01679
Austria	0	0.13081	-0.14839	-0.13081	0.02202	0.01711
Bangladesh	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Belarus	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Belgium	0.54	0.29295	0.38140	0.24705	0.14547	0.06104
Brazil	0.032	0.12957	-0.11639	-0.09757	0.01355	0.00952
Bulgaria	0.326	0.12957	0.17761	0.19643	0.03155	0.03859
Canada	0.153	0.13222	0.00461	0.02078	0.00002	0.00043
Chile	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
China	0.019	0.12958	-0.12939	-0.11058	0.01674	0.01223
Colombia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Congo-Kinshasa	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Croatia	0	0.13284	-0.14839	-0.13284	0.02202	0.01765
Cuba	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Czech Republic	0.33	0.16862	0.18161	0.16138	0.03298	0.02604
Denmark	0	0.12962	-0.14839	-0.12962	0.02202	0.01680
Egypt	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Ethiopia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Finland	0.316	0.12957	0.16761	0.18643	0.02809	0.03476
France	0.777	0.66250	0.11864	0.11450	0.01408	0.01311
Germany	0.178	0.14899	0.02961	0.02901	0.00088	0.00084
Ghana	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Greece	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Guatemala	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Hong Kong	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Hungary	0.433	0.12957	0.28461	0.30343	0.08100	0.09207
India	0.037	0.00589	0.03111	0.03111	0.00097	0.00097
Indonesia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Iran	0	0.00589	-0.00589	-0.00589	0.00003	0.00003
Iraq	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Ireland	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Israel	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Italy	0	0.12960	-0.14839	-0.12960	0.02202	0.01680
Japan	0.181	0.13465	0.03261	0.04635	0.00106	0.00215
Kazakhstan	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Kenya	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Kuwait	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Lebanon	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Libya	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Lithuania	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Malaysia	0	0.00594	-0.00589	-0.00594	0.00003	0.00004
Mexico	0.036	0.12957	-0.11239	-0.09357	0.01263	0.00876
Morocco	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Myanmar	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Nepal	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Netherlands	0.036	0.24694	-0.11239	-0.21094	0.01263	0.04450
New Zealand	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Nigeria	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
North Korea	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Norway	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Pakistan	0.038	0.00589	0.03211	0.03211	0.00103	0.00103
Peru	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Philippines	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Poland	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Portugal	0	0.13071	-0.14839	-0.13071	0.02202	0.01708
Qatar	0	0.00610	-0.00589	-0.00610	0.00003	0.00004

Table 5.9: Continued

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Romania	0.19	0.12957	0.04161	0.06043	0.00173	0.00365
Russia	0.176	0.27655	-0.09491	-0.10055	0.00901	0.01011
Saudi Arabia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Serbia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Singapore	0	0.00619	-0.00589	-0.00619	0.00003	0.00004
Slovakia	0.54	0.12957	0.39161	0.41043	0.15336	0.16845
Slovenia	0.417	0.12957	0.26861	0.28743	0.07215	0.08262
South Africa	0.052	0.13165	-0.09639	-0.07965	0.00929	0.00634
South Korea	0.346	0.13203	0.19761	0.21397	0.03905	0.04578
Spain	0.195	0.17005	0.04661	0.02495	0.00217	0.00062
Sri Lanka	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Sudan	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Sweden	0.396	0.45966	0.24761	-0.06366	0.06131	0.00405
Switzerland	0.409	0.12970	0.26061	0.27930	0.06792	0.07801
Syria	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Taiwan	0.1902	0.00588	0.18431	0.18432	0.03397	0.03397
Tanzania	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Thailand	0	0.00590	-0.00589	-0.00590	0.00003	0.00003
Tunisia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Turkey	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
UAE	0	0.00590	-0.00589	-0.00590	0.00003	0.00003
UK	0.178	0.37633	-0.21542	-0.19833	0.04641	0.03933
Ukraine	0.472	0.12981	0.32361	0.34219	0.10472	0.11709
USA	0.193	0.13335	0.04461	0.05965	0.00199	0.00356
Uzbekistan	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Venezuela	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Vietnam	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Yemen	0	0.00591	-0.00589	-0.00591	0.00003	0.00003

Table 5.10 is the residual table for the step in which the variable $NUEXP$ is selected for addition, which results in the model (4.4). There are several states that have non-zero nuclear technology export; Sweden, Belgium, and Netherlands are the top three exporters per capita respectively. The largest decrease in the magnitude of square residual is seen in Belgium and Sweden, both states have highly aggressive nuclear reliance which was seen form Table 5.9, and also discussed previously in Figure 5.7.

Table 5.10: Residual table after the addition of $NUEXP$

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Algeria	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Argentina	0.05	0.12957	-0.09839	-0.07957	0.00968	0.00633
Australia	0	0.12959	-0.14839	-0.12959	0.02202	0.01679
Austria	0	0.13081	-0.14839	-0.13081	0.02202	0.01711
Bangladesh	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Belarus	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Belgium	0.54	0.29295	0.38140	0.24705	0.14547	0.06104
Brazil	0.032	0.12957	-0.11639	-0.09757	0.01355	0.00952
Bulgaria	0.326	0.12957	0.17761	0.19643	0.03155	0.03859
Canada	0.153	0.13222	0.00461	0.02078	0.00002	0.00043
Chile	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
China	0.019	0.12958	-0.12939	-0.11058	0.01674	0.01223
Colombia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Congo-Kinshasa	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Croatia	0	0.13284	-0.14839	-0.13284	0.02202	0.01765
Cuba	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Czech Republic	0.33	0.16862	0.18161	0.16138	0.03298	0.02604
Denmark	0	0.12962	-0.14839	-0.12962	0.02202	0.01680
Egypt	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Ethiopia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Finland	0.316	0.12957	0.16761	0.18643	0.02809	0.03476
France	0.777	0.66250	0.11864	0.11450	0.01408	0.01311
Germany	0.178	0.14899	0.02961	0.02901	0.00088	0.00084
Ghana	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Greece	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Guatemala	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Hong Kong	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Hungary	0.433	0.12957	0.28461	0.30343	0.08100	0.09207
India	0.037	0.00589	0.03111	0.03111	0.00097	0.00097
Indonesia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Iran	0	0.00589	-0.00589	-0.00589	0.00003	0.00003
Iraq	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Ireland	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Israel	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Italy	0	0.12960	-0.14839	-0.12960	0.02202	0.01680
Japan	0.181	0.13465	0.03261	0.04635	0.00106	0.00215
Kazakhstan	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Kenya	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Kuwait	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Lebanon	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Libya	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Lithuania	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Malaysia	0	0.00594	-0.00589	-0.00594	0.00003	0.00004
Mexico	0.036	0.12957	-0.11239	-0.09357	0.01263	0.00876
Morocco	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Myanmar	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Nepal	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Netherlands	0.036	0.24694	-0.11239	-0.21094	0.01263	0.04450
New Zealand	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Nigeria	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
North Korea	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Norway	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Pakistan	0.038	0.00589	0.03211	0.03211	0.00103	0.00103
Peru	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Philippines	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Poland	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
Portugal	0	0.13071	-0.14839	-0.13071	0.02202	0.01708
Qatar	0	0.00610	-0.00589	-0.00610	0.00003	0.00004

Table 5.10: Continued

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Romania	0.19	0.12957	0.04161	0.06043	0.00173	0.00365
Russia	0.176	0.27655	-0.09491	-0.10055	0.00901	0.01011
Saudi Arabia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Serbia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Singapore	0	0.00619	-0.00589	-0.00619	0.00003	0.00004
Slovakia	0.54	0.12957	0.39161	0.41043	0.15336	0.16845
Slovenia	0.417	0.12957	0.26861	0.28743	0.07215	0.08262
South Africa	0.052	0.13165	-0.09639	-0.07965	0.00929	0.00634
South Korea	0.346	0.13203	0.19761	0.21397	0.03905	0.04578
Spain	0.195	0.17005	0.04661	0.02495	0.00217	0.00062
Sri Lanka	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Sudan	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Sweden	0.396	0.45966	0.24761	-0.06366	0.06131	0.00405
Switzerland	0.409	0.12970	0.26061	0.27930	0.06792	0.07801
Syria	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Taiwan	0.1902	0.00588	0.18431	0.18432	0.03397	0.03397
Tanzania	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Thailand	0	0.00590	-0.00589	-0.00590	0.00003	0.00003
Tunisia	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Turkey	0	0.12957	-0.14839	-0.12957	0.02202	0.01679
UAE	0	0.00590	-0.00589	-0.00590	0.00003	0.00003
UK	0.178	0.37633	-0.21542	-0.19833	0.04641	0.03933
Ukraine	0.472	0.12981	0.32361	0.34219	0.10472	0.11709
USA	0.193	0.13335	0.04461	0.05965	0.00199	0.00356
Uzbekistan	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Venezuela	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Vietnam	0	0.00588	-0.00589	-0.00588	0.00003	0.00003
Yemen	0	0.00591	-0.00589	-0.00591	0.00003	0.00003

Table 5.11 is the residual table for the step in which the variable nIC is selected for addition to the model, which leads to the model (4.5). There are several states that are involved in supplying nuclear reactor and other supporting components for civil nuclear power units. The U.S., Russia, and France are the top three exporters for reactor core and major components, while Czech Republic, Sweden, and Canada are the top exporters per capita respectively [23] (Refer to Appendix B for raw nIC data).

Table 5.11: Residual table after the addition of nIC

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Algeria	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Argentina	0.05	0.11778	-0.07957	-0.06778	0.00633	0.00459
Australia	0	0.11779	-0.12959	-0.11779	0.01679	0.01388
Austria	0	0.11874	-0.13081	-0.11874	0.01711	0.01410
Bangladesh	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Belarus	0	0.11778	-0.12957	-0.11778	0.01679	0.01387
Belgium	0.54	0.24565	0.24705	0.29435	0.06104	0.08664
Brazil	0.032	0.11778	-0.09757	-0.08578	0.00952	0.00736
Bulgaria	0.326	0.11778	0.19643	0.20822	0.03859	0.04336
Canada	0.153	0.20954	0.02078	-0.05654	0.00043	0.00320
Chile	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
China	0.019	0.11818	-0.11058	-0.09918	0.01223	0.00984
Colombia	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Congo-Kinshasa	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Croatia	0	0.12032	-0.13284	-0.12032	0.01765	0.01448
Cuba	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Czech Republic	0.33	0.35706	0.16138	-0.02706	0.02604	0.00073
Denmark	0	0.11782	-0.12962	-0.11782	0.01680	0.01388
Egypt	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Ethiopia	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Finland	0.316	0.11778	0.18643	0.19822	0.03476	0.03929
France	0.777	0.65817	0.11450	0.11883	0.01311	0.01412
Germany	0.178	0.15160	0.02901	0.02640	0.00084	0.00070
Ghana	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Greece	0	0.11778	-0.12957	-0.11778	0.01679	0.01387
Guatemala	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Hong Kong	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Hungary	0.433	0.11778	0.30343	0.31522	0.09207	0.09936
India	0.037	0.00589	0.03111	0.03111	0.00097	0.00097
Indonesia	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Iran	0	0.00589	-0.00589	-0.00589	0.00003	0.00003
Iraq	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Ireland	0	0.11779	-0.12957	-0.11779	0.01679	0.01387
Israel	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Italy	0	0.12215	-0.12960	-0.12215	0.01680	0.01492
Japan	0.181	0.13038	0.04635	0.05062	0.00215	0.00256
Kazakhstan	0	0.11778	-0.12957	-0.11778	0.01679	0.01387
Kenya	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Kuwait	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Lebanon	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Libya	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Lithuania	0	0.11778	-0.12957	-0.11778	0.01679	0.01387
Malaysia	0	0.00593	-0.00594	-0.00593	0.00004	0.00004
Mexico	0.036	0.11778	-0.09357	-0.08178	0.00876	0.00669
Morocco	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Myanmar	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Nepal	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Netherlands	0.036	0.20869	-0.21094	-0.17269	0.04450	0.02982
New Zealand	0	0.11778	-0.12957	-0.11778	0.01679	0.01387
Nigeria	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
North Korea	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Norway	0	0.11779	-0.12957	-0.11779	0.01679	0.01387
Pakistan	0.038	0.00589	0.03211	0.03211	0.00103	0.00103
Peru	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Philippines	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Poland	0	0.11778	-0.12957	-0.11778	0.01679	0.01387
Portugal	0	0.11866	-0.13071	-0.11866	0.01708	0.01408
Qatar	0	0.00605	-0.00610	-0.00605	0.00004	0.00004
Romania	0.19	0.11778	0.06043	0.07222	0.00365	0.00522
Russia	0.176	0.33443	-0.10055	-0.15843	0.01011	0.02510

Table 5.11: Continued

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Saudi Arabia	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Serbia	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Singapore	0	0.00612	-0.00619	-0.00612	0.00004	0.00004
Slovakia	0.54	0.11778	0.41043	0.42222	0.16845	0.17827
Slovenia	0.417	0.11778	0.28743	0.29922	0.08262	0.08953
South Africa	0.052	0.11939	-0.07965	-0.06739	0.00634	0.00454
South Korea	0.346	0.11969	0.21397	0.22631	0.04578	0.05122
Spain	0.195	0.14914	0.02495	0.04586	0.00062	0.00210
Sri Lanka	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Sudan	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Sweden	0.396	0.47073	-0.06366	-0.07473	0.00405	0.00558
Switzerland	0.409	0.15261	0.27930	0.25639	0.07801	0.06573
Syria	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Taiwan	0.1902	0.00588	0.18432	0.18432	0.03397	0.03397
Tanzania	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Thailand	0	0.00590	-0.00590	-0.00590	0.00003	0.00003
Tunisia	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Turkey	0	0.11778	-0.12957	-0.11778	0.01679	0.01387
UAE	0	0.00590	-0.00590	-0.00590	0.00003	0.00003
UK	0.178	0.35834	-0.19833	-0.18034	0.03933	0.03252
Ukraine	0.472	0.17690	0.34219	0.29510	0.11709	0.08709
USA	0.193	0.17463	0.05965	0.01837	0.00356	0.00034
Uzbekistan	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Venezuela	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Vietnam	0	0.00588	-0.00588	-0.00588	0.00003	0.00003
Yemen	0	0.00590	-0.00591	-0.00590	0.00003	0.00003

Table 5.12 is the residual table for the step in which the variable $MING$ is selected for addition to the model, which leads to the model (4.11). The only states having nonzero values for this variable are, in order of decreasing value, Canada, Kazakhstan, Australia and South Africa [22] (See data for $MING$ in Appendix B). These correspond respectively to the following *civil* mining and milling facilities:

1. Key Lake/McArthur River,Saskatchewan, Canada

2. McClean Lake, Saskatchewan, Canada

3. Rabbit Lake, Saskatchewan, Canada

4. Appak LLP, Chimkent Region, Kazakhstan

5. Betpak-Dala JV LLP, South-Kazakhstan Oblast, Kazakhstan

6. Centralnoye (Taukent), Chimkent Region, Kazakhstan

7. JV Inkai, Chimkent Region, Kazakhstan

8. JV Katco (Moynkum), Chimkent Region, Kazakhstan

9. KenDala.kz JSC, Chimkent Region, Kazakhstan

10. Mining Group 6 LLP, Chimkent Region, Kazakhstan

11. Stepnoye Mining Group LLP, Chimkent Region, Kazakhstan

12. Beverley, South Australia, Australia

13. Olympic Dam, South Australia, Australia

14. Ranger, Northern Territory, Australia

15. Nuclear Fuels Corporation (NUFCOR), Gauteng, South Africa

16. Uranium One, Doornfontain, South Africa

17. Uranium One, Dominion, South Africa

18. Vaal Reefs - 2, Gauteng, South Africa

Table 5.12 shows the residual table after the addition of $MING$. The changes in estimated NR and residuals occasion by addition of $MING$ to the model have already been illustrated in Figures 5.4 and 5.9, which showed some improvement in the “overestimated” cluster of states discussed in Section 5.1.

Table 5.12: Residual table after the addition of $MING$

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Algeria	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Argentina	0.05	0.12950	-0.06778	-0.07950	0.00459	0.00632
Australia	0	0.04458	-0.11779	-0.04458	0.01388	0.00199
Austria	0	0.13032	-0.11874	-0.13032	0.01410	0.01698
Bangladesh	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Belarus	0	0.12950	-0.11778	-0.12950	0.01387	0.01677
Belgium	0.54	0.23990	0.29435	0.30010	0.08664	0.09006
Brazil	0.032	0.12644	-0.08578	-0.09444	0.00736	0.00892
Bulgaria	0.326	0.12950	0.20822	0.19650	0.04336	0.03861
Canada	0.153	0.10712	-0.05654	0.04588	0.00320	0.00211
Chile	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
China	0.019	0.12189	-0.09918	-0.10289	0.00984	0.01059
Colombia	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Congo-Kinshasa	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Croatia	0	0.13167	-0.12032	-0.13167	0.01448	0.01734
Cuba	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Czech Republic	0.33	0.40781	-0.02706	-0.07781	0.00073	0.00605
Denmark	0	0.12954	-0.11782	-0.12954	0.01388	0.01678
Egypt	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Ethiopia	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Finland	0.316	0.12950	0.19822	0.18650	0.03929	0.03478
France	0.777	0.65910	0.11883	0.11790	0.01412	0.01390
Germany	0.178	0.16538	0.02640	0.01262	0.00070	0.00016
Ghana	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Greece	0	0.12950	-0.11778	-0.12950	0.01387	0.01677
Guatemala	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Hong Kong	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Hungary	0.433	0.12950	0.31522	0.30350	0.09936	0.09211
India	0.037	0.00493	0.03111	0.03207	0.00097	0.00103
Indonesia	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Iran	0	0.00651	-0.00589	-0.00651	0.00003	0.00004
Iraq	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Ireland	0	0.12950	-0.11779	-0.12950	0.01387	0.01677
Israel	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Italy	0	0.13484	-0.12215	-0.13484	0.01492	0.01818
Japan	0.181	0.14348	0.05062	0.03752	0.00256	0.00141
Kazakhstan	0	0.01971	-0.11778	-0.01971	0.01387	0.00039
Kenya	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Kuwait	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Lebanon	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Libya	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Lithuania	0	0.12950	-0.11778	-0.12950	0.01387	0.01677
Malaysia	0	0.00654	-0.00593	-0.00654	0.00004	0.00004
Mexico	0.036	0.12950	-0.08178	-0.09350	0.00669	0.00874
Morocco	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Myanmar	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Nepal	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Netherlands	0.036	0.20737	-0.17269	-0.17137	0.02982	0.02937
New Zealand	0	0.12950	-0.11778	-0.12950	0.01387	0.01677
Nigeria	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
North Korea	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Norway	0	0.12950	-0.11779	-0.12950	0.01387	0.01677
Pakistan	0.038	0.00651	0.03211	0.03149	0.00103	0.00099
Peru	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Philippines	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Poland	0	0.12950	-0.11778	-0.12950	0.01387	0.01677
Portugal	0	0.13025	-0.11866	-0.13025	0.01408	0.01697
Qatar	0	0.00665	-0.00605	-0.00665	0.00004	0.00004
Romania	0.19	0.12311	0.07222	0.06689	0.00522	0.00447
Russia	0.176	0.31896	-0.15843	-0.14296	0.02510	0.02044

Table 5.12: Continued

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Residual Squared	Post Residual Squared
Saudi Arabia	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Serbia	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Singapore	0	0.00671	-0.00612	-0.00671	0.00004	0.00004
Slovakia	0.54	0.12950	0.42222	0.41050	0.17827	0.16851
Slovenia	0.417	0.12950	0.29922	0.28750	0.08953	0.08266
South Africa	0.052	0.05374	-0.06739	-0.00174	0.00454	0.00000
South Korea	0.346	0.13114	0.22631	0.21486	0.05122	0.04617
Spain	0.195	0.15636	0.04586	0.03864	0.00210	0.00149
Sri Lanka	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Sudan	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Sweden	0.396	0.46655	-0.07473	-0.07055	0.00558	0.00498
Switzerland	0.409	0.17213	0.25639	0.23687	0.06573	0.05611
Syria	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Taiwan	0.1902	0.00650	0.18432	0.18370	0.03397	0.03375
Tanzania	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Thailand	0	0.00652	-0.00590	-0.00652	0.00003	0.00004
Tunisia	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Turkey	0	0.12950	-0.11778	-0.12950	0.01387	0.01677
UAE	0	0.00652	-0.00590	-0.00652	0.00003	0.00004
UK	0.178	0.36324	-0.18034	-0.18524	0.03252	0.03431
Ukraine	0.472	0.19283	0.29510	0.27917	0.08709	0.07793
USA	0.193	0.13587	0.01837	0.05713	0.00034	0.00326
Uzbekistan	0	-0.02050	-0.00588	0.02050	0.00003	0.00042
Venezuela	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Vietnam	0	0.00650	-0.00588	-0.00650	0.00003	0.00004
Yemen	0	0.00652	-0.00590	-0.00652	0.00003	0.00004

5.5 Summary

As speculated at the end of Section 5.1 (pages 29-30), some of the variables chosen in the model seem to be influenced by the a few states, which was repeatedly seen throughout our analyses from confusion matrix to tabular residual analysis. Relating back to the desirable model properties discussed in Section 1, the model evolution shows stability in each step from Eq. (4.1)-(4.11) is studied, as the coefficient of the variables do not increase or decrease quite dramatically. The statistics (shown in Table 4.1) are very close to Nelson-Sprecher model but slightly less than $R^2 \approx .53$ and adj. $R^2 \approx .50$ obtained in Nelson-Sprecher model in reference [1]. Category representation, as discussed in Chatper 1, in model (4.11) is slightly less wide than desired. Out of five categories the model represents international commerce (represented by NSG, NUEXP and nIC), nuclear material processing (represented by RePro and MING), but Population (POP) also has strong influence in the selected variables because of per capita standardization; therefore "people and economy" also is concealed in the obtained model (4.11). All the models for forward stepwise regression are obtained using STEPWISEFIT function of MatLab [15] and backward elimination model is obtained using BACKWARD under GLMSELECT of Statistical Analysis System (SAS) [18]. STEPWISE graphical interface discussed in Section 3.2 is further used to select candidate independent variable representing any missing category that came closest to being included in Section 4.4. Following residual summary in Table 5.13 provides insight of the largest contribution from one state and contributions from severely overestimated and underestimated states, following each step in the evolution of the model in Eq. (4.11). Table 5.13 shows continues reduction in the Sum of squared residuals, barring the addition of NSG? into the model, the most significant reduction in the sum of squared residual can be in seen when $RePro$ ($\approx 18\%$ reduction) is added into the basic linear model in Eq. (4.3).

Table 5.13: Residual summary

Last variable added	Sum of squared residuals	Largest single-state contributor (state)	Contribution from severely overestimated states	Contribution from underestimated states
Constant	2.84	16.9% (France)	10.6%	60.7%
NSG?	1.65	22.3% (France)	26.3%	65.1%
RePro	1.34	15.3% (Slovakia)	22.0%	62.1%
NUEXP	1.20	14.0%(Slovakia)	21.4%	57.0%
nIC	1.14	15.6% (Slovakia)	18.4%	65.0%
MING	1.12	15.0% (Slovakia)	19.7%	62.2%

A brief history of Slovakian civil nuclear history could help in understanding the largest residual contribution from Slovakia in Table 5.13 when four of the five estimators were added to the model. Slovakia imports more than 78 percent of its power sources from Russia. Nuclear energy is mostly used for electricity generation. The following information is adapted from [24]. Slovakia has the history of nuclear program since 1956 when the contemporary governments of USSR and CSSR agreed to build industrial and research nuclear power plant on the CSSR territory. In 1958, contemporary Czechoslovak government started to built their first gas-cooled heavy water reactor. Since then there have been five major nuclear programs since then: A-1 Bohunice, V-1 Bohunice, V-2 Bohunice, Mochovce 1&2 Mochovce 3&4 respectively. Currently there are five operating nuclear reactor and two are under construction units present in Slovak Republic which are as follows [23]:

1. Bohunice Unit 3 and Unit 4
2. Mochovce Unit 1, Unit 2, Unit 3 (under construction), Unit 4 (under construction)
3. Krsko Unit 1.

Slovakia's long term energy policy includes economically feasible, safe, reliable electric power production, and to decrease the ratio of gross domestic consumption to the gross domestic product. Slovakian nuclear energy policy listed by International Atomic Energy Agency (IAEA) is to have modern technology, and improved safety and increase capacity of NPP V-2 unit, economic and timely management of spent fuel, prudent decision and implementation of new nuclear plants at Mochive, and create suitable condition for "Nuclear Forum" to conduct smoothly [24]. Slovakia has shown strong commitment to nuclear power in the future [25].

Section 6 will study the model when all the natural categories are forced to fit into the basic linear model. The next section will also study a new model, with a crude time series statistical analysis, where the dependent variable is changed by subtracting NR from 1980 from the current NR. Moreover, it will try to address the issue of inaccuracy from "underestimated" states that accounted for two-third of aggregated square residual.

6. INITIAL EXPLORATION OF POSSIBLE ALTERNATIVE MODELS

The focus in Section 6 will be upon an initial exploration of models that use nearly the same independent variables as in the preceding Section 5, but are more satisfactory in some respect or other. Section 6 will focus on searching for the model with minimize the residual contribution from “severely overestimated” and “underestimated” states, first discussed in Section 5.1. Section 6.1 concerns the behavior of the model when the improvement is sought from adding the most likely variables from each natural category listed earlier in Section 2.2. Next, Section 6.2 will seek improvements in the goodness of fit by introducing a “Persistence” ($PERSIST$) variable. Superficial analysis of the Basic Linear Model is conducted with the new ($PERSIST$) variable. Section 6.3 is devoted to a state-by-state residual analysis of the model obtained in Section 6.2. This section is intended to suggest possible alternatives that could be implemented for further study in the future.

6.1 Model with Variables from all Natural Categories

This section is concerned with study of the nature of linear regression model when one mostly likely variable is added from each of the natural five categories (see Section 2.2). The linear regression model obtained in this section is built upon the model in Eq. (4.11). That model already contains two of the five categories listed in Section 2.2. The objective here is to add only three more variable, one from each of the remaining categories, to explore how the model evolves. Each new variable is selected as that having the smallest p-value in a given category. The following evolutionary steps, in Eqs. (6.1)-(6.2), which were built upon Eq. (4.11), were generated using the STEPWISE graphical interface (see Section 3.2).

$$\begin{aligned}\widehat{NR} =&(.13 \pm .028)NSG? + (.57 \pm .14)RePro + (.20 \pm .11)NUEXP \\ &+ (.21 \pm .11)nIC - (.12 \pm .08)MING - (.11 \pm .08)NWS? \\ &+ .006\end{aligned} \tag{6.1}$$

$$\begin{aligned}\widehat{NR} =&(.12 \pm .028)NSG? + (.81 \pm .20)RePro + (.20 \pm .11)NUEXP \\ &+ (.28 \pm .11)nIC - (.13 \pm .08)MING - (.31 \pm .14)NWS? \\ &+ (.29 \pm .17)EGEN + .001\end{aligned} \tag{6.2}$$

$$\begin{aligned}\widehat{NR} =&(.12 \pm .028)NSG? + (.85 \pm .20)RePro + (.20 \pm .11)NUEXP \\ &+ (.28 \pm .11)nIC - (.14 \pm .08)MING - (.31 \pm .14)NWS? \\ &+ (.44 \pm .22)EGEN - (.14 \pm .13)POP + .006\end{aligned} \tag{6.3}$$

The resulting statistical data from each step are listed in Table 6.1. Comparing with three desired properties in Section 1, Eq. (6.3) shows only a slightly better R^2 value of $\approx .54$ when compared to Nelson-Sprecher model ($R^2 \approx .53$)[1]. The model (6.3) selects variables from all five natural category, which is another desired property for the new model. Also looking at all the steps discussed in this section, there are large jumps in $RePro$ when new variables are added to the model. Model in Eq. 4.11 had .44 as coefficient for $RePro$ which increases to .57 (13% increase) when $NWS?$ was added. Moreover, when $EGEN$ is added in Eq. 6.2, coefficient for $RePro$ jumps to .81, which is again a significant increase of 24%. Comparing Eq. (6.2) -(6.3) there is no significant increase in the coefficient of $RePro$, but coefficient of $EGEN$ increases by 15%. These two erratic changes in the coefficient suggest multicollinearity (significant correlation) between estimators in the model, especially $RePro$, NWS, and $EGEN$ (Scope of present work does not permit further pursuit of this interesting possibility). Furthermore, the final model in Eq. (6.3) also has a large p-value of .2857 for POP, as shown in Table 6.1, and also the relative uncertainty in this coefficient is very high.

Table 6.1: Statistical data for model obtained from STEPWISE regression

Step	**p-value(s)**	R^2	**Adj** R^2	**f-stat**	**RSME**
7	$[1.602e-5]$, $[1.46e-4]$, [.0742], [.0113], [.1447], [.1568]	.5176	.4809	14.12	.1167
8	$[3.56e-5]$, $[1.45e-4]$, [.0727], [.0082], [.1051], [.0336], [.1019]	.5340	.4921	12.77	.115
9	$[9.89e-5]$, $[9.05e-5]$, [.0733], [.0105], [.0766], [.0208], [.0529], [.2857]	.5408	.4931	11.33	.115

This superficial analysis intended to produce a model with at least one variables from each of the five categories of Section 2.2 resulted in some interesting statistical observation when the significance level of variables such as NWS? increased from .1447 to .1051 and for $EGEN$ it increased significantly from .1568 to .0336. One of the interesting features to be noted in this analysis was that the statistical significance of NWS? variable increases from p-value of .16 in step 7 to p-value of .02 in step 9.

6.2 Persistence Variable

Focus of this section is to introduce an alternative, which seeks better statistics by fitting the same independent variables to change in NR from the baseline year of 1980, rather than current NR (2011) itself. In the past, some states such as Italy had civil nuclear energy, but the nuclear energy programs were dissolved in it entirety because of major nuclear power plant accidents such as Three Mile Island (TMI) or Chernobyl, and those states have moved towards alternative energy sources [26]. Due to the recent Fukushima nuclear power plant disaster, nuclear phaseout is still in process as states such as Germany are shutting down their current operations of their civil nuclear power plants until in-depth research is conducted [27]. Newly created "Persistence" $PERSIST$ variable collects nuclear reliance data of states from 1980. The year 1980 was precisely chosen because of the 1979 TMI nuclear disaster, and also the Nuclear Non-Proliferation Act, purposed by John Glenn was passed by congress and President Carter on 1978. Data were collected from The World Bank [28]. 1980 NR data for Soviet Union was collected from IAEA Country Profile [29]. Data obtained for Soviet Union had to be deaggregated in order to fit the current list of 86NC. NR data for Soviet was then broken down based on their population of 1980 [30]. Likewise Czechoslovakia's NR was also deaggregrated to Czech Republic and Slovakia based on their population in 1980 [31, 32]. Yugoslavia did not have operating nuclear reactor until 1984 [33]. Table 6.2 provides nuclear reliance data [28] for states having civil nuclear power production in 1980, and states not listed in the table had no nuclear reliance in 1980.

The dependent variable in this case is basically the change during one 30 year time interval ΔNR. The hypothesis here is with the inclusion of "persistence" variable the residual nuclear reliances will decrease in magnitude. Following Eqs. (6.4)-(6.9) are the evolutionary steps of the basic linear model obtained from this analysis.

Table 6.2: States, and their respective nuclear reliance in 1980

State	Nuclear Reliance [%]	State	Nuclear Reliance [%]
Argentina	5.90	Netherlands	6.50
Belgium	23.60	Russia	3.52
Bulgaria	17.70	Slovakia	7.42
Canada	10.20	South Korea	9.30
Czech Republic	15.28	Spain	4.70
Finland	17.20	Sweden	27.50
France	23.80	Switzerland	29.80
India	2.50	UK	13.00
Italy	1.20	Ukraine	1.27
Japan	14.40	USA	11.00
Kazakhstan	0.38	Uzbekistan	0.39
Lithuania	0.09		

$$\widehat{NR} = PERSIST + .055 \tag{6.4}$$

$$\widehat{NR} = (.10 \pm .02)NSG? + PERSIST + .005 \tag{6.5}$$

$$\widehat{NR} = (.09 \pm .028)NSG? + (.34 \pm .09)RePro + PERSIST + .005 \tag{6.6}$$

$$\begin{aligned}\widehat{NR} =& (.10 \pm .022)NSG? + (.41 \pm .09)RePro - (.067 \pm .03)PFPS? \\ &+ PERSIST + .01\end{aligned} \tag{6.7}$$

$$\begin{aligned}\widehat{NR} =& (.09 \pm .02)NSG? + (.38 \pm .01)RePro - (.07 \pm .03)PFPS? \\ &+ (.14 \pm .08)nIC + PERSIST + .01\end{aligned} \tag{6.8}$$

$$\begin{aligned}\widehat{NR} =& (.08 \pm .02)NSG? + (.38 \pm .01)RePro - (.07 \pm .03)PFPS? \\ &+ (.14 \pm .08)nIC + (.014 \pm .009)EI + PERSIST + .02\end{aligned} \tag{6.9}$$

The resulting statistical data from each step are listed in Table 6.3. The model in Eq. (6.5) is obtained at default p-values, which are set at 0.05 for inclusion and

0.1 for expulsion. Remaining models in the evolutionary process, however, were not at default p-value. STEPWISE interface was used to select the next most likely variable. Three new variables were selected: $PFPS$? at p-value of 0.05 in step 4, and nIC in step 5 and and EI in step 6 at p-value of 0.09 and 0.11, respectively. Statistics for the final model in Eq. (6.9) show that this variable with $R^2 \approx .65$ is considerably higher compared to $R^2 \approx .53$ obtained in Nelson and Sprecher model in [1]. There are five variables from four (*People and Economy* category is missing) of the five natural categories of Section 2.2, which is one of the desired properties discussed in Section 1. The variables do not show linear dependence among each other as the estimators' coefficient estimate remain stable in each step as shown in Eq. (6.4)-(6.9). Even though NSG? is selected as one of the significant variables in Eq. (6.9) , its influence is slightly less (about 4% less) in predicting nuclear reliance when compared to Eq. (4.11). This study, again, is a superficial study that can be explored further in the future.

Table 6.3: Statistical data for model obtained from STEPWISE regression after adding *PERSIST*

Step	**p-value(s)**	R^2	**Adj** R^2	**f-stat**	**RSME**
1	–	–	–	NaN	.1215
2	$[2.46e-5]$	.5503	..5450	19.95	.1093
3	$[1.08e-4]$ $[4.71e-4]$	.6127	.6034	18.06	.1014
4	$[2.69e-5]$ $[7.39e-5]$ [.0504]	.6305	.6170	13.78	.0991
5	[.0002], [.0002], [.0338], [.1016]	.6426	.6250	11.24	.0974
6	[.0007], [.0002], [.0262], [.0926], [.1068]	.6540	.6324	9.71	.0959

Addition of $PFPS?$ in model (6.9) shows a negative contribution to civil nuclear power production. This variable measures a state's ability to produce sensitive nuclear material (mainly enriched and reprocessed nuclear fuel). This is contrary to the hypothesis presented by Nelson and Sprecher [1], where this variable was expected to increase civil nuclear power program of a state. From proliferation standpoint, states with fuel production capability face about 7% reduction in nuclear reliance, which is similar to the Nelson-Sprecher model 2.2 when $FCS?$ showed reduction in states NR. The negative relation to the Nuclear Reliance suggests that states will falter in their ambition to have successful civil nuclear program if they are also try to acquire sensitive material for military purposes, which could be perceived from model (6.9) and also Nelson-Sprecher model.

The R^2 obtained in models (4.11) and (6.9) are .51 and .65 respectively. The two models (4.11) and (6.9) have the same number of predictors, yet model (6.9) has better R^2 compared to model (4.11). Better R^2 suggests the model created here is better fit; however, future analysis should use residual plot to see if the model is constantly over or under estimating the Nuclear Reliances.

For each variable added into the model the corresponding confusion matrices in Tables 6.4-6.9. Table 6.4 shows the confusion matrix for step (6.4) when the constant is added to the model. The average difference between current NR and 1980 NR was 5.5%. All 57 states that have no nuclear reliance either in 2011 and 1980 are predicted to have estimated NR of 5.5%. It is seen that more diagonal elements of the matrix in Table 6.4 are non-zero when compared to Table 5.1.

Table 6.5 is the confusion matrix when first variable $NSG?$ is added to the model, which results in 6.5. Comparing two models in Eqs. (4.2) and (6.5), $NSG?$ is the first variable added in both models, but the coefficient of NSG? in the newer model (6.4) is reduced by about 4%; therefore removes states with no NR to from severely

Table 6.4: Confusion matrix after constant term is added to the model in (6.4).

	Actual				
	NR	0%	0-10%	10-25%	> 25%
Estimated	0%	0	0	0	0
	$0-10\%$	57	7	4	3
	$10-25\%$	0	1	5	5
	$> 25\%$	0	0	0	4

"overestimated" cell, as previously discussed in Section 5. However, the coefficient for $NSG?$ is 10% plus constant term, which makes the estimated NR over 10%, hence all 15 states (Australia, Austria, Belarus, Croatia, Denmark, Greece, Ireland, Italy, Kazakhstan, Lithuania, New Zealand, Norway, Poland, Portugal, and Turkey) with no NR move to severely overestimated cell. For most states with NR $\geq$ 10% are estimated to be in correct range with the addition of $NSG?$.

Table 6.5: Confusion matrix after $NSG?$ is added to the model in (6.5).

	Actual				
	NR	0%	0-10%	10-25%	> 25%
Estimated	0%	42	1	1	0
	$0-10\%$	0	1	1	1
	$10-25\%$	15	6	7	4
	$> 25\%$	0	0	0	7

Table 6.6 shows the confusion matrix when $RePro$ is added to the model, which results in (6.6), With the addition of reprocessing 13 out of 15 "overestimated" states move to NR estimated-threshold NR actual-*de minimis* cell. The movement of large number of states here can be considered borderline movement as the coefficient of

NSG is 0.09; thus states producing no reprocessed material have about 9% estimated nuclear reliance. Four states, Slovakia, Ukraine, South Korea, and Czech Republic, are underestimated because these states are NSG member states but are not producing reprocessed nuclear material. Among these two states, South Korea has ≈ 9%, and Czech Republic had ≈ 15.5% NR in 1980. Czech Republic is in the borderline of moderate and aggressive estimated NR with 24.5%.

Table 6.6: Confusion matrix after $RePro$ is added to the model in (6.6).

		Actual			
	NR	0%	0-10%	10-25%	> 25%
Estimated	0%	42	1	1	0
	$0-10\%$	14	5	2	2
	$10-25\%$	1	2	5	4
	> 25%	0	0	1	6

Table 6.7 shows the confusion matrix after the addition of $PFPS$? in model (6.7). Primary Fuel Production State ($PFPS$?) is a first new variable obtained compared to model (4.11); however, similar variable FCS? was a selected independent variable in the Nelson-Sprecher model in Eq. (2.2). A complete description for this variable is provided in Appendix A. With the addition of $PFPS$ the coefficient of NSG? jumps up to 0.10. None of the 15 overestimated states were valuated under primary fuel production state. Therefore, all states jump to from estimated-threshold and actual-*de minimis* cell to estimated-moderate , and NR actual-*de minimis* cell. While on the other end Slovakia, Ukraine, South Korea, and Slovenia remain as four underestimated states, and Hungary is the severely underestimated state.

Table 6.8 shows the confusion matrix for model (6.8) when nIC is added. At this

Table 6.7: Confusion matrix after $PFPS$? is added to the model in (6.7).

	Actual				
	NR	0%	0-10%	10-25%	> 25%
Estimated	0%	42	2	1	0
	$0-10\%$	0	3	0	1
	$10-25\%$	15	3	7	4
	$>25\%$	0	0	1	7

level of analysis very minor changes results in the confusion matrix, when compared to Table 6.7. All 15 "overestimated" states remain at the actual $NR =$ *de minimis* and estimated $NR =$ moderate cell. Two changes are observed in the confusion matrix in Table 6.8. One state moves from actual-threshold estimated-moderate cell to actual and estimated threshold cell, and another state moves from actual-moderate predicted-*de minimis* celll to actual-moderate predicted-threshold cell. Again at this level it shows that improvement in residual of a few states can influence the selection of variable.

Table 6.8: Confusion matrix after nIC are added to the model in (6.9).

	Actual				
	NR	0%	0-10%	10-25%	> 25%
Estimated	0%	42	2	0	0
	$0-10\%$	0	4	1	1
	$10-25\%$	15	2	7	4
	$>25\%$	0	0	1	7

In step 6 of the evolutionary model, EI is added into the model. Table 6.9 shows the corresponding confusion matrix for model in Eq. (6.9). Energy insecurity has both positive and negative data. If a state has a negative energy insecurity value, it signifies that the state is energy sufficient by itself and also is exporting energy. Addition of this variable results in substantial changes to the confusion matrix in Table 6.9, as compared to the previous confusion matrices in Tables 6.4 - 6.7. The cell boundary between *de minimis* and threshold in this case was set at 1.5% because of high constant term of $\approx 2\%$. Lowest actual NR was 1.6% for China in the dataset, therefore, the boundary was chosen to be at 1.5%. Most of the changes are observed in states in the actual and estimated *de minimis* cell. In the previous three confusion matrices the majority of states with no nuclear reliance were estimated to be in actual and estimated *de minimis*; however, addition of EI has moved twenty-three such states from actual and estimated-*de minimis* cell to actual-*de minimis* and estimated-threshold cell. Five states with overestimated nuclear reliance were also moved up a row to actual-*de minimis* estimated-threshold cell. Addition of EI makes a slight improvement to correct the states with low nuclear reliance, but no changes are observed for the underestimated states when compared to Table 6.8.

Table 6.9: Confusion matrix after EI are added to the model in (6.9).

		Actual			
	NR	0%	0-10%	10-25%	> 25%
Estimated	0%	19	2	0	0
	$0-10\%$	28	6	2	1
	$10-25\%$	10	0	6	4
	$> 25\%$	0	0	1	7

It is also informative to compare briefly the confusion matrices for the steps of the model evolution considered in this section to those for the corresponding steps in the evolution of the model (4.11). For example the confusion matrix of Table 6.5 is somewhat more strongly diagonally dominant, especially in the lower right entries (larger actual and estimated nuclear reliances), than that of Table 5.2, notwithstanding that both are for models with only one nonconstant (and nonpersistence) independent variable, and further that variable is the same (NSG?) for both evolutionary sequences. This tendency toward somewhat stronger clustering in the diagonal of the confusion matrices of this section, as compared to those of Section 5.1, carries through the entire evolutionary sequences. But the two model evolutions are similar, in that changes from one step to the next are small, after the initial step in which NSG? is selected for the model.

One suspects that, similarly to the evolution (4.1)-(4.5) and (4.11), this means the addition of each of the variables past NSG? largely is driven by the properties of a small number of states, perhaps even one state (e.g., France, in the case of $RePro$ as the added variable). A thorough exposition of the properties of the evolution (6.4)-(6.9) would explore that by means of scatter plots such as employed in Sections 5.2 and 5.3. For the initial exploration of that evolution intended here that exploration will be limited to the residual analysis that is the subject of the following section. Moreover, following section compares the improvements, if any, against Table 5.13 to consider the impact of all variables in the model.

6.3 Residual Analysis of Model (6.4)-(6.9)

As in Section 5, the purpose of Tables 6.10 - 6.15 is to provide residual data after the addition of each variables into the basic linear model (BLM). However, major difference between tables in this sections in tables in Section 5.4 are that in this section only first 25 states with largest absolute maximum change between prior and post squared residuals are presented. Tables 6.10- 6.15 are sorted from

Table 6.10: Residual table after the addition of the constant term

State	NR	Estimated NR	Posterior Residual	Prior Squared Residual	Post Squared Residual	Absolute Difference between Prior and Post Squared Residual
France	0.7770	0.2930	0.4840	0.6037	0.2343	0.3695
Belgium	0.5400	0.2910	0.2490	0.2916	0.0620	0.2296
Switzerland	0.4090	0.3530	0.0560	0.1673	0.0031	0.1641
Sweden	0.3960	0.3300	0.0660	0.1568	0.0044	0.1525
Slovakia	0.5400	0.1292	0.4108	0.2916	0.1688	0.1228
Bulgaria	0.3260	0.2320	0.0940	0.1063	0.0088	0.0974
Czech Republic	0.3300	0.2078	0.1222	0.1089	0.0149	0.0940
Finland	0.3160	0.2270	0.0890	0.0999	0.0079	0.0919
South Korea	0.3460	0.1480	0.1980	0.1197	0.0392	0.0805
Ukraine	0.4720	0.0677	0.4043	0.2228	0.1634	0.0593
Hungary	0.4330	0.0550	0.3780	0.1875	0.1429	0.0446
Slovenia	0.4170	0.0550	0.3620	0.1739	0.1310	0.0428
USA	0.1930	0.1650	0.0280	0.0372	0.0008	0.0365
Japan	0.1810	0.1990	-0.0180	0.0328	0.0003	0.0324
UK	0.1780	0.1850	-0.0070	0.0317	0.0000	0.0316
Spain	0.1950	0.1020	0.0930	0.0380	0.0086	0.0294
Russia	0.1760	0.0902	0.0858	0.0310	0.0074	0.0236
Canada	0.1530	0.1570	-0.0040	0.0234	0.0000	0.0234
Taiwan	0.1902	0.0550	0.1352	0.0362	0.0183	0.0179
Romania	0.1900	0.0550	0.1350	0.0361	0.0182	0.0179
Germany	0.1780	0.0550	0.1230	0.0317	0.0151	0.0166
Netherlands	0.0360	0.1200	-0.0840	0.0013	0.0071	0.0058
Italy	0.0000	0.0670	-0.0670	0.0000	0.0045	0.0045
Uzbekistan	0.0000	0.0589	-0.0589	0.0000	0.0035	0.0035
Kazakhstan	0.0000	0.0588	-0.0588	0.0000	0.0035	0.0035

largest to smallest for “Absolute difference between prior and post squared residual”. Moreover, these tables also show states contributing the highest square residual does not undergo largest change in squared residual when compared to the preceding step. Table 6.10 presents the tabular residual when the null hypothesis is replaced by the mean of ΔNR (difference between 2011 NR and 1980 NR) plus the 1980 NR, as in model (6.4). In this step, France, Belgium, Switzerland, and Sweden have larger contribution to reduce the square residual of the model, but France still has the largest square residual.

Table 6.11 shows the residual table when $NSG?$ is added to the model in Eq. (6.5). Similar to Section 5.4, if these data are sorted on actual Nuclear Reliances, then the 15 severely overestimated states in the corresponding confusion matrix (Table

Table 6.11: Residual table after the addition of NSG?

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Squared Residual	Post Squared Residual	Absolute Difference between Prior and Post Squared Residual
France	0.7770	0.3430	0.4840	0.4340	0.2343	0.1884	0.0459
Slovakia	0.5400	0.1792	0.4108	0.3608	0.1688	0.1302	0.0386
Ukraine	0.4720	0.1177	0.4043	0.3543	0.1634	0.1255	0.0379
Hungary	0.4330	0.1050	0.3780	0.3280	0.1429	0.1076	0.0353
Slovenia	0.4170	0.1050	0.3620	0.3120	0.1310	0.0973	0.0337
Belgium	0.5400	0.3410	0.2490	0.1990	0.0620	0.0396	0.0224
South Korea	0.3460	0.1980	0.1980	0.1480	0.0392	0.0219	0.0173
Taiwan	0.1902	0.0050	0.1352	0.1852	0.0183	0.0343	0.0160
Romania	0.1900	0.1050	0.1350	0.0850	0.0182	0.0072	0.0110
Netherlands	0.0360	0.1700	-0.0840	-0.1340	0.0071	0.0180	0.0109
Germany	0.1780	0.1050	0.1230	0.0730	0.0151	0.0053	0.0098
Czech Republic	0.3300	0.2578	0.1222	0.0722	0.0149	0.0052	0.0097
Italy	0.0000	0.1170	-0.0670	-0.1170	0.0045	0.0137	0.0092
Argentina	0.0500	0.1640	-0.0640	-0.1140	0.0041	0.0130	0.0089
Kazakhstan	0.0000	0.1088	-0.0588	-0.1088	0.0035	0.0118	0.0084
Lithuania	0.0000	0.1059	-0.0559	-0.1059	0.0031	0.0112	0.0081
Australia	0.0000	0.1050	-0.0550	-0.1050	0.0030	0.0110	0.0080
Austria	0.0000	0.1050	-0.0550	-0.1050	0.0030	0.0110	0.0080
Belarus	0.0000	0.1050	-0.0550	-0.1050	0.0030	0.0110	0.0080
Croatia	0.0000	0.1050	-0.0550	-0.1050	0.0030	0.0110	0.0080
Denmark	0.0000	0.1050	-0.0550	-0.1050	0.0030	0.0110	0.0080
Greece	0.0000	0.1050	-0.0550	-0.1050	0.0030	0.0110	0.0080
Ireland	0.0000	0.1050	-0.0550	-0.1050	0.0030	0.0110	0.0080
New Zealand	0.0000	0.1050	-0.0550	-0.1050	0.0030	0.0110	0.0080
Norway	0.0000	0.1050	-0.0550	-0.1050	0.0030	0.0110	0.0080

6.4) are revealed as: Australia, Austria, Belarus, Croatia, Denmark, Greece, Ireland, Italy, Kazakhstan, Lithuania, New Zealand, Norway, Poland, Portugal and Turkey. These states are Nuclear Suppliers Group member states but have no actual nuclear reliance. Likewise four underestimated NR states in Slovakia, Ukraine, Hungary, and Slovenia. France has the most reduction in the squared residual, but still has the largest square residual contribution to the model.

Table 6.12 shows the residual table when $RePro$ is added in the model in Eq. (6.6). As the coefficient for NSG decreases Italy the only state that remain in the overestimated range, while 14 other states reduce the residual slightly. Same four states mentioned about remain in the underestimated range which also contribute to increase the squared residual of the model.

Table 6.13 shows the residual table when $PFPS$? is added in the model in Eq.

Table 6.12: Residual table after the addition of $RePro$

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Squared Residual	Post Squared Residual	Absolute Difference between Prior and Post Squared Residual
France	0.7770	0.6730	0.4340	0.1040	0.1884	0.0108	0.1775
UK	0.1780	0.3884	-0.0570	-0.2104	0.0032	0.0443	0.0410
Slovakia	0.5400	0.1692	0.3608	0.3708	0.1302	0.1375	0.0073
Ukraine	0.4720	0.1077	0.3543	0.3643	0.1255	0.1327	0.0072
Hungary	0.4330	0.0950	0.3280	0.3380	0.1076	0.1142	0.0067
Slovenia	0.4170	0.0950	0.3120	0.3220	0.0973	0.1037	0.0063
South Korea	0.3460	0.1880	0.1480	0.1580	0.0219	0.0250	0.0031
Netherlands	0.0360	0.1600	-0.1340	-0.1240	0.0180	0.0154	0.0026
Italy	0.0000	0.1070	-0.1170	-0.1070	0.0137	0.0114	0.0022
Argentina	0.0500	0.1540	-0.1140	-0.1040	0.0130	0.0108	0.0022
Kazakhstan	0.0000	0.0988	-0.1088	-0.0988	0.0118	0.0098	0.0021
Lithuania	0.0000	0.0959	-0.1059	-0.0959	0.0112	0.0092	0.0020
Australia	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Austria	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Belarus	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Croatia	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Denmark	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Greece	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Ireland	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
New Zealand	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Norway	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Poland	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Portugal	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Turkey	0.0000	0.0950	-0.1050	-0.0950	0.0110	0.0090	0.0020
Romania	0.1900	0.0950	0.0850	0.0950	0.0072	0.0090	0.0018

(6.7). This variable has negative correlation with NR. With the addition of $PFPS$ variable the coefficient of NSG? moves over 10%; therefore all '15' severely overestimated states move to overestimated range, hence, increasing the residual created by the states in this range. Most of the primary fuel cycle state show reduction in square residual, except India and Pakistan where the posterior square residual has increased when compared to previous step. Slovakia has the largest square residual contribution. The reduction residual from four underestimated states decreases from 56.7% to 54.5%.

Table 6.14 shows the residual table when nIC is added in the model in Eq. (6.8). No change in the confusion matrix in Table 6.8 was observed. The residual from the underestimated states increases about 2% from the previous step, but the residual from the overestimated state decreases about 2%.

Table 6.13: Residual table after the addition of $PFPS$?

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Squared Residual	Post Squared Residual	Absolute Difference between Prior and Post Squared Residual
Slovakia	0.5400	0.1842	0.3708	0.3558	0.1375	0.1266	0.0109
Ukraine	0.4720	0.1227	0.3643	0.3493	0.1327	0.1220	0.0107
Netherlands	0.0360	0.1050	-0.1240	-0.0690	0.0154	0.0048	0.0106
Hungary	0.4330	0.1100	0.3380	0.3230	0.1142	0.1043	0.0099
Slovenia	0.4170	0.1100	0.3220	0.3070	0.1037	0.0942	0.0094
UK	0.1780	0.3670	-0.2104	-0.1890	0.0443	0.0357	0.0085
Pakistan	0.0380	-0.0600	0.0330	0.0980	0.0011	0.0096	0.0085
Argentina	0.0500	0.0990	-0.1040	-0.0490	0.0108	0.0024	0.0084
Belgium	0.5400	0.3542	0.2022	0.1858	0.0409	0.0345	0.0064
China	0.0190	0.0400	-0.0760	-0.0210	0.0058	0.0004	0.0053
India	0.0370	-0.0350	0.0070	0.0720	0.0000	0.0052	0.0051
South Korea	0.3460	0.2030	0.1580	0.1430	0.0250	0.0204	0.0045
Brazil	0.0320	0.0400	-0.0630	-0.0080	0.0040	0.0001	0.0039
Israel	0.0000	-0.0600	-0.0050	0.0600	0.0000	0.0036	0.0036
Italy	0.0000	0.1220	-0.1070	-0.1220	0.0114	0.0149	0.0034
Japan	0.1810	0.1840	-0.0580	-0.0030	0.0034	0.0000	0.0034
Kazakhstan	0.0000	0.1138	-0.0988	-0.1138	0.0098	0.0129	0.0032
Lithuania	0.0000	0.1109	-0.0959	-0.1109	0.0092	0.0123	0.0031
Australia	0.0000	0.1100	-0.0950	-0.1100	0.0090	0.0121	0.0031
Austria	0.0000	0.1100	-0.0950	-0.1100	0.0090	0.0121	0.0031
Belarus	0.0000	0.1100	-0.0950	-0.1100	0.0090	0.0121	0.0031
Croatia	0.0000	0.1100	-0.0950	-0.1100	0.0090	0.0121	0.0031
Denmark	0.0000	0.1100	-0.0950	-0.1100	0.0090	0.0121	0.0031
Greece	0.0000	0.1100	-0.0950	-0.1100	0.0090	0.0121	0.0031
Ireland	0.0000	0.1100	-0.0950	-0.1100	0.0090	0.0121	0.0031

Table 6.14: Residual table after the addition of nIC

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Squared Residual	Post Squared Residual	Absolute Difference between Prior and Post Squared Residual
Ukraine	0.4720	0.1508	0.3493	0.3212	0.1220	0.1032	0.0188
Slovakia	0.5400	0.1742	0.3558	0.3658	0.1266	0.1338	0.0072
Hungary	0.4330	0.1000	0.3230	0.3330	0.1043	0.1109	0.0066
Slovenia	0.4170	0.1000	0.3070	0.3170	0.0942	0.1005	0.0062
UK	0.1780	0.3521	-0.1890	-0.1741	0.0357	0.0303	0.0054
Belgium	0.5400	0.3436	0.1858	0.1964	0.0345	0.0386	0.0041
South Korea	0.3460	0.1930	0.1430	0.1530	0.0204	0.0234	0.0030
Kazakhstan	0.0000	0.1038	-0.1138	-0.1038	0.0129	0.0108	0.0022
Lithuania	0.0000	0.1009	-0.1109	-0.1009	0.0123	0.0102	0.0021
Australia	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Austria	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Belarus	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Croatia	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Denmark	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Greece	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Ireland	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
New Zealand	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Norway	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Poland	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Portugal	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Turkey	0.0000	0.1000	-0.1100	-0.1000	0.0121	0.0100	0.0021
Italy	0.0000	0.1148	-0.1220	-0.1148	0.0149	0.0132	0.0017
Romania	0.1900	0.1000	0.0800	0.0900	0.0064	0.0081	0.0017
Sweden	0.3960	0.4378	0.0110	-0.0418	0.0001	0.0017	0.0016
Netherlands	0.0360	0.0930	-0.0690	-0.0570	0.0048	0.0032	0.0015

Table 6.15 shows the residual table when EI is added in the model in Eq. (6.9). Substantial changes were observed in the confusion matrix in Table 6.9 was observed. Even though the overall residual for this model decrease from .8 to .77 the contribution from the "overestimated" states increased from .13 to .16, and residual from the "underestimated" states also increased about 2%. Most states with no nuclear reliance were estimated to have nuclear reliance of at least more than 1.5%. Notably in residual Table 6.15, Norway shows relatively a large reduction in its squared residual when EI was added in the model (6.9).

Table 6.15: Residual table after the addition of EI

State	NR	Estimated NR	Prior Residual	Posterior Residual	Prior Squared Residual	Post Squared Residual	Absolute Difference between Prior and Post Squared Residual
Norway	0.0000	0.0181	-0.1000	-0.0181	0.0100	0.0003	0.0097
Taiwan	0.1902	0.0302	0.1802	0.1600	0.0325	0.0256	0.0069
Australia	0.0000	0.0749	-0.1000	-0.0749	0.0100	0.0056	0.0044
Slovakia	0.5400	0.1800	0.3658	0.3600	0.1338	0.1296	0.0043
Kazakhstan	0.0000	0.0837	-0.1038	-0.0837	0.0108	0.0070	0.0038
Hungary	0.4330	0.1049	0.3330	0.3281	0.1109	0.1077	0.0032
Belgium	0.5400	0.3507	0.1964	0.1893	0.0386	0.0358	0.0027
Qatar	0.0000	-0.0513	-0.0100	0.0513	0.0001	0.0026	0.0025
South Korea	0.3460	0.2014	0.1530	0.1446	0.0234	0.0209	0.0025
Slovenia	0.4170	0.1039	0.3170	0.3131	0.1005	0.0980	0.0025
Pakistan	0.0380	-0.0497	0.1000	0.0877	0.0100	0.0077	0.0023
Israel	0.0000	-0.0415	0.0620	0.0415	0.0038	0.0017	0.0021
Italy	0.0000	0.1234	-0.1148	-0.1234	0.0132	0.0152	0.0021
Ireland	0.0000	0.1091	-0.1000	-0.1091	0.0100	0.0119	0.0019
Ukraine	0.4720	0.1537	0.3212	0.3183	0.1032	0.1013	0.0019
Belarus	0.0000	0.1088	-0.1000	-0.1088	0.0100	0.0118	0.0018
India	0.0370	-0.0245	0.0740	0.0615	0.0055	0.0038	0.0017
Portugal	0.0000	0.1074	-0.1000	-0.1074	0.0100	0.0115	0.0015
Turkey	0.0000	0.1068	-0.1000	-0.1068	0.0100	0.0114	0.0014
France	0.7770	0.6932	0.0910	0.0838	0.0083	0.0070	0.0013
Austria	0.0000	0.1061	-0.1000	-0.1061	0.0100	0.0113	0.0013
Greece	0.0000	0.1061	-0.1000	-0.1061	0.0100	0.0113	0.0013
Denmark	0.0000	0.0944	-0.1000	-0.0944	0.0100	0.0089	0.0011
UK	0.1780	0.3552	-0.1741	-0.1772	0.0303	0.0314	0.0011
Croatia	0.0000	0.1043	-0.1000	-0.1043	0.0100	0.0109	0.0009

Table 6.16 shows the summary table similar to that of obtained in Section 5.5. A comparison of the two Tables 6.16 and 5.13 shows the sum of squared residual in each evolutionary step of model in this section is smaller. The residual contribution

from a single state, as shown in the third column of Table 6.16, is also smaller when compared to residual obtained from models in Eqs. (4.1)-(4.5) and Eq. (4.11) in Table 5.13. The number of severely overestimated and underestimated states in between Table 5.6 and Table 6.9 have been reduced from 13 states to 10 states and 9 states to 4 states respectively; therefore the corresponding residual contributions were lower here, compared to the models in Section 4. However, the single major residual contribution in each step comes from same states, which are France in the first two steps and Slovakia in the remaining four, as shown in Table 6.16. There has been some improvement in the model obtained here when compared to previous model; therefore it can one of the directions to pursue for future study.

Table 6.16: Residual summary

Last variable added	Sum of squared residuals	Largest single-state contributor (state)	Contribution from severely overestimated states	Contribution from underestimated states
Constant	1.24	23.4% (France)	2.4%	48.9%
NSG?	1.00	18.8% (France)	11.3%	46.1%
RePro	0.86	13.7% (Slovakia)	10.8%	56.7%
PFPS?	0.82	12.6%(Slovakia)	15.1%	54.5%
nIC	0.80	13.4% (Slovakia)	13.0%	56.2%
EI	0.77	12.9% (Slovakia)	16.1%	58.1%

7. CONCLUSION

In this final thesis section, Section 7.1 provides summary of the work conducted, and Section 7.2 points to alternative methodologies that could be implemented for related future study.

7.1 Summary

The basis of this thesis has been to improve on the work done by Nelson and Sprecher in [1] using multivariate linear regression technique. Reiterating, the objective of the present work has been to replace dichotomous and subjectively defined predictors from Nelson-Specher model (2.2), with more objectively defined variables. MatLab statistical function STEPWISEFIT was useful for initial explorations and MatLab statistical graphical interface STEPWISE and backward elimination in SAS (Statistical Analysis System) were useful to obtain improved basic linear models. The research conducted was successful in replacing aforementioned dichotomous and subjective predictors with more objectively defined variables from Nelson-Sprecher model, but the newly obtained model had less natural category (discussed in Section 1) representation, and slightly degraded statistics. Following conclusion could be made from the basic linear model (4.11) obtained in Section 4:

1. States who are involved in international commerce of nuclear technology and nuclear material have greater tendency of depending on civil nuclear energy. Three of five explanatory variables that are positively influencing Nuclear Reliance come from *International Commerce* category. International commerce gives of about 60% contribution to a state's Nuclear Reliance when adding the coefficients of variables from International commerce category in the model (4.11).

2. Having natural resources tends to have negative influence on a state's Nuclear Reliance. The coefficient of the mining and milling ($MING$) independent variable shows up to a 13% reduction of Nuclear Reliance when a state has natural uranium resources.

3. Variables were included in the model to decrease the square residual influenced by only a few states. Step-by-step evolutionary process in Section 5 gave better insight on selection some of the variables, such as new International Commerce (nIC).

7.2 Related Future Work

A possible future work can be creating a model with square root of nuclear reliance as dependent variable. Use of Square root of nuclear reliance is hypothesized to give a better model as a larger separation of dependent variable between the *de minimis* states and *threshold* states is created. Another approach would be create a new dependent variables, which the current nuclear power plant capacity based on the net capacity (MWe) of electricity from nuclear source. This "*gridwise*" analysis instead of "*statewise*" analysis will provide better understanding of nuclear reliance because some states that do not have nuclear power plants in their state do rely on the electricity generated from nuclear sources in a different state.

Another possible methodology would be to use multinomial logistic regression. As two-third of the states under consideration have zero Nuclear Reliance, logistic regression can be useful to estimate possible outcomes of discrete distribution for different dependent variables [34]. Moreover, dataset used for this study still contains seven independent variables with dichotomous data (*ALGN*, *ENR*, *FCS*, *NSG*, *NWS*, *NWs*, *PFPS*). Improved model could be estimated using all objectively defined variables. Furthermore, independent variables seem to be left disaggregated, and some method of aggregating the independent variable might improve the basic linear model quite substantially. For this, principle component analysis (PCA) could be used, which can reduce the number of variables while still retaining the information of the original database. Addition of new variables in the dataset can improve the model. Variables defining the degree of democracy, and prices of natural gas and coal can major implications for the nuclear power industry in the future.

REFERENCES

[1] P. Nelson and C. M. Sprecher, “What determines the extent of national reliance on civil nuclear power?,” *Nuclear Security Science and Policy Institute, Texas A&M University, College Station, Texas, Report No. NSSPI-08-014*, vol. 14, 2008.

[2] J. Li, M.-S. Yim, and D. N. McNelis, “Model-based calculations of the probability of a countrys nuclear proliferation decisions,” *Progress in Nuclear Energy*, vol. 52, no. 8, pp. 789–808, 2010.

[3] D. E. Farrar and R. R. Glauber, “Multicollinearity in regression analysis: the problem revisited,” *The Review of Economic and Statistics*, vol. 49, no. 1, pp. 92–107, 1967.

[4] P. Nelson and C. M. Sprecher, “Are sensitive technologies enablers of civil nuclear power? An empirical study,” *Atoms for Peace: an International Journal*, vol. 3, no. 2, pp. 93 – 112, 2010.

[5] P. Nelson and C. M. Sprecher, “Where will new nuclear power plants be constructed? Validation of a predictive model,” in *Proc. 31st Annual Meeting of ESARDA*, (Vilnius, Lithuania), May 2009.

[6] M.-S. Yim and J. Li, “Examining relationship between nuclear proliferation and civilian nuclear power development,” *Progress in Nuclear Energy*, vol. 66, pp. 108 – 114, 2013.

[7] S. Singh and C. R. Way, “The correlates of nuclear proliferation: A quantitative test,” *Journal of Conflict Resolution*, vol. 48, no. 6, pp. 859–885, 2004.

[8] D.-J. Jo and E. Gartzke, "Determinants of nuclear weapons proliferation," *Journal of Conflict Resolution*, vol. 51, no. 1, pp. 167–194, 2007.

[9] R. Rauchhaus, M. Kroenig, and E. Gartzke, *Causes and Consequences of Nuclear Proliferation*. New York, NY: Routledge, 2011.

[10] U.N. Comtrade, "*United Nations Commodity Trade Statistics Database*." [Online]. Available: http://comtrade.un.org/ [Accessed: Sep 15, 2013].

[11] World Nuclear Association, "*The Economics of Nuclear Power*," [Online]. Available: http://www.world-nuclear.org/info/Economic-Aspects/Economics-of-Nuclear-Power/ [Accessed: Oct. 15, 2013].

[12] Central Intelligence Agency, "*The World Factbook: GDP*," [Online]. Available: https://www.cia.gov/library/publications/the-world-factbook/fields/2195.html#af [Accessed: Oct. 15, 2013].

[13] United Nations, "*Member States of the United Nations*," 2014. [Online]. Available: http://www.un.org/en/members/index.shtml [Accessed: Apr. 20, 2014].

[14] IAEA, "Nuclear power reactors in the world," Tech. Rep. IAEA-RDS-2/31, International Atomic Energy Agency, Vienna, Austria, June 2012. Available: http://www-pub.iaea.org/books/IAEABooks/8954/Nuclear-Power-Reactors-in-the-World-2012-Edition.

[15] MATLAB, *version 8.2 (R2013b)*. Natick, MA: The MathWorks Inc., 2013.

[16] R. Wieland, "Stepwise linear regression terms and implementation," 2007, Available: www.math.byu.edu/~matt/byu/math410_2007fall/projects/rosewieland.pdf.

[17] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*. Hoboken, NJ: John Wiley & Sons, 4th ed., 2006.

[18] *SAS/STAT 9.22 Users Guide*. Cary, NC: SAS Institute, 2010.

[19] M. Humphreys, J. D. Sachs, and J. E. Stiglitz, *Escaping the Resource Curse (Initiative for Policy Dialogue at Columbia)*. New York, NY: Columbia University Press, 2007.

[20] A. R. Sofia Visa, Brian Ramsay and E. van der Knaap, "Confusion matrix-based feature selection," *CEUR Workshop Proceedings*, vol. 710, 2011. Available: http://ceur-ws.org/Vol-710/paper37.pdf.

[21] World Nuclear Association, "Emerging nuclear energy countries," [Online]. Available: http://www.world-nuclear.org/info/Country-Profiles/Others/Emerging-Nuclear-Energy-Countries/ [Accessed: Mar. 07, 2014].

[22] International Atomic Energy Agency, "Nuclear fuel cycle information system," [Online]. Available: http://infcis.iaea.org/NFCIS/NFCISMain.asp?Order=1&Scale=Commercial&RPage=1&Page=1&RightP=CountryReport [Accessed: Sep. 07, 2012].

[23] "World list of nuclear power plants," *Nuclear News*, vol. 55, no. 3, pp. 57–88, March 2012.

[24] International Atomic Energy Agency, "*IAEA Country Nuclear Power Profile 2013: Slovakia*," [Online]. Available: http://www-pub.iaea.org/MTCD/Publications/PDF/CNPP2013_CD/countryprofiles/Slovakia/Slovakia.htm [Accessed: Mar. 07, 2014].

[25] World Nuclear Association, "*Nuclear Power in Slovakia*," [Online]. Available: http://www.world-nuclear.org/info/Country-Profiles/Countries-O-S/Slovakia/ [Accessed: Mar. 07, 2014].

[26] International Atomic Energy Agency, "*IAEA Country Nuclear Power Profile 2013: Italy*," [Online]. Available: http://www-pub.iaea.org/MTCD/Publications/PDF/CNPP2013_CD/countryprofiles/Italy/Italy.htm [Accessed: Apr. 02, 2014].

[27] B. B. Wittneben, "The impact of the fukushima nuclear accident on european energy policy," *Environmental Science & Policy*, vol. 15, no. 1, pp. 1–3, 2012.

[28] The World Bank, "*Electricity production from nuclear sources (kWh)*," [Online]. Available http://data.worldbank.org/indicator/EG.ELC.NUCL.KH [Accessed: Jun. 03, 2014].

[29] V. Korobeynikov, "IAEA Country Nuclear Power Profile 2013: Russian Federation," 2013. Available: http://www-pub.iaea.org/MTCD/Publications/PDF/CNPP2013_CD/countryprofiles/Russia/Russia.htm [Accessed: Apr. 02, 2014].

[30] R. Sakwa, *Soviet politics in perspective.* New York, NY: Routledge, 2nd ed., 1998.

[31] Infostat, "Slovakia population data," Demographic Research Center, 2013, [Online]. Available: http://www.infostat.sk/vdc/en/index.php?option=com_wrapper&Itemid=35 [Accessed: Apr. 02, 2014].

[32] Czech Statistical Office (2009), "*Czech Demographic Handbook*," [Online]. Available: http://www.czso.cz/csu/2008edicniplan.nsf/engt/24003E05E3/$File/4032080106a.pdf [Accessed: Apr. 02, 2014].

[33] B. Obelic, B. I. Krajcar, D. Srdoc, and N. Horvatincic, "Environmental (super 14) c levels around the 632 MWe nuclear power plant Krsko in Yugoslavia.," *Radiocarbon*, vol. 28, no. 2A, pp. 644–648, 1986.

[34] W. H. Greene, *Econometric analysis.* Upper Saddle River, NJ: Pearson Education, Inc., 7th ed., 2012.

[35] M. G. Marshall and T. R. G. (2010), "*Polity IV Project: Political Regime Charactersitcs and Transitions, 1800-2013*," [Online]. Available: http://www.systemicpeace.org/polity/polity4.htm [Accessed: Nov 20, 2013].

[36] Nuclear Suppliers Group, "*Nuclear Suppliers Group (NSG)*," [Online]. Available: http://www.nuclearsuppliersgroup.org/A_test/01-eng/index.php [Accessed: Nov. 20, 2013].

[37] H. Tulsidas. (2012, Sep. 17). "RE: Date for currently available data under Country Reports presented in Nuclear Fuel Cycle Information System." Personal e-mail.

[38] O. Authier, E. Thunin, P. Plion, C. Schnnenbeck, G. Leyssens, J.-F. Brilhac, and L. Porcheron, "Kinetic study of pulverized coal devolatilization for boiler {CFD} modeling," *Fuel*, vol. 122, no. 0, pp. 254 – 260, 2014.

[39] U.S. Energy Information Administration, "*International Energy Statistics: Coal*," 2008. [Online]. Available: http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=1&pid=7&aid=6 [Accessed: Sep 07, 2012].

[40] U.S. Energy Information Administration, "*International Energy Statistics: Electricity*," [Online]. Available: http://www.eia.gov/cfapps/ipdbproject/iedindex3.cfm?tid=2&pid=2&aid=12&cid=TZ,&syid=2008&eyid=2012&unit=BKWH [Accessed: Jun 05, 2013].

[41] The World Bank, "*Energy imports, net (% of energy use)*," [Online]. Available: http://data.worldbank.org/indicator/EG.IMP.CONS.ZS [Accessed Aug. 28, 2012].

[42] A. M. Jaffe, "Shale gas will rock the world," *Wall Street Journal*, vol. 10, 2010.

[43] U.S. Energy Information Administration, "*International Energy Statistics: Reserves of natural gas*," [Online]. Available: http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=3&pid=3&aid=6 [Accessed: Aug 29, 2012].

[44] IAEA, "Country nuclear fuel cycle profile," Tech. Rep. ISSN 0074-1914, International Atomic Energy Agency, Vienna, Austria, 2005. Available: http://www-pub.iaea.org/MTCD/publications/PDF/TRS425_web.pdf.

[45] URENCO, "Uranium Hexafluoride (UF6)," [Online]. Available: http://www.urenco.com/page/58/Uranium-Hexafluoride-UF6.aspx [Accessed: Apr. 25, 2014].

[46] A. Petoe, "Safeguards obligation related to uranium/thorium mining and processing," IAEA, 2009. Available: http://www-pub.iaea.org/mtcd/meetings/PDFplus/2009/cn175/URAM2009/Session%201/4_114_Petoe_IAEA.pdf.

[47] Nuclear Threat Initiative, "*Country Profiles*," [Online]. Available: http://www.nti.org/country-profiles/ [Accessed: Nov. 20,2013].

[48] Central Intelligence Agency, "*The World Factbook: GDP per capita purchasing power parity*," [Online]. Available: https://www.cia.gov/library/publications/the-world-factbook/fields/2004.html#85 [Accessed: Feb. 15, 2013].

[49] Central Intelligence Agency, "*The World Factbook: Population*," [Online]. Available: https://www.cia.gov/library/publications/the-world-factbook/fields/2119.html#af [Accessed: Feb. 15, 2013].

APPENDIX A

CANDIDATE CORRELATES

A description of all candidate independent variables in more detail than the list by name and symbol that appears in Section 2.2 is provided in this Appendix. Details include descriptions of the precise measures associated to each of these dependent variables, and a brief explanation of the hypothetical rationale for a correlation with nuclear reliance underlying each of these correlates. Section A.1 includes the variables that appear in model (3.1), Section A.2 includes the variables that appear in models (4.1)-(4.5) and (4.11); and all other candidate independent variables are described in Section A.3, which is divided into subsections according to the five natural categories listed in Section 2.2. Within each subsection variables are listed alphabetically. Subsections *A.3.1* and *A.3.5* include descriptions for EI and $PFPS$?, respectively that were model (6.9) of Section 6.

A.1 Independent Variables in Eq. (3.1)

1. nIC (new International Commercialization): As mentioned in the thesis in Section 1, the data for the older version of this variable was subjective and dichotomous and needed to be changed to an objectively quantifiable variable; therefore "n" in nIC represent the newly collected objective data. Data for this variable was obtained only from operating power units from Ref. [23]. For each currently (end of 2011) operating nuclear power plant, the reactor supplier and a list of major participants (e.g. for Unit 1 of Monticello Nuclear Plant in Monticello, MN, General Electric (GE) is the reactor supplier, and Bechtel Corporation is a major participant which provides plant recovery support,

plant license renewal, steam generator replacement). Raw data were created as follows. One point each was given to each state where an external reactor supplier was located, and likewise one point was prorated across the major participants, but awarded only to the states where the external major participant was based. The accumulated points were then prorated on a per capita basis by dividing by the population of the state and then unit standardized by dividing the maximum value over all states. The result then is presumably a measure of the net per capita benefit seen by the citizens of states housing enterprises engaged in international commerce in nuclear reactors. This variable was intended as a more objective replacement for the very subjective version of International Commercialization employed in previous work [1]. This dependent variable played a very significant role, arguably the most significant role of all candidate dependent variables in the basic linear model developed in Nelson-Sprecher model in Ref. [1].

2. $PLTY$? (Polity IV): Polity IV is a political science measure that examines "concomitant quality of democracy or autocratic authority" of different states. The polity values range from -10 to 10. The lowest end of the spectrum represents "hereditary monarchy" and the highest end represents "consolidated democracy" [35]. Level of democracy can be hypothesized as a important factor to promote the use of civil nuclear power; therefore, polity is regarded to be a key political variable [1]. Unit standardization was conducted by adding 10 to each value and then dividing the sum by 20.

A.2 Independent Variables in Eq. (4.1)-(4.5) and (4.11)

1. $MING$ (Mining and Milling): Mining and milling was encoded as the uranium production data (metric tons of uranium per year), as obtained from the

Integrated Nuclear Fuel Cycle Information Systems [22]. Similar to the Reprocessing independent variable, amount of currently mined and milled nuclear fuel material was aggregated for each state. There are several states, such as Gabon, Malawi, Namibia, and Niger, that are involved in mining and milling, but did not satisfy the research 20-20 criterion. The supposition underlying inclusion of this attribute as a candidate variable is that indigenous supplies of uranium ore will make a state more likely to rely on nuclear power to generate electricity than it might otherwise be.

2. NSG? (Members of Nuclear Suppliers Group): The Nuclear Suppliers Group website [36] defines the group as: "The Nuclear Suppliers Group (NSG) is a group of nuclear supplier countries which seeks to contribute to the non-proliferation of nuclear weapons through the implementation of guidelines for nuclear exports and nuclear related exports." As of September 13, 2012 there were forty six members of the Nuclear Suppliers Group (NSG?). Six of those (Cyprus, Estonia, Iceland, Latvia, Luxembourg and Malta) were excluded from our NC86 dataset because these states did not meet the criterion of 20 million population or 20 billion GDP criteria. For the forty NSG? members in this dataset the variable NSG? was assigned a value 1, and for the remaining NC86 states it were assigned a value 0. Membership in the NSG? serves as a virtual license to engage in legitimate international commerce in nuclear materials and technology. It is intended here as a measure of the extent to which a state is engaged in such international commerce, which is an attribute that can reasonably be expected to correlate with nuclear reliance. (States that develop a significant domestic reliance on nuclear energy can reasonably be expected to capitalize on the expertise that necessarily accompanies such

development by exporting the products of that expertise, to the benefit of both themselves and the importers.) As a dummy (bivariate) variable, NSG? certainly is an imperfect measure of international commercialization of nuclear activities, because membership covers a wide range of degrees of commercial activity in the international nuclear market. Membership in the NSG? is necessary to significant commercial activities in the international market for nuclear materials and technology, but it is by no means sufficient.

3. $NUEXP$ (Nuclear Export): Inclusion of the variable Nuclear Export in the dataset is to provide economic value to nuclear technology and material export. This variable provides the monetary value for the export of nuclear material and parts for nuclear reactors and other nuclear facilities. The idea is that states with larger export of the nuclear material will rely more on nuclear power than the states with minimal nuclear technology and material export. The data for this variable are obtained from UN Comtrade database. Code 8401 in the database contains nuclear technology parts such as nuclear reactors, fuel elements (cartridges), non-irradiated, for nuclear reactors, machinery, and apparatus for isotopic separation [10]. This variable was prorated on per capita basis and unit standardized by the maximum value over all states.

4. $REPRO$ (Reprocessing): Reprocessing is a key variable in the backend of the fuel cycle. Civil nuclear reactors that can use reprocessed material as fuel can be hypothesized to increase nuclear reliance because advanced level of technology and knowledge exists in the state. Inclusion of frontend and backend variables (such as $ENRCH$, $MING$ and so forth) appeared advisable as part of a replacement for the old international commercialization data, because nIC only includes information about operating nuclear reactors. The data for re-

processing was obtained from Integrated Nuclear Fuel Cycle Information Systems website, maintained by International Atomic Energy Agency (IAEA) [22]. These data are based on officially nominated country coordinators and open sources of information. About 25 countries actively update the information at least once in a year. In other cases updates are done to the information several times a year [37]. For a given state, numbers of tons-heavy metal produced per year for all the operating facilities were added to get a total.

A.3 Remaining Independent Variables (Correlates)

Description of all candidate variables below do not get selected to either in model (3.1) or in the final model (4.11). All the variables description are listed alphabetically within their respective natural categories.

A.3.1 Energy

1. *COAL* (Reserves of Coal): This attribute was one of the significant variable in Nelson-Sprecher model in [1]. Nelson and Sprecher found coal as a major alternative to nuclear fuel for electrical production. Coal with around 40% of electricity worldwide still consists of considerable share in global electricity production, when compared to 12.3 % electricity generation from nuclear energy [38]. Coal is a substantial alternative to nuclear fuel because it is reliable, inexpensive, and readily available source of electricity generation. This attribute was prorated again on per capita basis then was unit standardized over all states. Most recent data for coal reserves was collected in million short ton units from the U.S. Energy Information Administration, circa 2008 [39].

2. *EGEN* (Electricity Generation): Electricity generation attribute was used by Nelson and Sprecher in [1] as a measure of electricity demand within a state.

It is hypothesized that a state might consider investing in civil nuclear power if there is an increase in electricity demand. 2011 electricity generation were data were obtained from U.S. Energy Information Agency (EIA), which was unit standardized by dividing by the maximum electricity generation value over all states [40].

3. *EI* (Energy Insecurity): This attribute is included in the dataset on the basis that states having high energy imports might consider nuclear energy in order to restrict their dependency on imported energy. Nelson and Spercher rational to include this attribute in Reference [1] was for the fact that the nuclear energy, with high energy density and longer refueling period, would work as an incentive for states looking to minimize foreign dependence, immunity to short-term market or political fluctuations on energy. It was also hypothesized states with higher imports that in state production will be inclined to built nuclear power plants, if other factors remain constant. Year 2009 energy insecurity data was collected from The World Bank for most states and 2010 for some states. A negative value in the data means the given state is a net exporter of energy [41].

4. *GAS* (Reserves of Natural Gas): This attribute was previously selected as a candidate variable in [1] by Nelson and Sprecher. The relative reserves of natural gas is defined as the ratio of national reserves of natural gas to electricity generated within the state [1]. At present times natural gas seems to be in dominance over nuclear energy as some of the supporter are moving their attention away from nuclear to cheap natural gas supplies to produce electricity [42]. 2008 reserves of natural gas data was obtained from EIA [43]. This attribute was unit standardized by the maximum value among all 86 NC states.

A.3.2 International Commerce

All variables in International Commerce of nuclear fuel are hypothesized to increase the nuclear reliance of the supplier state. States having technology and knowledge fuel-cycle are expected to utilize the resource for both civil nuclear power generation and export for profit and jobs.

1. IC_{COVR} (International Commerce of Conversion): This attributes measures trade in nuclear material that was converted to UF_6. Data for commerce in converted material were obtained from Table 5 of Country Nuclear Fuel Cycle Profile reported by IAEA [44]. Canada, France, Russia, UK, and USA are the suppliers for this attribute. Every instances of 'x', where each 'x' represent domestic or international trade. The instances of 'x' that represent international trade of conversion material were added first for a given state and prorated on the basis of per capita. Unit standardization for this variable was adopted for this variable.

2. IC_{ENRCH} (International Commercialization of Enrichment): International commercialization of enriched nuclear fuel provides export data for enriched uranium produced by a given state. The trading of enriched fuel is considered to have direct impact on nuclear electricity generation process. The data were collected from Country Nuclear Fuel Cycle Profiles, published in 2005 [44]. From Table 6 of this IAEA document, each 'x' under a given enrichment company was given one point. A 'x' represent both domestic and international trade, but only international trades were accounted. In this national category IC_{ENRCH} is the only variable in [44] where company names were provided instead of a state; therefore, corresponding state was identified depending on the location of enrichment companies' headquarters. Every instances of 'x' conducting in-

ternational trade were added together, and the corresponding numerical value was assigned to the state. The total value was prorated per capita and unit standardized over all states. Old international commercialization variable in [1] rendered information of nuclear fuel cycle subjectively. The addition of variables under this category along with other nuclear fuel material commerce is employed to rectify the missing fuel cycle information from nIC. This variable was eliminated during a search of a better model Section 4.

3. IC_{FUFAB} (International Commerce of fuel fabrication): International commerce in fabricated fuel is another measure of nuclear material trade. Data collected from Table 7 of Country Nuclear Fuel Cycle Profile report many states being involved in international trade of fabricated material. Belgium, France, Germany, Russian Federation, Spain, Sweden, UK, and USA are among the suppliers for this attribute [44]. Again similarly to IC_{ENRCH} every instances of 'x' corresponding to each international supplier state were added and prorated on per capita basis, the unit standardization was done by dividing by the maximum value over all state.

4. IC_{MING} (International Commerce of Milling and Mining): Mining and milling commerce data were also collected using the same IAEA report mentioned above [44]. Table 4 of the IAEA report was used to obtain the supplier state information. Data for this attribute has *Others* as supplier, which has a total of 15 instances of 'x' that could not be assigned to any given state. Among the suppliers Australia, Canada, Former Soviet Union, and USA were listed by name, but states such as Kazakhstan could not be identified individually, and were not valuated. Data obtained were prorated on per capita basis, and then unit normalized by the maximum value over all states.

5. IC_{REPRO} (International Commerce of Reprocessing): Similarly to previous variables in this subsection, data for international commerce in reprocessing were collected using Table 8 of Country Nuclear Fuel Cycle Profile reported by IAEA [44]. According to the IAEA report, only three states (France, UK, and Russia) are involved in international trade of reprocessed fuel. According to the data presented in Integrated Nuclear Fuel Cycle Information Systems [22], Belgium also has an operating reprocessing facility which produces about 100 ton HM per year; however, it is not mentioned in [44] either as a domestic supplier or international supplier, and no points were given for such states. This variable was eliminated during a search of a better model in Section 4.

A.3.3 Nuclear Material Processing

States with advanced nuclear fuel processing facilities are hypothesized to have increased nuclear reliance. States with such knowledge and technology are expected to have most of the necessary resources to build civil nuclear power plants, and could also lead to military application of nuclear technologies.

1. $COVR$ (Conversion): Similarly to other attributes from this natural category, conversion was also obtained from the Integrated Nuclear Fuel Cycle Information Systems [22], and was coded for each state as the total amount of UF_6 produced in metric tons of uranium per year. Conversion is the process conducted in the frontend of nuclear fuel cycle after milling and mining to convert U_3O_8 into UF_6. Commercial enrichment facility require a gaseous process medium; therefore conversion of uranium oxide to uranium hexaflouride is important because of its high vapor at room temperature [45]. Materials from this step and beyond in a nuclear fuel cycle are under detailed nuclear accounting and verification for safeguards purposes [46].

2. *ENR?* (Enrichment and Reprocessing): Enrichment and reprocessing were coded in aggregate as a subjective and dichotomous variable,to measure either physical presence of an enrichment or reprocessing facility. States country profiles were studied in Nuclear Threat Initiative [47] website to obtain their history of enrichment and reprocessing. A value one was given to states that are thought to have knowledge from [47] of either facility can build them if necessary, and states that have commercial enrichment or reprocessing facility were given a value 1. Remaining states were given a value zero. In addition to separately defining enrichment and reprocessing this subjective attribute was included because some states (e.g. India and Iran) have not declared the presence of any existing nuclear facility; therefore no data for such states were collected in *ENRCH* and *RePro*.

3. *ENRCH* (Enrichment): Enrichment was encoded using enriched fuel production data for each state (metric tons of uranium per year), collected from the Integrated Nuclear Fuel Cycle Information Systems [22]. Similarly to *RePro*, the amount of currently enriched nuclear fuel material was aggregated for each state. The three major states with largest production of enriched material are France, United States, and Russia respectively. There are total of nine states that have a currently operating uranium enrichment facility. Enrichment is a key frontend process, which is a part of sensitive technology for weapons proliferation. Unit standardization is obtained by dividing by the maximum value over all states for this attribute.

4. *FUFAB* (Fuel fabrication): Fuel fabrication was encoded similarly to other variables in this category. Data collected from Integrated Nuclear Fuel Cycle Information System were recorded [22]. Of 86 states under consideration, 17

states had an operating fuel fabrication facility. This facility is the final step before the nuclear fuel is prepared to use in a nuclear reactor. Different types of fuel assemblies are produced in this facility depending on the type of reactor.

A.3.4 People and Economy

1. *GDP* (Gross Domestic Product): GDP is one of the two attributes that has been used to select states that are included in the dataset. Nelson and Specher included this variable in the prior study [1] because of the large capital requirement for a state to have the capability to build a nuclear power plants (typically of $\approx$ 1000 MWe). Gross domestic product data were collected from [12] and were unit standardized by the maximum value over all states.

2. *pcGDPppp* (per capita Gross Domestic Product purchasing power parity): This attribute was hypothesized to have positive correlation with nuclear reliance. Higher per capita GDP tends to correspond to technological education and training by a large fraction of population, as required to work in nuclear power plants. This attribute also correlates to increase in electricity demand [1]. Data for this variable were obtained from [48], and were unit normalized by the maximum among all states included in the dataset.

3. *MLES*? (Mid-Level Economic State): Mid-level economic state variable was created to see if the states with "*medium GDP*," which implies to have a unit-standardized *pcGDPppp* between 0.1 and 0.4, can be involved in high nuclear reliance. A value of 1 was given to states within that range, and zero for rest of the states. Hypothesis underlining this variable was that the states with low aggregate wealth are not capable of investing in nuclear power plants, and states with high aggregate wealth could invest in other forms of alternative

energy leaving states with medium level aggregate wealth uniquely drawn to investing in known technology with reliable base load power.

4. *POP* (Population): As exercised by Nelson and Sprecher in [1], population was one of the attributes used to select states that are included in the dataset. This attribute is also significant as it has been used to prorate several independent variable in the current dataset. Additionally this variable was included as an independent variable because states with large population are hypothesized to pursue civil nuclear power to meet the increase energy demand. Population data were obtained as reported in [49]. Population data were unit normalized by dividing by the maximum value over all states.

A.3.5 Political

1. *ALGN*? (Historic Alignment): Historic alignment was one of the significant variable in Nelson-Sprecher model in Eq. (2.2). No changes were made in this variable for this thesis project. Nelson and Sprecher implemented this variable as a dichotomous variable that will show material and technology "assurance" of supply for states that unable to produce such technology and material domestically. This variable hypothesized such dependency would make a major impact in "policy decisions, electricity generations, and building civil nuclear power plants." From Reference [1] states that are "neither fuel-cycle states nor *de jure* nuclear weapon states, but are successor states of the former Soviet Union or have at one time been member of NATO, SEATO or the Warsaw Pact, and Pakistan and Taiwan" were given value 1 and the rest of the states were assigned a value 0 [1].

2. *FCS*? (Fuel Cycle State): Nelson and Sprecher in Reference [1] defined FCS as "A Fuel Cycle State is a state that is not a nuclear-weapon state, under the

provisions of the Nuclear Nonproliferation Treaty, but nonetheless attempts to attain some level of nuclear material & technology (M&T) assurance through maintaining some indigenous capability for the relatively difficult technologies required to produce material that can help to sustain a chain reaction (i.e., enriched uranium, recycled plutonium or heavy water)." This dicohotomous variable assigned value of 1 to *de facto* nuclear-weapon states (India, Israel, and Pakistan) and additional states Argentina, Brazil, Canada, Japan, Netherlands, and remaining states were assigned a value 0. States assigned value one under this variable is hypothesized to have greater difficulty to have a successful civil nuclear energy program, when compared to *de jure* nuclear-weapon states[1].

3. *NWs*? (*de facto* Nuclear Weapon State): This attribute was not updated for current thesis project. Nelson and Sprecher in [1] assigned a value 1 for all *de jure* nuclear-weapons states under the Nonproliferation Treaty and addition three states; India, Israel and Pakistan (More recently such states have been described as "nuclear-armed"). The hypothesis behind inclusion of this variable was existence of some relationship between civil and military nuclear power programs [1].

4. *NWS*? (*de jure* Nuclear Weapon State): This attribute was not changed for the current thesis. For this is dichotomous measure, a value 1 was assigned to the five recognized nuclear-weapon states under the Nuclear Nonproliferation Treaty, and 0 to the remaining states. The idea for addition of this variable as a candidate independent variable was to identify the impact of being a *de jure* nuclear-weapons state on the civil nuclear program of a state [1].

5. *PFPS*? (Primary Fuel Production State): This dichotomous variable was included to measure a state's ability to produce "sensitive nuclear material."

Nelson and Sprecher hypothesized that this variable will increase the civil nuclear program of a state. All states in FCS and all *de jure* nuclear-weapons states were assigned value 1 and all other states were assigned a value 0. The three variables FCS?, NWS?, and $PFPS$? are linearly dependent [1].

APPENDIX B

DATASET FOR VARIABLES IN MODELS (3.1), (4.11), AND (6.9)

All the raw data for variables included in models (3.1), (4.11), and (6.9) are listed below in Table B.1. Model (3.1) in Section 3 contains variables nIC and $PLTY$, model (4.11) in Section 4 contains variables $NSG?$, $RePro$, nIC, $NUEXP$, and $MING$, and model (6.9) in Section 6 contains variables $NSG?$, $RePro$, $PFPS$ nIC, and EI.

Table B.1: Dataset for variables appearing in models (3.1), (4.11), and (6.9)

State	nIC	$PLTY$	$NSG?$	$RePro$ [tU/yr]	$NUEXP$ [$]	$MING$ [tU/yr]	$PFPS?$	EI [%]
Algeria	0	2	0	0	0.00	0	0	-283
Argentina	0	8	1	0	12007.00	0	1	-9
Australia	0	10	1	0	137789.00	9438	0	-157
Austria	0	10	1	0	3607513.00	0	0	66
Bangladesh	0	5	0	0	0.00	0	0	16
Belarus	0	-7	1	0	0.00	0	0	85
Belgium	0	8	1	100	564695115.00	0	0	73
Brazil	0	8	1	0	255114.00	340	1	4
Bulgaria	0	9	1	0	0.00	0	0	44
Canada	11.57	10	1	0	32181798.00	14890	1	-55
Chile	0	10	0	0	0.00	0	0	69
China	2	-7	1	0	4315512.00	900	1	8
Colombia	0	7	0	0	0.00	0	0	-211
Congo-Kinshasa	0	5	0	0	0.00	0	0	-2
Croatia	0	9	1	0	5184510.00	0	0	53
Cuba	0	-7	0	0	0.00	0	0	52
Czech Republic	8	8	1	0	140355447.00	400	0	26
Denmark	0	10	1	0	107339.00	0	0	-18
Egypt	0	-3	0	0	7519.00	0	0	-22
Ethiopia	0	1	0	0	0.00	0	0	7
Finland	0	10	1	0	5972.00	0	0	52
France	15.27	9	1	4995	549497297.00	0	1	49
Germany	5.74	10	1	0	557767225.00	0	0	61
Ghana	0	8	0	0	2394.00	0	0	24
Greece	0	10	1	0	1147.00	0	0	66
Guatemala	0	8	0	0	0.00	0	0	39
Hong Kong	0	0	0	0	0.00	0	0	100
Hungary	0	10	1	0	0.00	0	0	57
India	0	9	0	0	3842142.00	175	1	26
Indonesia	0	8	0	0	29584.00	0	0	-74
Iran	0	-6	0	0	399507.00	0	0	-62
Iraq	0	3	0	0	0.00	0	0	-272
Ireland	0	10	1	0	11179.00	0	0	87
Israel	0	10	0	0	0.00	0	1	83
Italy	1	10	1	0	720269.00	0	0	83
Japan	4.15	10	1	0	228508908.00	0	1	81
Kazakhstan	0	-6	1	0	0.00	12200	0	-121
Kenya	0	8	0	0	0.00	0	0	17
Kuwait	0	-7	0	0	0.00	0	0	-332
Lebanon	0	7	0	0	0.00	0	0	97
Libya	0	-7	0	0	0.00	0	0	-327
Lithuania	0	10	1	0	0.00	0	0	50

Table B.1: Continued

State	nIC	$PLTY$	NSG?	$RePro$ [tU/yr]	$NUEXP$ [$]	$MING$ [tU/yr]	$PFPS$?	EI [%]
Malaysia	0	6	0	0	644149.00	0	0	-34
Mexico	0	8	1	0	127499.00	0	0	-28
Morocco	0	-6	0	0	0.00	0	0	95
Myanmar	0	-6	0	0	0.00	0	0	-48
Nepal	0	6	0	0	0.00	0	0	11
Netherlands	0	10	1	0	693496281.00	0	1	16
New Zealand	0	10	1	0	0.00	0	0	10
Nigeria	0	4	0	0	800.00	0	0	-111
North Korea	0	-9	0	0	0.00	0	0	-5
Norway	0	10	1	0	10891.00	0	0	-563
Pakistan	0	6	0	0	673745.00	0	1	24
Peru	0	9	0	0	0.00	0	0	4
Philippines	0	8	0	0	0.00	0	0	40
Poland	0	10	1	0	0.00	0	0	33
Portugal	0	10	1	0	4330893.00	0	0	75
Qatar	0	-10	0	0	153621.00	0	0	-487
Romania	0	9	1	0	9.00	710	0	18
Russia	45.18	4	1	1200	1201804182.00	4300	1	-83
Saudi Arabia	0	-10	0	0	0.00	0	0	-235
Serbia	0	8	0	0	0.00	0	0	35
Singapore	0	-2	0	0	590381.00	0	0	100
Slovakia	0	10	1	0	96.00	0	0	64
Slovenia	0	10	1	0	0.00	0	0	50
South Africa	0	9	1	0	35865542.00	8572	0	-12
South Korea	0	8	1	0	42525886.00	0	0	82
Spain	0	10	1	0	672567288.00	0	0	74
Sri Lanka	0	4	0	0	0.00	0	0	45
Sudan	0	-2	0	0	0.00	0	0	-123
Sweden	3.33	10	1	0	1061266866.00	120	0	36
Switzerland	1	10	1	0	344141.00	0	0	52
Syria	0	-7	0	0	0.00	0	0	-5
Taiwan	0	10	0	0	0.00	0	0	95
Tanzania	0	-1	0	0	76911.00	0	0	8
Thailand	0	4	0	0	622032.00	0	0	40
Tunisia	0	-4	0	0	0.00	0	0	15
Turkey	0	7	1	0	10740.00	0	0	71
UAE	0	-8	0	0	42679.00	0	0	-183
UK	4.23	10	1	2400	46544987.00	0	1	27
Ukraine	9.94	6	1	0	3810889.00	1000	0	33
USA	63.64	10	1	0	418859769.00	6909	1	22
Uzbekhistan	0	-9	0	0	0.00	3000	0	-24
Venezuela	0	-3	0	0	0.00	0	0	-204
Vietnam	0	-7	0	0	0.00	0	0	-20
Yemen	0	-2	0	0	255060.00	0	0	-101