2014

Evaluating Alternative Techniques for Forecasting Industrial and Occupational Employment

Drew A. Varnado
Louisiana State University and Agricultural and Mechanical College, dvarnado@agcenter.lsu.edu

EVALUATING ALTERNATIVE TECHNIQUES FOR FORECASTING INDUSTRIAL AND OCCUPATIONAL EMPLOYMENT

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Agricultural Economics

by
Drew A. Varnado
B.S., Millsaps College, 2005
August 2014

Acknowledgements

I would like to acknowledge the help I've received during my work. The work would not have been possible without the help of my advisor, Dr. Matt Fannin. His guidance, patience, persistence were, and remain, invaluable to me. I would like to especially thank my committee for their efforts and Dr. Paul Darby for his help.

Mostly, I would like to thank my family. My wife, Sarah, your support is the foundation upon which I stand, and without you I would truly be lost. I hope this work proves to my sons, Jude and Anders, that hard work and persistence is worth the investment. To my parents, Scott and Lindsay, and my brother, William, I say this achievement belongs as much to you as to anyone. Without the love, guidance, and motivation of my family, none of this would be possible; anything I achieve in life is a testament to them.

Table of Contents

Acknowledgements ii

Abstract 1

1.1 Framing the Policy Questions 1

1.2 Objectives 6

1.3 Guidelines for the Creation of Industry/Occupation Employment Forecasts 10

1.4 State of the Art 12

2 Review of Literature 17

2.1 Introduction 17

2.2 Bureau of Labor Statistics Employment Projections 17

2.3 Evaluating Employment Forecasts 20

2.4 Comparative Analysis 25

2.5 Alternative Industrial Output Forecasting Techniques 27

2.6 Meta-Analysis Regression 34

3 Methods and Models 37

3.1 Introduction 37

3.2 Industrial Output Models 37

3.3 Occupation-by-Industry Matrices 65

4 Data and Descriptive Statistics 71

4.1 Introduction 71

4.2 Louisiana State Output-by-Industry Estimates Data and Summary Statistics 71

4.3 Occupation-by-Industry Matrix Data and Summary Statistics 76

4.4 Industrial and Occupational Employment Data and Summary Statistics 82

5 Results and Statistical Analysis 88

5.1 Introduction 88

5.2 Louisiana Industrial Output Forecast Estimates and Statistical Analysis 88

5.3 Louisiana Industrial and Occupational Employment Forecast Estimates Statistical Analysis 96

5.4 Summary and Analysis of Results 105

6 Conclusions, Contributions, and Suggested Extensions 109
6.1 Introduction 109
6.2 Summary 109
6.3 Contributions 111
6.4 Conclusion 115

Cited Works 117

Appendix A: Computable General Equilibrium 126

Appendix B: GAMS Code 144

Appendix C: Full Descriptive Statistics 174

The Vita 184

Abstract

This paper offers three different regional output-by-industry forecasting techniques (time series, Social Accounting Matrix (SAM)-based, and Computable General Equilibrium (CGE)-based) and two different occupation-by-industry matrices (national and state geographies) for use in the creation of industry/occupation employment forecasts. Estimates are compared to actual data from eight years for 2001 to 2010. OLS regressions are run to determine how well modeled employment estimates fit actual employment for the state of Louisiana. A meta-analysis-style regression of the R-squared values on model characteristics (accounted for using Boolean dummy-variables) determines that industrial output forecasting techniques do not provide statistically different R-squared values, but that models which use the state level occupation-by-industry matrix constructed for this paper should expect a statistically higher (by about 3.5%) R-squared value. Theil inequality coefficient decomposition analysis indicates that the assumed direct link between output and employment present in many projection methodologies may need further consideration.

1 Introduction

1.1 Framing the Policy Question

As regional and state policy agencies address workforce and education issues that accompany macroeconomic structural shifts, accurate regional and state industry/occupation employment projections are a vital input to policy functions at regional and state levels. Sweeney (2004) asserts that "detailed industry/occupation employment forecasts are an important class of regional labor market information." Using these projections, policymakers determine state funding for various educational and training programs that will produce the necessary workers to fill future employment needs.

With regional and state policies frequently informed by industry/occupation employment projections, the methods and techniques used to produce such forecasts are crucial to the efficacy of the policies they influence. Inaccurate, biased projections could cause significant mismatching in the efficiencies of labor markets (Sweeney, 2004). Therefore, it becomes increasingly important that economists strive to provide projections using the most appropriate processes to policymakers during this transitional period in the U.S. and world economy.

One of the most pressing concerns regarding regional industry/occupation forecasting is determining which techniques, when used within the BLS recommended process, best serve the needs of state policymakers, workers, and industries. There are several different methods that vary in approach, technique, structure, complexity of theory, complication in construction, and data requirements. Thus, an assessment of the practical costs and benefits to the state of

Louisiana of each model, or type of model, is valuable in hopes of finding the appropriate place in the policy function for each variation.

Sweeney (2004) asserts that "in theory, the [industry/occupation] forecasts should improve the national, interregional, and intertemporal matching efficiency of labor markets." However, this statement assumes that the projections are accurate or unbiased, hence the precursor "in theory." If the projections used to craft educational policies are not reasonably accurate, then those policies will not produce the desired economic effects, even if they create the intended amount of trained workers. That is, without a solid industry/occupation forecasting method, there is no guarantee that the specially trained workers will be in the correct industries or in the correct amounts. It then becomes critical that processes for producing such forecasts be examined at both the theoretical and practical levels, and further, that alternative projection models be considered.

This research will focus on two specific steps in the forecasting procedure for the state of Louisiana for years from 2001 to 2010 (which will be elaborated upon in the following section). The first is output-by-industry forecasts, which can be conducted using methods that vary in the level of complication. The second stage is to convert these output projections into employment counterparts. This second stage is a much simpler process both theoretically and technically but requires data granularity that can be difficult to obtain. However it maintains a key role in the creation of industry/occupation projections and deserves investigation by any paper discussing industry/occupation style labor projection techniques and processes due to some existing assumptions that are forced due to availability.

Tiebout (1969) asserts that in terms of industrial output estimates, the movement from econometric analysis toward new Input-Output (I/O) or Leontief-style models during the 1940's represented a significant step forward in the ability to model economies, and in a 1957 paper asserts that "it's not too much of an overstatement to say that post World War II regional research has been almost completely dominated by regional applications of Input-Output models." While academic regional economists have begun to move beyond Input-Output modeling after the analysis of ready-made secondary data-based Input-Output models, Input-Output analysis still dominates private sector consulting and government regional economic research (Partridge, 2007). Input-Output style impact analysis has been used to evaluate regional policy in almost every economic field: labor, growth and development, agriculture, and general industry, to name a few. Impact analysis in particular benefitted from the adoption of Input-Output techniques because, together, they provided a more realistic economic model in which policy alternatives could be tested.

Input-Output (or the expanded version, Social Accounting Matrix (SAM)) analyses improved the science of impact analysis by providing a more realistic view of the economy which includes previously ignored inter-industry linkages. That is, prior models had largely ignored or were incapable of accounting for such linkages, effectively ignoring indirect effects. Nicholson (1995) asserts that pricing outcomes in one market create ripples that affect other markets, possibly such that the ripples then affect the originating market. The existence of these rippled, indirect effects is ignored completely in partial equilibrium models, but de Melo and Tarr (1992) argue that a general equilibrium model, such as a model based on Input-Output framework, has the capability capture these inter-industry linkages. Input-Output models

essentially improved the glass through which economists and policymakers viewed the economy. By providing a clearer, more detailed and intricate picture, Input-Output analysis increased the efficiency of the projections, thus theoretically improving both policy efficiency and efficacy.

Just as Input-Output and Social Accounting Matrix analyses improved realism over partial equilibrium models, so Computable General Equilibrium (CGE) models are improving on Input-Output theory by relaxing many of the rigid assumptions that make Input-Output models unrealistic or problematic in evaluating regional policies. Waters, Holland, and Weber (1997) conclude that "compared with fixed-price I/O and econometric forecasting models, Computable General Equilibrium models can better address the implications for efficiency and equity of alternative public policies because the underlying assumptions regarding economic behavior are more tenable." CGE is a system of simultaneously solved equations that govern the actions of economic agents of a variety, thus it could be classified as a Walrasian, neoclassical, general equilibrium approach. The basic argument for CGE implementation over I/O or SAM is similar to the arguments originally made by I/O over econometric models: by including more economic information in the model one improves the ability of the model to react as markets do in reality.

With the idea that CGE models represent a theoretical improvement over Leontief structures (Menezes et. al., 2006), the policy question then becomes: Are the theoretical superiorities of I/O and SAM models over econometric models and CGE over I/O or SAM models apparent in their production of industry/occupation forecasts for the state of Louisiana?

Though improvements in economic understanding allows for the modeling of more complex ideas, the more pertinent question is, are any practical improvements worth the costs incurred in their acquisition at the regional level? CGE models are larger, more complex to construct (and understand) and easily misspecified. Are the improvements in projections significant, providing real-world improvements validating any increase in costs and effort spent by state agencies?

Regardless of output estimation method chosen, the next task is the conversion of those output rates into occupational units. Most state agencies follow BLS recommendations (Franklin, 2007) which suggest applying fixed staffing patterns of occupation-by-industry directly to output projections. This information is contained in occupational matrices at the national level without regional consideration, which may be biasing results (Vargas et.al., 1999). The possible biases may be the result of a geographical information mismatching. That is, this process contains the rather untenable assumption that national staffing patterns are an appropriate proxy for regional patterns. However, as regional staffing pattern matrices do not exist for the state of Louisiana it remains unknown how this unification of geographical information may improve estimates of industry/occupation projections.

Current methods of application in which these staffing patterns are applied to industrial output estimates will be spelled out in future sections along with several alternatives, but the basic approach remains the same. Each different method will apply labor market information directly to output estimates to obtain the estimate of the number jobs by occupation and by

industry. Variance in this particular application of the process comes from geographic nature rather than the structure of data.

A more thorough understanding of these two particular steps within the projection process, the theory that each is built on, their sensitivity, their structure, and the ways in which they can be used most appropriately is important for policymakers in creating policies that shape educational funding and therefore workforce development and efficient economic growth at the national, regional, and local levels. Knowing how these models perform empirically is equally as important as understanding theoretical underpinnings. Further, empirical testing of this nature has been ignored in labor economics to this point.

1.2 Objectives

1.2.1 General Objective

The primary objective of this paper is to analyze the procedure for the creation of industry/occupation employment forecasts for the state of Louisiana, considering alternative industrial output forecasting techniques and occupation-by-industry matrix geographies. This paper will focus on two primary mathematical pieces of the projection procedure published by BLS via the *BLS Handbook of Methods*, and thus will include evaluation of these individual parts as well as the combinations thereof. The structure of the industry/occupation forecasts as presented by BLS is given in Figure 1.1. The two steps for which alternative techniques are considered in this dissertation are enclosed in hexagons rather than boxes.

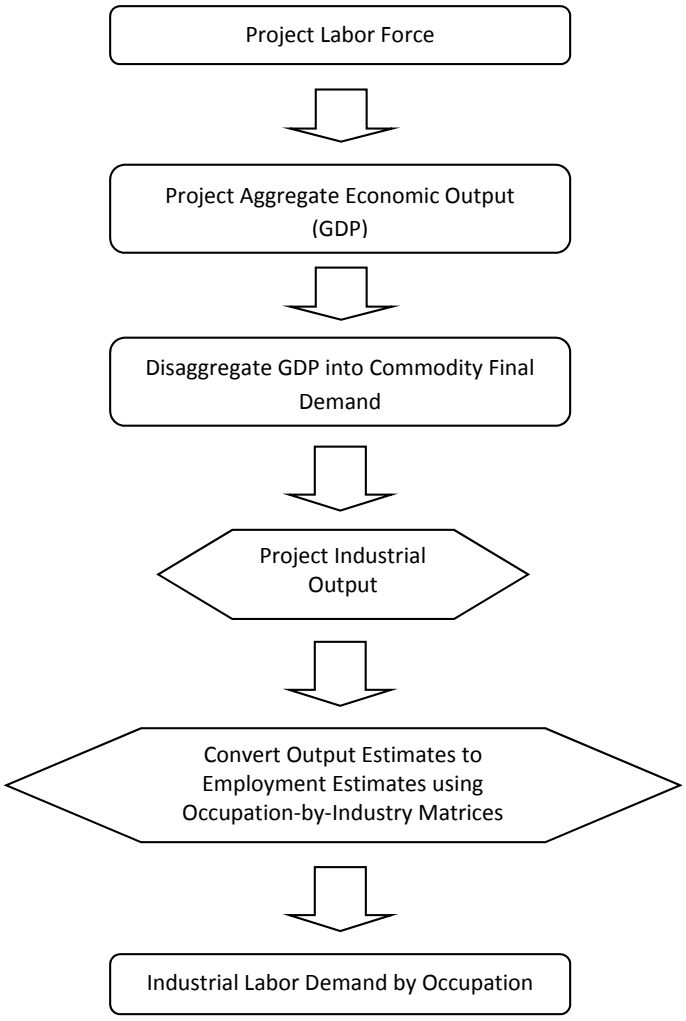


Figure 1.1: Industry/Occupation Forecast Procedure Flow Chart

First, this paper seeks to determine if the theoretical superiorities of Input-Output style models over partial equilibrium models and of Computable General Equilibrium over I/O or SAM models produce different industrial output forecasts and thus different industrial/occupational forecasts, statistically and functionally. The distinction between statistical and functional differences is crucial. Both will be addressed and both have valuable interpretation. However, following Friedman's instrumentalist advice (Hausman, 2008), functionality will take precedent in this paper. Should a simpler, less rigorous method provide

more accurate estimates when used as a prediction tool than a more theoretically sound model, the first model would be preferred. Simply, forecast accuracy is preferred to theoretical or interpretive solidarity.

Though there have been several published works covering the theoretical underpinnings and comparisons of Input-Output and CGE models, there have not been any that measured these differences empirically. There has been piece-meal discussion including a comparison of Input-Output and CGE models in theory (Partridge and Rickman, 1997), countless empirical Input-Output models, and slightly fewer empirical CGE's. However, this paper represents the first empirical effort to estimate any superiorities of CGE over Input-Output (and of I/O over econometric based models) in labor market forecasting. This paper will focus on the specific application of Input-Output and CGE to Louisiana industry/occupation projections, but if the models produce significantly different results those results could speak to a variety of other applications. Regional development groups, education boards, workforce development groups, and private industrial firms in the state of Louisiana, as well as nation-wide, all use output by industry projections and thus would stand to benefit from a quantified, empirical comparison of methods.

The second part of the evaluation pertains to the application of labor market information to these output estimates. Current methods use national staffing patterns to link industry output to labor. This paper will construct a matrix of staffing patterns for the state of Louisiana for comparison to the national level data to determine if proper scope provides more reliable labor estimates. Matrices of this variety do not currently exist for the state of Louisiana;

thus, it is constructed. This is the first set of state industry/occupation matrices constructed for the state of Louisiana for use in employment projections.

After the methods for output estimation and the varying occupation-by-industry matrices have been evaluated and compared to one another, projections will be produced using all available combinations of output estimates and occupation-by-industry matrices. Resulting comparisons should provide an outline and optimal projection techniques in terms of output estimates and application of staffing pattern data to such estimates within the industry/occupation employment forecasting framework.

1.2.2 Specific Objectives

This general objective will be accomplished via the following specific objectives:

1) Evaluate current and relevant projection procedures. (Chapter 3)
2) Evaluate the theoretical underpinnings of current projection procedures, as well an Input-Output and Social Accounting Matrix modeling, Computable General Equilibrium modeling, and alternative occupation-by-industry matrices. (Chapter 3)
3) Build SAM and CGE models of industrial output for the state of Louisiana as a single region. (Chapter 3)
4) Construct alternative occupation-by-industry matrices. (Chapter 3)
5) Test the industrial output techniques, the occupation-by-industry matrices, and all combinations thereof. (Chapter 4)

These specific objectives are discussed in detail, constructed, and tested in their designated chapters, so the rest of this chapter will outline for the reader some of the details of the BLS guidelines for industry/occupation employment forecasting, the current state of the art in the production of industry/occupation projections, and some of the motivation behind the alternative techniques considered in this dissertation.

1.3 Guidelines for the Creation of Industry/Occupation Employment Forecasts

BLS publishes guidelines for the creation of industry/occupation employment forecasts. Franklin (2007) outlines this six step process. Contained among the six steps are two of the primary interests of this dissertation: (1) industrial output projections and (2) the conversion of output forecasts to employment forecasts via occupation-by-industry matrices. Each set of projections made, regardless of alternative techniques for each individual step, will follow the general structure outlined below which is a discussion of the process described in Figure 1.1.

I. Project Labor Force

The first step towards industry/occupation projections is to estimate the future supply of labor. This is accomplished by applying labor force participation rate projections published by BLS to population projections published by the Census Bureau. This process projects the gross change in the labor force, which then is adjusted for various workforce reductions (military workers, prison inmates, etc.) and for exit trends. The data are smoothed and then checked for consistency with past trends.

II. Project Aggregate Economic Output

Aggregate output is projected vis-à-vis Gross Domestic Product (GDP) and primary demand and income sectors. These projections are published by BLS and are available at national and state levels and include assumptions about monetary and fiscal policy, energy prices and supply, world economic growth, and demographic changes. (Bureau of Labor Statistics)

III. Commodity Final Demand

Commodity final demand projections are the product of disaggregating the macro economic projections from the previous section. The aggregate demand is first divided into national income and product account categories, and then into the types of commodities purchased by those categories.

IV. Industrial Output Forecasts

This section generally disaggregates industrial output given the projected GDP from above. Each method, of which there are many, distributes the projected aggregate output over the industry categories according to the rules which accompany that particular technique. The most common techniques are partial-equilibrium, econometric models, but more sophisticated models can also be used.

The econometric model will have an equation for each industry that will project output dependent on past trends and the aggregate GDP projection. These equations may also include some adjustments for regional deviation from national trends when appropriate. The SAM and

CGE models, on the other hand, are general equilibrium and involve the creation of an input-output table. Regardless of the method used, the results are projected output by industry category. This will be discussed further under the specific objectives section.

V. Employment

The industry output derived above is then used to solve for industry-level employment. Data from the Current Employment Statistics (CES) and the Current Population Survey (CPS) are used to model industrial employment as a function of output, time, wages, and prices. These employment projections must be consistent with the aggregate employment projected above.

VI. Occupational Employment

The industry-level employment projections are then converted into occupational projections via an occupation-by-industry matrix. There will be two different types of matrices in use, national matrices produced by BLS and Louisiana state matrices which are constructed for this study. Each matrix consists of staffing patterns that distribute industry employment over the occupations within that industry, resulting in employment projections by industry and occupation. The data are then aggregated by occupation resulting in the desired industry/occupation projections.

1.4 State of the Art

When Louisiana state, regional, and local political agencies construct industry/occupation employment projections, most use some general version of the BLS recommended process. They may gather data from other sources rather than, say, projecting

the regional labor force internally; they use various techniques within the guidelines most commonly econometric (partial-equilibrium) and Input-Output style models to project industrial output and national occupation-by-industry matrices (Franklin, 2007).

Econometric estimation of industrial output is conducted using a variety of econometric techniques (panel-data regression analysis, time-series regression analysis, shift-share analysis, etc.) (Franklin, 2007). The future production of each industry is modeled as a function of previous industry output, some industry-specific measurements, and many policymakers insert variables to account for various expected trends over the projection period. The last of these inputs are, in many cases, attempts to correct for things that are not accounted for within the structure of the model; like inter-industry linkages, inter-regional linkages, national and/or regional economic trends, or industrial trends (Franklin, 2007). These models are partial equilibrium models, thus the accuracy of each industry's projections depends on granularity of industrial segregation, as well as specificity of data and the "tweaks" inserted asystematically by policy-makers.

In an attempt to improve output estimates some policymakers have used Input-Output or Leontief style models (Tiebout, 1969). These models are general equilibrium models and account for inter-industry linkages within the structure of the model. Thus, they do not require variable insertion to account for these linkages, as in econometric models. These models are often adapted for state and regional use from national models because of their intense mathematical structures, and they must be expanded in order to account for inter-regional linkages that are more prominent between states or regions than between countries. Vargas et

al (1999) claim that regional economies are more open than national economies and that frameworks designed for national economies may not correctly handle interregional openness, thus regional econometric models adapted from national models may be misspecified.

Today, most states use econometric models to estimate industrial output. Several regional industry/occupation forecast producers have moved to Input-Output or Social Accounting Matrix style models, but of those, many are developed by software packages that can be easily misused or adopted from national models. No states, to my knowledge, currently use Computable General Equilibrium modeling within the industry/occupation projection framework.

Regardless of output projection strategy, workforce projections are acquired by applying labor market information or staffing patterns (via a fixed occupation-by-industry matrix) to the industry output forecasts to determine expected output within each occupation-by-industry. These estimates can then be aggregated by occupation to predict future workforce needs. These projections of occupational needs by industry allow policymakers to fund educational programs according to the state workforce needs.

As a crude example, assume projections for five years from now indicate that Louisiana healthcare industry will need 1,000 more nurses than are currently projected as available. Policymakers might offer tuition incentives to nursing programs in order to create the trained workforce needed to meet that requirement. Obviously the reality is more complex, but the example illustrates the usefulness of industry/occupation projections and their effect on educational policy.

Currently, strategies for applying labor market information vary. Many state agencies use national occupation-by-industry averages while others use national averages with regional adjustments. The state of Louisiana does not produce a state-wide occupation-by-industry employment matrix. However, several economic processes may indicate that labor distribution by occupation over a given industry could differ across regions, and those differences may occur between the national and regional labor distributions (Goodrich, 1936). A region's endowments of natural resources, human capital, and infrastructure may cause production functions to differ slightly (Pindyck and Rubinfeld, 2001). That is, varying factor inputs could change the production function in a particular region to the point where labor requirements in that region are structurally different than the national industrial trends.

Further, differences between state and national matrices may stem from product life cycle variations. If production of a specific product in a specific region is at a different level of maturity than the national industry, it would have a different ratio of capital to labor due to differing stages in the product life cycle (Vernon, 1966). That is, if the region's production process is more mature than the national, that region would be using less high-skill labor, possibly more low-skill labor, and possibly less total labor.

As a simple example, consider the real estate industry. In the southeast portion of the US, most real estate transactions are brokered by real estate agents. However, in other parts of the country the same business transaction would be brokered by a lawyer. This difference may be a result of the institutional environment, inherited tradition, or possible policy differences between regions. Regardless, differences of this sort could cause the national occupation by

industry matrix to differ from a regional or state matrix. This paper, in part, will create these two matrices in order to empirically determine if they are different, how they are different, and possibly to offer some intuition as to why they might be different.

Though some states have attempted to use regional staffing pattern data, there exists no such matrix for the state of Louisiana. Basic economic geography reasoning would indicate that when creating projections for a specific region, input data from that region is preferred to information from larger or smaller segments of the population. Assuming local markets act like national markets could cause bias in the estimation process, leading to bias in the policy function which might be easily avoidable by matching geographic information between output estimates and labor market information.

It would benefit policymakers and economists to know whether alternate projection techniques provide benefits in the instrumentalist sense. That is, would alternate methods of projecting output and/or applying labor market information to estimates of output produce markedly better results not just in terms of theory, but in real world application?

2 Review of Literature

2.1 Introduction

This chapter sets out to provide a review of literature on several topics that will be prominent in the development of the ideas presented in this dissertation. The introduction is followed by a discussion of industry/occupation employment forecast procedures as laid out by the Bureau of Labor Statistics at both the national and regional levels. The next section addresses literature which compares employment estimates from projection models to actual employment statistics. The fourth section presents literature relative to comparative analysis, which helps economists choose between models which represent similar economic phenomena. The fifth section provides literature on the alternative industrial output forecasting techniques chosen for this study, while the sixth section presents literature on Meta-Analysis Regression which will be used in the fourth chapter. The final section offers some concluding remarks.

2.2 Bureau of Labor Statistics Employment Projections

2.2.1 National Framework for Employment Forecasts

The Bureau of Labor Statistics (BLS) publishes the *BLS Handbook of Methods*, in which it devotes an entire chapter (Chapter 13: Employment Projections) to the procedure used to create industry/occupation employment projections. This chapter begins by laying out a basic history of employment projection procedures, and state that “since the late 1970’s…, the basic

methodology has remained largely the same," and that the methodology is built on an Input-Output framework and a national occupation-by-industry matrix (Franklin, 2007).

BLS produces national industry/occupation employment projections using the framework from Figure 1.1 and has done so regularly since late 1960's (Fullerton, 2003). These forecasts are published by BLS and available for a variety of base years and projections years, with a variety of time-lengths (Bureau of Labor Statistics). These projections are used widely and have been tested.

2.2.2 Regional Framework for Employment Forecasts

Regional frameworks for producing employment forecasts generally follow the outlines set by the national procedures (Bell, 1981). Most State Employment Security Agencies use the general framework laid out by the *BLS Handbook of Methods* discussed in the previous section modified to fit regional projects where possible, or they contract out employment projections (Goldstein and Cruze, 1987). Companies like Moody's Analytics and ProjectionsCentral.com sell industry/occupation employment forecasts and analyses which are often purchased by regional economic agencies. While there are many topic areas that could be covered in regional employment projection analyses, this literature review will focus on three aspects of regional employment projections, each with a seminal paper: (1) the adoption of the general national framework to regional economic analysis (Goldstein and Cruze, 1987); (2) the outsourcing of regional employment projections to proprietary entities; and (3) the use of regional parameters to relax assumptions that national parameters are proxies for regional parameters.

Goldstein and Cruze (1987) calculated average errors for projections made by BLS in 1972 using essentially the same framework (including I/O techniques and advanced regression analysis) projecting 1982 industry/occupation employment. They cite problems with comparisons due to changes in aggregations, definitions, and classification systems, as in Wyatt (2010). Goldstein and Cruze (1987) suggest that the following improvements be made to state industry/occupation employment forecast framework: improvements in system design and data collection that would reflect available data and statistical tools, and improvements in the documentation of the creation and evaluation processes. These results bring forth the two primary branches of literature based on these regional projection frameworks.

The first branch is concerned with the technical aspects of creating better employment projections by using the best available data and statistical tools. Oosterhaven and Stelder (2007), Partridge and Rickman (1998), and Vargas et. al. (1999) represent different ways in which this branch has developed. Oosterhaven and Stelder (2007) outline the development of interregional I/O models from closed national models. They present, among other things, the use of regional I/O tables, the disaggregation from a national model to an interregional I/O model, and different ways in which the interregional models can be extended and applications in multiplier and impact analysis.

Partridge and Rickman (1998) introduce regional I/O coefficients instead of assuming that national coefficients are sufficient proxies for regional coefficients (this paper also addresses come comparative analysis and will come up again during a later section devoted to that topic). They evaluate five different models for employment forecasting in the state of

Georgia, ranging from a simple auto-regressive model to a general Bayesian Vector Autoregression approach, and find that econometric-style projections "were comparatively more accurate in the short run, but the most input-output-oriented models were comparatively more accurate in the longer run," and that "more accurate forecasts may be obtained in applications to regional economies that are less dependent on the national economy" (Partridge and Rickman, 1998). For other examples of studies that study improvements to regional frameworks by introducing more geography-appropriate parameters, see Vargas et. al. (1999) who discuss the relative openness of regional economies, Adkins et. al. (2002) who use Computable General Equilibrium (CGE) modeling for regional production, Giescke (2002) who uses a dynamic multi-regional CGE of regional Australian economies, and Kim and Kim (2002) who evaluates several regional development strategies.

The second branch mentioned above is the improvement in documentation of the creation and evaluation of these regional projections. This branch turns out to be rather large and encompasses how employment forecasts have historically been evaluated. This branch deserves its own dedicated discussion and is presented in the following section.

2.3 Evaluating Employment Forecasts

The evaluation of employment forecasts literature can most easily be categorized into two groups: (1) literature evaluating BLS projections and (2) literature that suggests improvements for, and tests assumptions and implications of, the generally accepted employment projection methodology in general (without producing any specific new

employment forecast estimates, but rather evaluating an aspect of the methodology in some other way).

2.3.1 Evaluating BLS Projections

Rosenthal (1999), Fullerton (2003), Stekler and Thomas (2005), and Wyatt (2010) all published studies which evaluate various BLS employment projections. These studies represent a large literature and were chosen to outline that literature and show how the BLS projections have been evaluated over the past 20 years, but these papers should by no means be considered exhaustive, merely representative.

Rosenthal (1999) discussed the general history of projection procedures used by BLS to forecast occupational employment and analyzed projections from five time periods ranging from 1960 to 1995. The study uses only descriptive statistics in its analysis (no regressions or detailed statistical tools), but finds that occupational employment projections predicted actual employment reasonably well. Projections are better for total employment estimates than for major occupational groups and better for major occupational groups than minor occupational groups. For other information on BLS forecasting history and review, see Toossi (2006).

Fullerton (2003) and Stekler and Thomas (2005) evaluate BLS employment projections for 2000. Fullerton (2003) evaluated five sets of labor force projections made by BLS between 1986 and 1994 that projected the labor force to 2000. Using a descriptive statistical analysis, the study suggests that projections of labor force participation rate projections overestimated actual rates in four of the five projection series and projections become less reliable as the granularity of the projections increases (Fullerton, 2003).

Stekler and Thomas (2005) study labor force, employment, and occupation projections for 2000 evaluating projections using a variety of descriptive statistical analyses, even going so far as to reframe some of the accuracy measures. The study finds that BLS projections "were comparable to estimates obtained from naïve extrapolative models" (Stekler and Thomas, 2005).

Wyatt (2010) follows Fullerton (2003) and Stekler and Thomas (2005) but evaluates employment projections for years 1996-2006. This paper is perhaps the most appropriate citation for this dissertation; Wyatt (2010) asserts that projections should be evaluated relative to alternative projections available (this segues nicely to section 2.4 which discusses comparative analysis). Wyatt (2010) finds that each of four different series of employment estimates produced by BLS are more accurate than projections made by alternative methods; however, the author identifies a problem that hampers all in depth, long-term employment forecasts (like the ones in this dissertation): the fact that changes in "data series, definitions, and classification systems hamper [an] article's analysis by decreasing the number of occupations available for analysis and creating substantial data comparison problems" (Wyatt, 2010). For other evaluations of BLS employment projections, see, most notably Alpert and Auyer (2003), but also Rosenthal (1997), and Hecker (2005) among others.

Though all of these studies evaluate BLS employment projections in some way, they all use simple descriptive statistical analyses rather than in-depth regression analysis or other more complex forms of analysis to determine how well the projections fit actual data. In many of these studies, descriptive statistics are sufficient, because there is only one set of projections

to evaluate. That is, there are no alternative models to be considered, no model characteristics to evaluate; the forecasts have errors when compared to actual data that are either small or large. There are few nuances that require more sensitive statistical tools as there might be if multiple alternative models are presented that each have relative strengths and weaknesses (Goldstein and Cruze, 1987; Vargas et. al., ; Partridge and Rickman, 2007).

2.3.2 Literature on the Indirect Aspects of the General Employment Projections Framework

The literature concerning the improvements in the general industry/occupation employment projections framework can be viewed in general as non-empirical. Though some works in this area do present empirical results, many do not, and those that do, often present empirical results that test a specific assumption or implication of the existing framework in an indirect way. For example, Bezdek (1984) tests three different implications of the Leontief I/O model which are central to the employment forecasting framework, but the paper never creates new employment projections nor evaluates any existing projections. This branch of the literature covers vast theoretical ground and an entire literature review of this kind could be produced for each of the six steps in the projection framework. Thus, to provide sufficient literature review in this area, examples and suggestions for further reading are presented very briefly.

Literature regarding the evaluation of actual labor force projections was discussed in the previous section addressing the evaluation of BLS employment projections. However literature regarding assumptions about labor force projections can be found in Durand et. al. (1996) who decomposes labor force growth into effects from population change and effects from changes

in the activity rates. Parsons (1980) discusses assumptions about labor force participation concerning shifts in participation rates of specific demographics.

For literature regarding the testing of employment migration and relative wage assumptions, see Blanchard and Katz (1992) and Barro and Sala-i-Martin (1991). Blanchard and Katz discuss regional evolutions in terms of regional employment, regional wages, and the linkages between the two, while Barro and Sala-i-Martin (1991) discuss how relative wages converge across regions, implying effects this convergence may have on regional worker migration. The reader may also see Treyz et. al. (1993) for discussion of specific U.S. migration.

There is a wealth of literature on regional employment multipliers including the work of Leontieff (1936, 1941), Pyatt and Round (1978), and Miller and Blair (1985), who lay out the foundations of the I/O framework and the multipliers. Early empirical studies of these multipliers were conducted in New Hampshire (Weiss and Gooding, 1968) and there are many modifications to the framework for specific situations: Mathur and Rosen (1974) presented a modified framework for regional employment multipliers which separated localized and non-localized employment and was critiqued by Isserman (1975) and Park (1970) suggested a regression approach to calculating employment multipliers.

The literature regarding alternative industrial output forecasting will be presented in section 2.5. It is separated from this discussion because the topic is central to this dissertation and the literature on specific alternatives deserves in-depth presentation.

2.3.3 Non-BLS Employment Projections

Since the 1970's, the general framework for the production of industry/occupation employment forecasts has not undergone significant structural change (Bureau of Labor Statistics), and thus the creation of actual projections has mainly been left to BLS, state agencies follow guidelines laid out by the BLS, and proprietary groups who produce economic statistics; thus, recent studies which produce new employment forecasts are rather rare, with the majority of recent literature focused on the evaluation of BLS projections or on assumptions and implications contained within the minutia of the projection framework. This leaves no recent studies that produce employment projections internally and empirically test those projections against actual data. Thus, this dissertation will take its testing guidance from the comparative analysis literature and from common econometric practice theory.

2.4 Comparative Analysis

The literature on comparative analysis, or the comparison of several models which purport to account for the same phenomena is presented (in a field called encompassing tests) as a theoretical problem in Cox (1961) who considered "comparing separate families of hypotheses" (Cox, 1961). Pesaran (1974) expanded the idea to testing linear single-equation econometric models, and Pesaran and Deaton (1978) expanded further to include comparative analysis for multivariate nonlinear models by deriving a test that uses alternative models to reject hypotheses about one another (forming the Cox-Perasan-Deaton, or CPD, encompassing test) by comparing five models of the consumption-income relationship.

Chong and Hendry (1986) discuss the distinction between theoretical solidarity and functional performance in the empirical evaluation of systems as a whole. That is, “since system characteristics are the prime concern of economy-wide models, it might be the case that the validity of every individual component is not essential to adequate overall performance” (Chong and Hendry, 1986). This question is particularly important to this dissertation in that empirical performance is preferred to solidarity with economic principle, ceteris paribus.

Regressions of actual data on projected values of that data were used by Fair and Shiller (1990) to compare real Gross National Product (GNP) growth rates from various models. Fair and Shiller (1990) cite the CPD tests, as well as Chong and Hendry (1986), in their discussion of these regression equations, but Fair and Shiller note that their tests do not provide insight as to whether a model, as a whole, contains “useful information” (a question that this dissertation addresses in a later section discussing Meta-Analysis Regression). The testing process, presented by Fair and Shiller (1990) and advanced by Diebold and Mariano (1995), and Diebold and Lopez (1996) (among others), is a relative test; that is, there is not a general goodness-of-fit calculated for each alternative model that has meaning outside of the relative study but rather the models are evaluated relative to one another. To include more models in the analysis, one cannot simply calculate a new test statistic for the new models, new regression equations must be run.

Fair and Shiller (1990) conclude that this type of comparative analysis seems useful in determining how alternative models perform relative to one another. For more detailed, mechanical referential material on this type of testing, the reader is directed to West and

McCracken (1998) who “develop regression-based tests of hypotheses about out of sample prediction errors” and find that simulations of these types of testing procedures work well and to Vuong (1989) and Rivers and Vuong (2002) who develop relative performance tests for nonlinear dynamic models.

This dissertation will rely heavily on this type of testing, as will become apparent in Chapter four. The general strategy of running regressions of actual data on projected data and comparing regression statistics and goodness-of-fit measures is a strategy that allows for multiple models to be considered for the same economic phenomena and evaluates them strictly on how well they fit the data, rather than on how well the structures fit economic theory and principle (Diebold and Mariano, 1995).

2.5 Alternative Industrial Output Forecasting Techniques

2.5.1 Input-Output (I/O)

Input-Output (I/O) and Social Accounting Matrix (SAM) analysis began its ascendency with Leontief’s work in the 1930’s which laid an analytical framework designed to describe the transactions and, more importantly, the inter-industry linkages within an economy (Miller and Blair, 1985). Though Quesnay published “Tableau Economique” in the 1750’s which contained some of Leontief’s structures and basic theories, it is Leontief that is referred to as the “father” of Input-Output analysis. The fundamental breakthrough that Quesnay and Leontief share is the separate descriptions of agents in the economy as producer and consumer. Leontief’s transaction table, which is at the center of both I/O and SAM analyses, is a clever representation of this theoretical separation.

A transactions table, at its core, is a double-entry bookkeeping of accounts, in which "each transaction appears in the table as an output or sale [row] and simultaneously as an input or purchase [column], and when factor incomes are included as inputs, then the sum of the outputs from each industry is observed to balance, in an accounting sense, with the sum of inputs" (Adams and Stewart, 1956). Leontief's Input-Output analysis structure consists of *n* linear equations with *n* unknowns, and is therefore solvable. The relations can also be represented in matrix form. The model is constructed from observable data that can cover regions of various sizes (nation, states, counties, etc.). Because much of the model is based around the linkages between industries, industrial clarity is required. The extent to which the sectors are divided can vary between models; industrial granularity should be determined with respect to the research question requirements and data availability.

The monetary flows from one industry to another representing production make up the primary data and are usually gathered over a specific, non-recurring time period. That is, the data are static. Note that the data are in monetary terms, thus this model is not able to account for changes in price within the specified time period. Although physical goods can be used as units, this is a much less common practice due to complex comparability issues across industry.

Miller and Blair (1985) provide a full review of input-output techniques, and the following basic structures follow directly from that review. Denote the flow of monetary value from industry *i* to industry *j* as z_{ij}. The amount of input *i* that will be used by industry *j* will be closely tied to the amount of output produced by industry *j*. In addition to providing industry *j* with an intermediate good, industry *i* may also sell its commodity in the final goods market. The

final demand for a product is the monetary value of the good in industry *i* sold to final goods markets, denoted as Y_i. Let the total production of industry *i*'s product be X_i, and be represented as

$$X_i = \Sigma_{j=1,2,3\ldots n}(z_{i,j}) + Y_i$$

This equation represents the total production for industry *i*, and there exists one for each industry in the model. Consider the compilation of these equations in matrix form.

$$X_1 = \Sigma_{j=1,2,3\ldots n}(z_{1,j}) + Y_1$$

$$X_2 = \Sigma_{j=1,2,3\ldots n}(z_{2,j}) + Y_2$$

…

$$X_n = \Sigma_{j=1,2,3\ldots n}(z_{n,j}) + Y_n$$

Note that, the matrix made by stacking these equations has two nice properties: (1) columns represent the sum of all inputs used by each industry in dollars, and (2) rows represent the sum of production from each industry in dollars, hence the name Input-Output. (Miller and Blair, 1985)

The information contained within this system of equations, combined with industrial data on value added and import/export activities, make up the transactions table. The industry of origin is represented by row information and purchasing industry by column in a square matrix. Alternative sectors of the economy, including value-added and imports, are represented by one row and one column each, representing production and inputs of the sectors

respectively. The number and specific characteristics of sectors added to the base production information can vary depending on the scope of the research question and data availability. (Pyatt and Round, 1978)

The use of transactions tables in I/O and SAM analyses allow the systems to account for inter-industry linkages in a much more specific, accurate manner than with a single equation model. In the transaction table, if output in one industry rises, it is represented as an increase in the column (input) entries for that industry (since the columns represent the input mix). Each input increase is, following the double-entry bookkeeping idea, also seen as an increase in the output of the intermediary good since more inputs are required to produce the extra output. This process is repeated for the intermediate good's industry, thus revealing the depth of inter-industry linkages that are overlooked in a single equation model.

These transaction tables are used in lieu of econometric equations to predict industry output. However, once these output projections are obtained, the process in which those forecasts are used to produce industry/occupation forecasts is the same in this paper as in standard state agency practice (using staffing patterns along with duration and turnover statistics).

2.5.2 Social Accounting Matrices

A Social Accounting Matrix (SAM) is a larger model produced by expanding an I/O model structure to include income, demand, and factors of production such that it focuses on more detailed descriptions of institutions and inter-institutional linkages. That is to say an I/O transactions table is a subset of the larger SAM model. By including information on income and

demand changes, a SAM can track economic linkages outside of production much the same way I/O models track them inside production sectors.

The SAM is formed by simply adding sectors for income, demand, and factors of production in dollar terms to the I/O transactions table. This expansion does not change the zero-sum, double-entry bookkeeping nature of the I/O model, while allowing policy makers to see effects beyond changes in production. In essence, because SAM models are a simple addition to any I/O model and allow for deeper interpretation and application, a SAM model will be used for this study. Also note that the theory, interpretation, and content are very much the same, which makes the above substitution of SAM for I/O quite natural (Pyatt and Round, 1979). In fact, most models created in policy work today are called I/O but in fact are SAM; thus, the two terms are often used interchangeably (even if technically incorrect).

2.5.3 Computable General Equilibrium

Computable General Equilibrium models do not contain rigid price and wage assumptions, instead prices (including wages, which are just the price of labor) are the solution set in which the system of equations is at general equilibrium. Still, Computable General Equilibrium models have not seen wide adoption in the regional policy world in part because they are complex to construct and require large amounts of regional data. As these problems become less restrictive with increases in regional data collection and storage and availability of computers with high processing power and large storage capabilities, Computable General Equilibrium models are becoming more feasible at the regional level (Vargas et al, 1999).

Robinson, Kilkenny, and Hanson (1990) summarized the debate between the theory of I/O and CGE models:

> The CGE framework offers an alternative for regional analysis. It encompasses both the I/O and SAM frameworks by making demand and supply of commodities and factors dependent on prices. A CGE model simulates the working of a market economy in which prices and quantities adjust to clear all markets. It specifies the behavior of optimizing consumers and producers while including the government as an agent capturing all transactions in circular flow of income.

A CGE model, at its core, is a "Walrasian neoclassical general equilibrium approach" (Vargas et al, 1999). The solution is a vector of prices which satisfies a system of equations that govern agents' actions. Producers maximize profits, while consumers maximize utility, both subject to constraints. Production factors are paid the value of their marginal productivities, and the solution prices clear commodity and factor markets.

In general, CGE frameworks offer several key improvements over Leontief structures which will be enumerated rather than expounded upon in this draft. As stated, CGE models do not contain fixed wage or price assumptions. CGE models contain inter-industry linkage information via the inclusion of a SAM, and thus contain many of the novelties of I/O-SAM frameworks, without some of the drawbacks (fixed price and wage assumptions or non-substitutability of inputs).

Vargas et al (1999) and Partridge and Rickman (1998) provide a rigorous review of CGE modeling which begins with assumptions and ends with solution methods. These papers provide a solid literature review for anyone unfamiliar with either national or regional CGE structures. Robinson, Kilkenny, and Hanson (1990) present a CGE of the US developed by the

Economic Research Service, but note that since then CGE modeling has expanded applications throughout domestic alternative policy analysis, especially regional CGE modeling.

2.5.4 Empirical Studies

Empirical papers using I/O-SAM models tend to focus on the empirics of testing intersectoral relationships (Cella, 1984; Tiebout, 1969; Forni and Paba, 2002), economic impact analysis (Fletcher, 1989; Psacharopoulos, 1973), forecasting structural change (Israilevich et. al., 1997), or estimating multiplier effects (Kilkenny, 1999; Frechtling and Horvath, 1999).

Being that most of the employment forecast estimates produced are done by BLS, as discussed earlier in this chapter, existing studies that examine, empirically, the difference between employment forecasts and actual data are limited to the literature which evaluates employment projections made by BLS, discussed in section 2.3.1.

Similarly to I/O-SAM models, CGE models have been used to investigate inter-industry linkages (Goulder and Eichengreen, 1992), alternative policy impact analysis (Kilkenny, 1993; Prescott, 1995), international trade (De Melo and Robinson, 1989; Shoven and Whalley, 1992), and structural change (Kilkenny and Otto, 1994). Though Johansen (1960) is generally given credit as the first attempt to use CGE to study actual economies, CGE models were not adopted regionally until much later with studies like Norrie and Percy (1983), Kimbell and Harrison (1984), and Liew (1984), which all looked at regional CGE models.

Though there are numerous papers which construct, discuss, and dissect I/O-SAM models and CGE models, there are relatively few which discuss them both (Partridge and

Rickman, 1997; Rose, 1995; Partridge and Rickman, 2008), and even fewer which empirically compare the two directly.

Seung et al (1997) compared a modified "supply-determined" SAM (SDSAM) to a CGE approach. They found that SDSAM models overestimate, relative to CGE, policy impacts on factor income and output and recommend CGE for "impact analysis where productive capacity of rural sectors is reduced," but results are restricted to impacts of surface water reallocation and have little input as to how CGE models might perform in terms of employment projections.

Cardenete and Sancho (2004) addressed the sensitivity of CGE models to SAMs of various aggregations and structures and found little sensitivity. Unfortunately, the study did not compare estimates of any kind from a model built on I/O-SAM framework to estimates built on a CGE framework.

Partridge and Rickman (2007) evaluated theoretical underpinnings of regional CGE modeling, compared CGE to I/O in terms of economic theory, and built a regional CGE, but did not directly compare to a similarly specified I/O model. Other papers have addressed related issues, but an empirical comparison that might apply to labor markets, much less regional labor markets, does not exist.

2.6 Meta-Analysis Regression

Meta-Analysis Regression (MAR) involves using regressions which use Boolean (dummy) independent variables to account for model characteristics (in possible conjunction with non-Boolean explanatory variable). Though some independent variables are discrete in MAR, the

dependent variable is continuous. That is, independent variables can be either discrete or continuous while the dependent variable is continuous. This allows for the regression coefficients from the MAR regression equation to indicate if the model characteristics (represented by a series of dummy variables) are associated with statistically significant changes in the goodness-of-fit of the previous regression equations (Stanley and Jarrell, 1989; Judge et. al., 1988).

In seeking to evaluate existing literature for unobserved patterns and to provide consensus, Stanley and Jarrell (1989) sought to create a testing structure which "provides a framework for replication and offers a sensitivity analysis for model specification." In this paper, they present an outline for Meta-Regression Analysis (MRA). This dissertation is not interested in any sort of meta-analysis, per se, but econometric testing which "offers a sensitivity analysis for model specification" is quite pertinent. By applying the econometric structure presented by Stanley and Jarrell (1989) to the question of model specification within the industry/occupation employment estimate procedure, this paper seeks the same sort of sensitivity analysis for model specification.

MAR provides a unique solution to the comparative analysis problem of choosing among available alternative methodologies; further, it echoes the "performance" ideas put forth by Chong and Hendry (1986) who place preference on statistical performance rather than theoretical solidarity or elegance. MAR analyses are most often used in, not surprisingly, meta-analyses. Various statistics have been used as the dependent variable in a wide variety of fields, such as medical treatment effects (van Houwelingen et. al., 1993; van Houwelingen et. al.,

2002), factors influencing share-holder wealth (Datta et. al., 1992), and new product performance (Montoya-Weiss and Calantone, 1994). The nature of MAR implies that virtually any statistic can be used as a dependent variable with various discrete and continuous variables used as independent variables (Stanley and Jarrell, 1989).

The only field in which MAR has an existing literature that relates to the topic of this dissertation is in employment turnover. Cotton and Tuttle (1986) used independent variables accounting for population, nationality, and other business characteristics to predict employment turnover rates. Griffeth et. al. (2000) follow Hom and Griffeth (1995) evaluating the "predictive strength of numerous turnover antecedents" (Griffeth et. al., 2000).

Though there are no guides for using MAR in employment projection analysis, Judge et. al. (1988) provide general economic theory for regression analysis using Boolean independent variables to explain continuous dependent variables.

3 Method and Models

3.1 Introduction

The primary objective of this chapter is to evaluate alternative methods for projecting workforce needs for the state of Louisiana. This chapter will focus on two primary functional pieces of the projection method, and thus will include evaluation of these individual parts as well as the combinations thereof. Section 3.2 will discuss three methods of producing industry output estimates; Section 3.3 will discuss occupation-by-industry matrices, and section 3.4 will provide a summary and conclusion.

3.2 Industrial Output Models

The general structure of this study pairs three different industry output projection methods with two different occupation-by-industry matrices within the structure of industry/occupation employment projection, as discussed in the previous section. The structure of industry/occupation projection method featured in the BLS guidelines (Franklin, 2007) is given in Table 3.1, annotated for chapter sections.

The following sections will discuss the three different output projection methods, followed by a discussion of the occupation-by-industry matrices, and will conclude with a discussion of industry/occupation projections using the various combinations thereof.

3.2.1 SAM

3.2.1.1 Theory

This section considers the expansion of the Input-Output framework already developed to a Social Accounting Matrix framework. A social accounting matrix (SAM) has two primary objectives: (1) the organization of economic and social information in a specific time period and (2) to provide a basis for the creation of a plausible economic model which can simulate reactions in economic markets of exogenous changes (Pyatt and Round, 1978).

The first objective according to King (Pyatt and Round, 1978) is the organization of data. The anatomy of a SAM is most easily presented in the form of an example. Consider a simple theoretical economy whose SAM is presented in the table below. First consider the accounts: (1) production, (2) factors of productions, (3) institutions, (4) the rest of the world (ROW), and (5) regional balances. This presentation and discussion is a paraphrasing of discussions written by King and featured in Pyatt and Round (1978) and Miller and Blair (1985).

Table 3.1: Example SAM

			Expenditures					
			Production	**Factors**	**Institutions**	**ROW**	**Regional Balances**	**Total**
Receipts	**1**	**Production**			1,000	250		1,250
	2	**Factor of Production**	1,025					1,025
	3	**Institutions**	100	1,000		25		1,125
	4	**ROW**	125	25	450			600
	5	**Regional Balances**			(325)	325		0
	6	**Total**	1,250	1,025	1,125	600	0	

Consider the five accounts. Production requires factors to produce value-added that is sold in domestic and foreign markets. Factors of production consist primarily of capital and

labor that are used in, and receive income from, the production process. The factors are drawn from institutions such as households, private and public firms, and government institutions who supply labor and capital and are represented in the second account. The ROW account represents the transactions with accounts outside the region, and the final account, "regional balance," is a direct result of these transactions. Transactions with the rest of the world rarely balance and the difference must be made up through lending (Pyatt and Round, 1978) which is represented in this final account.

In this example there are, excluding the border cells, 11 cell entries. A brief cell-by-cell analysis begins by considering row (1). The production process receives sale of product of $1,000 from domestic institutional demand and $250 from demand from foreign markets. Notice that these payments are offset exactly by the expenditures of production in column (1), which is comprised of the value-added by factors of production ($1025), institutional payments like profits and taxes ($100), and payments for imported materials ($125) (Pyatt and Round, 1978).

Looking at the factors of production account, consider row (2). The $1,025 payment from production is the only payment to factors of production. In column (2) income is distributed among institutions such as households and taxes ($1000) and to ROW ($25) for payments to foreign factors (Pyatt and Round, 1978).

With production and factors of production entries covered, consider the institutions account in row (3). Household income from factors is received in the second column ($1000), direct tax payments from production are in the first column ($100), and transfers from the

ROW account is in the fifth column ($25). Column (3) shows how institutional income is distributed among consumption of domestic final goods ($1000) and imported final goods ($450); the remaining payment entry is a negative payment to the regional balance (-$325) that is the result of the imbalance of domestic income and domestic consumption, which in this case produces a deficit (one might also think of this sector as investment or as a regional account). The final unmentioned cell is the ROW surplus to regional balances ($325). This also is a result of the regional trade; if the domestic region runs a $325 deficit with the ROW, then the ROW account must be running a $325 surplus with the domestic region which is why the surplus in the ROW account exactly balances the deficit run by the domestic region in the regional balances account (Pyatt and Round, 1978).

This completes discussion of data organization. Pyatt and Round (1978) assert that this organization should be used a basis upon which to build a plausible model. The SAM structure in and of itself does not produce any industrial output estimates, much less projections thereof. However, this framework will be the basis upon which a model to produce industrial output estimates and projections is built.

Since the purpose of this process is to provide industrial output estimates, the rest of the discussion will focus on the disaggregated production sector of the SAM (also called the transactions table or Input-Output Table). This process is valid for all of the accounts included in the SAM, but the terminology, interpretation, and notation can get vague when speaking in generalities. To be clear, the inclusion of the other sectors of the economy in the SAM provides access to all of the linkages in the economy and is vital to include in the model from the

beginning, but interest from this point forward is to extract industry output estimates from the SAM which already account for the linkages; thus, I restrict this discussion to the production account. Regardless, it is important to note that this process is valid for all sectors.

The production sector, though presented as a single row-column combination in the example above, is comprised of a row-column combination for each industry (i) in the model. The $i \times i$ matrix is inserted into the top left cell of the example SAM and is often referred to as the make matrix.

Consider an economy with n industries each producing output of x_i. Each industry i requires $a_{i,j}$ units from industry j. Industries sell their production to other sectors (as intermediate goods) as well as to consumer via final demand. Thus output equals the sum of intermediate and final demands:

Equation (3.1)

$$x_i = [a_{1,i}x_1 + a_{2,i}x_2 + \cdots + a_{n,i}x_n] + d_i$$

where d_i is final demand and $a_{1,i}x_1 + a_{2,i}x_2 + \cdots + a_{n,i}x_n$ is the summation of intermediate demand to all other industries from industry i. This calculation may also be represented in matrix form as

Equation (3.2)

$$X = AX + D$$

where X is the vector of total outputs, A is the matrix of $a_{i,j}$ coefficients, and D is the vector of final demands. This notation becomes useful in the discussion of how these output measures are calibrated and projected which occurs in future sections.

This process of calibrating and projecting the output calculated via the SAM can be used in lieu of econometric equations. However, once output projections are obtained the process in which those estimates are converted into industry/occupation forecasts is the same in this paper as in standard state agency practice (using staffing patterns) (Franklin, 2007).

3.2.1.2 Assumptions and Shortcomings of SAM:

Though SAM models are general equilibrium models, they are not without drawbacks. Further, because a SAM is used in the construction of any CGE model, some of the faults in the model are carried over. Drawbacks in this type of modeling typically either stem from assumptions or from scope issues.

Input-Output and Social Accounting Matrix models assume wage fixity, yet Tokle and Huffman (1991) claim that regional wage disparities exist for a variety of reasons. Along similar lines, SAM analysis assumes that prices are fixed, but there are numerous articles suggesting this is not entirely true. Jedidi et. al. (2003) suggest that fixed-price models produce "biased demand estimates" when predicting product line pricings. If one expands this idea to the labor market, biased demand estimates equate to inaccurate work force projections, which bias the policy function.

On a theoretical level, it is assumed the prices move proportionally. Nicholson (2005) claims that this only holds because of zero-transaction cost and perfect information assumptions, which are unrealistic in most if not all regional economic settings. Thus, if the underlying assumptions to price fixity are unrealistic, then fixed-price models may result in biased estimates even in the event they are specified correctly.

The reality of fixed wages comes into question when considering alternative functional forms like Cobb-Douglas (Chand and Kaul, 1986). Flexible functional forms imply that labor and capital are substitutable and each is paid equal to the value of its marginal product. This substitutability suggests that wages may differ in regions where transportation costs, capital production, or technology levels differ. Thus, correctly specified models should account for wage differences across regions.

A further issue with Leontief style models is that they "hinge on the crucial assumption that sectoral production is completely demand driven." More specifically, the problem is that Input-Output type models ignore capacity constraints and assume that any demand will be met instantaneously by production, which makes them more useful as "guidelines to potential induced linkage effects, and as indicators of likely bottlenecks," but less useful as predictive models (Sadoulet and de Janvry, 1995).

It is important during times of structural change that economists and policymakers evaluate policies and the tools used to craft them. When looking forward, models that contain restrictive, unrealistic assumptions may not be able to handle impending structural changes. During such shifts in economic organization, a projection methodology with such restrictions

may in fact inhibit policy efficacy. That is, the use of such models lowers the quality of projections, a primary input in the policy function, thus decreasing policies' effectiveness.

The adoption of national models to the regional setting is also problematic. Regional models "differ from their national counterparts in several aspects. Most of these differences stem from the fact that regions are relatively more open economies compared to nations. Because of this regional openness, commodity trade and resource migration are more important" (Vargas et al, 1999). This statement also speaks to the rigid assumptions above and their inability to cope with structural change and wage or price disparities.

3.1.1.3 Data

In order to construct the SAMs for this study, data on industrial output for the state of Louisiana was purchased from Impact Analysis for Planning (IMPLAN). Due to financial restrictions only data for 2001, 2002, 2006, 2007, and 2008 were available. However, the data from 1998 and 1999 used industry classification codes that require 3 different sets of disaggregation and reaggregation in order to be compatible with the post-2000 data in terms of NAICS codes. These bridge tables are large, cumbersome, and contain a significant amount of overlap at the two-digit NAICS level; therefore data from 1998 and 1999 are excluded.

The data from 2001, 2002 and 2006 also required a bridge table in order to merge the industry classification codes with the data from 2007 and 2008. However there is very little overlap at the two-digit level and the crosswalk is far less extensive. Thus there will be five SAMs constructed in this dissertation, one for each post-2000 IMPLAN data set: 2001, 2002, 2006, 2007, and 2008.

Data are organized into the five SAM structures using the IMPLAN software. Trade data, or inter-institutional transfer data, from each year is used to track linkages between industries that feed into one another so that these linkages are not ignored when forecasting output by industry. Data about other aspects of the economy (consumption, taxes, etc.) are used in the IMPLAN construction of SAMs to dictate linkages between institutions. However much of the actual mechanics are protected as proprietary information used within the IMPLAN software packages and are thus unavailable.

Further details on relevant calibration, simulation, and projection techniques for industry output are discussed in the section of this chapter devoted specifically to those topics.

3.2.2 Computable General Equilibrium Models

3.2.2.1 Theory

The following specifics are excerpts from the CGE literature that specifically apply to this paper. For expansive CGE models and comprehensive literature reviews, see Partridge and Rickman (1997 and 2007), Vargas et al (1999), Lofgren et al (2002), and/or De Menezes et al (2006).

The theory presented here are sections pertinent to the models that will be constructed for this dissertation following the Standard Computable General Equilibrium (CGE) Models in GAMS produced by the International Food Policy Research Institute (IFPRI) (Lofgren, 2002). The IFPRI model is published with code for a solver named GAMS (General Algebraic Modeling System) which is used for CGE construction in this chapter. The IFPRI model is not entirely

adopted, but rather adjusted in terms of parameters, regionality, and structure in order to adhere to tenable assumptions given the nature of this project. The IFPRI model was designed to be a national model for use in developing countries, so adjustments are needed to make the structure suitable for regional modeling. This section will describe the primary sections of the CGE along with changes that were made and rationale for those changes.

Perhaps the easiest way to conceptualize a Computable General Equilibrium (CGE) is in two stages. The first stage is intuitive while the second is organizational and formulaic. The intuition of the CGE can be summarized in how certain aspects of the economy act and interact within the CGE: (1) production and factor markets; (2) institutions; (3) commodity markets; and (4) macroeconomic balances. These concepts and the specifics thereof are at the core of the model and must be understood in order to grasp the general structures of the CGE. However, the actual formalized model of equations is not so neatly organized; for example, a single equation that addresses institutional expenditures on foreign commodities could easily fall into any of the four categories listed above. For this reason, the equations are divided up into slightly different categories that make their presentation more coherent for the reader: (A) price equations; (B) production and trade equations; (C) institutional equations; and (D) system constraint equations. With this designation, the institutional expenditures on foreign commodities equation would squarely fall under the trade equations.

Begin with how the CGE structure in this dissertation deals with the four intuitive concepts. Production in the CGE is driven by the neoclassical specifications. Each producer can make several commodities within the production function, but the proportion of the

commodities produced by each firm's production process is fixed. To produce these commodities, firms demand factors until their marginal revenue equals their marginal cost; thus factor prices may differ across industries and commodities (Partridge and Rickman, 1998). The quantity of each factor is taken from the observed data and an "economy-wide wage variable is free to vary to assure that the sum of demands from all activities equals the quantity supplied" (Lofgren et al, 2002), which closes the factor markets.

Technology is specified as a Leontief function of value-added and aggregate intermediate input functions. The value-added is described by a constant elasticity of substitution (CES) function while aggregate intermediate inputs are a Leontief function of disaggregated intermediate inputs (Lofgren et al., 2002).

Institutions include households, government accounts, enterprises (corporate profits), and the rest of the world (ROW) account. Households receive incomes from the factors they provide to the production process and from other institutions. They distribute that income to direct taxes, savings, transfers to other institutions, and consumption. Household consumption distribution over commodities is driven by a linear expenditure (LES) demand function which comes from utility maximization. Commodity prices adjust so that markets clear (Partirdge and Rickman, 1998).

The government collects taxes and institutional transfers and distributes that income over consumption and transfers to other institutions. Governmental consumption is held as exogenous (Lofgren et al, 2002).

Instead of being paid directly to households factor payments may be paid to enterprises, which may also receive other institutional transfers. Enterprises distribute their income over direct taxes, savings, and institutional transfers. Enterprises do not consume; however "apart from this, the payments to and from enterprises are modeled in the same way as the payments to and from households" (Lofgren et al, 2002).

The ROW account receives and makes payments to other institutions and factors, but those transactions are fixed in foreign currency. The ROW account receives transfers from domestic institutions and foreign factors in domestic use while spending that income on institutional transfers within the region, domestic factors, and consumption of domestic goods. The difference between spending and receipts is foreign savings. Trade and closure of markets concerning the ROW account are covered in the following discussion of commodity markets and macroeconomic closures (Partridge and Rickman, 1998).

Commodity market discussion begins with aggregating commodity output from all producers using a CES function. This aggregated commodity output is then distributed between exports and domestic sales using a Constant Elasticity of Transformation (CET) function. At this point, domestic supplies are joined with imports to form a composite good. This composite good supplies household and government consumption as well as investment and intermediate markets. Commodity prices fluctuate to clear markets (Lofgren et al, 2002).

There are two particular assumptions used here that warrant discussion. "The assumptions of imperfect transformability (between exports and domestic sales of domestic output) and imperfect substitutability (between imports and domestically sold domestic

output) permit the model to better reflect empirical realities" instead of forcing the alternative assumptions of perfect substitutability and perfect transformation (Lofgren et al, 2002).

The CGE model includes three primary macroeconomic balances, or closures, pertaining to government balances, foreign markets, and savings-investment. The government balance is the difference between revenues and expenditures from governmental institutions and is treated as a flexible residual. In other words, government consumption and tax rates are fixed while the government savings or deficit fluctuates to balance the accounts (Lofgren et al, 2002). International models may use flexible tax income models, but those models are inappropriate for regional use within the U.S.

Though foreign markets are of relatively little importance to a study of Louisiana, in order to close the system macroeconomic balances are required (Lofgren et. al., 2002), and thus are presented here. The foreign market balances are held in foreign currencies and the real exchange rate is allowed to fluctuate while foreign savings is fixed. Further, all transactions between domestic institutions and ROW are fixed via the SAM which fixes the trade balance.

The investment-savings balance is probably the most simplistic closure intuitively. Balances are investment driven in that real investment is fixed and savings rates adjust to create savings that equal investment. It is assumed that governmental policy can influence savings rates enough to balance the fixed investment (Vargas et. al, 1999).

There are alternative CGE macroeconomic balance closure methods mentioned in Lofgren et al. (2002), Partridge and Rickman (1998), and Vargas et al (1999), but these methods were most commonly associated with the U.S. and world markets in all three papers. For a

reader needing specifics on alternative closure methods, Lofgren et al. (2002) is the most useful reference.

This concludes the discussion of how the CGE intuitively models primary aspects of the economy. However, a slightly different organization will improve readability when discussing the CGE structure in terms of equations. These equations can be categorized into price equations, production and trade equations, institutional equations, and system constraint equations. The formal presentation of the equations can be found in Appendix A, but a brief discussion of those equations is presented here.

The price block of equations contains equations for export and import prices, market output valuation, domestic price vectors, absorption, aggregate intermediate input prices, consumer and producer price indices, demand prices of domestic non-traded goods, activity prices, and producer revenue and cost equations. These equations link endogenous prices to one another as well as linking prices of other non-price variables, both endogenous and exogenous (Lofgren et al., 2002). This block is particularly rigorous "primarily because of the assumed quality differences among commodities of different origins and destinations" (Lofgren et al., 2002).

The production and trade equations cover domestic production, allocation of output to domestic and foreign markets (including consumption), aggregate supply to domestic markets, and demand for trade inputs. This equation block includes production functions, value-added-intermediate-input ratios, aggregated and disaggregated demand for value-added, aggregated and disaggregated demand for intermediate inputs, factor demand, commodity production and

allocation, output aggregation and first order conditions, output transformation equations, export-domestic supply ratios, composite supply (Armington) functions, import-domestic demand ratios, and transaction services demand equations (Lofgren et al., 2002).

Equations for the institution block deal with factor incomes, institutional incomes, infra-institutional transfers, household consumption expenditures, investment demand, government consumption, government revenue, and government expenditures. These equations dictate how institutions respond to one another within the framework of the CGE: how income is gained, how transfers are made, and how institutions distribute their income in consumption (Lofgren et al., 2002).

The final block of equations governs system closures and constraints. Factor market closure, composite commodity market closure, foreign market closure, government balances, direct institutional tax rates, institutional savings rates, the savings-investment balance, total absorption, investment-absorption ratios, and government consumption-absorption ratios equations are all in this block of equations that force markets to close, providing the impetus for prices to adjust and equilibrating the model as a whole.

These blocks of equations are presented in detail in Appendix A along with a visual representation and discussion of the CGE equation structure.

3.2.2.2 Data

CGE specifications begin with the creation of the five SAMs, which are the same structures from which the SAM estimates are based. These five SAMs are the primary

ingredients for the construction of a CGE. The structural equations were covered in the theoretical model section. The additional specifications include the exogenous variable list: consumer price indices and elasticites of several varieties. Tax rates are endogenous and calculated from the payments to the tax sectors within the SAM (which are then paid in full to the government accounts). The consumer price index numbers were taken from BEA.gov.

The approach taken to parameterization of the CGE models in this dissertation was driven by the overall instrumental view of this policy question. Friedman asserts "The only relevant test of the validity of a hypothesis is comparison of its predictions with experience" (Housman, 2008). Being that many policymakers are not economists and may not have professional economists on staff, it is unlikely that policymakers are using the up-to-date, "best practices" of parameterization when constructing economic models for use in policy research. It is far more likely that models will be parameterized by choosing among preset parameterizations recommended by economists for general situations (international, national, regional, state, etc.) or by statistical economic software packages.

Thus, in this dissertation, CGE models are parameterized via suggestions from the GAMS user guide that accompanies the IFRPI code (Lofgren et al, 2002) and is verified to some degree using other sources (Partridge and Rickman (1997, 2007), Vargas et al (1999), the readme files and parameter sets that accompany the Lofgren et al (2002) chapters on CGE specification). These parameters are identified as appropriate for regional use in the U.S. by Lofgren et. al. (2002), which is a standard regional CGE model (Partridge and Rickman, 2007).

Elasticities are the primary parameters of CGE models. The elasticity of substitution between imports and domestic output in domestic demand and the elasticity of transformation for domestic marketed output between exports and domestic supplies are both set to equal 1.5. Lofgren et al (2002) suggest this as an appropriate number for regional (sub-national) use within the U.S. There are also several production elasticities recommended by Lofgren et. al. (2002): elasticities between factors tend to be around 0.8 for U.S. industry; the output aggregation elasticity for commodity is set to four; and the elasticity of market demand for commodity by household is equal to 1. All of these measurements come from the IFPRI model discussion in Lofgren et al (2002) and from the GAMS code and "readme" files which accompanies the Lofgren publication. In those files are suggested levels of elasticities for different simulation scenarios. These parameters have been cited and discussed by many previous regional CGE creators and that information and discussion was taken into account during parameterization of these CGE's (Shoven and Whalley (1992), Watson et al (2012), and Robinson, Kilkenny, and Hanson (1990)).

3.2.3 Calibration/Simulation for SAM and CGE Comparison

3.2.3.1 SAM

The SAM models for this paper are calibrated according to SAM simulation processes from Lofgren et al (2002). To calibrate the model to the actual data, real GSP numbers are used to repopulate the SAM; this calibrates the model to fit the actual data rather than the data estimated by IMPLAN while maintaining the economic structure and linkages given by the observed data.

To do this, after the SAM-based model arrives at estimates of industry output, set the ratio of industry output to the industry's gross state product (GSP) equal to the observed (or actual) ratio:

Equation (3.3)

$$\frac{x^t_{i,SAM(b)}}{GSP^t_{SAM(b)}} = \frac{\hat{x}^t_{i,SAM(b)}}{\overline{GSP}^t}$$

where $x^t_{i,SAM(b)}$ is output for industry i in time period t from the SAM built on data from base year b, $GSP^t_{SAM(b)}$ is the gross state product for time period t from the SAM built with data from base year b, a hat ($\widehat{\ }$) over any variable means that it is a calibrated estimate, and a bar ($\overline{\ }$) over any variable means that it is actual data (the bar notation makes the subscript indicating model type irrelevant). Thus, $\hat{x}^t_{i,SAM(b)}$ is the calibrated estimate of industry output for industry i in time period t from the SAM built on data from base year b, and $\overline{GSP}^t$ is the actual gross state product for industry i from time period t.

Using the actual GSP data with the SAM estimates of industry output and GSP, rearrange equation 3.3 to get a calibrated industry output estimate:

Equation (3.4)

$$\hat{x}^t_{i,SAM(b)} = \overline{GSP}^t * \frac{x^t_{i,SAM(b)}}{GSP^t_{SAM(b)}}$$

A process similar to calibrating the SAM estimates can be used to project the economic anatomy of the SAM to future estimates of GSP. That is, changing the assumption to

Equation (3.5)

$$\frac{x_{i,SAM(b)}^{t}}{GSP_{SAM(b)}^{t}} = \frac{\hat{x}_{i,SAM(b)}^{t'}}{\overline{GSP}^{t'}}$$

or

Equation (3.6)

$$\hat{x}_{i,SAM(b)}^{t'} = \overline{GSP}^{t'} * \frac{x_{i,SAM(b)}^{t}}{GSP_{SAM(b)}^{t}}$$

would project the anatomy of the SAM from base year b to industry i's gross state product from some future time period, year t', yielding calibrated output for industry i in time period t' based on the SAM built on data from base year b. This effectively projects the GSP from projection year t' using the structure of the SAM from base year b in order to arrive at projections of future industry outputs.

3.2.3.2 CGE

The CGE model is calibrated using a process that is fundamentally identical to the calibration process for the SAM models, if different in application. In the same way the industry-output-to-GSP ratio was held constant and used to repopulate the SAM for different levels of GSP, the CGE equation parameters are held constant and the GSP data are used to simulate how these equations would distribute production given the structure of the CGE.

In addition to the CGE code itself, simulation code allows the user to shock the economic system from the demand side while holding the relationships, variables, and parameters from the base model constant. Here, insert actual gross state product, $\overline{GSP}^{t}$, and

the simulation process distributes that GSP over the economy according to the relationships, parameters, and solutions found in solving the base CGE model yielding estimates of $\hat{x}^t_{i,CGE(b)}$: the calibrated output for industry i for time period t estimated by the CGE built on data from base year b.

This exact mathematical process can be used to project industry output for future time periods according to the structures of the economy discovered through the CGE built on base year b. That is, by running the simulation program and shocking the CGE equations with $\overline{GSP}^{t'}$ rather than $\overline{GSP}^{t}$, one can arrive at the calibrated output for industry i for some future time period t' estimated by the CGE base on data from base year b.

The CGE projection process, as with the process for SAM-based models, distributes GDP according to the anatomy of the economy estimated by its base model. However, there are differences in the results of the projections processes worthy of discussion. The SAM-based models project linearly. That is, the relationship between industry output and GSP is held constant for each base model-industry pairing (only scaled by the constant ratio multiplier $\frac{x^t_{i,SAM(b)}}{GSP^t_{SAM(b)}}$), and each model will therefore produce industry output projections that are linear transformations of the base-year data.

In contrast, the CGE simulation process will adjust its output estimates in a non-linear manner by allowing the exogenous variables and parameters in the CGE to re-equilibrate after the GSP shock (Lofgren et al, 2002). This results in the ability of CGE models to adjust to shifts in GSP in terms of capital/labor ratios, input mixes, labor migration, and price changes instead of

linearly scaling industrial output to GDP. This approach allows for a more neoclassical reaction to changes in GSP within the CGE structure than is capable in the SAM structure (Lofgren et al, 2002).

3.2.4 Time Series

3.2.4.1 Theory

The time series models use the actual data on industry output from 1993-2010. The data are aggregated by 2-digit NAICS to fit with the scope of the model, and then imported into STATA. The data are then declared in the software as yearly time series of data with observations of actual output by industry for years from 1993 up to the base year for each series ($\bar{x}_i^t$, for $1993 \leq t \leq b$). There are a total of five different time series. The models are constructed for $b =$ 2001, 2002, 2006, 2007, and 2008 in order to match the base years available for SAM and CGE model construction for comparison.

Each of the five time series models are considered individually. Since time series models contain the actual data for any years prior to their calibration year, backcasting can get statistically complex (as well as interpretively unreliable) in terms of information and bias so the focus is on the ability of time series to project forward only (Judge et al., 1988).

When considering the form of the structural equations for each time series, start with the general autoregressive integrated moving average (ARIMA) model as a template, which regresses the dependent variable on a series of lagged dependent variables and moving average calculations (Judge et al., 1988). When considering the general model, these three

parameters are needed to specify the ARMA model equations: the integrated difference order (d), the Moving-average order (q), and the Autoregressive order (p) (Judge et al., 1988). The equations and variable structures differ for each data series according to data trends and parameter values, thus an ARIMA model is generally identified as ARIMA(p,d,q) with a different equation structure for each parameter combination. The equations for each combination used in this dissertation will be presented fully after the parameters and their values are discussed individually.

The autoregressive order (p) describes the number of lags that are statistically significant in predicting industry output in any particular year.

The integrated difference order (d) determines the difference degree needed to assure trendless observations in the regression analysis.

The moving-average order (q) determines the number of lags that should be included in a moving-average calculation. The moving-average is then used as an independent variable in regression analysis.

The decision-making process for specifying each time series model considers each of the three primary parameters in turn. That is, each industry's data series is evaluated independently to allow each industry to be specified according to its own data and trends.

The integrated difference order (d) is approached first. Time series modeling requires that data have a white noise error term, or lack any significant consistent trends. The process of determining the integrated difference order (d) begins with a simple line plot of the data. If the

data appear to have a white-noise error the data are difference-appropriate for time series use and the independent variables will take the form of X_i (Judge et al., 1988). That is, the regression equations will take the general form

Equation (3.7)

$$\hat{x}_i^t = \beta_0 + \beta_1 \bar{x}_i^{t-1} + \varepsilon$$

However, if there is a trend in the data, the first adjustment considered is to use the natural logarithm of the observations. If the data exhibit an exponential trend then often the natural logarithm of the variable will exhibit white-noise characteristics. This was considered for many models, but never achieved the desired result and therefore was not used, which was expected because of the low likelihood that output for any industry will expand exponentially.

The next adjustment in search for the appropriate integrated difference order (d) was to plot first order differences ($\hat{x}_i^t - \bar{x}_i^{t-1}$) rather than the observations themselves in hopes that the trend might disappear. If the line plot of the differences exhibited no trend where the original data had provided one, the difference order parameter was set to one (d=1), and the data were considered adequate for the ARMA model following

Equation (3.8)

$$[\hat{x}_i^t - \bar{x}_i^{t-1}] = \beta_0 + \beta_1 [\bar{x}_i^{t-1} - \bar{x}_i^{t-2}] + \varepsilon$$

or, rearranging to

Equation (3.9)

$$\hat{x}_i^t = \beta_0 + \beta_1[\bar{x}_i^{t-1} - \bar{x}_i^{t-2}] + \bar{x}_i^{t-1} + \varepsilon$$

In this case, knowing $\bar{x}_i^{t-1}$ and $\bar{x}_i^{t-2}$ (because they are actual data) allows one to solve easily for $\hat{x}_i^t$ given the β's from the regression equations. If the first order difference also exhibited a trend, then the second order difference (the difference of the differences) was considered. If this series was without trend, then the difference operator was set to two (d=2) and the left hand side of the equation would have taken the form $\left[[\hat{x}_i^t - \bar{x}_i^{t-1}] - [\bar{x}_i^{t-1} - \bar{x}_i^{t-2}]\right]$ and the dependent variables in the regression would also be the second order difference (Judge et. al., 1988). A more complex structure is needed to solve for $\hat{x}_i^t$, however none of the models required a difference operator larger than one so this method was not required for any models in this dissertation. For the detailed equations based on varying integrated difference operator (d) values see the table 3.2.

The reason the integrated difference order is considered first is that if it is determined that the difference order is greater than zero then each of the other parameters must be determined by taking the appropriate differences of the variable rather than the original variable series.

The moving-average order (q) is approached next and begins with looking at the correlogram. The correlogram indicates how many lags are correlated with any particular observation. A t-test of each autocorrelation indicates if the lag's correlation with the observation is significantly larger than zero. None of the time series data that were used for this study exhibited more than one statistically significant lagged term. Further, because moving averages require three consecutive significant terms to calculate, all models have a moving-

average order of zero (Judge et. al., 1988) (q=0 for all times series models in this paper, in which case the general ARMA structure collapses to a simple Autoregressive (AR) Model) .

The autoregressive order (p) is addressed last. The partial correlogram shows the correlation between an observation and any other lagged observation of the same variable (not just the immediate lag) having adjusted for the linear relationship between the observation and its' lags discovered through the correlogram (Judge et. al., 1988). This parameter often allows for cyclical adjustments when the moving-average order is greater than zero, but since q=0 for all models, the partial correlogram is appropriate for determining which lags should be included in the ARMA model (Judge et. al., 1988).

The partial correlogram coefficients for each lagged term are t-tested to determine if they are statistically different from zero. In all cases the first autoregressive term is significant, thus all models have p≥1, so the term $\bar{x}_i^{t-1}$ is used as an independent variable for all regressions of $\hat{x}_i^t$ (or in the case where $d = 1$ and , $[\bar{x}_i^{t-1} - \bar{x}_i^{t-2}]$ would be used as an independent variable in the regressions with $[\hat{x}_i^t - \bar{x}_i^{t-1}]$ as the dependent variable). In some cases the second lagged term also tested as significant, in which case p=2 and $\bar{x}_i^{t-2}$ is included in addition to $\bar{x}_i^{t-1}$ as independent variables for the regression of $\hat{x}_i^t$ (or in the case where $d = 1$ and , $[\bar{x}_i^{t-2} - \bar{x}_i^{t-3}]$ would be used along with $[\bar{x}_i^{t-1} - \bar{x}_i^{t-2}]$ as an independent variables in the regressions of $[\hat{x}_i^t - \bar{x}_i^{t-1}]$). None of the third lags were ever significant, so the autoregressive order (p) has a maximum value of two in this dissertation.

The structural equations for the time series models for each parameter combination used in this dissertation are listed below.

Table 3.2: Time Series Equations

Specifications	Equations
D=0 P=1	$\hat{x}_i^t = \beta_0 + \beta_1 \bar{x}_i^{t-1} + \varepsilon$
D=0 P=2	$\hat{x}_i^t = \beta_0 + \beta_1 \bar{x}_i^{t-1} + \beta_2 \bar{x}_i^{t-2} + \varepsilon$
D=1 P=1	$\hat{x}_i^t = \beta_0 + \beta_1 \left[\bar{x}_i^{t-1} - \bar{x}_i^{t-2}\right] + \bar{x}_i^{t-1} + \varepsilon$
D=1 P=2	$\hat{x}_i^t = \beta_0 + \beta_1 \left[\bar{x}_i^{t-1} - \bar{x}_i^{t-2}\right] + \beta_2 \left[\bar{x}_i^{t-2} - \bar{x}_i^{t-3}\right] + \bar{x}_i^{t-1} + \varepsilon$

The ARMA models will produce coefficients ($\beta's$) for the specified terms in each model. The β's from the regressions provide a basis for which the time series model can be extrapolated or projected to future time periods. The error terms are assumed to take a normal or white-noise distribution and $\hat{x}_i^t$ can then be easily solved. Then, $\hat{x}_i^{t+1}$ can be found using that estimate of $\bar{x}_i^t$ and the β's from the structural equation. Further, $\bar{x}_i^{t+2}$ using the estimate of $\bar{x}_i^{t+1}$, and so on. This iteration allows us to extrapolate the trends found in the time series data to the desired projection year.

3.2.4.2 Data

The Time Series model requires industry level output as compared to GSP data for the SAM and CGE. The data used are acquired from Moody's Analytics and contains industry level output from 1993-2010 and projected industry level output from 2011-2023. These data are aggregated to the 2-digit NAICS codes to match the industry detail of this project. These data are all that is required, which will later be argued as a distinct advantage, both structurally and financially, of this technique.

3.2.4.3 Specification

Each time series model starts with annual industrial output data from 1993 up through the base year. That is, the 2001 calibrated model uses the data from 1993-2001 and the 2008 calibrated model uses the 1993-2008. Each of the five models is specified according to the decision making process laid out in the previous discussion of the theoretical models. For each of the five models, the two primary defining parameters for the ARMA model are presented in Table 3.4.

Table 3.3: Time Series Specification

	2001	2002	2006	2007	2008
Agriculture/ Forestry	d = 0 p = 1	d = 0 p = 1	d = 0 p = 1	d = 0 p = 1	d = 0 p = 2
Prob > Chi2	0.68	0.62	0.49	0.28	0.21
Mining	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1
Prob > Chi2	0.86	0.708	0.807	0.806	0.78
Utilities	d = 0 p = 1	d = 0 p = 1	d = 0 p = 1	d = 0 p = 1	d = 0 p = 1
Prob > Chi2	0.809	0.99	0.95	0.98	0.92
Construction	d = 1 p = 2	d = 1 p = 2	d = 1 p = 1	d = 1 p = 2	d = 1 p = 1
Prob > Chi2	0.52	0.48	0.38	0.31	0.26
Manufacturing	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1	d = 0 p = 1	d = 0 p = 1
Prob > Chi2	0.489	0.6049	0.61	0.38	0.198
Wholesale	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1
Prob > Chi2	0.94	0.96	0.93	0.95	0.78
Transport/Warehousing	d = 1 p = 2	d = 1 p = 2	d = 1 p = 2	d = 1 p = 2	d = 1 p = 2
Prob > Chi2	0.93	0.61	0.401	0.299	0.15
Retail Services	d = 0 p = 1	d = 0 p = 1	d = 0 p = 1	d = 0 p = 1	d = 0 p = 1
Prob > Chi2	0.52	0.56	0.75	0.92	0.75
Professional Services	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1
Prob > Chi2	0.97	0.77	0.504	0.33	0.15
Education/Health Services	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1	d = 1 p = 1
Prob > Chi2	0.708	0.85	0.96	0.85	0.84
Government/Non-NAICS	d = 0 p = 2	d = 0 p = 2	d = 0 p = 2	d = 0 p = 2	d = 0 p = 2
Prob > Chi2	0.78	0.83	0.47	0.41	0.29

The data show remarkable consistency over time with only the Agriculture/Forestry, Construction, and Manufacturing industries showing any specification change over the models. The Agriculture/Forestry, Retails Services, and Government/Non-NAICS sectors were the only industries for which the source data did not require a difference operator.

The transportation/warehousing and government/non-NAICS sectors are the only two industries that exhibited an autoregressive order of two for each year, while construction has p=2 for three of the five models and agriculture/forestry exhibited p=2 for only the 2008 model. Expectations might have been to have more lagged observations be significant more often, but with the exclusion of significance of any moving-average coefficients it could be that the first lag is explaining a large amount of the data trend and the second lag is consequently is not adding much information that is not inherited by the inclusion of the first lag.

The Chi-squared statistics are generally good. There are several aspects that warrant discussion however. First, agriculture/forestry, construction, manufacturing, transportation/warehousing, education/health services, and government/non-NAICS see significant drops in the chi-squared statistic as the models incorporate more data. This would indicate that as the data from 2006-2008 are added to the time series models, the models do not fit the actual data as well. When considering the time period of the data, economic intuition might lead one to expect this result. The data from 1993 through 2005 show relatively smooth growth as exhibited by the need for difference operators on most of the industrial models. However in 2006 the Louisiana economy started showing effects of a slowing economy. That is, data starting in 2006 began to deviate from the nice smooth economic growth path set out by

the precious 13 years of data. This deviation makes the linear time series estimates adjust downward, or away from previous trends, thus the overall performance of the model decreases as sum-squared-errors rise.

3.3 Occupation-by-Industry Matrices

Bureau of Labor Statistics (BLS) projection method guidelines recommend using a static, fixed occupation-by-industry matrix: $OCC_G^{b'}$,where $G\epsilon(nat, state)$ is geographic distinction between state and national employment data, and b' designates the base year from which the employment data were observed. These matrices are produced by BLS and available for national trends for various years (Franklin, 2007). This matrix is $n \times m$, where n and m are the number of industries and occupations in the model. The industries are aggregated to two-digit NAICS codes, as mentioned earlier, and occupations are aggregated to two-digit Standard Occupational Classification (SOC) codes.

The BLS guidelines recommend directly multiplying this matrix by the vector of industry outputs to estimate the number of workers in each occupation within each industry needed to meet production levels, or

Equation (3.10)

$$X_M^t \cdot OCC_G^{b'} = E_{M,G}^{b',t}$$

where X_M^t is the vector of industry outputs for time period t estimated by model $M\epsilon[SAM(b), CGE(b), TS(b)]$, $OCC_G^{b'}$ is the industry-by-occupation matrix of geography G built

on employment data from year b', and $E_{M,G}^{b',t}$ is the matrix of employment for time period t arrived at using model M, geography G, employment data from time period b'.

This method implicitly assumes that industrial employment is tied directly to current industrial output (Franklin, 2007). However, many characteristics of an economy could contribute to the level of employment including but in no way limited to levels of human capital (Simon, 1998), commuting and transportation costs (Hendrickson, 1986), innovation (Brouwer et al., 1993), past employment (Boarnet, 2005), and past production (Carlino and Mills, 1987). Regardless, the BLS guidelines issued to state agencies recommend this method for creating occupation/industry estimates, and this dissertation will follow those guidelines seeking to determine if the geographic nature of the employment data make a significant difference when constructing occupation-by-industry estimates. The structural application of the employment data is kept constant in order to tease out the effects of changing employment geography.

Louisiana state development agencies primarily use employment by occupation and industry numbers for the U.S. which are adjusted to represent state trends and expectations; however in some cases direct national averages are used without adjustments. Based on a review of the literature, no state agency has been found to construct a state-specific matrix. Although there is documentation of different approaches, no statewide or agency protocols have been found to indicate under which circumstances national patterns are sufficient and which require adjustment (Louisiana Workforce Commission, 2010).

This approach has some distinct advantages, most of them in application rather than in theory. These national matrices are easy to find and are low-cost. Once output projections by

industry are provided, the application of the occupation-by-industry matrix calculations is mechanically simple. This, combined with the a simple industry output projection model, makes it simple for policymakers and state development agencies to quickly, cheaply, and easily follow a reasonable method when producing industry/occupation projections (BLS.gov).

3.3.1 State Matrix

The above approach may be reasonable, but maintains a problematic issue: regional or state staffing patterns do not always adhere to national trends. Forcing national patterns on a regional market may bias estimates. Vargas et al. (1999) suggest that regional investment shifts and a lack of regional control over national monetary policy may make certain regions more or less volatile in terms of investment and production, and therefore, employment. Further, Partridge and Rickman (1998) argue that regional economic reactions to exogenous changes may not be proportional to changes at the national level "depending upon the various elasticities of supply and demand." Ultimately, this dissertation hopes to determine if the suggested differences between national and regional employment data drive significant differences in industry/occupation projections for the state of Louisiana in the early years of the 2000's.

Using staffing patterns specifically from the region seems to be a much more economically sound principle. However, no such matrix exists for the state of Louisiana, so one is created. By aggregating staffing patterns by industry and by occupation, for the state of Louisiana, the first state occupation-by-industry matrix is constructed for use in occupational growth projections. Thus, by using state-specific data to construct an occupation-by-industry

matrix this paper hopes to improve occupational growth estimates and thus improve state-wide policy functions.

The construction of the industry/occupation matrix begins with data on employment-by-occupation for each industry. Each of the 11 industries (listed in the table below) has

Table 3.4: Industry and Occupation Categories

Industry Categories
Agriculture and Forestry
Mining
Utilities
Construction
Manufacturing
Wholesale Trade
Retail Trade
Transportation and Warehousing
Professional Services
Education and Health Services
Government and Non-NAICS industries

Occupation Categories
Management
Business and Financial Operations
Computer and Mathematical
Architecture and Engineering
Life, Physical, and Social Sciences
Community and Social Services
Legal
Education, Training, and Library
Arts, Design, Entertainment, Sports, and Media
Healthcare Practitioners and Technical Support
Healthcare Support
Protective Services
Food Preparation and Serving
Building, Grounds Cleaning, and Maintenance
Personal Care and Services
Sales and Related Services
Office and Administrative Support
Farming, Fishing, and Forestry
Construction and Extraction
Installation, Maintenance, and Repair
Production
Transportation and Material Moving

employment in each of the 22 occupational categories, resulting in a 22×11 matrix of actual employment by occupation for the state of Louisiana: $OCC_{state}^{b'}$. Entries in the matrix $occ_{i,soc,state}^{b'}$ represent the actual employment needed in occupation soc to create an actual

dollar of output in industry i during time period b', or $occ_{i,soc,state}^{b'} = \frac{\overline{emp}_{i,soc,state}^{b'}}{\bar{x}_{i,state}^{t=b'}}$. [Note that the output on the bottom of the ratio is the state output data; in order to maintain geographic matching within the data, regional output must be paired with regional employment. In the construction of a national matrix, national industrial output by industry ($\bar{x}_{i,nat}^{t=b'}$) is paired with national occupational employment by industry ($\overline{emp}_{i,soc,nat}^{b'}$).]

3.3.2 National Matrix

The national industry employment-by-occupation data comes from BLS staffing patterns. The data are for industrial employment by occupation. In order to create the $OCC_{nat}^{b'}$ matrix, the national industry output for the time period b' is required exogenously (the data come from BLS). Employment by occupation by industry is then divided by the industrial output to arrive at the individual matrix entries: $occ_{i,soc,nat}^{b'} = \frac{\overline{emp}_{i,soc,nat}^{b'}}{\bar{x}_{i,nat}^{t=b'}}$. These entries fill the matrix $OCC_{nat}^{b'}$, and the construction of the national matrix is complete.

3.3.3 Data

The state data are available in some sense, but lack the desired structure. Thus, the data are compiled by industry to the two-digit code. However, some of the industries, for example construction and utilities, use many of the same workers and are therefore combined within the state of Louisiana employment data. This aggregation seems logical as many utilities projects involve construction and vice versa, but it does require a bit of adjustment to the structure.

Though the output numbers are separated by industry, at the point just before using the matrix to designate occupations by industry, the output for certain industries are combined: utilities with construction; agriculture with mining; and retail trade with wholesale trade. These merges are required because state data on employment by category does not aggregate further. To clarify, the industrial output of these industries are projected independently, but in order to avoid some disaggregation challenges among occupations at the state level, aggregation of those output numbers is required before those numbers are converted to employment. Also, because this process is required for the state level data, the aggregation is performed with the national matrix as well in order to maintain consistency throughout the process.

4 Data and Descriptive Statistics

4.1 Introduction

Previous chapters discussed the motivation and methods for the creation of employment forecast estimates. This chapter will discuss the data used for the construction of these estimates, the data to which these estimates will be compared (evaluated against), and the descriptive statistics of these two data sets. The rest of this chapter addresses these data issues for industrial output estimates, occupation-by-industry matrices, and employment estimates, respectively. This chapter will not present any statistical analysis or hypothesis testing; such analyses are in the fifth chapter.

4.2 Louisiana State Output-by-Industry Estimates Data and Summary Statistics

4.2.1 Data for Model Construction

Industrial output estimates for the state of Louisiana were created using the data from each base year as the inputs to the three industrial output estimation techniques (CGE, SAM, and time series). Base model years are years for which data are gathered and used in the construction of each of the alternative methods for projecting industrial output to the chosen projection years, as mentioned in Chapter 3. The series of projection years represents years for which each model, regardless of base year, produces estimates of industrial output for use in the creation of industry/occupation estimates.

The model base years were chosen based largely on data availability. Since each model must be constructed using primary data from each model base year and a SAM is required for

all non-econometric models, the project is therefore limited to model base years for which sufficient data are available to construct a proper SAM for Louisiana. The data are available through the IMPLAN group's proprietary data products line. Due to financial restrictions, however, only data from 1998, 1999, 2001, 2002, 2006, 2007, and 2008 were available for consideration in this dissertation.

The data for 2001 and 2002 required re-aggregation as well, but the process of converting previous versions of NAICS classification to a newer version is significantly easier and more reliable that the conversion from SIC to NAICS codes. For this reason, 2001 and 2002 were included as model base years after the NAICS codes were updated to the 2007 NAICS series. The 2006, 2007, and 2008 data were industrially classified with the 2007 NAICS codes and required no re-aggregation. This completes the model base year series as $b \in (2001, 2002, 2006, 2007, 2008)$.

The projection year series represent the years for which estimates will be produced from each model in this dissertation. The in-sample years were all chosen to begin with. In addition to the five model base years, 2005, 2010, 2015, and 2020 were chosen to create a consistent five-year span series of estimates. The 2009 year was added to complete the 2005-2010 series, and 2023 was also added as it was the last year in which exogenous industrial output data were projected. This leaves the final projection year series as $t \in (2001, 2002, 2005, 2006, 2007, 2008, 2009, 2010, 2015, 2020, 2023)$.

The time series base model years and projection years have, however, a small caveat that causes them to be a subset of the series just discussed. Since actual data are used in their

construction, time series models have problems with endogeneity if used to backcast (Judge et. al., 1988). That is, when testing predictions made using actual historical data against that same historical data, one cannot separate the effects on prediction accuracy of the model and methods from the internal effects of having built those estimates on the knowledge of the actual data to which estimates are being compared (Judge et. al., 1988). For this reason, backcasts are not included in the evaluation of the time series models, and each time series model is only used to project forward (or for years later than the model's base year).

4.2.2 Data for Estimate Comparison

The industry output projections are compared to actual data to determine how accurate they are as an individual component of the larger method. For historical data, data can be re-aggregated from BEA data on state industrial output for comparison to projections using the various models. However, no federal or state agencies produce industrial output projections in a series, with consistent structures and aggregations, for a consistent group of projection years. That is, there is no consistency in foundation and structure over base years or projection years. Thus, a projection series that is documented to predict actual data reasonably well is needed to which alternative projection series can be compared.

Moody's projects annual state industrial output by industry to 2023. The projection methods used are proprietary, but Cochrane's 2011 paper "The Moody's Analytics U.S. State Economic Model System" discusses their process without providing technical details. Cochrane's 2011 paper states that Moody's uses "a system of simultaneous econometric state models that have enhanced simulation properties." These models form an output-based

industry model system that includes shift-share analyses, time series regression analyses, and panel data regression analyses among other econometric techniques (Cochrane, 2011).

Moody's uses data from the Bureau of Labor Statistics (BLS), the Bureau of Economic Analysis (BEA), the Current Employment Survey (CES), the Quarterly Census of Employment and Wages (QCEW), the U.S. Census Bureau, and other national and state agencies to fill primary and historical data requirements while using their proprietary structural equations to calculate additional economic and demographic data as well as to project current data into future time periods (Cochrane, 2011). This means that the Moody's data equals the actual data for years prior to and including 2012, but that Moody's projected data makes up the industrial output data for years past 2012. Thus, this series is suitable for comparison to alternative methods for years which historical data are available, since the Moody's data equals the actual data, and that the Moody's data is a reasonable approximation for all projection years past 2012. Further, since Moody's data series are used, they are aggregated and structured consistently.

Though the data are collected for future time periods (years beyond 2012) and testing will be done using all available estimates and data, much of the results section will focus on comparing estimates from different models against the actual data before for projection years up to 2010, excluding the years for which the Moody's data are projected.

4.2.3 Descriptive Statistics

With three different model types producing forecast estimates for five base years and eight projection years, the descriptive statistics can be quite expansive. While a more complete

table of descriptive statistics can be found in Appendix B, a brief summary of these statistics is presented in Table 4.1. These statistics are for aggregated forecast estimates and provide a general sense of the magnitude of the estimates, more thorough trend analyses are held for the regressions analysis sections later in this chapter.

There are several trends that deserve noting, not least of which is the upward bias present in CGE (an over estimation of 109% on average) and SAM (overestimation of 69% on average) models when compared to the actual output data. Though there is sizeable bias in the CGE and SAM estimates, there is very little upward bias in the time series estimates, 12% on average.

Table 4.1: Industrial Output Forecast Descriptive Statistics by Model and Base Year

Actual 2008 Output (all sectors)	186,836					
	CGE 2001	**CGE 2002**	**CGE 2006**	**CGE 2007**	**CGE 2008**	**Avg**
Total Output	309,726	294,042	420,725	450,655	474,905	390,011
Percent of Actual	166%	157%	225%	241%	254%	209%
	SAM 2001	**SAM 2002**	**SAM 2006**	**SAM 2007**	**SAM 2008**	
Total Output	267,424	225,401	338,108	340,197	403,724	314,971
Percent of Actual	143%	121%	181%	182%	216%	169%
	TS 2001	**TS 2002**	**TS 2006**	**TS 2007**	**TS 2008**	
Total Output	198,869	193,468	224,511	209,877	220,230	209,391
Percent of Actual	106%	104%	120%	112%	118%	112%

This upward bias comes with a tradeoff however. The average standard deviation of estimates from CGE models are the lowest ($2,016,000). Time series ($2,417,000) have a slightly larger standard deviation and SAM models ($4,945,000) have an average standard deviation nearly double CGE and time series models, implying that estimates from CGE models are relatively more consistent than time series models within industries.

Thus, it seems that CGE models tend to overestimate industrial output relative to SAM and time series models; they do provide forecasts with smaller standard deviations. SAM models also show upward bias, though not as large as CGE models, but have by far the largest standard deviations. Time series models show little upward bias relative to CGE and SAM models and have standard deviations that are comparable to those from CGE models.

4.3 Occupation-by-Industry Matrix Data and Summary Statistics

The output-by-industry estimates are converted to employment estimates using unitized occupation-by-industry matrices. These occupation-by-industry matrices are constructed using actual data on employment for each two-digit SOC code within each two-digit NAICS industry sector. This is an occupation-by-industry-sized matrix of employment, which has element entries as the annual number of employees in the designated industry-occupation combination, hence the title occupation-by-industry employment matrix.

In order to use this matrix to distribute employment, the employment matrix is converted to a matrix of employment-to-output ratios using the industrial output and the employment for each industry-occupation combination, with each element taking the form $\frac{emp_{i,soc}}{x_i}$ where i and $emp_{i,soc}$ is the employment in industry i and occupation soc, and x_i is the output for industry. This gives a matrix of the form $emp = \begin{bmatrix} \frac{emp_{1,1}}{x_1} & \frac{emp_{2,1}}{x_2} & \cdots \\ \frac{emp_{1,2}}{x_1} & \ddots & \\ \vdots & & \end{bmatrix}$. This matrix is multiplied by the industrial output estimates (scalar multiplication) to arrive at employment

estimates for each industry-occupation combination. This process is discussed in further details in Chapter 3.

When estimating industrial employment, the number of employees in each occupation within the industry is divided by the industrial output of that industry to create an employment-to-output ratio: $\frac{\widehat{emp}^{b'}_{i,soc,G}}{\hat{x}^{b'}_{i,M(b)}}$. This ratio is assumed to be constant from year to year: $\frac{\widehat{emp}^{b'}_{i,soc,G}}{\hat{x}^{b'}_{i,M(b)}} = \frac{\widehat{emp}^{t}_{i,soc,G}}{\hat{x}^{t}_{i,M(b)}}$. Rearranged, this assumption can be used to solve for the industrial employment estimates for year t:

Equation (4.1)

$$\frac{\widehat{emp}^{b'}_{i,soc,G}}{\hat{x}^{b'}_{i,M(b)}} \times \hat{x}^{t}_{i,M(b)} = \widehat{emp}^{t}_{i,soc,G}$$

(Franklin, 2007). The results are employment estimates for each industry-occupation pairing, so they can easily be aggregated to industry level (including all occupations) or to the occupation level (including all industries).

State and national level data from 2006, 2008, and 2010 were used to construct a total of six alternative occupation-by-industry matrices. The state and national matrices are measurably different from one another, but the differences between the 2006, the 2008, and the 2010 matrices were negligible at both the state and national levels.

Consider the matrix of these fractions by industry. Each industry has in its section of the matrix the list of occupations that are employed by that industry. Each occupation has an employment-to-output ratio attached to it, which is multiplied by the industrial output to arrive

at employment-by-industry-by-occupation. The sum of these ratios by industry is the industry's total employment coefficient.

The total employment coefficients for industries over the 2006, 2008, and 2010 matrices differ by an average of less than two percent at the state level and less than one percent at the national level. Since these matrices are fixed coefficients of employment-to-output ratios, estimates that are one percent apart would convert any particular output number to employment estimates that would differ by roughly one percent.

It could be the case however, that industrial employment is relatively constant from year to year, but that the distribution of employment over occupations differs, as discussed in Chapter 3. To investigate, the matrix was reaggregated by occupation (instead of by industry). In order to do this, industrial output was no longer sufficient because the aggregation by occupation collects people who work in similar types of jobs for all types of industry. This makes the employment/industrial output ratio irrelevant. Instead, each occupation's employment-to-output coefficient used GDP as its output.

Regardless, when these occupational coefficients were summed, the difference between 2006, 2008, and 2010 state models was less than two percent on average and the difference among the national matrices was less than half of a percent on average.

These two reasons made using multiple years of data to build these matrices redundant for a project whose base level scope is a six-year window. If the scope of data was longer there may be more significant differences in these matrices, but for this project, the matrices are too

similar to predict employment levels that are statistically different from one another, and two of the observation years are dropped. The matrices from 2008 are chosen without loss of generality.

Though the differences between the 2006, 2008, and 2010 matrices at both the national and state level were negligible, the differences between 2008 national and state matrices are sufficiently large in terms of both magnitude and distribution structure. The industry employment-to-output coefficients are summed by industry and presented in Table 4.2.

Table 4.2: Matrix Industrial Coefficient Comparison	State Matrix Aggregated Industry Coefficients	National Matrix Aggregated Industry Coefficients
Agriculture/Forestry/Mining	2.80	3.07
Construction/Utility	9.64	5.62
Manufacturing	4.07	4.55
Wholesale/Retail Trade	18.45	9.45
Transportation/Warehousing	22.40	17.50
Professional Services	3.90	3.43
Education and Health Services	13.44	5.47
Government and Non-NAICS sectors	2.43	0.63

Each ratio represents the annual number of workers needed to produce $1,000,000 worth of output in each industry. For instance, the agriculture/forestry/mining sectors only require 2.8 workers to produce the $1,000,000 worth of production for the state of Louisiana while businesses nationwide require slightly more than three workers to get that same production. These simple ratios show how capital or labor intensive these industries are at the state and national levels. For example, Louisiana construction industries are probably more labor intensive than the industry nationally, as are education and health services. On the other

hand, the Louisiana aggregate agricultural industry is probably less labor intensive (therefore more capital intensive) than the national average. This might be due to the fact that national agriculture sector average includes labor intensive fruit and vegetable farming, of which Louisiana has very little.

Most of the industries have larger state coefficients than national coefficients. Only the agriculture/forestry/mining and manufacturing sectors have larger national coefficients. The government and non-NAICS sector have the largest difference in coefficient values with the state coefficient roughly four times larger than the national estimate, while transportation/warehousing has the largest magnitude coefficient in both the state and national matrices. If these coefficients are very rough measures of labor intensiveness, then one might say that the state of Louisiana is more labor-intensive than the nation in education and health services and government/nonNAICS sectors, but less labor-intensive in wholesale and retail trade when compared to the national averages.

It should be noted however than no adjustments are made for differing wage rates at the regional and national levels in these matrix ratios. That is, Louisiana may have more employees per dollar of output, but these employees may cost less. That is, one may be getting less productivity out of each employee causing you to need more employees, but if wages are lower in a particular region than national average wages this may not necessarily by less cost effective.

Aggregation with respect to occupations yields the coefficients in Table 4.3. The employment-to-output ratios for occupations are calculated in the same manner as for the industrial aggregations and are presented in Table 4.2.

Table 4.3: Matrix Occupational Coefficient Comparison

OCC Code	State Matrix Aggregated Occupation Coefficients	National Matrix Aggregated Occupation Coefficients
11 – Management Occupations	0.0732	0.0409
13 – Business and Financial Operations	0.0269	0.0391
15 – Computer and Mathematical	0.0035	0.0346
17 – Architecture and Engineering	0.0508	0.0363
19 – Life, Physical, and Social Science	0.0165	0.0133
21 – Community and Social Services	0.0167	0.0067
23 – Legal	0.0016	0.0018
25 – Education, Training, and Library	0.0696	0.0775
27 – Arts, Design, Entertainment, Sports, and Media	0.0070	0.0158
29 – Healthcare Practitioners and Technicians	0.0801	0.0810
31 – Healthcare Support	0.0410	0.0457
33 – Protective Services	0.0186	0.0072
35 – Food Preparation and Service	0.1286	0.1336
37 – Building and Grounds Cleaning and Maintenance	0.0288	0.0128
39 – Personal Care and Service	0.0425	0.0806
41 – Sales and Related	0.2278	0.1339
43 – Office and Administrative Support	0.1029	0.0476
45 – Faming, Fishing, and Forestry	0.0048	0.0049
47 – Construction and Extraction	0.1409	0.0915
49 – Installation, Maintenance, and Repair	0.0699	0.0441
51 – Production	0.0368	0.0365
53 – Transportation and Material Moving	0.0865	0.0373
Sum	**1.2749**	**1.0226**

Some differences are very small, such as occupation codes 29, 35, 45, and 51 which are virtually the same, while occupation codes 11, 21, 33, 37, 41 exhibit state estimates more than twice as large as the national counterparts. Only codes 15, 27, and 39 have national coefficients that are significantly larger than the state coefficients, while 13 of the 22 occupations have smaller national coefficients compared to their state coefficients.

One might expect that the sum for the state matrix would be higher than the sum of coefficients for the national matrix due to economies-of-scale and regional comparative

advantages within the country. That is the case for the data used in this study. The sum of the coefficients for the state matrix is about 20% higher than that of the national matrix.

4.4 Industrial and Occupational Employment Data and Summary Statistics

4.4.1 Data

For both industry and occupation aggregations, projections for all models are compared to actual employment data from the BLS and from LaWorks.net, the website of the Louisiana Workforce Commission responsible for compiling labor market information in Louisiana. The actual data are industry level data for 2005, 2006, 2007, 2008, 2009, and 2010 and occupational level data for 2001, 2002, 2005, 2006, 2007, 2008, 2009, and 2010. The industry level data prior to 2005 use SIC classification codes that are unable to be reliably converted to current classification codes because of ambiguity in the original datasets.

The industry level data are state of Louisiana employment figures by industry at the four digit NAICS code level. However, that level of granularity becomes irrelevant because the employment projections are limited by the data that supports the state occupation-by-industry matrix, which aggregates industries only to two digits NAICS code. This being the case, the industry employment figures were aggregated to the two-digit level in order to make an apples-to-apples comparison.

The occupational employment data are available at the six-digit SOC level for the state of Louisiana. However, in much the same way as the industry level data, aggregation to the

two-digit level is required for a proper comparison because the model is limited to two-digit granularity by the state occupation-by-industry data.

4.4.2 Descriptive Statistics

The descriptive statistics for employment forecasts are divided, as in previous discussions, into industrial employment estimates and occupational estimates. Thus, this section is divided as well to discuss the descriptive statistics of each separately.

Industrial Employment Descriptive Statistics

A complete table of industrial employment descriptive statistics is presented in full in Appendix B, but an overview of those statistics is presented here, in Table 4.4. These statistics

Table 4.4: Industrial Employment Descriptive Statistics

Actual Employment	2,449,537					
	CGE 2001	**CGE 2002**	**CGE 2006**	**CGE 2007**	**CGE 2008**	**All State CGE**
Total Employment	3,138,146	3,009,635	3,635,959	4,207,892	4,276,692	3,653,665
Percent Change	49%	23%	48%	72%	75%	49%
	SAM 2001	**SAM 2002**	**SAM 2006**	**SAM 2007**	**SAM 2008**	**All State SAM**
Total Employment	2,793,942	2,353,238	2,944,988	3,330,628	3,639,448	3,012,449
Percent Change	14%	-4%	20%	36%	49%	23%
	TS 2001	**TS 2002**	**TS 2006**	**TS 2007**	**TS 2008**	**All State TS**
Total Employment	1,852,888	1,831,064	1,940,100	1,891,159	1,886,796	1,880,401
Percent Change	-24%	-25%	-21%	-23%	-23%	-23%
	CGE 2001 Nat	**CGE 2002 Nat**	**CGE 2006 Nat**	**CGE 2007 Nat**	**CGE 2008 Nat**	**All Nat CGE**
Total Employment	1,527,992	1,441,711	2,054,158	2,275,103	2,307,393	1,921,271
Percent Change	-38%	-41%	-16%	-7%	-6%	-22%
	SAM 2001 Nat	**SAM 2002 Nat**	**SAM 2006 Nat**	**SAM 2007 Nat**	**SAM 2008 Nat**	**All Nat SAM**
Total Employment	1,358,488	1,161,162	1,715,127	1,846,370	2,017,259	1,619,681
Percent Change	-45%	-53%	-30%	-25%	-18%	-34%
	TS 2001 Nat	**TS 2002 Nat**	**TS 2006 Nat**	**TS 2007 Nat**	**TS 2008 Nat**	**All Nat TS**
Total Employment	894,998	883,274	991,234	939,976	944,098	930,716
Percent Change	-63%	-64%	-60%	-62%	-61%	-62%

show continued trends from the descriptive statistics of the industrial output forecast estimates. That is, ceteris paribus, CGE models have the largest employment estimates followed by SAM and time series estimates, respectively. However, with three industrial output projection techniques combining with two occupation-by-industry matrices for five base years and multiple projection years, it can be difficult to sort out what trends, if any, are present in the descriptive statistics of industrial employment forecast estimates.

Table 4.5 aggregates some of the model estimates to more clearly demonstrate descriptive statistic trends.

Table 4.5: Aggregated Industrial Employment Descriptive Statistics

All CGE	
Avg Tot Emp	2,787,468
Percent Change	14%

All State	
Avg Tot Emp	2,848,838
Percent Change	16%

All SAM	
Avg Tot Emp	2,316,065
Percent Change	-5%

All National	
Avg Tot Emp	1,490,556
Percent Change	-39%

All TS	
Avg Tot Emp	1,405,559
Percent Change	-43%

Though their relative relationships are the same (in that CGE forecast estimates are higher than SAM estimates, which are in turn higher than the time series estimates), SAM estimates show the least bias by underestimating industrial employment by 5%. CGE models are second best in terms of bias, overestimating industrial employment by an average of 14%, and time series models underestimate by an average of 43%.

Further, the models which employ the state occupation-by-industry matrix have an upward bias of 16% while models using the national matrix underestimate industrial employment by an average of 39%. This indicates that national matrix models are forecasting less than optimal industrial employment and that possible improvements in industrial employment forecasts may be achieved through the use of a regional occupation-by-industry matrix.

Occupational Employment Descriptive Statistics

Again, a complete table of occupational employment descriptive statistics is available in Appendix B, but a brief overview is included here. Table 4.6 shows some general descriptive

Table 4.6: Occupational Employment Descriptive Statistics

Actual Employment	1,845,343					
	CGE 2001	**CGE 2002**	**CGE 2006**	**CGE 2007**	**CGE 2008**	**All State CGE**
Total Employment	3,070,774	2,962,288	3,558,698	4,117,768	4,166,902	3,110,974
Percent Change	66%	61%	93%	123%	126%	86%
	SAM 2001	**SAM 2002**	**SAM 2006**	**SAM 2007**	**SAM 2008**	**All State SAM**
Total Employment	2,733,685	2,313,249	2,878,325	3,255,303	3,546,194	3,069,493
Percent Change	48%	25%	56%	76%	92%	66%
	TS 2001	**TS 2002**	**TS 2006**	**TS 2007**	**TS 2008**	**All State TS**
Total Employment	1,726,612	1,742,971	1,594,424	1,600,641	1,575,821	2,042,168
Percent Change	-6%	-6%	-14%	-13%	-15%	11%
	CGE 2001 Nat	**CGE 2002 Nat**	**CGE 2006 Nat**	**CGE 2007 Nat**	**CGE 2008 Nat**	**All Nat CGE**
Total Employment	1,527,992	1,441,711	2,054,158	2,275,103	2,307,393	1,774,957
Percent Change	-17%	-22%	11%	23%	25%	-4%
	SAM 2001 Nat	**SAM 2002 Nat**	**SAM 2006 Nat**	**SAM 2007 Nat**	**SAM 2008 Nat**	**All Nat SAM**
Total Employment	1,358,488	1,229,055	1,715,127	1,846,370	2,017,259	1,691,287
Percent Change	-26%	-33%	-7%	0%	9%	-8%
	TS 2001 Nat	**TS 2002 Nat**	**TS 2006 Nat**	**TS 2007 Nat**	**TS 2008 Nat**	**All Nat TS**
Total Employment	894,998	871,131	932,830	899,414	897,745	1,123,126
Percent Change	-51%	-53%	-49%	-51%	-51%	-39%

statistics. As with industrial output and industrial employment, the CGE forecast estimates are the highest, with SAM and time series following, respectively. The magnitude of the bias in occupational employment estimates, though the relationship between CGE, SAM, and time series remains the same, is somewhere between the overestimations demonstrated in the industrial output descriptive statistics and the biases displayed by the models when looking at industrial employment. That is, all models which use the state occupation-by-industry matrix overestimate occupational employment, while all models that use the national matrix underestimate employment.

However, as earlier, it may be easier to digest some of these statistics if they are aggregated to model and occupation-by-industry matrix geography separately, as in Table 4.7.

Table 4.7: Aggregated Occupational Employment Descriptive Statistics

All CGE

Avg Tot Emp	2,442,966
Percent Change	36%

All State

Avg Tot Emp	2,740,878
Percent Change	52%

All SAM

Avg Tot Emp	2,380,390
Percent Change	29%

All National

Avg Tot Emp	1,529,790
Percent Change	-17%

All TS

Avg Tot Emp	1,582,647
Percent Change	-14%

This aggregation of the statistics indicates that, although the models maintain their relative order in terms of bias, the range is smaller, indicating that the models produce more similar estimates for occupational employment than for industrial output or industrial employment.

Further, models which use the state occupation-by-industry matrix tend to overestimate occupational employment by approximately one-half, while models that use the national matrix tend to underestimate employment by about 17%. This implies that although forecast estimates might be improved by the use of a regional occupation-by-industry matrix in terms of industrial employment, those gains may be at the cost of overestimating occupational employment. The balance of this tradeoff is beyond the scope of descriptive statistics and the regression analysis will shed more light.

5 Results and Statistical Analysis

5.1 Introduction

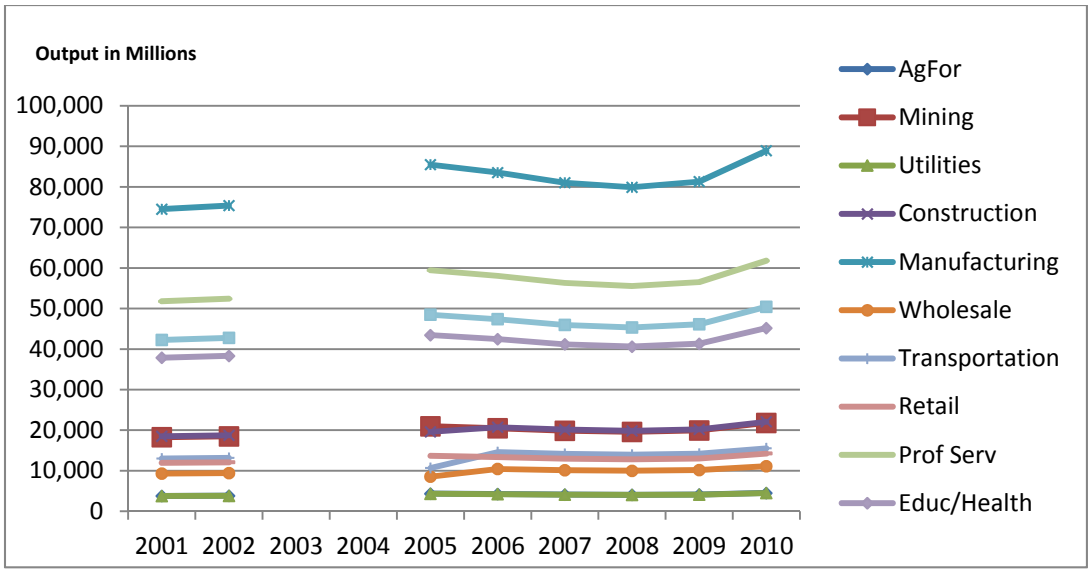
This chapter presents statistical analyses of industrial output and industry/occupation employment forecast estimates. The first section presents a two different regression analyses of the industrial output estimates against actual data, as well as a Theil Statistic analysis. The second section repeats this series of tests for both industrial and occupational employment.

5.2 Louisiana Industrial Output Forecast Estimates Statistical Analysis

5.2.1 Tests of Industrial Output Forecast Estimates Grouped by Model Type (M), Base Year (b), and Industry (i)

The first regression analysis aggregates the industrial output forecast estimates as shown in the graph below. These estimates are grouped for a specific model type-base year combination; for example, Graph 5.1 contains industrial output estimates from only the CGE model with 2001 as a base year for data. Further, Graph 5.1 compiles the forecast estimates for all industries and projection years 2001, 2002, 2005, 2006, 2007, 2008, 2009, and 2010. This aggregation creates a total of 15 (three estimation techniques by five base years) data series (graphs) each with 88 observations (11 industries by eight projection years).

For the first series of tests, all of the point estimates from Graph 5.1 will be used as the dependent variables in a single OLS regression which use the corresponding actual output data as the independent variable series. This style of regression is run for every model with each output projection technique-base year combination, resulting in 15 regressions.



Graph 5.1: 2001 CGE Model Estimates (All)

Because each industry has its own series within each larger series, as demonstrated by the lines connecting the individual data points for each industry in Graph 5.1, an intercept dummy variable will be included as an independent variable for each industry. Equation (5.1) gives the basic form of the regression.

Equation (5.1)

$$\hat{x}_i^t = \beta_0 + \beta_1 \bar{x}_{i,M(b)}^t + \beta_2 AG + \beta_3 Mining + \beta_4 Utilities + \beta_5 Construction + \beta_6 Manufacturing + \beta_7 Wholesale + \beta_8 Transportation + \beta_9 Retail + \beta_9 ProfServ + \beta_{10} EducHealth + \varepsilon$$

where AG, $Mining$, $Utilities$, $Construction$, $Manufacturing$, $Wholesale$, $Transportation$, $Retail$, $ProfServe$, and $EducHealth$ are Boolean intercept dummy variables which indicate the specific industry series. The variable for the non-NAICS industry is left out to avoid a singular matrix (due to perfect multicollinearity).

The use of intercept dummy variables should lower the range of the confidence intervals due to the increased observations (Judge et. al., 1988). Using these intercept dummy variables assumes that the individual industrial output series from Graph 5.1 follow similar trends; more specifically, each industry's data will follow a similarly shaped best-fit line. Though the best-fit lines need to be similarly shaped, they need not have different y-axis intercepts (Judge et. al., 1988). Looking at the lines for each industry in Graph 5.1 confirms that the trends of the data follow very similar paths. Further, it is reasonable to expect that industrial output at the two-digit NAICS level aggregation, each of which entails large groups of industries, would have similar trends over time following general macroeconomic trends, regardless of the specific industry classification.

The typical statistical analysis using R-squared, adjusted R-squared, F, and Durbin-Watson statistics will be used to determine how well each output-by-industry estimation technique/base year combination fits the data. Coefficient analysis and hypothesis testing on β_1 will reveal how well the estimates fit the actual data, while hypothesis testing on the other Booelan variables' coefficients will indicate if estimates for particular industries predict actual output better than other industries.

This regression analysis is completed for each output-by-industry estimation technique/base year combination (M_b) for a total of 15 regressions in the first series. All regression equations for models that use SAM or CGE based industrial output estimation methods have 88 observations: one for each industry (11) in each of the projection years (8). Since the time series models do not backcast, they have diminishing numbers of observations.

The time series models for 2001, 2002, 2006, 2007, and 2008 base years have 80, 72, 56, 48, and 40 observations respectively, one for each industry (11) in each projection year for which time series estimates are made for that particular base year.

In all regressions, the forecast estimates were significant (at the one percent level) predictors of the actual output data. The goodness-of-fit measures from each regression can be analyzed and compared to one another to determine which, if any, models fit closest to the actual data. The adjusted-R-squared and F-statistics for the first series of fifteen regressions are shown in Table 5.1 below.

Table 5.1: Industrial Output Regression Results

	Base years						
	adj R^2	**2001**	**2002**	**2006**	**2007**	**2008**	**AVG**
Model type	**CGE**	0.9313	0.9249	0.939	0.9306	0.9284	0.9308
	SAM	0.9308	0.9169	0.9384	0.9284	0.9306	0.929
	TS	0.9119	0.9265	0.9371	0.9288	0.9388	0.9286
		Base years					
	F	**2001**	**2002**	**2006**	**2007**	**2008**	**AVG**
Model type	**CGE**	148.97	135.32	168.9	147.34	142.35	148.576
	SAM	147.79	121.35	167.19	142.37	147.32	145.204
	TS	103.51	113.37	103.96	78.11	76.36	95.062

First note that all of the models explain large amounts of the variance in the actual data (all above 90%). The CGE models have a slightly higher average R-squared statistic (followed by SAM and then time series), but all of the models' averages are within one percent.

Further, all of the F-statistics are significant at the one percent level, which implies that the independent variables have a correlation with the dependent variable that is statistically different than zero.

5.2.2 Meta-Analysis Regression (MAR) Tests of Industrial Output Forecast Estimates Grouped by Model Type, Base Year, and Industry

To determine if any of the model characteristics significantly affect the R-squared value, the R-squared values are collected for use as the dependent variable in OLS regression. The series of independent variables are Boolean variables for model characteristics. The regression equation is

Equation (5.2)

$$R^2 = \beta_0 + \beta_1 C2001 + \beta_2 C2002 + \beta_3 C2006 + \beta_4 C2007 + \beta_5 CGE + \beta_6 SAM + \varepsilon$$

where R^2 is the adjusted R-squared value from regressions equation (5.1), $C2001$, $C2002$, $C2006$, and $C2007$ are dummy variables for the base year of the model ($C2008$ is left out to avoid perfect multicollinearity), CGE and SAM are dummy variables for the model's industrial output estimation technique (TS, the time series dummy, is left out to avoid perfect multicollinearity), and ε is the regression error term.

Hypothesis testing on the variable coefficients indicates whether a variable is significantly correlated with the R-squared value, all else equal. The regression results are presented in Table 5.2.

Table 5.2: Industrial Output MAR Results

R2	0.5767		**F**	1.82
Adj R2	0.2592		**Prob > F**	0.2129

Variable	**Coeff**	**t**
C2001	-0.0079	-1.47
C2002	-0.0098	-1.83
C2006	0.0056	1.03
C2007	-0.0033	-0.62
CGE	0.0022	0.53
SAM	0.0004	0.10

The individual variable hypothesis testing reveals that none of the independent variables are statistically significant. The CGE and SAM variables' insignificance implies that neither method provides a statistically significant improvement in a model's adjusted R-squared value over the excluded time series dummy variable. That is, none of the models have a statistically significant advantage over the others in terms of R-squared values. However, the F-statistic indicates that the regression as a whole is insignificant, implying that the ability of these independent variables to explain variance in the dependent variable is not statistically different from zero. This could be because the number of observations is low (15: 5 base years x 3 models), but regardless, the results lack robustness.

5.2.3 Theil Statistic Analysis

In an attempt to provide depth to the statistical analysis, a Theil Statistical Decomposition Analysis is performed. The Theil statistic is a measure of inequality that can be used to determine differences among data sets (Theil, 1967). The Theil Inequality Coefficient (U) is calculated as $U = \frac{\sqrt{\frac{1}{n}\sum(X_i - \bar{X}_i)^2}}{\sqrt{\frac{1}{n}\sum {X_i}^2} + \sqrt{\frac{1}{n}\sum {\bar{X}_i}^2}}$, which returns a value between zero and one. A U-statistic of zero implies that the estimates and the actual data are exactly the same. A U-statistic of one implies that the estimates are no better than a naïve guess (Pyndyck and Rubenfeld, 1981). Theil Statistical Analysis "also assesses the mode's ability to duplicate turning points or rapid changes in data" (Tijskens et. al., 2001) using a decomposition of the standard U-statistic.

The Theil Inequality Coefficient can be decomposed into three indices that account for bias, variance, and covariance respectively (Pyndyck and Rubenfeld, 1981). The bias decomposition (U_{bias}) is an index which addresses possible bias between datasets on the whole, regardless of any variance within either of the datasets. The variance decomposition (U_{var}) "represents the ability of the model to replicate the degree of variability in the observed data" (Tijskens et. al., 2001). The covariance decomposition (U_{cov}) index represents the remaining error after bias and variance are removed (random error). In general, the bias and variance coefficients will be low for estimates which replicate data well. The covariance coefficient is widely considered to have less value, in terms of goodness-of-fit, than the other two decompositions, but nonetheless, a perfectly correlated set of estimates would yield zero for bias and variance coefficients and one for the covariance decomposition (Pyndyck and Rubenfeld, 1981).

Table 5.3 has the average U-statistics for industrial output forecast estimates from each model type as well as a total statistic for all model types.

Table 5.3: Average U-Statistics

	CGE	**SAM**	**TS**	**Total**
U	**0.41**	**0.35**	**0.32**	**0.36**
Ubias	0.36	0.26	0.00	0.21
Uvar	0.46	0.43	0.02	0.30
Ucov	0.18	0.31	0.39	0.29

These Theil coefficients indicate that CGE models have higher U-statistics than either SAM or time series models, indicating that CGE estimates are performing worst among the three options. In the decomposition, it is apparent that the upward bias in the CGE and SAM

forecast estimates seen in the descriptive statistics is present here. Further, the variance coefficients for the CGE and SAM models are significantly higher than the coefficient for the time series models. These results, measured against the perfect correlation coefficients (bias = 0, var = 0, cov = 1), indicate that time series models significantly outperform both SAM and CGE models with negligible bias and variance coefficients. SAM models also slightly outperform CGE models, though the difference between SAM and CGE models is relatively small.

One might have expected CGE and SAM models to overestimate output. There is literature supporting the theory the I/O-style models may overestimate production and/or income (Simonovitz, 1975; Lahiri, 1983; Lahiri and Satchel, 1985; Bullard and Sebard, 1988). However, the poor variance decomposition coefficients indicate that the SAM and CGE models are somehow less sensitive to the movements in the actual data than are the time series estimates. In theory, the opposite should occur: CGE and SAM models should capitalize on the additional information included in their structural models in terms of increased sensitivity and detail.

In order to investigate the effects of the strong upward bias known to be present, and any possible spillover effects to the variance coefficient, all output forecast estimates were calibrated to the actual data from the base year upon which the producing model was based. That is, the upward bias was systematically removed while maintaining the variant relationship of the original estimates. The U-statistics were recalculated and the results are in Table 5.4

With the upward bias corrected, the models converge. The CGE models jump from worst to best, followed by time series and SAM models respectively. However, the U-statistics

for all three model types are within 8% of one another. Regardless, it seems that correcting the upward bias in the CGE and SAM model estimates also corrected the high variance coefficients.

Table 5.4: Average U-statistics

	CGE	SAM	TS	All
U	0.44	0.52	0.46	0.47
Ubias	0.00	0.01	0.00	0.00
Uvar	0.00	0.01	0.01	0.01
Ucov	0.25	0.09	0.16	0.17

Note that the upward bias in the output estimates was only removed for the purposes of investigating the unanticipated variance coefficient results. These upward biases will not be corrected moving forward from this point. This decision is made with knowledge of results to come, and will speak to final conclusions. Full results, all the way through to employment forecast estimates and statistical analyses thereof, were calculated and analyzed for both corrected and uncorrected estimates, but the corrected estimates are left out of this dissertation in part because they are not necessary or additive to the research narrative (and partly to spare the pages).

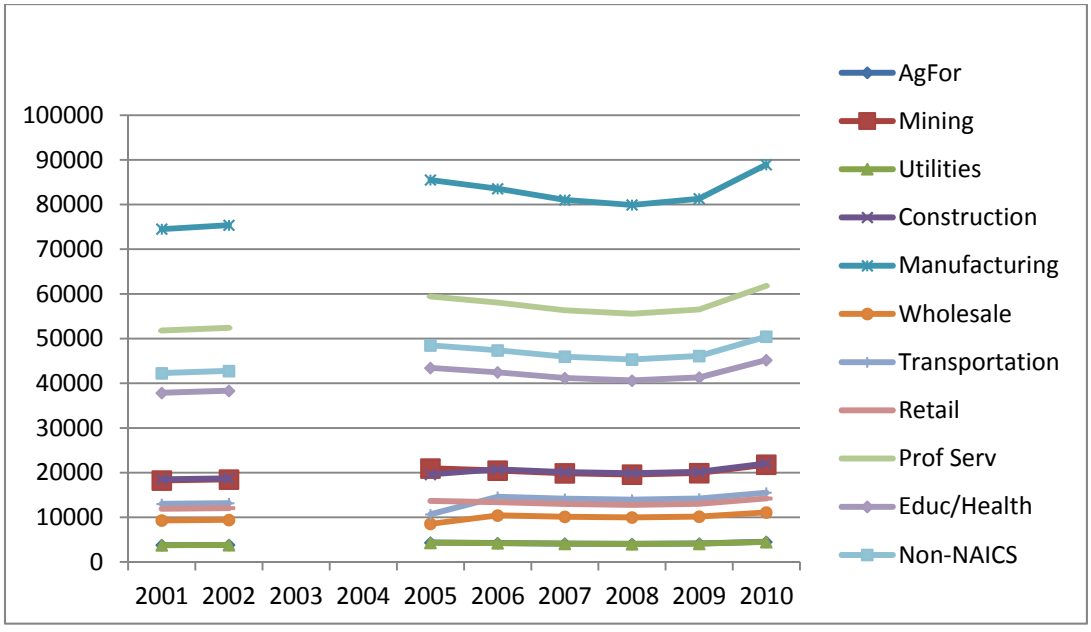
5.3 Louisiana Industrial and Occupational Employment Forecast Estimates Statistical Analysis

5.3.1 Tests of Industry/Occupation Employment Forecast Estimates Grouped by Model Type, Base Year, and Industry

For the first employment estimate series, the data observations are grouped such that estimates in each regression are from the same combination of the following parameters: model base year, industrial output estimation technique, and occupation-by-industry matrix

geography, expressed in matrix notation as $\left[\begin{array}{c}[\quad]_{i,t}\\ \\ \end{array}\right]_{M(b)}$. That is, there is a series of estimates for all projection years for each industry, as in the figure below which shows the employment estimates by industry for the model using CGE industrial output estimates and the state occupation-by-industry matrix. This mirrors the structure of the first series of tests for industrial output estimates. Each of the data points on the graph will be an observation in the regression that fills each block of the matrix $\left[\begin{array}{c}[\quad]_{i,t}\\ \\ \end{array}\right]_{M(b)}$.

With the industrial output estimates, there were 15 of these graphs (CGE, SAM, and time series models for each of the five base years), and hence, 15 different regression equations. With the employment data these same 15 regressions equations are doubled because there are estimates using each of these regressions with both the national and state occupation-by-industry matrices, bringing the total number of regressions to 30



Graph 5.1:2001 CGE Estimates of Industrial Employment

The specific regression equations take the form

Equation (5.3)

$$\widehat{emp}_{i,M(b),G}^{t} = \beta_0 + \beta_1 \overline{emp}_i^t + \beta_2 AG + \beta_3 Mining + \beta_4 Utilities + \beta_5 Construction + \beta_6 Manufacturing + \beta_7 Wholesale + \beta_8 Transportation + \beta_9 Retail + \beta_9 ProfServ + \beta_{10} EducHealth + \beta_{11} STATE + \varepsilon$$

where $\widehat{emp}_{i,M(b),G}^{t}$ is the estimated employment by industry i (or occupation soc) for a model using industrial output estimation technique M, base year b, and occupation-by-industry matrix geography $G \in (National, State)$ indicates the geography of the occupation-by-industry matrix, $\overline{emp}_i^t$ is the actual employment by industry (or occupation), indicates which occupation-by-industry matrix was used to create the estimate, AG, $Mining$, $Utilities$, $Construction$, $Manufacturing$, $Wholesale$, $Transportation$, $Retail$, $ProfServe$, and $EducHealth$ are Boolean intercept dummy variables which indicate the specific industry series, and $STATE$ is a Boolean intercept dummy which equals one if the observation uses the state occupation-by-industry matrix and zero otherwise.

Within each dependent variable series, there are estimates for each industry in each projection year. The Boolean variables that account for industry allow each industrial series to have its own intercept.

The results from these regressions are presented in Tables 4.11 (industrial employment) and 4.12 (occupational employment).

The first impression of the industrial employment results is that all the models have very high adjusted R-squared values and very high F-statistics. This means that combinations of base year, industrial output estimation technique, and occupation-by-industry matrix geography explain variation in actual employment data generally well.

Table 5.5: Industrial Employment Regression Results

State Matrix						
adj R^2	**2001**	**2002**	**2006**	**2007**	**2008**	**Averages**
CGE	0.9732	0.976	0.9712	0.9688	0.9724	0.9723
SAM	0.9773	0.9622	0.9628	0.9596	0.9661	0.9656
TS	0.952	0.959	0.9633	0.9683	0.9623	0.9609
F	**2001**	**2002**	**2006**	**2007**	**2008**	
CGE	396.57	443.34	367.83	339.11	384.51	386.272
SAM	469.22	469.22	282.13	259.1	311.25	358.184
TS	216.57	232.09	207.82	211.17	150.86	203.702
National Matrix						
adj R^2	**2001**	**2002**	**2006**	**2007**	**2008**	
CGE	0.9365	0.9395	0.918	0.9238	0.9233	0.9282
SAM	0.9394	0.9267	0.9124	0.916	0.9167	0.9222
TS	0.945	0.9545	0.9296	0.9328	0.9266	0.9377
F	**2001**	**2002**	**2006**	**2007**	**2008**	
CGE	161.47	169.88	122.74	132.93	131.95	143.794
SAM	169.61	138.49	114.21	119.54	120.64	132.498
TS	187.86	208.01	105.02	96.36	71.01	133.652

For the models that use the state occupation-by-industry matrix, the CGE models have the highest average R-squared value with SAM and time series models averaging about 0.5% and 1% respectively. For national occupation-by-industry matrix, time series models average the highest adjust-R-squared value with CGE and SAM models each about 0.5% lower.

The models which use the state occupation-by-industry matrix average about 3.5% higher R-squared values than the models that use the national matrix. This evidence might suggest that geographical matching may improve a model's ability to predict industrial employment.

Table 5.6: Occupational Employment Regression Results

State Matrix						
adj R^2	**2001**	**2002**	**2006**	**2007**	**2008**	**Averages**
CGE	0.9755	0.9737	0.9672	0.9705	0.9716	0.9717
SAM	0.972	0.9633	0.9687	0.9686	0.9702	0.96856
TS	0.9612	0.9647	0.966	0.9653	0.9684	0.96512
F	**2001**	**2002**	**2006**	**2007**	**2008**	
CGE	313.94	292.1	233.18	259.63	270.44	273.858
SAM	274.06	207.45	244.59	244.17	257.01	245.456
TS	196.06	188.61	139.69	108.91	89.15	144.484
National Matrix						
adj R^2	**2001**	**2002**	**2006**	**2007**	**2008**	
CGE	0.9679	0.968	0.9646	0.9653	0.9645	0.96606
SAM	0.9679	0.9595	0.9629	0.9622	0.9626	0.96302
TS	0.9604	0.9639	0.966	0.9653	0.9679	0.9647
F	**2001**	**2002**	**2006**	**2007**	**2008**	
CGE	238.72	239.36	215.3	220.19	215.22	225.758
SAM	238.28	187.38	205.26	201.24	203.85	207.202
TS	192.18	184.74	139.67	108.89	87.61	142.618

Looking at occupation employment, models that use the state occupation-by-industry matrix, the CGE models have the highest average R-squared value (0.9717) with SAM (0.9686) and time series (0.9651) within one percent. For models that use the national matrix, CGE (0.9661) is still the highest, but time series (0.9647) slightly outperforms SAM (0.9630). The F-

statistics for all models are significant at the one percent level, which implies that the characteristics of the model have significant predictive power in terms of adjusted R-squared.

The state matrix models, as with industrial employment, show a slightly higher average R-squared value (0.9685) than the national matrix models (0.9645). However, this difference is less than half of one percent.

5.3.2 Meta-Analysis Regression (MAR) Tests of Industrial Output Forecast Estimates Grouped by Model Type, Base Year, and Industry

Though all the models have high adjusted R-squared statistics, this series should help reveal which, if any, model characteristics produce higher R-squared values statistically. The R-squared values from the regressions in section 4.4.3 above are used as the dependent variables in regressions which use dummy variables for various model characteristics as independent variables. The equations take the form

Equation (5.4)

$$R^2 = \beta_0 + \beta_1 CGE + \beta_2 SAM + \beta_3 State + \beta_4 C2001 + \beta_5 C2002 + \beta_6 C2006 + \beta_7 C2007 + \varepsilon$$

where SAM (SAM based industrial output estimates), TS (time series base industrial output estimates), NAT (national occupation-by-industry matrix), $C2001$ (model base year of 2001), $C2002$ (model base year of 2002), $C2006$ (model base year of 2006), and $C2007$ (model base year of 2007) are Boolean variables equal to one if the model has the particular characteristic, Note that variables for time series based industrial output estimates, national occupation-by-industry matrices, and 2008 model base years are excluded from the regression equation to avoid singularity. The results of this regression equation are presented in the table below.

The only variable that is significant is the dummy variable for the state occupation-by-industry matrix in the industrial employment estimates. The results indicate that models which use the state occupation-by-industry matrix should expect a 1% higher R-squared value than models which use the national matrix to predict industrial employment, ceteris paribus. The model for industrial employment has an F-statistic that is statistically significant at the 10% level.

Table 5.7: Industrial and Occupational 1st MAR Results

Industrial Employment				
R2	0.065		**F**	1.95
Adj R2	0.0317		**Prob > F**	0.0759
	Variable	**Coeff**	**t**	
	C2001	0.0204	2.21	
	C2002	0.0066	0.71	
	C2006	-0.0063	-0.67	
	C2007	0.0181	1.89	
	CGE	0.0016	0.21	
	SAM	0.0004	0.05	
	State	0.0287	5.14*	
Occupational Employment				
R2	0.6986		**F**	73.49
Adj R2	0.6891		**Prob > F**	0
	Variable	**Coeff**	**t**	
	C2001	0.0253	1.23	
	C2002	0.0215	1.07	
	C2006	0.016	1.35	
	C2007	0.0195	1.14	
	CGE	0.0199	0.21	
	SAM	-0.0167	-0.17	
	State	0.0073	0.97	

* indicates significance at the 1% level

The occupational employment regression is much more statistically significant in general, with an F-statistic significant at the 0.1% level, but the regression finds no significant variables. The state matrix has a positive coefficient, but is not statistically significant.

Both the industrial and occupational employment models find no statistical difference between the alternative industrial output forecasting methods. This is evidenced by the lack of significance of their corresponding variables in both models.

This is a similar problem that was encountered with the industrial output tests. Now, as then, robustness is needed in these results. In efforts to bolster results, employment estimates are reorganized in order to more evenly spread the observation between the first and second series of tests.

5.3.3 Employment Forecast Theil Statistical Analysis

The Theil Inequality Coefficient Decomposition Analysis from section 5.2.3 is applied here to both industrial and occupational employment forecast estimates. The U-statistics for industrial employment forecast estimates are in Table 5.5. The U-statistic and each of its decompositions are averaged for each model type and geography and a total average is also included.

The U-statistics show that CGE and SAM models are performing very similarly, as might have been expected, but that the time series models have a U-statistic about 50% higher than CGE and SAM. Further, the models which used the state occupation-by-industry matrix have slightly better averages than models which used the national matrix.

Table 5.8: Industrial Employment

U

	CGE	SAM	TS	All
State	0.45	0.42	0.57	**0.48**
National	0.43	0.45	0.63	**0.50**
Both	**0.44**	**0.43**	**0.60**	

Ubias

State	0.11	0.33	0.04	**0.16**
National	0.09	0.16	0.40	**0.22**
Both	**0.10**	**0.25**	**0.22**	

Uvar

State	0.11	0.06	0.01	**0.06**
National	0.07	0.08	0.23	**0.13**
Both	**0.09**	**0.07**	**0.12**	

Ucov

State	0.78	0.89	0.78	**0.82**
National	0.84	0.76	0.31	**0.63**
Both	**0.81**	**0.82**	**0.55**	

The bias coefficients show distinct advantages for the CGE model type and the state occupation-by-industry matrix, indicating that the upwardly biased output numbers from the CGE and SAM models produced the least biased industrial employment forecast estimates. Conversely, the least biased industrial output numbers (time series) have produced industrial employment forecast estimates with relatively large bias.

This contradictory result is supported by the variance coefficients where SAM and CGE models again outperform the time series models which had much better industrial output variance coefficients. These results are troublesome and more discussion about them will follow.

The U-statistic coefficients for occupational employment forecast estimates are in Table 5.6.

The occupational employment forecast estimates mimic the results from the industrial employment estimates. The CGE and SAM models outperform the time series models in the general U-statistic as well as in the variance decomposition coefficients. All three models are within 2% in the variance coefficients. Though the bias coefficients favored the models which used the state occupation-by-industry matrix, the models which used the national occupation-by-industry matrix outperformed state models in terms of the U-statistics and the variance coefficients.

Table 5.9: Occupational Employment

		CGE	SAM	TS	All
U					
	State	0.61	0.54	0.44	**0.53**
	National	0.33	0.36	0.59	**0.43**
	Both	**0.47**	**0.45**	**0.52**	
Ubias					
	State	0.10	0.07	0.00	**0.06**
	National	0.06	0.06	0.24	**0.12**
	Both	**0.08**	**0.07**	**0.12**	
Uvar					
	State	0.51	0.45	0.32	**0.43**
	National	0.15	0.16	0.31	**0.21**
	Both	**0.33**	**0.31**	**0.32**	
Ucov					
	State	0.39	0.48	0.61	**0.49**
	National	0.79	0.77	0.22	**0.59**
	Both	**0.59**	**0.63**	**0.41**	

5.4 Summary and Analysis of Results

The summary statistics from chapter 4 indicated that the CGE and SAM models were overestimating industrial output. This overestimation was not entirely unexpected and follows trends from existing literature. This overestimation is confirmed by the Theil Statistical Decomposition analysis of the industrial output forecast estimates. The CGE and SAM estimates

of industrial output have significantly worse bias and variance coefficients than did the time series models. However, despite this upward bias, when output estimates were regressed against actual data, the CGE, SAM, and time series models have R-squared statistics that are statistically indistinguishable, implying that each of the estimation techniques replicate variation in the actual data with roughly equal efficiency (this implication was confirmed by the investigation using bias-corrected estimates).

The statistical analysis of the employment forecast estimates seems to contradict the results from the output estimates. Though all three model types again have high R-squared values when regressing employment estimates against actual industrial and occupational employment figures, employment estimates made using CGE and SAM-produced output estimates outperformed models using time series-produced output estimates in terms of U-statistics. CGE and SAM estimates of industrial employment have better U, Ubias, Uvar, and Ucov coefficients than time series estimates. The three models had roughly equivalent Uvar coefficients for occupational employment, but the overall U-statistic and the Ubias coefficient both favored CGE and SAM models.

These results indicate that though the CGE and SAM industrial output forecast estimates are upwardly biased (in some cases significantly), those estimates are nonetheless producing the best industrial and occupational employment forecast estimates. Conversely, the most accurate industrial output estimates (time series) are producing the least accurate employment estimates.

There are at least two possible reasons for this result. The first possible reason for this could lie in the distribution of industrial output to employment by the occupation-by-industry matrix. This matrix is the data bottleneck for this dissertation. As discussed in Chapter 3, regional employment data marked by industrial and occupational indices is often redacted for firm privacy reasons. Thus, there were only three years of data available (2006, 2008, and 2010).

In a preliminary analysis of these state occupation-by-industry matrices, the occupational employment ratios were compared across years. That is, the ratio of employees in a particular occupation within an industry to the total employment in that industry. These ratios were found to be nearly identical across the three years of data. Since the process of distributing the industrial output into employment is a linear transformation and the transformers (the matrices) were almost identical, it stood to reason that the choice between matrices would be trivial. Thus, the 2008 matrix was chosen and the others discarded.

This decision may have biased the data in an unanticipated manner. In choosing the 2008 state occupation-by-industry matrix for use in the creation of industry/occupation employment forecast estimates, the calibration of output-to-employment for the base model is fixed to the 2008 level. That is, by choosing to fix this matrix, the ratio of output-to-employment is also fixed. However, that ratio may have been skewed by the 2008 financial crisis. Schaal (2012) asserts that employment was abnormally low and worker productivity was abnormally high surrounding the 2008 recession. If the worker productivity was artificially high during 2008, this would possible explain why the most accurate industrial output forecast

estimates were producing too little employment. Further, the overestimated CGE and SAM estimates would have been artificially “deflated” and would appear to be unbiased, when in fact their existing upward bias was merely being offset by the downward bias presented by the artificially high worker productivity.

The second possible reason for the discrepancies is a generalization of the first. That is, it could be that the specific year used for this study was a bad year, but it could be that the relationship between output and employment is not properly characterized within the industry/occupation employment projection strategy and requires more specific attention. That is, the process explicitly assumes that output in a given period is directly related to employment in that same period. In reality, employment decisions are based on a myriad of variables including, but not limited to, national or regional economic trends, future expectations, wage trends, and characteristics of the workforce.

The reason these two possibilities are separated is that I view the first as a somewhat misfortunate choice, a poor choice in application of a valid model to be learned from in future model applications. The second possibility I believe is more indicative of a poor modeling strategy that would hinder the applicable use of this model moving forward. The determination of which possibility (if either) is true is left for future research.

6 Conclusions, Contributions, and Suggested Extensions

6.1 Introduction

This chapter provides a summary of the previous four chapters which outlines the primary contents of each. Next, the four contributions identified in the literature review are addressed individually along with results and implications. A final section provides some concluding remarks.

6.2 Summary

In the first chapter, the motivations and backgrounds of industry/occupation employment forecasting were introduced. The Bureau of Labor Statistics (BLS) recommended process for producing such forecasts (Franklin, 2007) was presented and two specific stages of the process were identified as possible points of improvement by the use of modern economic methods: the production of industrial output forecasts and the conversion of output forecasts to employment figures via the use of an occupation-by-industry matrix.

Within the guidelines recommended by the BLS for the creation of industry/occupation employment projections, various methods may be used to produce output-by-industry forecast estimates, including but not limited to econometric analysis, Input-Output or Social Accounting Matrix (I/O-SAM) analysis, and Computable General Equilibrium (CGE) analysis. However, empirical evidence of their comparative performance in the industry/occupation employment forecasting methodology does not exist. Thus, this dissertation set out to provide such an

empirical study of alternative techniques' ability to create reliable employment forecasts, and provides these specific objectives which guide this dissertation.

1) Evaluate current and relevant projection procedures for creating regional industry/occupation employment forecasts.
2) Evaluate the theoretical underpinnings of current forecast procedures, as well an Input-Output and Social Accounting Matrix modeling, Computable General Equilibrium modeling, and alternative occupation-by-industry matrices.
3) Build SAM and CGE models of industrial output for the state of Louisiana as a single region.
4) Construct alternative occupation-by-industry matrices.
5) Test the industrial output techniques, the occupation-by-industry matrices, and all combinations thereof.

The second chapter provides a brief literature review of the existing body of academic research around employment forecasts, alternative industrial output forecasting techniques (I/O-SAM and CGE), occupation-by-industry matrices, and meta-regression analysis (MRA). These sections draw together a basic history of each topic and seek to motivate the study of these topics by illustrating how this study might address unanswered questions in existing literature. Appropriately, the chapter ends by identifying four specific contributions to which this dissertation aspires.

The third chapter (and Appendix A) offers a formal presentation of the alternative models for consideration, including the three alternative industrial output forecasting

techniques (time series, I/O-SAM, and CGE), the construction of the regional occupation-by-industry matrix, and mechanics of how each of these alternatives were combined to create industry/occupation employment forecasts.

In the fourth chapter, the forecasts produced by each model, as outlined by the third chapter, are tested against actual employment data. Some typical regression analysis finds that, in general, no significant difference is found between the alternative industrial output forecast techniques. This analysis is supplemented by a MAR analysis that adds depth to the general analyses of results and their implications, confirming the lack of statistical variation in industry/occupation employment forecasts due to alternative industrial output forecasting technique and finding a statistically significant improvement when using the state occupation-by-industry matrix rather than the national matrix.

6.3 Contributions

To conclude the literature review, this dissertation aspired to four contributions to existing literatures that are outlined above. First, this dissertation provides an empirical study of the performance of alternative industrial output forecasting techniques used within the industry/occupation employment forecasting procedure recommended by BLS (Franklin, 2007). Second, the dissertation provides a deliverable, an occupation-by-industry employment matrix for the state of Louisiana, which is not currently available from national or state sources. Third, the performance of the new regional occupation-by-industry matrix compared to that of the national occupation-by-industry matrix is tested. I find that the regional matrix does outperform the national matrix in terms of improved occupational and industrial employment

forecasts. Finally, I use a unique approach to testing which employs Meta-Analysis Regression (MAR) analysis. This method takes advantage of the large numbers of models with many combinations of model characteristics in this dissertation and uses this analysis to determine which, if any, model characteristics are correlated with improved industry/occupation employment forecasts.

I believe each of the contributions is achieved in this dissertation, and each provides insight to future research and/or policy development. This section devotes a brief discussion to each contribution, the results of this paper which apply, and the implications to research or policy that result.

6.3.1 The First Contribution

The first contribution is providing the literature with an empirical study of alternative industrial output forecasting techniques within the industry/occupation employment guidelines recommended by BLS. This study found that the alternative techniques provided results that were not statistically different for any models in this dissertation. That is, the integration of more complex, general equilibrium models into the industry/occupation employment forecasting does not improve the employment forecasts in this study under likely model specification.

The implication to policy development is not absolute, but rather, as many economic decisions do, it depends on a simple cost benefit analysis. The evidence from this research suggests that investing increased resources in the insertion of a CGE into the industry/occupation employment forecasting procedure does not provide returns by way of

improved forecasts. However, the choice between using a time series econometric technique or an I/O-SAM technique is not as straightforward. Since these methods seem to have no significant differences, they should be viewed, generally, as substitutes in terms of benefits in the policy function. Thus, the choice of which is more appropriate for use depends on the cost of production of each, which can change depending on who is commissioning the forecasts, scope and granularity of the desired forecasts, econometric forecasting experience of forecasters, and other situational variables.

As a brief example, consider a state agency in Louisiana that would like to create industry/occupation employment forecasts for the state of Louisiana and is choosing between time series regression analysis and I/O-SAM analysis for an industrial output forecasting technique. If there is an economist on staff, it may be easy for him or her to create a reliable time series estimate. However, if the agency does not have someone on staff that can do this, it may be more cost efficient to contract out the one-time creation of an I/O-SAM model rather than hire someone specifically to create and maintain a set of time series equations.

Thus, this research suggests that regional, state, and local agencies should be wary of investing in more complex methods of forecasting regional industrial output. However, this research is by no means deterministic or absolute. I believe that more research is necessary to determine the ability of CGE frameworks in forecasting industrial output. Issues of aggregation and granularity that are present in this research due to employment data and aggregation restrictions would not be present in research devoted to industrial output forecasting outside

of the industry/occupation employment forecasting framework. This would allow for a more detailed account of possible benefits of using more complex structures like I/O-SAM and CGE.

6.3.2 The Second and Third Contributions

The second and third contribution, the creation of a deliverable in the form of a previously unavailable occupation-by-industry matrix for the state of Louisiana and the testing of that matrix against its alternative, are inextricably linked, and are thus addressed together. This state-level occupation-by-industry matrix, the construction of which is discussed in the third chapter, however, is prevented from having high-levels of industrial and occupational granularity. Regional employment data is often not collected with both industry and occupation identifiers, and the data that does contain both identifiers can have privacy disclosure issues and is thus of low granularity (or high aggregation).

This research suggests that the state-level occupation-by-industry matrix outperformed the national matrix in terms of improved employment forecasts. This research, as discussed, is at high levels of aggregation (low levels of granularity), but the data are unavailable at the moment to determine if these results hold true as granularity is increased. If data collection methods were available to the public, or to economists through non-disclosure agreements of some kind, matrices of higher granularity could be constructed, tested, and possibly implemented in the policy functions as improvements of their own accord, but also as improvements within the forecasting of industry/occupation employment.

6.3.3 The Fourth Contribution

The last contribution sought to sort through the many models created by the alternative techniques in combination, as well as for the number of years for which each model was based, by using MAR analysis to determine which model characteristics predicted high R-squared values when a model's forecast estimates were regressed on actual data. This method is typically used in meta-analysis literature, but is used here to identify trends in goodness-of-fit measures associated with specific model characteristics.

This application is attractive because, with as many as 144 different sets of employment forecasts to test, individual goodness-of-fit measures, descriptive statistics, and hypothesis testing can easily become extensive and trends can be difficult to spot amongst the statistical noise. Further, as computing power increases, economists are able to create studies which have many more alternative methods than the 5 varying model characteristics in this dissertation. MAR could provide a tool, or perhaps the basis of a larger evaluation method, to determine which models characteristic combinations perform best.

6.4 Conclusion

This dissertation set out to evaluate the industry/occupation forecasting guidelines recommended by BLS, set out by Franklin (2007). The guidelines were evaluated, possible alternative techniques identified, and models were constructed. There were a total of 32 industry/occupation employment models constructed for 9 industry groups and 22 occupation groups.

The estimates from each model were compared to actual employment data and regression analysis (including MRA) revealed that none of the alternative industrial output forecasting techniques produced significantly better employment forecasts. However, the results did show that the state occupation-by-industry matrix, constructed in this paper, outperformed the standard national matrix, implying that the assumption that national staffing patterns are a proxy for regional staffing patterns for purposes of employment forecasting may need to be scrutinized further.

This research also suggests that a proper industry/occupation employment forecast procedure should forecast industrial output using an econometric or I/O-SAM-based technique (depending on cost efficiency) in combination with an occupation-by-industry matrix that is geographically matched with the scope of the project rather than assuming that national staffing patterns are a proxy for regional patterns.

Cited Works

Adams, A. A. and Stewart I. G. “Input-Output Analysis: An Application.” The Economic Journal Vol.66. (1956): 263, 442-454.

Adkins, Lee C., Dan S. Rickman, and Abid Hameed. "Bayesian estimation of regional production for CGE modeling." Journal of Regional Science 43.4 (2003): 641-661.

Alpert, Andrew, and Jill Auyer. "Evaluating the BLS 1988-2000 employment projections." Monthly Lab. Rev. 126 (2003): 13.

America’s Labor Market Information System (ALMIS), (May, 2001). Projecting Long-Term Industry and Occupational Employment for States and Areas.

Barkoff, Kendra “Salazar Calls for New Safety Measures for Offshore Oil and Gas Operations; Orders Six Month Moratorium on Deepwater Drilling.” www.DOI.gov. The U.S. Department of the Interior, May 27, 2010. Web. Aug 23, 2012.

Barro, Robert J., et al. "Convergence across states and regions." Brookings papers on economic activity (1991): 107-182.

Bell, David. "Regional Output. Employment and Unemployment Fluctuations."Oxford Economic Papers (1981): 42-60.

Bell, C. L. G. and Hazell, P. B. R. (Feb., 1980). Measuring the Indirect Effects of an Agricultural Investment Project on Its Surrounding Region, American Journal of Agricultural Economics, 62 (1), 75-86.

Bezdek, R. H. (1984). Test of Three Hypotheses Relating to the Leontief Input-Output Model, Journal of the Royal Statistical Society. Series A (General). 147 (3), 499-509.

Blanchard, Olivier Jean, and Lawrence F. Katz. "Regional evolutions."Brookings papers on economic activity 1.1992 (1992): 1-61.

Boarnet, Marlon G., Saksith Chalermpong, and Elizabeth Geho. "Specification issues in models of population and employment growth*." Papers in regional Science 84.1 (2005): 21-46.

Boisvert, R. N. (Feb., 1984). Decomposing the Induced Income Changes in Input-Output Models, American Journal of Agricultural Economics, 66 (1), 99-103.

Bon, R. (Nov., 1984). Comparative Stability Anlysis of Multiregional Input-Output Models: Column, Row, and Leontief-strout Gravity Coefficient Models, The Quarterly Journal of Economics, 99 (4), 791-815.

Brouwer, Erik, Alfred Kleinknecht, and Jeroen ON Reijnen. "Employment growth and innovation at the firm level." Journal of Evolutionary Economics 3.2 (1993): 153-159.

Buckley, P. H. (1992). A Transportation-oriented Interregional Computable General Equilibrium Model of the United States, The Annals of Regional Science, 26, 331-348.

Bureau of Labor Statistics, U.S. Department of Labor, 2014–2015 Occupational Outlook Handbook, [http://www.bls.gov/ooh/].

Burfisher, M. E., Robinson, S., Thierfelder, K. (Aug., 2002). Developing Countries and the Gains from Regionalism: Links between Trade and Farm Policy Reforms in Mexico, American Journal of Agricultural Economics, 84 (3), 736-748.

Cardonete, M. A. and Sancho, F. (2004). Sensitivity of CGE Simulation Results to Competing SAM Updates, The Review of Regional Studies, 34 (1), 37-56.

Carlino, Gerald A., and Edwin S. Mills. "THE DETERMINANTS OF COUNTY GROWTH*." Journal of Regional Science 27.1 (1987): 39-54.

Cella, Guido. "THE INPUT-OUTPUT MEASUREMENT OF INTERINDUSTRY LINKAGES*." Oxford Bulletin of Economics and Statistics 46.1 (1984): 73-84.

Chand, R. and Kaul, J. L. (Feb., 1986). A Note on the Use of the Cobb-Douglas Profit Function, American Journal of Agricultural Economics, 68 (1), 162-4.

Chong, Yock Y., and David F. Hendry. "Econometric evaluation of linear macro-economic models." The Review of Economic Studies 53.4 (1986): 671-690.

Cotton, John L., and Jeffrey M. Tuttle. "Employee turnover: A meta-analysis and review with implications for research." Academy of management Review 11.1 (1986): 55-70.

Cox, Robert W. "Social forces, states and world orders: beyond international relations theory." Millennium: journal of international studies 10.2 (1981): 126-155.

Datta, Deepak K., George E. Pinches, and V. K. Narayanan. "Factors influencing wealth creation from mergers and acquisitions: A meta-analysis."Strategic management journal 13.1 (1992): 67-84.

De Melo, Jaime, and Sherman Robinson. "Product differentiation and the treatment of foreign trade in computable general equilibrium models of small economies." Journal of International Economics 27.1 (1989): 47-67.

Diebold, Francis X., and Jose A. Lopez. "Forecast evaluation and combination." (1996).

Diebold, Francis X., and Robert S. Mariano. "Comparing predictive accuracy."Journal of Business & economic statistics 20.1 (2002).

United States Department of the Interior. (2010) “Salazar: Deepwater Drilling May Resume for Operators Who Clear Higher Bar for Safety, Environmental Protection” [Press Release]. Retrieved from http://www.doi.gov/news/pressreleases/Salazar-Deepwater-Drilling-May-Resume-for-Operators-Who-Clear-Higher-Bar-for-Safety-Environmental-Protection.cfm on 7/15/11.

Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010. Public L. No. 111-203 (2010).

Durand, Jorge, et al. "International migration and development in Mexican communities." Demography 33.2 (1996): 249-264.

Fair, Ray C., and Robert J. Shiller. "Comparing information in forecasts from econometric models." The American Economic Review (1990): 375-389.

Forni, M. and Paba, S. (2002). Spillovers and the Growth of Local Industries, The Journal of Industrial Economics, 50 (2), 151-71.

Fletcher, John E. "Input-output analysis and tourism impact studies." Annals of tourism research 16.4 (1989): 514-529.

Francois, J. F. (Nov., 1996). Labour Force Growth and Wages, The Economic Journal, 106 (439), 1586-1609.

Franklin, J. C. (Nov., 2007). An Overview of BLS Projections to 2016, Monthly Labor Review.

Frechtling, Douglas C., and Endre Horvath. "Estimating the multiplier effects of tourism expenditures on a local economy through a regional input-output model." Journal of travel research 37.4 (1999): 324-332.

Fullerton, Howard N. "Evaluating the BLS labor force projections to 2000."Monthly Labor Review 126.10 (2003): 3-12.

General Algebraic Modeling Systems. An Introduction to GAMS. Retrieved from http://www.gams.com/ on 7/15/11

Giesecke, J. (2002). Explaining Regional Economic Performance: An Historical Application of a Dynamic Multi-Regional CGE Model, Papers in Regional Science, 81, 247-278.

Goldstein, Harvey A., and Alvin M. Cruze. "An evaluation of State projections of industry, occupational employment." Monthly Labor Review (1987): 29-38.

Gomes de Menezes, A., Fortuna, M., Silva, F., Vieira, J. C. (2006). Computable General Equilibrium Models: A Literature Review, Centro de Estudos de Economia Aplicada do Atlantico (Working Paper Series), 6.

Goodrich, Calvin. "Goniobasis of the Coosa River, Alabama." (1936).

Goulder, L. H. and Eichengreen, B. (May, 1992). Trade Liberization in General Equilibrium: Intertemporal and Inter-Industry Effects, The Canadian Journal of Economics, 25 (2), 253-280.

Griffeth, Rodger W., Peter W. Hom, and Stefan Gaertner. "A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium." Journal of management 26.3 (2000): 463-488.

Hamilton, J. R., Whittlesey, N. K., Robinson, M. H., and Ellis, J. (May, 1991). Impacts, Value-added, and Benefits in Regional Project Analysis, American Journal of Agricultural Economics. 72 (2), 334-44.

Heady, E. O. and Schnittker, J. A. (Aug., 1957). Application of Input-Output Models to Agriculture, Journal of Farm Economics, 39 (3), 745-58.

Hecker, Daniel E. "Occupational employment projections to 2014." Monthly Lab. Rev. 128 (2005): 70.

Hendrickson, Chris. "A note on trends in transit commuting in the United States relating to employment in the central business district." Transportation Research Part A: General 20.1 (1986): 33-37.

Higgs, P. J., Parmenter, B. R., and Rimmer, R. J. (1988). A Hybrid Top-Down, Bottom Up Regional Computable General Equilibrium Model, International Regional Science Review, 11, 317.

Hom, Peter W., and Rodger W. Griffeth. Employee turnover. South-Western College Pub., 1994.

Horrell, S., Humphries, J., and Weale, M. (Aug., 1994). An Input-Output Table for 1841, The Economic History Review, New Series 47 (3), 545-66.

Impact Analysis for Planning (IMPLAN). Reference Manual (User's Guide to IMPLAN Version 3.0 Software). Retrieved from http://implan.com/V4/index.php?option=com_multicategories&view=categories&cid=222:referencemanualusersguidetoimplanversion30software&Itemid=10 on 7/15/11

Israilevich, Philip R., et al. "Forecasting Structural Change With a Regional Econometric Input-Output Model." Journal of Regional science 37.4 (1997): 565-590.

Isserman, Andrew M. "" Regional Employment Multiplier: A New Approach": Comment." Land Economics (1975): 290-293.

Jedidi, K., Jagpal, S., and Manchanda, P. (winter, 2003). Measuring Heterogeneous Reservation Prices for Product Bundles, Marketing Science, 22 (1), 107-30.

Johansen, Leif. "A multisectoral model of economic growth." (1960).

Jones, R., and Whalley, J. (Dec., 1986). A Canadian Regional General Equilibrium Model and Some Applications, Journal of Urban Economics, 25, 368-404.

Judge, George G. Introduction to the Theory and Practice of Econometrics. New York: Wiley, c1988, 1988.

Julia-Wise, R., Cooke, S. C., and Holland, D. (2002). A Computable General Equilibrium Analysis of a Property Tax Limitation Initiative in Idaho, Land Economics, 78 (2), 207-27.

Kilkenny, M. (Nov., 1993). Rural/Urban Effects of Terminating Farm Subsidies, American Journal of Agircultural Economics, 75 (4), 968-980.

Killkenny, M. (Fall, 1999). Interregional Fiscal Accounting, Growth and Change, 30 (4), 567-589.

Kilkenny, M. and Otto, D. (Dec., 1994). A General Equilibrium Perspective on Structural Change in the Rural Economy, American Journal of Agricultural Economics, 76 (5), 1130-7.

Kim, E. and Kim, K. (2002). Impacts of Regional Development Strategies on Growth and Equity of Korea: A Multiregional CGE Model, The Annals of Regional Science, 36, 165-180.

Kimbell, Larry J., and Glenn W. Harrison. General equilibrium analysis of regional fiscal incidence. Department of Economics, University of Western Australia, 1982.

Kraybill, D. S., Johnson, T. G., and Orden, D. (Aug., 1992). Macroeconomic Imbalances: a Multiregional General Equilibrium Analysis, American Journal of Agricultural Economics, 74 (3), 726-736.

Krueger, A. B. and Summers, L. H. (Mar., 1988). Efficiency Wages and the Inter-Industry Wage Stucture, Econometrica, 56 (2), 259-93.

Leontief, Wassily W. "Quantitative input and output relations in the economic systems of the United States." The review of economic statistics (1936): 105-125.

Leontief, Wassily W. "structure of American economy, 1919-1929." (1941).

Levernier, William, Mark D. Partridge, and Dan S. Rickman. "The Causes of Regional Variations in US Poverty: A Cross-County Analysis." Journal of Regional Science 40.3 (2000): 473-497.

Liew, L. H. (1984). A Johansen Model for Regional Analysis, Regional Science and Urban Economics, 14, 129-146.

Little, C. H. and Doeksen, G. A. (Nov., 1968). Measurement of Leakage by the Use of and Input-Output Model, American Journal of Agricultural Economics, 50 (4), 921-34.

Lofgren, H., Harris, R. L., and Robinson, S. (2002). A Standard Computable General Equilibrium (CGE) Model in GAMS, International Food Policy Research Institute.

Louisiana Workforce Commission. Louisiana Workforce Investment Council. WIC Strategic Plan. April 2010. Retrieved from http://www.laworks.net/PublicRelations/WIC_Main.asp on 7/15/11.

Martin, R. and Sunley, P. (Jul., 1996). Paul Krugman's Geographical Economics and Its Implications for Regional Development Theory: A Critical Assessment, Economic Geography, 72 (3), 259-92.

Mathur, Vijay K., and Harvey S. Rosen. "Regional employment multiplier: a new approach." Land Economics (1974): 93-96.

McCracken, Lance M. "Learning to live with the pain: acceptance of pain predicts adjustment in persons with chronic pain." Pain 74.1 (1998): 21-27.

McNicoll, I. H. and Baird, R. G. (Nov., 1980). Empirical Applications of Regional Input-Output Analysis: A Case Study of Shetland, The Journal of the Operational Research Society, 31 (11), 983-91.

de Melo, J. and Tarr, D. (1992). A General Equilibrium Analysis of US Foreign Trade Policy, The MIT Press: Boston.

Miller, Ronald E., and Peter D. Blair. Input-output analysis: foundations and extensions. Cambridge University Press, 2009.Nicholson, W. (1995) Microeconomic Theory: Basic Principles and Extensions (sixth edition), The Dryden Press: Fort Worth.

Montoya-Weiss, Mitzi M., and Roger Calantone. "Determinants of new product performance: a review and meta-analysis." Journal of product innovation management 11.5 (1994): 397-417.

Norrie, Kenneth Harold, and Michael Percy. Economic Rents, Province-building, and Interregional Adjustment: A Two-region General Equilibrium Analysis. Economic Council of Canada, 1983.

Oosterhaven, Jan, et al. "Evaluation of Non-Survey International IO Construction Methods with the Asian-Pacific Input-Output Table." (2007).

Park, Se-Hark. "LEAST SQUARES ESTIMATES OF THE REGIONAL EMPLOYMENT MULTIPLIER: AN APPRAISAL*." journal of Regional Science10.3 (1970): 365-374.

Parsons, Donald O. "The decline in male labor force participation." The Journal of Political Economy 88.1 (1980): 117.

Partridge, Mark D., and Dan S. Rickman. "The dispersion of US state unemployment rates: the role of market and non-market equilibrium factors."Regional Studies 31.6 (1997): 593-606.

Partridge, Mark D., and Dan S. Rickman. "Regional computable general equilibrium modeling: a survey and critical appraisal." International Regional Science Review 21.3 (1998): 205-248.

Partridge, M.D., and Rickman, D. S. (2007). CGE Modeling for Regional Economic Development Analysis, State of the Art in Regional and Urban Modeling, ECOMOD, Seville, Spain.

Pesaran, M. Hashem. "On the general problem of model selection." The Review of Economic Studies (1974): 153-171.

Pesaran, M. Hashem, and Angus S. Deaton. "Testing non-nested nonlinear regression models." Econometrica: Journal of the Econometric Society (1978): 677-694.

Pindyck, Robert S., and Daniel L. Rubinfeld. Econometric models and economic forecasts. Vol. 2. New York: McGraw-Hill, 1981.

Prescott, E. C. (Nov., 1995). Comment on the Computable General Equilibrium Models and Monetary Policy Advice, Journal of Money, Credit and Banking. 27 (4), Part 2: Liquidity, Monetary Policy, and Financial Intermediation, 1502-5.

Psacharopoulos, G. (Nov., 1973). An Input-Output Model for the Assessment of Labor Skill Requirements, Higher Education, 2 (4), 461-474.

Pyatt, F. Graham, and Jeffery I. Round. "Accounting and fixed price multipliers in a social accounting matrix framework." Economic Journal 89.356 (1979): 850-73.

Rickman, D. S. (2002). A Bayesian forecasting approach to constructing regional Input-Output based employment multipliers, Papers in Regional Science, 81, 483-498.

Rivers, Douglas, and Quang Vuong. "Model selection tests for nonlinear dynamic models." The Econometrics Journal 5.1 (2002): 1-39.

Reid, L. W. and Rubin, B. A. (summer, 2003). Integrating Economic Dualism and Labor Market Segmentation: The Effects of Race, Gender, and Structural Location on Earnings, 1974-2000, The Sociological Quarterly, 44 (3), 405-32.

Robinson, S, Kilkenny, M, and Hanson, K, The USDA/ERS Computable General Equilibrium (CGE) Model of the United States. Tech. rep., USDA/ERS, 1990.

Rose, Adam. "Input-output economics and computable general equilibrium models." Structural Change and Economic Dynamics 6.3 (1995): 295-304.

Rosen, H. S., and V. K. Mathur. "An econometric technique versus traditional techniques for obtaining regional employment multipliers: a comparative study."Environment and Planning 5 (1973): 273-282.

Rosenthal, Neal H., et al. "Evaluating the 1995 BLS Projections." Monthly Labor Review 120.9 (1997): 3-31.

Rosenthal, Neal H. "Quality of BLS Projections: A Historical Account, The."Monthly Lab. Rev. 122 (1999): 27.

Sadoulet, E., and A. de Janvry. Quantitative Development Policy Analysis. Baltimore: The Johns Hopkins University Press, 1995.

Saltelli, A., Tarantola, S., and Campolongo, F. (Nov., 2000). Sensitivity as an Ingredient of Modeling, Statistical Science, 15 (4), 377-95.

Scarf, H. (1973). The Computation of Economic Equilibria. New Haven and London, Yale University Press

Seung, C. K., and Kraybill, D. S. (2001). The Effects of Infrastructure Investment: A Two-Sector Dynamic Computable General Equilibrium Analysis for Ohio, International Regional Science Review, 24, 261.

Seung, K. S., Harris, T.R., and MacDiarmid, T.R. (1997). Economic impacts of surface water reallocation policies: A comparison of supply-determined SAM and CGE models, Journal of Regional Analysis & Policy, 27 (2), 55-76.

Shoven, John B. Applying general equilibrium. Cambridge university press, 1992.

Shoven, John B., and John Whalley. "Applied general-equilibrium models of taxation and international trade: an introduction and survey." Journal of Economic Literature (1984): 1007-1051.

Simon, Curtis J. "Human capital and metropolitan employment growth." Journal of Urban Economics 43.2 (1998): 223-243.

Simpson, J. R. and Adams, J. W. (Nov., 1975). Disaggregation of Input-Output Models into Product Lines as an Economic Development Policy Tool, American Journal of Agricultural Economics, 57 (4), 584-90.

Stanley, Tom D., and Stephen B. Jarrell. "Meta-Regression analysis: A quantitative method of literature surveys." Journal of Economic Surveys 3.2 (1989): 161-170.

Stekler, H. O., and Rupin Thomas. "Evaluating BLS labor force, employment, and occupation projections for 2000." Monthly Lab. Rev. 128 (2005): 46.

Steed, G. P. F. and Thomas, M. D. (Jun., 1971). Regional Industrial Change: Northern Ireland, Annals of the Association of American Geographers, 61 (2), 344-60.

Sweeney, S. H. (2004). Regional Occupational Employment Projections: Modeling Supply Constraints in the Direct-Requirements Approach, Journal of Regional Science, 44 (2), 263-288.

Sweeney, S. H., and Goldstein, H. A. (2005). Accounting for Migration in Regional Occupational Employment Projections, Annals of Regional Science, 39, 297-316.

Taylor, J. E., Yunez-Naude, A., and Dyer, G. (Aug., 1999). Agricultural Price Policy, Employment, and Migration in a Diversified Rural Economy: A Village-Town CGE Analysis from Mexico, American Journal of Agricultural Economics, 81 (3), 653-62.

Tiebout, C. M. (Oct., 1957). Regional and Interregional Input-Output Models: An Appraisal, Southern Economic Journal, 24 (2), 140-7.

Tiebout, C. M. (Aug., 1969). An Empirical Regional Input-Output Projection Model: The State of Washington 1980, The Review of Economics and Statistics, 51 (3), 334-340.

Tijskens, L. M. M., M. L. A. T. M. Hertog, and Bart M. Nicolaï, eds. Food process modelling. Vol. 59. Woodhead Publishing, 2001.

Tokle, J. G. and Huffman, W. E. (Aug., 1991). Local Economic Conditions and Wage Labor Decisions of Farm and Rural Nonfarm Couples, American Journal of Agricultural Economics, 73 (3), 652-70.

Toossi, Mitra. "New Look at Long-Term Labor Force Projections to 2050", A. Monthly Lab. Rev. 129 (2006): 19.

Tretyakova, A. and Birman, I. (Apr., 1976). Input-Output Analysis in the USSR, Soviet Studies, 28 (2), 157-86.

Treyz, George I. REGIONAL ECONOMIC MODELING. A SYSTEMATIC APPROACH TO FORECASTING AND POLICY ANALYSIS. 1993.

Treyz, George I., et al. "The dynamics of US internal migration." The review of Economics and Statistics (1993): 209-214.

Tyre, M. J. and Hauptman, O. (Aug., 1992). Effectiveness of Organizational Responses to Technological Change in the Production Process, Organizational Science, 3 (3), Focused Issue: Management of Technology, 301-20.

Van Houwelingen, Hans C., Koos H. Zwinderman, and Theo Stijnen. "A bivariate approach to meta-analysis." Statistics in medicine 12.24 (1993): 2273-2284.

Van Houwelingen, Hans C., Lidia R. Arends, and Theo Stijnen. "Advanced methods in meta-analysis: multivariate approach and meta-regression."Statistics in medicine 21.4 (2002): 589-624.

Vargas, E., Schreiner, D., Tembo, G., and Marcouiller, D. (1999). Computable General Equilibrium Modeling for Regional Analysis, The Web Book of Regional Science, (http://rri.wvu.edu/WebBook/Schreiner/contents.htm).

Vernon, Raymond. "International investment and international trade in the product cycle." (1966).

Vuong, Quang H. "Likelihood ratio tests for model selection and non-nested hypotheses." Econometrica: Journal of the Econometric Society (1989): 307-333.

Waters, E. C., Holland, D. W., and Weber, B. A. (1997). Economic Impacts of a Property Tax Limitation: A Computable General Equilibrium Analysis of Oregon's Measure 5, Land Economics, 73 (1), 72-89.

Watson, Audrey L. "Implementing the 2010 Standard Occupational Classification in the Occupational Employment Statistics program." (2013).

Weiss, Steven J., and Edwin C. Gooding. "Estimation of differential employment multipliers in a small regional economy." Land Economics (1968): 235-244.

Wiese, A. M. (spring-summer, 2000). How Unilateral Economic Sanctions Affect the Farm Economy: An Inter-Industry Analysis, Review of Agricultural Economics, 22 (1), 55-76.

Wolsky, A. M. (May, 1984). Disaggregating Input-Output Models, The Review of Economics and Statistics, 66 (2), 283-91.

Wyatt, Ian D. "Evaluating the 1996-2006 employment projections." Monthly Labor Review 133.9 (2010).

Appendix A: Computable General Equilibrium

I) Visual Representation and Discussion

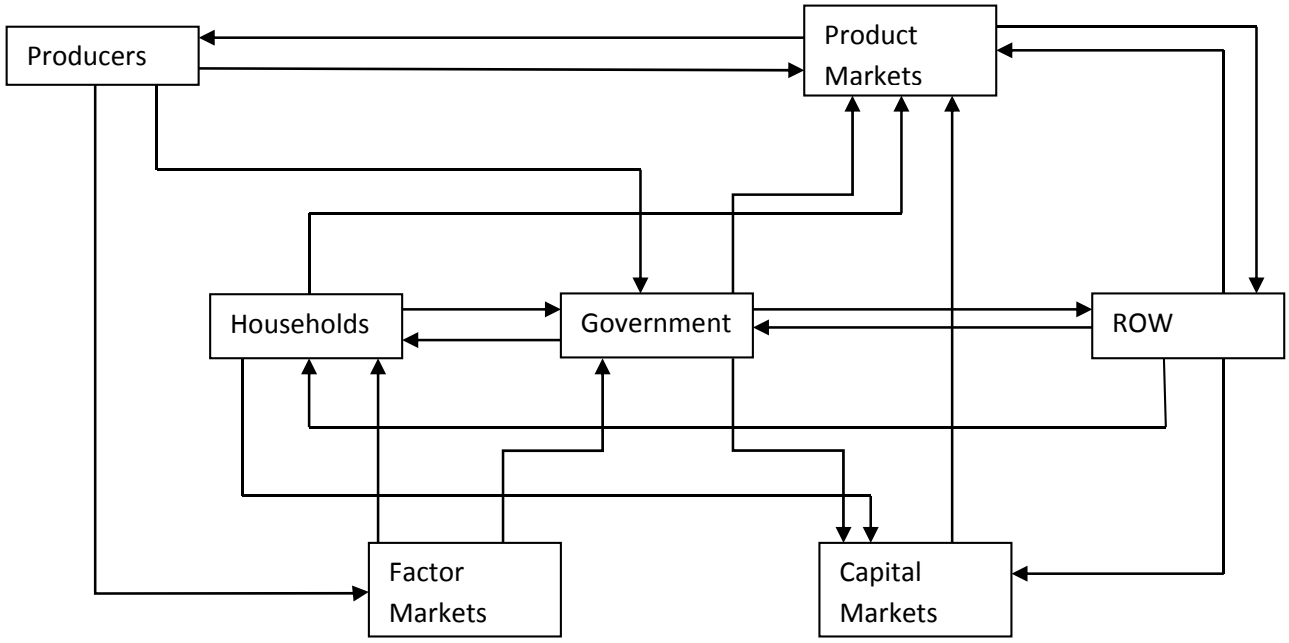


Figure A1: CGE Structure

Figure A1 is a visual representation of a Computable General Equilibrium (CGE) model. The boxes represent agents or markets and the lines represent transaction between the agents, between the markets, or between an agent and a market. The equations that comprise a CGE govern the transactions that occur between these agents and markets and are presented in the next section of the appendix.

The Producers use their production process to create commodities. These commodities can be sold in Product Markets which return to the producers their domestic sales. This transaction is noted by the lines going from Producers to Product Markets (intermediate demand) and from

Product Markets to Producers (domestic sales). The production process produces value-added which is used to acquire factors or production from Factor Markets and is represented by the flow from Producers to Factor Markets (Value-added). Producers also pay output taxes to Government.

Consider next the Product Markets. It has already been shown that the Product Market provides meets intermediate demands and facilitates domestic sales for Producers, but Households, Government, and Capital Markets also consume commodities which the purchase in the Product Markets. These transactions are represented by the lines from Households (private consumption), Government (gov't consumption), and Capital Markets (investment demand) respectively to Product Markets. Domestic supply and demand is supplemented by foreign producers who both demand and supply domestic products, as represented in the flows from the Rest of the World (ROW) accounts. Imports flow into Product Markets from the ROW while exports flow out.

Households receive income from Factor Markets and purchase goods from Product Markets for consumption, pay income taxes to the Government, and save in Capital Markets. Households also receive Government payments and payments from ROW for factors of production provided outside the region or for foreign investments.

Government receives taxes from Producers (output taxes), Factor Markets (factor taxes), Households (income taxes), and ROW (tariffs and foreign debt). The Government spends this income on government consumption in the Product Markets, government payments to Households, government savings in Capital Markets, and debt payments to the ROW.

Factor Markets are paid value-added from Producers and pay income to Households which provide the factors. Further, Factor Markets pay the factor taxes to Government.

Capital Markets receive investment from Households (private savings), Government (gov't davings), and ROW (foreign investment) while providing investment demand to Product Markets.

The ROW account exchanges imports and exports with the Product Markets, pays tariffs and debt payments to Government, makes foreign payments to Households, and provides foreign investment to Capital Markets.

This figure, and the accompanying discussion, covers the theoretical connections between agents and markets that are covered within the framework of the CGE. The equations that govern these transactions are presented in the next section of this appendix.

II) Equations

The equations presented here follow Lofgren et al. (2002) and are the structural equations for all CGE models in this dissertation.

The structure of this section follows the discussion in the primary body which separates the equations into blocks which contain equations that deal with specific transactions within the CGE: Prices, Production and Trade, Institutions, and Constraints. The equations will be presented in these blocks separately.

Price Block

Import Price $$PM_c = pwm_c \cdot (1+tm_c) \cdot EXR + \sum_{c' \in CT} PQ_{c'} \cdot icm_{c'c}$$

$$\begin{bmatrix} import \\ price \\ (LCU) \end{bmatrix} = \begin{bmatrix} import \\ price \\ (FCU) \end{bmatrix} \cdot \begin{bmatrix} tariff \\ adjust- \\ ment \end{bmatrix} \cdot \begin{bmatrix} exchange\ rate \\ (LCU\ per \\ FCU) \end{bmatrix} + \begin{bmatrix} cost\ of\ trade \\ inputs\ per \\ import\ unit \end{bmatrix} \qquad c \in CM \quad (1)$$

where

$c \in C$	=	a set of commodities (also referred to as c' and C'),
$c \in CM\ (\subset C)$	=	a set of imported commodities,
$c \in CT\ (\subset C)$	=	a set of domestic trade inputs (distribution commodities),
PM_c	=	import price in LCU (local-currency units) including transaction costs,
pwm_c	=	c.i.f. import price in FCU (foreign-currency units),
tm_c	=	import tariff rate,
EXR	=	exchange rate (LCU per FCU),
PW_c	=	composite commodity price (including sales tax and transaction costs), and
$icm_{c'c}$	=	quantity of commodity c' as trade input per imported unit of c.

Export Price $$PE_c = pwe_c \cdot (1-te_c) \cdot EXR - \sum_{c' \in CT} PQ_{c'} \cdot ice_{c'c}$$

$$\begin{bmatrix} export \\ price \\ (LCU) \end{bmatrix} = \begin{bmatrix} export \\ price \\ (FCU) \end{bmatrix} \cdot \begin{bmatrix} tariff \\ adjust- \\ ment \end{bmatrix} \cdot \begin{bmatrix} exchange\ rate \\ (LCU\ per \\ FCU) \end{bmatrix} - \begin{bmatrix} cost\ of\ trade \\ inputs\ per \\ export\ unit \end{bmatrix} \qquad c \in CE \quad (2)$$

where

$c \in CE\ (\subset C)$	=	a set of exported commodities (with domestic production),
PE_c	=	export price (LCU),
pwe_c	=	f.o.b. export price (FCU),
te_c	=	export tax rate,
$ice_{c'c}$	=	quantity of commodity c' as trade input per exported unit of c.

Demand Price of Domestic Non traded Goods

$$PDD_c = PDS_c + \sum_{c' \in CT} PQ_{c'} \cdot icd_{c'c}$$

$$\begin{bmatrix} \text{domestic demand price} \end{bmatrix} = \begin{bmatrix} \text{domestic supply price} \end{bmatrix} + \begin{bmatrix} \text{cost of trade inputs per unit of domestic sales} \end{bmatrix} \qquad c \in CD \quad (3)$$

where

$c \in CD\ (\subset C)$ = a set of commodities with domestic sales of domestic output,

PDD_c = demand price for commodity produced and sold domestically,

PDS_c = supply price for commodity produced and sold domestically, and

$icd_{c'c}$ = quantity of commodity c' as trade input per unit of c produced and sold domestically.

Absorption

$$PQ_c \cdot (1 - tq_c) \cdot QQ_c = PDD_c \cdot QD_c + PM_c \cdot QM_c$$

$$\begin{bmatrix} \text{absorption (at demand prices net of sales tax)} \end{bmatrix} = \begin{bmatrix} \text{domestic demand price times domestic sales quantity} \end{bmatrix} + \begin{bmatrix} \text{import price times import quantity} \end{bmatrix} \qquad c \in (CD \cup CM) \quad (4)$$

where

QQ_c = quantity of goods supplied to domestic market (composite supply),

QD_c = quantity sold domestically of domestic output,

QM_c = quantity of imports of commodity, and

tq_c = rate of sales tax (as share of composite price inclusive of sales tax).

Marketed Output Value

$$PX_c \cdot QX_c = PDS_c \cdot QD_c + PE_c \cdot QE_c$$

$$\begin{bmatrix} \text{producer price times marketed output quantity} \end{bmatrix} = \begin{bmatrix} \text{domestic supply price times domestic sales quantity} \end{bmatrix} + \begin{bmatrix} \text{export price times export quantity} \end{bmatrix} \qquad c \in CX \quad (5)$$

where

PX_c = aggregate producer price for commodity,

QX_c = aggregate marketed quantity of domestic output of commodity,

QE_c = quantity of exports, and

$c \in CX\ (\subset C)$ = a set of commodities with domestic output.

Activity Price

$$PA_a = \sum_{c \in C} PXAC_{a\,c} \cdot \theta_{a\,c}$$

$$\begin{bmatrix} \text{activity} \\ \text{price} \end{bmatrix} = \begin{bmatrix} \text{producer prices} \\ \text{times yields} \end{bmatrix} \qquad a \in A \quad (6)$$

where

$a \in A$ = a set of activities,
PA_a = activity price (gross revenue per activity unit),
$PXAC_{a\,c}$ = producer price of commodity c for activity a, and
$\theta_{a\,c}$ = yield of output c per unit of activity a.

Aggregate Intermediate Input Price

$$PINTA_a = \sum_{c \in C} PQ_c \cdot ica_{c\,a}$$

$$\begin{bmatrix} \text{aggregate} \\ \text{intermediate} \\ \text{input price} \end{bmatrix} = \begin{bmatrix} \text{intermediate input cost} \\ \text{per unit of aggregate} \\ \text{intermediate input} \end{bmatrix} \qquad a \in A \quad (7)$$

where

$PINTA_a$ = aggregate intermediate input price for activity a, and
$ica_{c\,a}$ = quantity of c per unit of aggregate intermediate input a.

Activity Revenue and Costs

$$PA_a \cdot (1 - ta_a) \cdot QA_a = PVA_a \cdot QVA_a + PINTA_a \cdot QINTA_a$$

$$\begin{bmatrix} \text{activity price} \\ \text{(net of taxes)} \\ \text{times activity level} \end{bmatrix} = \begin{bmatrix} \text{value-added} \\ \text{price times} \\ \text{quantity} \end{bmatrix} + \begin{bmatrix} \text{aggregate} \\ \text{intermediate} \\ \text{input price times} \\ \text{quantity} \end{bmatrix} \qquad a \in A \quad (8)$$

where

ta_a = tax rate for activity,
QA_a = quantity (level) of activity,
QVA_a = quantity of (aggregate) value-added,
$QINTA_a$ = quantity of aggregate intermediate input, and
PVA_a = price of (aggregate) value-added.

Consumer Price Index

$$\overline{CPI} = \sum_{c \in C} PQ_c \cdot cwts_c$$

$$\begin{bmatrix} \text{consumer} \\ \text{price index} \end{bmatrix} = \begin{bmatrix} \text{prices times} \\ \text{weights} \end{bmatrix} \qquad (9)$$

where

$cwts_c$ = weight of commodity c in the consumer price index, and
$\overline{CPI}$ = consumer price index (exogenous variable).

Producer Price Index for Nontraded Market Output

$$DPI = \sum_{c \in C} PDS_c \cdot dwts_c \qquad (10)$$

$$\begin{bmatrix} \text{producer price index} \\ \text{for non-traded outputs} \end{bmatrix} = \begin{bmatrix} \text{prices times} \\ \text{weights} \end{bmatrix}$$

where

$dwts_c$ = weight of commodity c in the producer price index, and
DPI = producer price index for domestically marketed output.

Production and Trade Block

CES Technology: Activity Production Fucntion

$$QA_a = \alpha_a^a \cdot \left(\delta_a^a \cdot QVA_a^{-\rho_a^a} + (1-\delta_a^a) \cdot QINTA_a^{-\rho_a^a} \right)^{-\frac{1}{\rho_a^a}} \qquad a \in ACES \quad (11)$$

$$\begin{bmatrix} \text{activity} \\ \text{level} \end{bmatrix} = CES \begin{bmatrix} \text{quantity of aggregate value added,} \\ \text{quantity of aggregate intermediate input} \end{bmatrix}$$

CES Technology: Value-Added–Intermediate-Input Ratio

$$\frac{QVA_a}{QINTA_a} = \left(\frac{PINTA_a}{PVA_a} \cdot \frac{\delta_a^a}{1-\delta_a^a} \right)^{\frac{1}{1-\rho_a^a}} \qquad a \in ACES \quad (12)$$

$$\begin{bmatrix} \text{value-added} - \\ \text{intermediate-} \\ \text{input quantity} \\ \text{ratio} \end{bmatrix} = f \begin{bmatrix} \text{intermediate-input :} \\ \text{value-added} \\ \text{price ratio} \end{bmatrix}$$

where

$a \in ACES (\subset A)$ = a set of activities with a CES function at the top of the technology nest,
a_a^a = efficiency parameter in the CES activity function,
δ_a^a = CES activity function share parameter, and
ρ_a^a = CES activity function exponent.

Leontief Technology: Demand for Aggregate Value-Added

$$QVA_a = iva_a \cdot QA_a \qquad a \in ALEO \quad (13)$$

$$\begin{bmatrix} \text{demand for} \\ \text{value added} \end{bmatrix} = f \begin{bmatrix} \text{activity} \\ \text{level} \end{bmatrix}$$

Leontief Technology: Demand for Aggregate Intermediate Input

$$QINTA_a = inta_a \cdot QA_a \qquad a \in ALEO \quad (14)$$

$$\begin{bmatrix} \text{demand for aggregate} \\ \text{intermediate input} \end{bmatrix} = f \begin{bmatrix} \text{activity} \\ \text{level} \end{bmatrix}$$

where

$a \in ALEO (\subset A)$ = a set of activities with a Leontief function at the top of the technology nest,
iva_a = quantity of value-added per activity unit, and
$inta_a$ = quantity of aggregate intermediate input per activity unit.

Value-Added and Factor Demands

$$QVA_a = \alpha_a^{va} \cdot \left(\sum_{f \in F} \delta_{fa}^{va} \cdot QF_{fa}^{-\rho_a^{va}} \right)^{-\frac{1}{\rho_a^{va}}} \qquad a \in A \quad (15)$$

$$\begin{bmatrix} \text{quantity of aggregate} \\ \text{value added} \end{bmatrix} = CES \begin{bmatrix} \text{factor} \\ \text{inputs} \end{bmatrix}$$

Factor Demand

$$WF_f \cdot \overline{WFDIST}_{fa} = PVA_a (1 - tva_a) \cdot QVA_a \cdot \left(\sum_{f \in F'} \delta_{fa}^{va} \cdot QF_{fa}^{-\rho_a^{va}} \right)^{-1} \cdot \delta_{fa}^{va} \cdot QF_{fa}^{-\rho_a^{va}-1} \qquad \begin{matrix} a \in A \\ f \in F \end{matrix} \quad (16)$$

$$\begin{bmatrix} \text{marginal cost of} \\ \text{factor } f \text{ in activity } a \end{bmatrix} = \begin{bmatrix} \text{marginal revenue product} \\ \text{of factor } f \text{ in activity } a \end{bmatrix}$$

where

$f \in F (= F')$ = a set of factors,
tva_a = rate of value-added tax for activity a,
α_a^{va} = efficiency parameter in the CES value-added function,
δ_{fa}^{va} = CES value-added function share parameter for factor f in activity a,
QF_{fa} = quantity demanded of factor f from activity a,
ρ_a^{va} = CES value-added function exponent,
WF_f = average price of factor, and
$\overline{WFDIST}_{fa}$ = wage distortion factor for factor f in activity a (exogenous variable).

Disaggregated Intermediate Input Demand

$$QINT_{c\,a} = ica_{c\,a} \cdot QINTA_a \qquad \begin{matrix} a \in A \\ c \in C \end{matrix} \quad (17)$$

$$\begin{bmatrix} \text{intermediate demand for commodity } c \text{ from activity } a \end{bmatrix} = f \begin{bmatrix} \text{aggregate intermediate input quantity for activity } a \end{bmatrix}$$

where

$QINT_{c\,a}$ = quantity of commodity c as intermediate input to activity a.

Commodity Production and Allocation

$$QXAC_{a\,c} + \sum_{h \in H} QHA_{a\,c\,h} = \theta_{a\,c} \cdot QA_a \qquad \begin{matrix} a \in A \\ c \in CX \end{matrix} \quad (18)$$

$$\begin{bmatrix} \text{marketed quantity of commodity } c \text{ from activity } a \end{bmatrix} + \begin{bmatrix} \text{household home consumption of commodity } c \text{ from activity } a \end{bmatrix} = \begin{bmatrix} \text{production of commodity } c \text{ from activity } a \end{bmatrix}$$

where

$QXAC_{a\,c}$ = marketed output quantity of commodity c from activity a, and

$QHA_{a\,c\,h}$ = quantity of household home consumption of commodity c from activity a for household h.

Output Aggregation Function

$$QX_c = \alpha_c^{ac} \cdot \left(\sum_{a \in A} \delta_{a\,c}^{ac} \cdot QXAC_{a\,c}^{\,-\rho_c^{ac}} \right)^{-\frac{1}{\rho_c^{ac}+1}} \qquad c \in CX \quad (19)$$

$$\begin{bmatrix} \text{aggregate marketed production of commodity } c \end{bmatrix} = CES \begin{bmatrix} \text{activity-specific marketed production of commodity } c \end{bmatrix}$$

where

α_c^{ac} = shift parameter for domestic commodity aggregation function,

$\delta_{a\,c}^{ac}$ = share parameter for domestic commodity aggregation function, and

ρ_c^{ac} = domestic commodity aggregation function exponent.

First-Order Condition for Output Aggregation Function

$$PXAC_{a\,c} = PX_c \cdot QX_c \left(\sum_{a \in A'} \delta_{a\,c}^{ac} \cdot QXAC_{a\,c}^{\,-\rho_c^{ac}} \right)^{-1} \cdot \delta_{a\,c}^{ac} \cdot QXAC_{a\,c}^{\,-\rho_c^{ac}-1} \qquad \begin{matrix} a \in A \\ c \in CX \end{matrix} \quad (20)$$

$$\begin{bmatrix} \text{marginal cost of commodity } c \text{ from activity } a \end{bmatrix} = \begin{bmatrix} \text{marginal revenue product of commodity } c \text{ from activity } a \end{bmatrix}$$

Output Transformation (CET) Function

$$QX_c = \alpha_c^t \cdot \left(\delta_c^t \cdot QE_c^{\rho_c^t} + (1-\delta_c^t) \cdot QD_c^{\rho_c^t}\right)^{\frac{1}{\rho_c^t}} \qquad c \in (CE \cap CD) \quad (21)$$

$$\begin{bmatrix}\text{aggregate marketed}\\ \text{domestic output}\end{bmatrix} = CET\begin{bmatrix}\text{export quantity, domestic}\\ \text{sales of domestic output}\end{bmatrix}$$

where

α_c^t = a CET function shift parameter,
δ_c^t = a CET function share parameter, and
ρ_c^t = a CET function exponent.

Export-Domestic Supply Ratio

$$\frac{QE_c}{QD_c} = \left(\frac{PE_c}{PDS_c} \cdot \frac{1-\delta_c^t}{\delta_c^t}\right)^{\frac{1}{\rho_c^t - 1}} \qquad c \in (CE \cap CD) \quad (22)$$

$$\begin{bmatrix}\text{export-domestic}\\ \text{supply ratio}\end{bmatrix} = f\begin{bmatrix}\text{export-domestic}\\ \text{price ratio}\end{bmatrix}$$

Output Transformation for Domestically Sold Outputs Without Exports and for Exports Without Domestic Sales

$$QX_c = QD_c + QE_c \qquad c \in (CD \cap CEN) \cup (CE \cap CDN) \quad (23)$$

$$\begin{bmatrix}\text{aggregate}\\ \text{marketed}\\ \text{domestic output}\end{bmatrix} = \begin{bmatrix}\text{domestic market}\\ \text{sales of domestic}\\ \text{output [for}\\ c \in (CD \cap CEN)]\end{bmatrix} + \begin{bmatrix}\text{exports [for}\\ c \in (CE \cap CDN)]\end{bmatrix}$$

where

$c \in CEN\ (\subset C)$ = non-exported commodities (complement of CE), and
$c \in CDN\ (\subset C)$ = commodities without domestic market sales of domestic output (complement of CD).

Composite Supply (Armington) Function

$$QQ_c = \alpha_c^q \cdot \left(\delta_c^q \cdot QM_c^{-\rho_c^q} + (1-\delta_c^q) \cdot QD_c^{-\rho_c^q}\right)^{-\frac{1}{\rho_c^q}} \qquad c \in (CM \cap CD) \quad (24)$$

$$\begin{bmatrix}\text{composite}\\ \text{supply}\end{bmatrix} = f\begin{bmatrix}\text{import quantity, domestic}\\ \text{use of domestic output}\end{bmatrix}$$

where

α_c^q = an Armington function shift parameter,
δ_c^q = an Armington function share parameter, and
ρ_c^q = an Armington function exponent.

Import-Domestic Demand Ratio

$$\frac{QM_c}{QD_c} = \left(\frac{PDD_c}{PM_c} \cdot \frac{\delta_c^q}{1-\delta_c^q} \right)^{\frac{1}{1+\rho_c^q}} \qquad c \in (CM \cap CD) \quad (25)$$

$$\begin{bmatrix} \text{import-domestic} \\ \text{demand ratio} \end{bmatrix} = f \begin{bmatrix} \text{domestic-import} \\ \text{price ratio} \end{bmatrix}$$

Composite Supply for Non-imported Outputs and Non-produced Imports

$$QQ_c = QD_c + QM_c \qquad c \in (CD \cap CMN) \cup (CM \cap CDN) \quad (26)$$

$$\begin{bmatrix} \text{composite} \\ \text{supply} \end{bmatrix} = \begin{bmatrix} \text{domestic use of} \\ \text{marketed domestic} \\ \text{output [for} \\ c \in (CD \cap CMN)] \end{bmatrix} + \begin{bmatrix} \text{imports [for} \\ c \in (CM \cap CDN)] \end{bmatrix}$$

where
$c \in CMN\ (\subset C)$ = a set of non-imported commodities.

Demand for Transactions Services

$$QT_c = \sum_{c' \in C'} \left(icm_{c\,c'} \cdot QM_{c'} + ice_{c\,c'} \cdot QE_{c'} + icd_{c\,c'} \cdot QD_{c'} \right)$$

$$\begin{bmatrix} \text{demand for} \\ \text{transactions} \\ \text{services} \end{bmatrix} = \begin{bmatrix} \text{sum of demands} \\ \text{for imports, exports,} \\ \text{and domestic sales} \end{bmatrix} \qquad c \in CT \quad (27)$$

where
QT_c = quantity of commodity demanded as transactions service input.

Institution Block

Factor Income

$$YF_f = \sum_{a \in A} WF_f \cdot \overline{WFDIST}_{f\,a} \cdot QF_{f\,a}$$

$$\begin{bmatrix} \text{income of} \\ \text{factor } f \end{bmatrix} = \begin{bmatrix} \text{sum of activity payments} \\ \text{(activity-specific wages} \\ \text{times employment levels)} \end{bmatrix} \qquad f \in F \quad (28)$$

where
YF_f = income of factor f.

Institutional Factor Incomes

$$YIF_{if} = shif_{if} \cdot \left[\left(1 - tf_f\right) \cdot YF_f - trnsfr_{row\,f} \cdot EXR \right]$$

$$\begin{bmatrix} \text{income of institution } i \text{ from factor } f \end{bmatrix} = \begin{bmatrix} \text{share of income of factor } f \text{ to institution } i \end{bmatrix} \cdot \begin{bmatrix} \text{income of factor } f \text{ (net of tax and tranzfer to RoW)} \end{bmatrix} \qquad \begin{matrix} i \in INSD \\ f \in F \end{matrix} \quad (29)$$

where

$i \in INS$ = a set of institutions (domestic and rest of the world),
$i \in INSD (\subset INS)$ = a set of domestic institutions,
YIF_{if} = income to domestic institution i from factor f,
$shif_{if}$ = share of domestic institution i in income of factor f,
tf_f = direct tax rate for factor f, and
$trnsfr_{if}$ = transfer from factor f to institution i.

Income of domestic, Nongovernment Institutions

$$YI_i = \sum_{f \in F} YIF_{if} + \sum_{i' \in INSDNG'} TRII_{ii'} + trnsfr_{i\,gov} \cdot \overline{CPI} + trnsfr_{i\,row} \cdot EXR$$

$$\begin{bmatrix} \text{income of institution } i \end{bmatrix} = \begin{bmatrix} \text{factor income} \end{bmatrix} + \begin{bmatrix} \text{tranzfers from other domestic non-government institutions} \end{bmatrix} + \begin{bmatrix} \text{tranzfers from government} \end{bmatrix} + \begin{bmatrix} \text{tranzfers from RoW} \end{bmatrix} \qquad i \in INSDNG \quad (30)$$

where

$i \in INSDNG (=INSDNG' \subset INSD)$
= a set of domestic nongovernment institutions,
YI_i = income of institution i (in the set INSDNG), and
$TRII_{ii'}$ = transfers from institution i' to i (both in the set INSDNG).

Infra-Institutional Transfers

$$TRII_{i\,i'} = shii_{i\,i'} \cdot \left(1 - MPS_{i'}\right) \cdot \left(1 - TINS_{i'}\right) \cdot YI_{i'}$$

$$\begin{bmatrix} \text{tranzfer from institution } i' \text{ to } i \end{bmatrix} = \begin{bmatrix} \text{share of net income of institution } i' \text{ tranzfered to } i \end{bmatrix} \cdot \begin{bmatrix} \text{income of institution } i', \text{ net of savings and direct taxes} \end{bmatrix} \qquad \begin{matrix} i \in INSDNG \\ i' \in INSDNG' \end{matrix} \quad (31)$$

where

$shii_{ii}$ = share of net income of i' to i ($i' \in INSDNG'$; $i \in INSDNG$),
MPS_i = marginal propensity to save for domestic nongovernment institution (exogenous variable), and
$TINS_i$ = direct tax rate for institution i ($i \in INSDNG$).

Household Consumption Expenditures

$$EH_h = \left(1 - \sum_{i \in INSDNG} shii_{i\,h}\right) \cdot \left(1 - MPS_h\right) \cdot \left(1 - TINS_h\right) \cdot YI_h \qquad h \in H \quad (32)$$

$$\left[\begin{array}{c}\textit{household income}\\ \textit{disposable for}\\ \textit{consumption}\end{array}\right] = \left[\begin{array}{c}\textit{household income, net of direct}\\ \textit{taxes, savings, and transfers to}\\ \textit{other non-government institutions}\end{array}\right]$$

where

$i \in H (\subset INSDNG)$ = a set of households, and

EH_h = household consumption expenditures.

Household Consumption Spending on Marketed Commodities

$$PQ_c \cdot QH_{c\,h} = PQ_c \cdot \gamma^m_{c\,h} + \beta^m_{c\,h} \cdot \left(EH_h - \sum_{c' \in C} PQ_{c'} \cdot \gamma^m_{c'\,h} - \sum_{a \in A} \sum_{c' \in C} PXAC_{a\,c'} \cdot \gamma^h_{a\,c'\,h}\right) \qquad \begin{array}{l} c \in C \\ h \in H \end{array} \quad (33)$$

$$\left[\begin{array}{c}\textit{household consumption}\\ \textit{spending on market}\\ \textit{commodity } c\end{array}\right] = f\left[\begin{array}{c}\textit{total household consumption}\\ \textit{spending, market price of c, and other}\\ \textit{commodity prices (market and home)}\end{array}\right]$$

where

$QH_{c\,h}$ = quantity of consumption of marketed commodity c for household h,

$\gamma^m_{c\,h}$ = subsistence consumption of marketed commodity c for household h,

$\gamma^h_{a\,c\,h}$ = subsistence consumption of home commodity c from activity a for household h, and

$\beta^m_{c\,h}$ = marginal share of consumption spending on marketed commodity c for household h.

Household Consumption Spending on Home Commodities

$$PXAC_{a\,c} \cdot QHA_{a\,c\,h} = PXAC_{a\,c} \cdot \gamma^h_{a\,c\,h} + \beta^h_{a\,c\,h} \cdot \left(EH_h - \sum_{c' \in C} PQ_{c'} \cdot \gamma^m_{c'\,h} - \sum_{a \in A} \sum_{c' \in C} PXAC_{a\,c'} \cdot \gamma^h_{a\,c'\,h}\right) \qquad \begin{array}{l} a \in A \\ c \in C \\ h \in H \end{array} \quad (34)$$

$$\left[\begin{array}{c}\textit{household consumption}\\ \textit{spending on home commodity}\\ \textit{c from activity a}\end{array}\right] = f\left[\begin{array}{c}\textit{total household consumption spending,}\\ \textit{producer price, and other}\\ \textit{commodity prices (market and home)}\end{array}\right]$$

where

$\beta^h_{a\,c\,h}$ = marginal share of consumption spending on home commodity c from activity a for household h.

Investment Demand

$$QINV_c = \overline{IADJ} \cdot \overline{qinv_c} \qquad c \in C \quad (35)$$

$$\begin{bmatrix} \text{fixed investment demand for commodity } c \end{bmatrix} = \begin{bmatrix} \text{adjustment factor times base-year fixed investment} \end{bmatrix}$$

where

$QINV_c$ = quantity of fixed investment demand for commodity,

$\overline{IADJ}$ = investment adjustment factor (exogenous variable), and

$\overline{qinv_c}$ = base-year quantity of fixed investment demand.

Government Consumption Demand

$$QG_c = \overline{GADJ} \cdot \overline{qg_c} \qquad c \in C \quad (36)$$

$$\begin{bmatrix} \text{government consumption demand for commodity } c \end{bmatrix} = \begin{bmatrix} \text{adjustment factor times base-year government consumption} \end{bmatrix}$$

where

QG_c = government consumption demand for commodity,

$\overline{GADJ}$ = government consumption adjustment factor (exogenous variable), and

$\overline{qg_c}$ = base-year quantity of government demand.

Government Revenue

$$\begin{aligned} YG = & \sum_{i \in INSDNG} TINS_i \cdot YI_i + \sum_{f \in F} tf_f \cdot YF_f + \sum_{a \in A} tva_a \cdot PVA_a \cdot QVA_a \\ & + \sum_{a \in A} ta_a \cdot PA_a \cdot QA_a + \sum_{c \in CM} tm_c \cdot pwm_c \cdot QM_c \cdot EXR + \sum_{c \in CE} te_c \cdot pwe_c \cdot QE_c \cdot EXR \\ & + \sum_{c \in C} tq_c \cdot PQ_c \cdot QQ_c + \sum_{f \in F} YIF_{gov\,f} + trnsfr_{gov\,row} \cdot EXR \end{aligned}$$

$$\begin{aligned} \begin{bmatrix} \text{government revenue} \end{bmatrix} = & \begin{bmatrix} \text{direct taxes from institutions} \end{bmatrix} + \begin{bmatrix} \text{direct taxes from factors} \end{bmatrix} + \begin{bmatrix} \text{value-added tax} \end{bmatrix} \\ & + \begin{bmatrix} \text{activity tax} \end{bmatrix} + \begin{bmatrix} \text{import tariffs} \end{bmatrix} + \begin{bmatrix} \text{export taxes} \end{bmatrix} \\ & + \begin{bmatrix} \text{sales tax} \end{bmatrix} + \begin{bmatrix} \text{factor income} \end{bmatrix} + \begin{bmatrix} \text{tranzfers from RoW} \end{bmatrix} \end{aligned} \qquad (37)$$

where

YG = government revenue.

Government Expenditure

$$EG = \sum_{c \in C} PQ_c \cdot QG_c + \sum_{i \in INSDNG} trnsfr_{i\ gov} \cdot \overline{CPI} \tag{38}$$

$$\begin{bmatrix} \text{government} \\ \text{spending} \end{bmatrix} = \begin{bmatrix} \text{government} \\ \text{consumption} \end{bmatrix} + \begin{bmatrix} \text{transfers to domestic} \\ \text{non-government} \\ \text{institutions} \end{bmatrix}$$

where

EG = government expenditures.

System Constraint Block

Factor Markets

$$\sum_{a \in A} QF_{f\ a} = \overline{QFS}_f \qquad f \in F \tag{39}$$

$$\begin{bmatrix} \text{demand for} \\ \text{factor } f \end{bmatrix} = \begin{bmatrix} \text{supply of} \\ \text{factor } f \end{bmatrix}$$

where

$\overline{QFS}_f$ = quantity supplied of factor (exogenous variable).

Composite Commodity Markets

$$QQ_c = \sum_{a \in A} QINT_{c\ a} + \sum_{h \in H} QH_{c\ h} + QG_c + QINV_c + qdst_c + QT_c \qquad c \in C \tag{40}$$

$$\begin{bmatrix} \text{composite} \\ \text{supply} \end{bmatrix} = \begin{bmatrix} \text{intermediate} \\ \text{use} \end{bmatrix} + \begin{bmatrix} \text{household} \\ \text{consumption} \end{bmatrix} + \begin{bmatrix} \text{government} \\ \text{consumption} \end{bmatrix} + \begin{bmatrix} \text{fixed} \\ \text{investment} \end{bmatrix} + \begin{bmatrix} \text{stock} \\ \text{change} \end{bmatrix} + \begin{bmatrix} \text{trade} \\ \text{input use} \end{bmatrix}$$

where

$qdst_c$ = quantity of stock change.

Current-Account Balance for the Rest of the World, in Foreign Currency

$$\sum_{c \in CM} pwm_c \cdot QM_c + \sum_{f \in F} trnsfr_{row\,f} = \sum_{c \in CE} pwe_c \cdot QE_c + \sum_{i \in INSD} trnsfr_{i\,row} + \overline{FSAV} \tag{41}$$

$$\begin{bmatrix} \text{import} \\ \text{spending} \end{bmatrix} + \begin{bmatrix} \text{factor} \\ \text{transfers} \\ \text{to RoW} \end{bmatrix} = \begin{bmatrix} \text{export} \\ \text{revenue} \end{bmatrix} + \begin{bmatrix} \text{institutional} \\ \text{transfers} \\ \text{from RoW} \end{bmatrix} + \begin{bmatrix} \text{foreign} \\ \text{savings} \end{bmatrix}$$

where

$\overline{FSAV}$ = foreign savings (FCU) (exogenous variable).

Government Balance

$$YG = EG + GSAV \tag{42}$$

$$\begin{bmatrix} \text{government} \\ \text{revenue} \end{bmatrix} = \begin{bmatrix} \text{government} \\ \text{expenditures} \end{bmatrix} + \begin{bmatrix} \text{government} \\ \text{savings} \end{bmatrix}$$

where

$GSAV$ = government savings.

Direct Institutional Tax Rates

$$TINS_i = \overline{tins_i} \cdot \left(1 + \overline{TINSADJ} \cdot tins01_i\right) + \overline{DTINS} \cdot t. \qquad i \in INSDNG \tag{43}$$

$$\begin{bmatrix} \text{direct tax} \\ \text{rate for} \\ \text{institution } i \end{bmatrix} = \begin{bmatrix} \text{base rate adjusted} \\ \text{for scaling for} \\ \text{selected institutions} \end{bmatrix} + \begin{bmatrix} \text{point change} \\ \text{for selected} \\ \text{institutions} \end{bmatrix}$$

where

$TINS_i$ = rate of direct tax on domestic institutions i,

$\overline{tins_i}$ = exogenous direct tax rate for domestic institution i,

$\overline{TINSADJ}$ = direct tax scaling factor (= 0 for base; exogenous variable),

$tins01_i$ = 0–1 parameter with 1 for institutions with potentially flexed direct tax rates, and

$\overline{DTINS_i}$ = change in domestic institution tax share (= 0 for base; exogenous variable).

Institutional Savings Rates

$$MPS_i = \overline{mps_i} \cdot \left(1 + \overline{MPSADJ} \cdot mps01_i\right) + DMPS \cdot mps01_i \qquad i \in INSDNG \quad (44)$$

$$\begin{bmatrix} \text{savings rate for institution } i \end{bmatrix} = \begin{bmatrix} \text{base rate adjusted for scaling for selected institutions} \end{bmatrix} + \begin{bmatrix} \text{point change for selected institutions} \end{bmatrix}$$

where

$\overline{mps_i}$ = base savings rate for domestic institution i,

$\overline{MPSADJ}$ = savings rate scaling factor (= 0 for base),

$MPS01_i$ = 0-1 parameter with 1 for institutions with potentially flexed direct tax rates, and

$DMPS$ = change in domestic institution savings rates (= 0 for base; exogenous variable).

Savings–Investment Balance

$$\sum_{i \in INSDNG} MPS_i \cdot \left(1 - TINS_i\right) \cdot YI_i + GSAV + EXR \cdot \overline{FSAV} = \sum_{c \in C} PQ_c \cdot QINV_c + \sum_{c \in C} PQ_c \cdot qdst_c \quad (45)$$

$$\begin{bmatrix} \text{non-government savings} \end{bmatrix} + \begin{bmatrix} \text{government savings} \end{bmatrix} + \begin{bmatrix} \text{foreign savings} \end{bmatrix} = \begin{bmatrix} \text{fixed investment} \end{bmatrix} + \begin{bmatrix} \text{stock change} \end{bmatrix}$$

Total Absorption

$$TABS = \sum_{h \in H} \sum_{c \in C} PQ_c \cdot QH_{c\,h} + \sum_{a \in A} \sum_{c \in C} \sum_{h \in H} PXAC_{a\,c} \cdot QHA_{a\,c\,h} + \sum_{c \in C} PQ_c \cdot QG_c + \sum_{c \in C} PQ_c \cdot QINV_c + \sum_{c \in C} PQ_c \cdot qdst_c \quad (46)$$

$$\begin{bmatrix} \text{total absorption} \end{bmatrix} = \begin{bmatrix} \text{household market consumption} \end{bmatrix} + \begin{bmatrix} \text{household home consumption} \end{bmatrix} + \begin{bmatrix} \text{government consumption} \end{bmatrix} + \begin{bmatrix} \text{fixed investment} \end{bmatrix} + \begin{bmatrix} \text{stock change} \end{bmatrix}$$

where

$TABS$ = total nominal absorption.

Ratio of Investment to Absorption

$$INVSHR \cdot TABS = \sum_{c \in C} PQ_c \cdot QINV_c + \sum_{c \in C} PQ_c \cdot qdst_c \quad (47)$$

$$\begin{bmatrix} \text{investment-absorption ratio} \end{bmatrix} \cdot \begin{bmatrix} \text{total absorption} \end{bmatrix} = \begin{bmatrix} \text{fixed investment} \end{bmatrix} + \begin{bmatrix} \text{stock change} \end{bmatrix}$$

where

$INVSHR$ = investment share in nominal absorption.

Ratio of Government Consumption to Absorption

$$GOVSHR \cdot TABS = \sum_{c \in C} PQ_c \cdot QG_c$$

$$\begin{bmatrix} \text{government consumption-absorption ratio} \end{bmatrix} \cdot \begin{bmatrix} \text{total absorption} \end{bmatrix} = \begin{bmatrix} \text{government consumption} \end{bmatrix} \quad (48)$$

where

$GOVSHR$ = government consumption share in nominal absorption.

Appendix B: GAMS Code

```
$ONSYMLIST ONSYMXREF OFFUPPER
*$OFFSYMLIST OFFSYMXREF
$ONEMPTY
*The dollar control option makes empty data initialization statements
*permissible (e.g. sets without elements or parameters without data)


*1. SET DECLARATIONS ###############################################

$ontext

In this section, all sets are declared. They are divided into the
following groups:
a. model sets (appearing in the model equations)
b. calibration sets (used to initialize variables and define model
   parameters)
c. report sets (used in report files)

$offtext

SETS

*a. model sets

 AC          global set for model accounts - aggregated microsam accounts
 ACNT(AC)    all elements in AC except TOTAL
 A(AC)       activities
 ACES(A)     activities with CES fn at top of technology nest
 ALEO(A)     activities with Leontief fn at top of technology nest

 C(AC)       commodities
 CD(C)       commodities with domestic sales of output
 CDN(C)      commodities without domestic sales of output
 CE(C)       exported commodities
 CEN(C)      non-export commodities
 CM(C)       imported commodities
 CMN(C)      non-imported commodities
 CX(C)       commodities with output

 F(AC)       factors
 INS(AC)     institutions
 INSD(INS)   domestic institutions
 INSDNG(INSD) domestic non-government institutions
 H(INSDNG)   households

*b. calibration sets
 CINV(C)     fixed investment goods
 CT(C)       transaction service commodities
 CTD(AC)     domestic transactions cost account
 CTE(AC)     export transactions cost account
```

```
 CTM(AC)     import transactions cost account

*c. report sets
 AAGR(A)     agricultural activities
 ANAGR(A)    non-agricultural activities
 CAGR(C)     agricultural commodities
 CNAGR(C)    non-agricultural commodities
 EN(INSDNG)  enterprises
 FLAB(F)     labor
 FLND(F)     land
 FCAP(F)     capital
 ;

*ALIAS statement to create identical cets
ALIAS
 (AC,ACP)  , (ACNT,ACNTP), (A,AP,APP)
 (C,CP,CPP), (CE,CEP)    , (CM,CMP)
 (F,FP)    , (FLAB,FLABP), (FCAP,FCAPP)    , (FLND,FLNDP)
 (INS,INSP), (INSD,INSDP), (INSDNG,INSDNGP), (H,HP)
 ;


*2. DATABASE ################################################################

PARAMETER
 SAM(AC,ACP)     standard SAM
 SAMBALCHK(AC)   column minus row total for SAM

 ;

*INCLUDE ONE COUNTRY DATA SET
*Remove asterisk in front of ONE (AND ONLY ONE) of the following lines
*or add a new line for new file with country data

$INCLUDE TEST.DAT
*$INCLUDE SWAZILAN.DAT
*$INCLUDE ZIMBABWE.DAT


$TITLE Core model files. Standard CGE modeling system, Version 1.01
$STITLE Input file: MOD101.GMS. Standard CGE modeling system, Version 1.01

*SAM adjustments ================================================

*In this section, some minor adjustments are made in the SAM (when
*needed) to fit the model structure.


*Adjustment for sectors with only exports and no domestic sales.
*If there is a very small value for domestic sales, add the discrepancy
*to exports.
 SAM(C,'ROW')$(ABS(SUM(A, SAM(A,C)) - (SAM(C,'ROW') - TAXPAR('EXPTAX',C)
                          - SUM(CTE, SAM(CTE,C))) ) LT 1.E-6)
```

```
            = SUM(A, SAM(A,C)) -  TAXPAR('EXPTAX',C)
                        - SUM(CTE, SAM(CTE,C)) ;

*Netting transfers between domestic institutions and RoW.
 SAM(INSD,'ROW')   = SAM(INSD,'ROW') - SAM('ROW',INSD);
 SAM('ROW',INSD)   = 0;

*Netting transfers between factors and RoW.
 SAM('ROW',F)  = SAM('ROW',F) - SAM(F,'ROW');
 SAM(F,'ROW')  = 0;

*Netting transfers between government and domestic non-
*government institutions.
 SAM(INSDNG,'GOV') = SAM(INSDNG,'GOV') - SAM('GOV',INSDNG);
 SAM('GOV',INSDNG) = 0;

*Eliminating payments of any account to itself.
 SAM(ACNT,ACNT) = 0;


*Checking SAM balance================================================
*Do NOT make any changes in the parameter SAM after this line!!!!!!!!!

*Account totals are recomputed. Check for SAM balance.
 SAM('TOTAL',ACNT) = SUM(ACNTP, SAM(ACNTP,ACNT));
 SAM(ACNT,'TOTAL') = SUM(ACNTP, SAM(ACNT,ACNTP));

 SAMBALCHK(AC)   = SAM('TOTAL',AC) - SAM(AC,'TOTAL');

 DISPLAY "SAM after final adjustments", SAMBALCHK;
 DISPLAY "SAM after final adjustments", SAM;


*Additional set definitions based on country SAM=======================

*CD is the set for commodities with domestic sales of domestic output
*i.e., for which (value of sales at producer prices)
*           > (value of exports at producer prices)
 CD(C)  = YES$
   (SUM(A, SAM(A,C)) GT (SAM(C,'ROW') - TAXPAR('EXPTAX',C)
                          - SUM(CTE, SAM(CTE,C))) );

 CDN(C) = NOT CD(C);

 CE(C)  = YES$(SAM(C,'ROW'));
 CEN(C) = NOT CE(C);

 CM(C)  = YES$(SAM('ROW',C));
 CMN(C) = NOT CM(C);

 CX(C) = YES$SUM(A, SAM(A,C));

 CT(C)
```

```
 $(SUM(CTD, SAM(C,CTD)) + SUM(CTE, SAM(C,CTE)) + SUM(CTM, SAM(C,CTM)))
  = YES;

 ALEO(A) = YES; ACES(A) = NO;

*If activity has no intermediate inputs, then Leontief function has to
*be used at the top of the technology nest
 ACES(A)$(NOT SUM(C, SAM(C,A))) = NO;
 ALEO(A)$(NOT ACES(A)) = YES;

DISPLAY
 CD, CDN, CE, CEN, CM, CMN, CX, CT, ACES, ALEO;


*Fine-tuning non-SAM data=============================================

*Generating missing data for home consumption====

*If SAM includes home consumption but NO data were provided for SHRHOME,
*data are generated assuming that the value shares for home consumption
*are identical to activity output value shares.

IF(SUM((A,H), SAM(A,H)) AND NOT SUM((A,C,H), SHRHOME(A,C,H)),

 SHRHOME(A,C,H)$SAM(A,H) = SAM(A,C)/SUM(CP, SAM(A,CP));

DISPLAY
 "Default data used for SHRHOME -- data missing"
 SHRHOME
 ;
*End IF statement
 );


*Eliminating superfluous elasticity data=========

 TRADELAS(C,'SIGMAT')$(CEN(C) OR (CE(C) AND CDN(C))) = 0;
 TRADELAS(C,'SIGMAQ')$(CMN(C) OR (CM(C) AND CDN(C))) = 0;

 PRODELAS(A)$(NOT SAM('TOTAL',A))     = 0;

 ELASAC(C)$(NOT SUM(A, SAM(A,C)))     = 0;

 LESELAS1(C,H)$(NOT SAM(C,H))        = 0;
 LESELAS2(A,C,H)$(NOT SHRHOME(A,C,H)) = 0;


*Diagnostics=====================================

*Include file that displays and generates information that may be
*useful when debugging data set.

$INCLUDE DIAGNOSTICS.INC
```

```
$STITLE Input file: MOD101.GMS. Standard CGE modeling system, Version 1.01

*Physical factor quantities======================

PARAMETER
 QF2BASE(F,A)  qnty of fac f employed by act a (extracted data)
 ;
*If there is a SAM payment from A to F and supply (but not
*demand) quantities have been defined in the country data file,
*then the supply values are used to compute demand quantities.
 QF2BASE(F,A)$(SAM(F,A)$((NOT QFBASE(F,A))$QFSBASE(F)))
   = QFSBASE(F)*SAM(F,A)/SUM(AP, SAM(F,AP));

*If there is a SAM payment from A to F and neither supply nor
*demand quantities have been defined in the country data file,
*then SAM values are used as quantities
 QF2BASE(F,A)$(SAM(F,A)$((QFBASE(F,A) EQ 0)$(QFSBASE(F) EQ 0)))
                                    = SAM(F,A);

*If there is a SAM payment from A to F and demand quantities have
*been defined in the country data file, then this information is used.
 QF2BASE(F,A)$QFBASE(F,A) = QFBASE(F,A);

DISPLAY QF2BASE, QFBASE, QFSBASE;


*3. PARAMETER DECLARATIONS ############################################
$ontext

This section is divided into the following subsections:
a. Parameters appearing in model equations
b. Parameters used for model calibration (to initialize variables and
   to define model parameters)

In each group, the parameters are declared in alphabetical order.

$offtext

PARAMETERS

*a. Parameters appearing in model equations================

*Parameters other than tax rates
 alphaa(A)      shift parameter for top level CES function
 alphaac(C)     shift parameter for domestic commodity aggregation fn
 alphaq(C)      shift parameter for Armington function
 alphat(C)      shift parameter for CET function
 alphava(A)     shift parameter for CES activity production function
 betah(A,C,H)   marg shr of hhd cons on home com c from act a
 betam(C,H)     marg share of hhd cons on marketed commodity c
 cwts(C)        consumer price index weights
 deltaa(A)      share parameter for top level CES function
```

deltaac(A,C)   share parameter for domestic commodity aggregation fn
deltaq(C)   share parameter for Armington function
deltat(C)   share parameter for CET function
deltava(F,A)   share parameter for CES activity production function
dwts(C)   domestic sales price weights
gammah(A,C,H) per-cap subsist cons for hhd h on home com c fr act a
gammam(C,H)   per-cap subsist cons of marketed com c for hhd h
ica(C,A)   intermediate input c per unit of aggregate intermediate
inta(A)   aggregate intermediate input coefficient
iva(A)   aggregate value added coefficient
icd(C,CP)   trade input of c per unit of comm'y cp produced & sold dom'ly
ice(C,CP)   trade input of c per unit of comm'y cp exported
icm(C,CP)   trade input of c per unit of comm'y cp imported
mps01(INS)   0-1 par for potential flexing of savings rates
mpsbar(INS)   marg prop to save for dom non-gov inst ins (exog part)
qdst(C)   inventory investment by sector of origin
qbarg(C)   exogenous (unscaled) government demand
qbarinv(C)   exogenous (unscaled) investment demand
rhoa(A)   CES top level function exponent
rhoac(C)   domestic commodity aggregation function exponent
rhoq(C)   Armington function exponent
rhot(C)   CET function exponent
rhova(A)   CES activity production function exponent
shif(INS,F)   share of dom. inst'on i in income of factor f
shii(INS,INSP) share of inst'on i in post-tax post-sav income of inst ip
supernum(H)   LES supernumerary income
theta(A,C)   yield of commodity C per unit of activity A
tins01(INS)   0-1 par for potential flexing of dir tax rates
trnsfr(INS,AC) transfers fr. inst. or factor ac to institution ins

*Tax rates
ta(A)   rate of tax on producer gross output value
te(C)   rate of tax on exports
tf(F)   rate of direct tax on factors (soc sec tax)
tinsbar(INS)   rate of (exog part of) direct tax on dom inst ins
tm(C)   rate of import tariff
tq(C)   rate of sales tax
tva(A)   rate of value-added tax

*b. Parameters used for model calibration==================

$ontext

For model calibration, one parameter is created for each model variable
with the suffix "0" added to the variable name. 0 is also added to the
names of parameters whose values are changed in experiments.

$offtext

PARAMETERS
*Parameters for definition of model parameters
alphava0(A)   shift parameter for CES activity production function

qdst0(C) stock change
qbarg0(C) exogenous (unscaled) government demand
gammah0(A,C,H) per-cap subsist cons for hhd h on home com c fr act a
gammam0(C,H) per-cap subsist cons of marketed com c for hhd h

ta0(A) rate of tax on producer gross output value
te0(C) rate of tax on exports
tf0(F) rate of direct tax on factors (soc sec tax)
tins0(INS) rate of direct tax on domestic institutions ins
tm0(C) rate of import tariff
tq0(C) rate of sales tax
tva0(A) rate of value-added tax

*Check parameters
cwtschk check that CPI weights sum to unity
dwtschk check that PDIND weights sum to unity
shifchk check that factor payment shares sum to unity

*Parameters for variable initialization
CPI0 consumer price index (PQ-based)
DPI0 index for domestic producer prices (PDS-based)
DMPS0 change in marginal propensity to save for selected inst
DTINS0 change in domestic institution tax share
EG0 total current government expenditure
EH0(H) household consumption expenditure
EXR0 exchange rate
FSAV0 foreign savings
GADJ0 government demand scaling factor
GOVSHR0 govt consumption share of absorption
GSAV0 government savings
IADJ0 investment scaling factor (for fixed capital formation)
INVSHR0 investment share of absorption
MPS0(INS) marginal propensity to save for dom non-gov inst ins
MPSADJ0 savings rate scaling factor
PA0(A) output price of activity a
PDD0(C) demand price for com'y c produced & sold domestically
PDS0(C) supply price for com'y c produced & sold domestically
PE0(C) price of exports
PINTA0(A) price of intermediate aggregate
PM0(C) price of imports
PQ0(C) price of composite good c
PVA0(A) value added price
PWE0(C) world price of exports
PWM0(C) world price of imports
PX0(C) average output price
PXAC0(A,C) price of commodity c from activity a
QA0(A) level of domestic activity
QD0(C) quantity of domestic sales
QE0(C) quantity of exports
QF0(F,A) quantity demanded of factor f from activity a
QFS0(F) quantity of factor supply
QG0(C) quantity of government consumption
QH0(C,H) quantity consumed of marketed commodity c by hhd h

```
 QHA0(A,C,H)    quantity consumed of home commodity c fr act a by hhd h
 QINT0(C,A)     quantity of intermediate demand for c from activity a
 QINTA0(A)      quantity of aggregate intermediate input
 QINV0(C)       quantity of fixed investment demand
 QM0(C)         quantity of imports
 QQ0(C)         quantity of composite goods supply
 QT0(C)         quantity of trade and transport demand for commodity c
 QVA0(A)        quantity of aggregate value added
 QX0(C)         quantity of aggregate marketed commodity output
 QXAC0(A,C)     quantity of ouput of commodity c from activity a
 TABS0          total absorption
 TINS0(INS)     rate of direct tax on domestic institutions ins
 TINSADJ0       direct tax scaling factor
 TRII0(INS,INSP) transfers to dom. inst. insdng from insdngp
 WALRAS0        savings-investment imbalance (should be zero)
 WF0(F)         economy-wide wage (rent) for factor f
 WFDIST0(f,A)   factor wage distortion variable
 YF0(f)         factor income
 YG0            total current government income
 YIF0(INS,F)    income of institution ins from factor f
 YI0(INS)       income of (domestic non-governmental) institution ins
 ;


*4. PARAMETER DEFINITIONS ##############################################


*All parameters are defined, divided into the same blocks as the
*equations.

*Price block====================================

$ontext

The prices PDS, PX, and PE  may be initialized at any desired price.
The user may prefer to initialize these prices at unity or, if
he/she is interested in tracking commodity flows in physical units, at
commodity-specific, observed prices (per physical unit). For any given
commodity, these three prices should be identical. Initialization at
observed prices may be attractive for disaggregated agricultural
commodities. If so, the corresponding quantity values reflect physical
units (given the initial price).

The remaining supply-side price, PXAC, and the non-commodity prices, EXR
and PA may be initizalized at any desired level. In practice, it may be
preferable to initialize PXAC at the relevant supply-side price and EXR
and PA at unity.

If physical units are used, the user should select the unit (tons vs.
'000 tons) so that initial price and quantity variables are reasonably
scaled (for example between 1.0E-2 and 1.0E+3) -- bad scaling may cause
solver problems. Initialization at unity should cause no problem as long
as the initial SAM is reasonably scaled.
```

```
$offtext

PARAMETER
 PSUP(C) initial supply-side market price for commodity c
;
 PSUP(C) = 1;

 PE0(C)$CE(C)       = PSUP(C);
 PX0(C)$CX(C)       = PSUP(C);
 PDS0(C)$CD(C)      = PSUP(C);
 PXAC0(A,C)$SAM(A,C) = PSUP(C);

 PA0(A)      = 1;

$ontext
The exchange rate may be initialized at unity, in which case all data are
in foreign currency units (FCU; e.g., dollars). Set the exchange rate at
another value to differentiate foreign exchange transactions, which will
be valued in FCU, and domestic transactions valued in local currency
units (LCU). The SAM is assumed to be valued in LCU, and the exchange rate
is then used to calculate FCU values for transactions with the rest of the
world.
$offtext

 EXR0        = 1 ;

*Activity quantity = payment to activity divided by activity price
*QA covers both on-farm consumption and marketed output
*output GROSS of tax
 QA0(A)       =  SAM('TOTAL',A)/PA0(A) ;

*Unit value-added price = total value-added / activity quantity
*define pva gross of tax
 QVA0(A)      =  (SUM(F, SAM(F,A))+ TAXPAR('VATAX',A)) ;
 PVA0(A)      =  (SUM(F, SAM(F,A))+ TAXPAR('VATAX',A))/QVA0(A);
 iva(A)       =  QVA0(A)/QA0(A) ;
 QXAC0(A,C)$SAM(A,C)
            = SAM(A,C) / PXAC0(A,C);

 QHA0(A,C,H)$SHRHOME(A,C,H) = SHRHOME(A,C,H)*SAM(A,H)/PXAC0(A,C);


*Output quantity = value received by producers divided by producer
*price
*QX covers only marketed output
 QX0(C)$SUM(A, SAM(A,C))
       =  SUM(A, SAM(A,C)) / PX0(C);

*Export quantity = export revenue received by producers
*(ie. minus tax and transactions cost) divided by
*export price.
 QE0(C)$SAM(C,'ROW')
```

```
   =  (SAM(C,'ROW') - TAXPAR('EXPTAX',C)
          - SUM(CTE, SAM(CTE,C)))/PE0(C);

*RoW export price = RoW export payment (in for curr) / export qnty
 PWE0(C)$QE0(C) = (SAM(C,'ROW')/EXR0) / QE0(C);

 te0(C)$SAM(C,'ROW') = TAXPAR('EXPTAX',C)/SAM(C,'ROW');
 te(C)            =  te0(C);


*Quantity of output sold domestically = output quantity less quantity
*exported = value of domestic sales divided by domestic supply price
*QD0 covers only marketed output
 QD0(C)$CD(C) =  QX0(C) - QE0(C);

*Domestic demander price = demander payment divided by quantity bought
 PDD0(C)$QD0(C)= (PDS0(C)*QD0(C) + SUM(CTD, SAM(CTD,C)))/QD0(C);

*Define import price to equal domestic price so that import and domestic
*units are the same to the purchaser. If no domestic good, set PM to 1.
 PM0(C)           = PDD0(C) ;
 PM0(C)$(QD0(C) EQ 0) = 1 ;

*Import quantity = demander payment for imports (including tariffs
*and marketing cost) divided by demander price.
 QM0(C)$CM(C) = (SAM('ROW',C) + TAXPAR('IMPTAX',C)
          + SUM(CTM, SAM(CTM,C)))/PM0(C);

*World price = import value (in foreign currency / import quantity
 PWM0(C)$QM0(C)= (SAM('ROW',C)/EXR0) / QM0(C);
 tm0(C)$SAM('ROW',C)
          = TAXPAR('IMPTAX',C) / SAM('ROW',C);
 tm(C)       = tm0(C);


*Composite supply is the sum of domestic market sales and imports
*(since they are initialized at the same price).
 QQ0(C)$(CD(C) OR CM(C)) = QD0(C) + QM0(C) ;
 PQ0(C)$QQ0(C) = (SAM(C,'TOTAL') - SAM(C,'ROW'))/QQ0(C);
 TQ0(C)$QQ0(C) = TAXPAR('COMTAX',C)/(PQ0(C)*QQ0(C)) ;
 TQ(C)       = TQ0(C) ;

*The following code works when for any number of sectors providing
*transactions services, as well as for the case when they are not
*in the SAM.

PARAMETERS
 SHCTD(C)  share of comm'y ct in trans services for domestic sales
 SHCTM(C)  share of comm'y ct in trans services for imports
 SHCTE(C)  share of comm'y ct in trans services for exports
  ;
 SHCTD(CT) = SUM(CTD, SAM(CT,CTD)/SAM('TOTAL',CTD)) ;
 SHCTM(CT) = SUM(CTM, SAM(CT,CTM)/SAM('TOTAL',CTM)) ;
```

```
SHCTE(CT) = SUM(CTE, SAM(CT,CTE)/SAM('TOTAL',CTE)) ;


*Transactions input coefficients
icd(CT,C)$QD0(c)
  = (shctd(ct)*SUM(CTD, SAM(CTD,C))/PQ0(ct)) / QD0(C);
icm(CT,C)$QM0(C)
 = (shctm(ct)*SUM(CTM, SAM(CTM,C))/PQ0(ct)) / QM0(C);
ice(CT,C)$QE0(C)
 = (shcte(ct)*SUM(CTE, SAM(CTE,C))/PQ0(ct)) / QE0(C);


*Indirect activity tax rate = tax payment / output value
*Tax is here applied to total output value (incl. on-farm cons.)
tva0(A)      = TAXPAR('VATAX',A) / (PVA0(A)*QVA0(A));
tva(A)       = tva0(A);

*QA is GROSS of tax, so base for ta is as well
ta0(A)       = TAXPAR('ACTTAX',A) / (SAM(A,'TOTAL'));
ta(A)        = ta0(A);

*Yield coefficient
* = quantity produced (including home-consumed output)
*   /activity quantity
theta(A,C)$PXAC0(A,C)
 = ( (SAM(A,C) + SUM(H, SHRHOME(A,C,H)*SAM(A,H)) ) / PXAC0(A,C) )
                                        / QA0(A);

*Intermediate input coefficient = input use / output quantity
QINTA0(A) = SUM(C$PQ0(C), SAM(C,A)  / PQ0(C)) ;

ica(C,A)$(QINTA0(A)$PQ0(C))
          = SAM(C,A)/PQ0(C) / QINTA0(A) ;

inta(A)      = QINTA0(A) / QA0(A) ;
pinta0(A)    = SUM(C, ica(C,A)*PQ0(C)) ;

*CPI weight by comm'y = hhd cons value for comm'y / total hhd cons value
*CPI does not consider on-farm consumption.
cwts(C)      = SUM(H, SAM(C,H)) / SUM((CP,H), SAM(CP,H));

*Domestic sales price index weight = dom sales value for comm'y
*/ total domestic salues value
*Domestic sales price index does not consider on-farm consumption.
dwts(C)      = (SUM(A, SAM(A,C)) - (SAM(C,'ROW') -
             SUM(cte, SAM(cte,C))))/
             SUM(CP, SUM(A, SAM(A,CP)) - (SAM(CP,'ROW') -
             SUM(cte, SAM(cte,CP))));

CWTSCHK      = SUM(C, cwts(C));
DWTSCHK      = SUM(C, dwts(C));

CPI0         = SUM(C, cwts(C)*PQ0(C)) ;
```

```
 DPI0        = SUM(CD, dwts(CD)*PDS0(CD)) ;

DISPLAY
 CWTSCHK, DWTSCHK;

*Production and trade block==========================

*Compute exponents from elasticites
 rhoq(C)$(CM(C) AND CD(C)) = (1/TRADELAS(C,'SIGMAQ')) - 1;
 rhot(C)$(CE(C) AND CD(C))  = (1/TRADELAS(C,'SIGMAT')) + 1;
 rhova(A)       = (1/PRODELAS(A)) - 1;
 rhoa(A)$ACES(A) = (1/PRODELAS2(A)) - 1;

*Aggregation of domestic output from different activities

 RHOAC(C)$ELASAC(C) = 1/ELASAC(C) - 1;

 deltaac(A,C)$ (SAM(A,C)$ELASAC(C))
           = (PXAC0(A,C)*QXAC0(A,C)**(1/ELASAC(C)))/
             SUM(AP, PXAC0(AP,C)*QXAC0(AP,C)**(1/ELASAC(C)));

 alphaac(C)$SUM(A,deltaac(A,C))
           = QX0(C)/
             (SUM(A$deltaac(A,C), deltaac(A,C) * QXAC0(A,C)
             **(-RHOAC(C))) )**(-1/RHOAC(C));

PARAMETERS
 WFA(F,A)        wage for factor f in activity a (used for calibration)
 ;

*Demand computations=====

*Defining factor employment and supply.
 QF0(F,A)  = QF2BASE(F,A);
 QFS0(F)   = SUM(A, QF0(F,A));

*Activity-specific wage is activity labor payment over employment
 WFA(F,A)$SAM(F,A) = SAM(F,A)/QF0(F,A);


*Economy-wide wage average is total factor income over employment
 WF0(F) = SUM(A, SAM(F,A))/SUM(A, QF0(F,A));

DISPLAY
"If the value of WF0 for any factor is very different from one (< 0.1"
"or >10) the user may consider rescaling the initial values for QFBASE"
"or QFSBASE for this factor to get a value of WF0 such that"
"0.1 < WF0 < 10"
 WF0
 ;

*Wage distortion factor
 wfdist0(f,A)$SAM(F,A) = WFA(F,A)/WF0(F);
```

```
*CES activity production function

 deltava(F,A)$SAM(F,A)
          = (wfdist0(F,A) * WF0(F)
            * (QF0(F,A))**(1+rhova(A)) )
           / SUM(FP, wfdist0(FP,A) * WF0(FP)*(QF0(FP,A))**(1+rhova(A)));

 alphava0(A)= QVA0(A)/( SUM(F$(QF0(F,A)), deltava(F,A)*QF0(F,A)
             **(-rhova(A))) )**(-1/rhova(A));

 alphava(A) = alphava0(A);


*CES top level production function
PARAMETER
  predeltaa(A)  dummy used to define deltaa
  ;

 predeltaa(A)  = 0 ;
 predeltaa(A)$(ACES(A) AND QINTA0(A))
             = (PVA0(A)/PINTA0(A))*(QVA0(A)/QINTA0(A))**(1+rhoa(A)) ;
 deltaa(A)$ACES(A) = predeltaa(A)/(1 + predeltaa(A)) ;
 alphaa(A)$deltaa(A)
             = QA0(A)/((deltaa(A)*QVA0(A)**(-rhoa(A))
              +(1-deltaa(A))*QINTA0(A)**(-rhoa(A)))**(-1/rhoa(A))) ;

*Intermediate demand
 QINT0(C,A)$PQ0(C) = SAM(C,A) / PQ0(C);

*Transactions demand
 QT0(CT) = ( SUM(CTD, SAM(CT,CTD)) + SUM(CTE, SAM(CT,CTE))
           + SUM(CTM, SAM(CT,CTM)) ) / PQ0(CT) ;

*CET transformation
 deltat(C)$(CE(C) AND CD(C))
   = 1 / (1 + PDS0(C)/PE0(C)*(QE0(C)/QD0(C))**(rhot(C)-1));

 alphat(C)$(CE(C) AND CD(C))
   = QX0(C) / (deltat(C)*QE0(C)**rhot(C) + (1-deltat(C))
              *QD0(C)**rhot(C))**(1/rhot(C));


*Armington aggregation

PARAMETER
 predelta(C)  dummy used to define deltaq
 ;

 predelta(C)$(CM(C) AND CD(C))
   = (PM0(C)/(PDD0(C)))*(QM0(C)/QD0(C))**(1+rhoq(C)) ;
```

```
deltaq(C)$(CM(C) AND CD(C))
  = predelta(C)/(1 + predelta(C)) ;

alphaq(C)$(CM(C) AND CD(C))
          = QQ0(C)/(deltaq(C)*QM0(C)**(-rhoq(C))
            +(1-deltaq(C))*QD0(C)**(-rhoq(C)))**(-1/rhoq(C)) ;


*Institution block================================

*Institutional income
 YI0(INSDNG) = SAM('TOTAL',INSDNG);

*Factor income by factor category
 YF0(F) = SUM(A, SAM(F,A));

*Institution income from factors
 YIF0(INSD,F) = SAM(INSD,F);

*Transfers to RoW from factors
 trnsfr('ROW',F) = SAM('ROW',F)/EXR0;

*Transfers from RoW to institutions
 trnsfr(INSD,'ROW') = SAM(INSD,'ROW')/EXR0;

*Government transfers
 trnsfr(INSD,'GOV') = SAM(INSD,'GOV')/CPI0;

*Factor taxes
 tf0(F)       = TAXPAR('FACTAX',F)/SAM('TOTAL',F);
 tf(F)        = tf0(F);

*Shares of domestic institutions in factor income (net of factor taxes
*and transfers to RoW).
 shif(INSD,F)  = SAM(INSD,F)/(SAM(F,'TOTAL') - TAXPAR('FACTAX',f)
             - SAM('ROW',F));

 SHIFCHK(F)    = SUM(INSD, shif(INSD,F));

DISPLAY
 SHIFCHK;

*Inter-institution transfers
 TRII0(INSDNG,INSDNGP) = SAM(INSDNG,INSDNGP);

*Share of dom non-gov institution in income of other dom non-gov
*institutions (net of direct taxes and savings).
 shii(INSDNG,INSDNGP)
  = SAM(INSDNG,INSDNGP)
   /(SAM('TOTAL',INSDNGP) - TAXPAR('INSTAX',INSDNGP) - SAM('S-I',INSDNGP));

*Scaling factors for savings and direct tax shares
 MPSADJ0      = 0;
```

```
 TINSADJ0     = 0;

*Savings rates
 MPS0(INSDNG)  = SAM('S-I',INSDNG)/(SAM('TOTAL',INSDNG) - TAXPAR('INSTAX',INSDNG));
 mpsbar(INSDNG) = MPS0(INSDNG);

*Direct tax rates
 TINS0(INSDNG)   = TAXPAR('INSTAX',INSDNG) / SAM('TOTAL',INSDNG);
 tinsbar(INSDNG) = TINS0(INSDNG);

*"Point" change in savings and direct tax shares
 DMPS0  = 0;
 DTINS0 = 0;


*Selecting institutions for potential "point" change in savings and tax rates

*If DMPS or MPSADJ is flexible, institutions with a value of 1 for mps01
*change their savings rates.
 mps01(INSDNG)  = 1;

*If DTIMS is flexible, institutions with a value of 1 for tins01 change
*their savings rates.
 tins01(INSDNG) = 1;

*Household consumption spending and consumption quantities.
 EH0(H)       = SUM(C, SAM(C,H)) + SUM(A, SAM(A,H));
 QH0(C,H)$PQ0(C) = SAM(C,H)/PQ0(C);

*Government indicators
 YG0          = SAM('TOTAL','GOV');
 EG0          = SAM('TOTAL','GOV') - SAM('S-I','GOV');
 QG0(C)$PQ0(C) = SAM(C,'GOV')/PQ0(C);

 qbarg0(C)     = QG0(C);
 qbarg(C)      = qbarg0(C);
 GADJ0         = 1;
 GSAV0         = SAM('S-I','GOV');

*LES calibration============================================

PARAMETERS
 BUDSHR(C,H)    budget share for marketed commodity c and household h
 BUDSHR2(A,C,H) budget share for home commodity c - act a - hhd h
 BUDSHRCHK(H)   check that budget shares some to unity
 ELASCHK(H)     check that expenditure elasticities satisfy Engel aggr
 ;

 BUDSHR(C,H)    = SAM(C,H)/(SUM(CP, SAM(CP,H)) + SUM(AP, SAM(AP,H)));

 BUDSHR2(A,C,H) = SAM(A,H)*SHRHOME(A,C,H)
              /(SUM(CP, SAM(CP,H)) + SUM(AP, SAM(AP,H)));
```

```
 BUDSHRCHK(H)   = SUM(C, BUDSHR(C,H)) + SUM((A,C), BUDSHR2(A,C,H));

 ELASCHK(H)     = SUM(C, BUDSHR(C,H)*LESELAS1(C,H))
              + SUM((A,C), BUDSHR2(A,C,H)*LESELAS2(A,C,H));

DISPLAY BUDSHR, BUDSHR2, BUDSHRCHK, LESELAS1, LESELAS2, ELASCHK;

 LESELAS1(C,H)   = LESELAS1(C,H)/ELASCHK(H);
 LESELAS2(A,C,H) = LESELAS2(A,C,H)/ELASCHK(H);

 ELASCHK(H)      = SUM(C, BUDSHR(C,H)*LESELAS1(C,H))
               + SUM((A,C), BUDSHR2(A,C,H)*LESELAS2(A,C,H));

DISPLAY ELASCHK, LESELAS1, LESELAS2;

 betam(C,H)   = BUDSHR(C,H)*LESELAS1(C,H);
 betah(A,C,H) = BUDSHR2(A,C,H)*LESELAS2(A,C,H);

 gammam0(C,H)$BUDSHR(C,H)
    =  ( (SUM(CP, SAM(CP,H)) + SUM(AP, SAM(AP,H))) / PQ0(C) )
                * ( BUDSHR(C,H) + betam(C,H)/FRISCH(H));

 gammah0(A,C,H)$BUDSHR2(A,C,H)
    =  ( (SUM(CP, SAM(CP,H)) + SUM(AP, SAM(AP,H))) / PXAC0(A,C) )
                * ( BUDSHR2(A,C,H) + betah(A,C,H)/FRISCH(H));

 gammam(C,H)   =  gammam0(C,H);

 gammah(A,C,H) =  gammah0(A,C,H);

*Checking LES parameters===================================
PARAMETERS
 SUBSIST(H)  subsistence spending
 FRISCH2(H)  alt. defn of Frisch -- ratio of cons to supernumerary cons
 LESCHK(H)   check on LES parameter definitions (error mssg if error)

 LESELASP(H,*,C,*,CP) price elasticity bt c and cp for h (with c and cp labeled by source)
*LESELASP defines cross-price elasticities when c is different from cp and
*own-price elasticities when c and cp refer to the same commodity.
*Source: Dervis, de Melo and Robinson. 1982. General Equilibrium Models
*for Development Policy. Cambridge University Press, p. 483
 ;
 SUPERNUM(H)  = SUM((A,C), gammah(A,C,H)*PXAC0(A,C))
             + SUM(C, gammam(C,H)*PQ0(C)) ;
 FRISCH2(H)   = -EH0(H)/(EH0(H) - SUPERNUM(H));
 LESCHK(H)$(ABS(FRISCH(H) - FRISCH2(H)) GT 0.00000001) = 1/0;


*Cross-price elasticities

 LESELASP(H,'MRK',C,'MRK',CP)
   $((ORD(C) NE ORD(CP)) AND LESELAS1(C,H) AND LESELAS1(CP,H))
 = -LESELAS1(C,H)
```

```
   * PQ0(CP)*gammam(CP,H) / (SUM(CPP, SAM(CPP,H)) + SUM(APP, SAM(APP,H)));

 LESELASP(H,A,C,'MRK',CP)
   $((ORD(C) NE ORD(CP)) AND LESELAS2(A,C,H) AND LESELAS1(CP,H))
 = -LESELAS2(A,C,H)
   * PQ0(CP)*gammam(CP,H) / (SUM(CPP, SAM(CPP,H)) + SUM(APP, SAM(APP,H)));

 LESELASP(H,'MRK',C,A,CP)
   $((ORD(C) NE ORD(CP)) AND LESELAS1(C,H) AND LESELAS2(A,CP,H))
 = -LESELAS1(C,H)
   * PXAC0(A,CP)*gammah(A,CP,H) /(SUM(CPP, SAM(CPP,H)) + SUM(APP, SAM(APP,H)));

*Own-price elasticities

 LESELASP(H,'MRK',C,'MRK',C)
  = -LESELAS1(C,H)
    *( PQ0(C)*gammam(C,H) / (SUM(CP, SAM(CP,H)) + SUM(AP, SAM(AP,H)))
                                    - 1/FRISCH(H));

 LESELASP(H,A,C,A,C)
  = -LESELAS2(A,C,H)
    *( PXAC0(A,C)*gammah(A,C,H) / (SUM(CP, SAM(CP,H)) + SUM(AP, SAM(AP,H)))
                                    - 1/FRISCH(H));

OPTION LESELASP:3:2:2;

DISPLAY
 SUPERNUM, FRISCH, FRISCH2, LESCHK, LESELASP
 ;


*System-constraint block ==========================

*Fixed investment
 qbarinv(c)$CINV(C) = SAM(C,'S-I')/PQ0(C);
 QINV0(C)          = qbarinv(C);
 IADJ0             = 1;

*Stock changes
 qdst0(C)$PQ0(C) = (SAM(C,'S-I')$(NOT CINV(C)) + SAM(C,'DSTK'))/PQ0(C);
 qdst(C)        = qdst0(C);

 FSAV0         = SAM('S-I','ROW')/EXR0;

 TABS0         = SUM((C,H), SAM(C,H)) + SUM((A,H), SAM(A,H))
              + SUM(C, SAM(C,'GOV')) + SUM(C, SAM(C,'S-I'))
              + SUM(C, SAM(C,'DSTK'));

 INVSHR0       = SAM('TOTAL','S-I')/TABS0;
 GOVSHR0       = SUM(C, SAM(C,'GOV'))/TABS0;

 WALRAS0       = 0;
```

*5. VARIABLE DECLARATIONS ##
*This section only includes variables that appear in the model.
*The variables are declared in alphabetical order.

VARIABLES
CPI consumer price index (PQ-based)
DPI index for domestic producer prices (PDS-based)
DMPS change in marginal propensity to save for selected inst
DTINS change in domestic institution tax share
EG total current government expenditure
EH(H) household consumption expenditure
EXR exchange rate
FSAV foreign savings
GADJ government demand scaling factor
GOVSHR govt consumption share of absorption
GSAV government savings
IADJ investment scaling factor (for fixed capital formation)
INVSHR investment share of absorption
MPS(INS) marginal propensity to save for dom non-gov inst ins
MPSADJ savings rate scaling factor
PA(A) output price of activity a
PDD(C) demand price for com'y c produced & sold domestically
PDS(C) supply price for com'y c produced & sold domestically
PE(C) price of exports
PINTA(A) price of intermediate aggregate
PM(C) price of imports
PQ(C) price of composite good c
PVA(A) value added price
PWE(C) world price of exports
PWM(C) world price of imports
PX(C) average output price
PXAC(A,C) price of commodity c from activity a
QA(A) level of domestic activity
QD(C) quantity of domestic sales
QE(C) quantity of exports
QF(F,A) quantity demanded of factor f from activity a
QFS(F) quantity of factor supply
QG(C) quantity of government consumption
QH(C,H) quantity consumed of marketed commodity c by household h
QHA(A,C,H) quantity consumed of home commodity c fr act a by hhd h
QINT(C,A) quantity of intermediate demand for c from activity a
QINTA(A) quantity of aggregate intermediate input
QINV(C) quantity of fixed investment demand
QM(C) quantity of imports
QQ(C) quantity of composite goods supply
QT(C) quantity of trade and transport demand for commodity c
QVA(A) quantity of aggregate value added
QX(C) quantity of aggregate marketed commodity output
QXAC(A,C) quantity of ouput of commodity c from activity a
TABS total absorption
TINS(INS) rate of direct tax on domestic institutions ins
TINSADJ direct tax scaling factor

```
 TRII(INS,INSP) transfers to dom. inst. insdng from insdngp
 WALRAS       savings-investment imbalance (should be zero)
 WALRASSQR    Walras squared
 WF(F)        economy-wide wage (rent) for factor f
 WFDIST(F,A)  factor wage distortion variable
 YF(F)        factor income
 YG           total current government income
 YIF(INS,F)   income of institution ins from factor f
 YI(INS)      income of (domestic non-governmental) institution ins
 ;

*6. VARIABLE DEFINITIONS ###############################################

*The initial levels of all model variables are defined in this file.
$INCLUDE VARINIT.INC

*Optional include file that imposes lower limits for selected variables
*The inclusion of this file may improve solver performance.
*$INCLUDE VARLOW.INC


$STITLE Input file: MOD101.GMS. Standard CGE modeling system, Version 1.01

*7. EQUATION DECLARATIONS ###############################################

EQUATIONS

*Price block================================================
 PMDEF(C)      domestic import price
 PEDEF(C)      domestic export price
 PDDDEF(C)     dem price for com'y c produced and sold domestically
 PQDEF(C)      value of sales in domestic market
 PXDEF(C)      value of marketed domestic output
 PADEF(A)      output price for activity a
 PINTADEF(A)   price of aggregate intermediate input
 PVADEF(A)     value-added price
 CPIDEF        consumer price index
 DPIDEF        domestic producer price index

*Production and trade block================================
 CESAGGPRD(A)   CES aggregate prod fn (if CES top nest)
 CESAGGFOC(A)   CES aggregate first-order condition (if CES top nest)
 LEOAGGINT(A)   Leontief aggreg intermed dem (if Leontief top nest)
 LEOAGGVA(A)    Leontief aggreg value-added dem (if Leontief top nest)
 CESVAPRD(A)    CES value-added production function
 CESVAFOC(F,A)  CES value-added first-order condition
 INTDEM(C,A)    intermediate demand for commodity c from activity a
 COMPRDFN(A,C)  production function for commodity c and activity a
 OUTAGGFN(C)    output aggregation function
 OUTAGGFOC(A,C) first-order condition for output aggregation function
 CET(C)         CET function
 CET2(C)        domestic sales and exports for outputs without both
 ESUPPLY(C)     export supply
```

```
 ARMINGTON(C)    composite commodity aggregation function
 COSTMIN(C)      first-order condition for composite commodity cost min
 ARMINGTON2(C)   comp supply for com's without both dom. sales and imports
 QTDEM(C)        demand for transactions (trade and transport) services

*Institution block ====================================
 YFDEF(F)        factor incomes
 YIFDEF(INS,F)   factor incomes to domestic institutions
 YIDEF(INS)      total incomes of domest non-gov't institutions
 EHDEF(H)        household consumption expenditures
 TRIIDEF(INS,INSP) transfers to inst'on ins from inst'on insp
 HMDEM(C,H)      LES cons demand by hhd h for marketed commodity c
 HADEM(A,C,H)    LES cons demand by hhd h for home commodity c fr act a
 INVDEM(C)       fixed investment demand
 GOVDEM(C)       government consumption demand
 EGDEF           total government expenditures
 YGDEF           total government income

*System constraint block==============================
 COMEQUIL(C)     composite commodity market equilibrium
 FACEQUIL(F)     factor market equilibrium
 CURACCBAL       current account balance (of RoW)
 GOVBAL          government balance
 TINSDEF(INS)    direct tax rate for inst ins
 MPSDEF(INS)     marg prop to save for inst ins
 SAVINVBAL       savings-investment balance
 TABSEQ          total absorption
 INVABEQ         investment share in absorption
 GDABEQ          government consumption share in absorption
 OBJEQ           Objective function
;


*8. EQUATION DEFINITIONS ##############################################
*Notational convention inside equations:
*Parameters and "invariably" fixed variables are in lower case.
*"Variable" variables are in upper case.

*Price block=================================================

 PMDEF(C)$CM(C)..
  PM(C) =E= pwm(C)*(1 + tm(C))*EXR + SUM(CT, PQ(CT)*icm(CT,C));

 PEDEF(C)$CE(C)..
  PE(C) =E= pwe(C)*(1 - te(C))*EXR - SUM(CT, PQ(CT)*ice(CT,C));

 PDDDEF(C)$CD(C).. PDD(C) =E= PDS(C) + SUM(CT, PQ(CT)*icd(CT,C));

 PQDEF(C)$(CD(C) OR CM(C))..
      PQ(C)*(1 - tq(c))*QQ(C) =E= PDD(C)*QD(C) + PM(C)*QM(C);

 PXDEF(C)$CX(C)..  PX(C)*QX(C) =E= PDS(C)*QD(C) + PE(C)*QE(C);
```

```
PADEF(A)..  PA(A) =E= SUM(C, PXAC(A,C)*theta(A,C));

PINTADEF(A).. PINTA(A) =E= SUM(C, PQ(C)*ica(C,A)) ;

PVADEF(A)..  PA(A)*(1-ta(A))*QA(A) =E= PVA(A)*QVA(A) + PINTA(A)*QINTA(A) ;

CPIDEF..  CPI =E= SUM(C, cwts(C)*PQ(C)) ;

DPIDEF..  DPI =E= SUM(CD, dwts(CD)*PDS(CD)) ;


*Production and trade block================================

*CESAGGPRD and CESAGGFOC apply to activities with CES function at
*top of technology nest.

CESAGGPRD(A)$ACES(A)..
  QA(A) =E= alphaa(A)*(deltaa(A)*QVA(A)**(-rhoa(A))
          + (1-deltaa(A))*QINTA(A)**(-rhoa(A)))**(-1/rhoa(A)) ;

CESAGGFOC(A)$ACES(A)..
  QVA(A) =E= QINTA(A)*((PINTA(A)/PVA(A))*(deltaa(A)/
                      (1 - deltaa(A))))**(1/(1+rhoa(A))) ;


*LEOAGGINT and LEOAGGVA apply to activities with Leontief function at
*top of technology nest.

LEOAGGINT(A)$ALEO(A)..  QINTA(A) =E= inta(A)*QA(A) ;

LEOAGGVA(A)$ALEO(A)..  QVA(A) =E= iva(A)*QA(A) ;


*CESVAPRD, CESVAFOC, INTDEM apply at the bottom of the technology nest
*(for all activities).

CESVAPRD(A)..
  QVA(A) =E= alphava(A)*(SUM(F,
               deltava(F,A)*QF(F,A)**(-rhova(A))) )**(-1/rhova(A)) ;


CESVAFOC(F,A)$deltava(F,A)..
  WF(F)*wfdist(F,A) =E=
  PVA(A)*(1-tva(A))
  * QVA(A) * SUM(FP, deltava(FP,A)*QF(FP,A)**(-rhova(A)) )**(-1)
  *deltava(F,A)*QF(F,A)**(-rhova(A)-1);

INTDEM(C,A)$ica(C,A).. QINT(C,A) =E= ica(C,A)*QINTA(A);


COMPRDFN(A,C)$theta(A,C)..
   QXAC(A,C) + SUM(H, QHA(A,C,H)) =E= theta(A,C)*QA(A) ;
```

```
OUTAGGFN(C)$CX(C)..
  QX(C) =E= alphaac(C)*SUM(A, deltaac(A,C)*QXAC(A,C)
                           **(-rhoac(C)))**(-1/rhoac(C));

OUTAGGFOC(A,C)$deltaac(A,C)..
  PXAC(A,C) =E=
  PX(C)*QX(C) * SUM(AP, deltaac(AP,C)*QXAC(AP,C)**(-rhoac(C)) )**(-1)
  *deltaac(A,C)*QXAC(A,C)**(-rhoac(C)-1);

CET(C)$(CE(C) AND CD(C))..
   QX(C) =E= alphat(C)*(deltat(C)*QE(C)**rhot(C) +
                (1 - deltat(C))*QD(C)**rhot(C))**(1/rhot(C)) ;

 ESUPPLY(C)$(CE(C) AND CD(C))..
  QE(C) =E=  QD(C)*((PE(C)/PDS(C))*
           ((1 - deltat(C))/deltat(C)))**(1/(rhot(C)-1)) ;

CET2(C)$( (CD(C) AND CEN(C)) OR (CE(C) AND CDN(C)) )..
  QX(C) =E= QD(C) + QE(C);


ARMINGTON(C)$(CM(C) AND CD(C))..
  QQ(C) =E= alphaq(C)*(deltaq(C)*QM(C)**(-rhoq(C)) +
               (1 -deltaq(C))*QD(C)**(-rhoq(C)))**(-1/rhoq(C)) ;

COSTMIN(C)$(CM(C) AND CD(C))..
  QM(C) =E= QD(C)*((PDD(C)/PM(C))*(deltaq(C)/(1 - deltaq(C))))
                                   **(1/(1 + rhoq(C)));

ARMINGTON2(C)$( (CD(C) AND CMN(C)) OR (CM(C) AND CDN(C)) )..
  QQ(C) =E= QD(C) + QM(C);

QTDEM(C)$CT(C)..
 QT(C) =E= SUM(CP, icm(C,CP)*QM(CP)+ ice(C,CP)*QE(CP)+ icd(C,CP)*QD(CP));


*Institution block ========================================

YFDEF(F)..  YF(F) =E= SUM(A, WF(F)*wfdist(F,A)*QF(F,A));

YIFDEF(INSD,F)$shif(INSD,F)..
 YIF(INSD,F) =E= shif(INSD,F)*((1-tf(f))*YF(F) - trnsfr('ROW',F)*EXR);

YIDEF(INSDNG)..
 YI(INSDNG) =E=
  SUM(F, YIF(INSDNG,F))  + SUM(INSDNGP, TRII(INSDNG,INSDNGP))
  + trnsfr(INSDNG,'GOV')*CPI + trnsfr(INSDNG,'ROW')*EXR;

TRIIDEF(INSDNG,INSDNGP)$(shii(INSDNG,INSDNGP))..
 TRII(INSDNG,INSDNGP) =E= shii(INSDNG,INSDNGP)
       * (1 - MPS(INSDNGP)) * (1 - TINS(INSDNGP))* YI(INSDNGP);

EHDEF(H)..
```

```
 EH(H) =E= (1 - SUM(INSDNG, shii(INSDNG,H))) * (1 - MPS(H))
                                * (1 - TINS(H)) * YI(H);

HMDEM(C,H)$betam(C,H)..
  PQ(C)*QH(C,H) =E=
   PQ(C)*gammam(C,H)
      + betam(C,H)*( EH(H) - SUM(CP, PQ(CP)*gammam(CP,H))
                  - SUM((A,CP), PXAC(A,CP)*gammah(A,CP,H))) ;

HADEM(A,C,H)$betah(A,C,H)..
  PXAC(A,C)*QHA(A,C,H) =E=
    PXAC(A,C)*gammah(A,C,H)
           + betah(A,C,H)*(EH(H) - SUM(CP, PQ(CP)*gammam(CP,H))
                - SUM((AP,CP), PXAC(AP,CP)*gammah(AP,CP,H))) ;

INVDEM(C)$CINV(C)..  QINV(C) =E= IADJ*qbarinv(C);

GOVDEM(C)..  QG(C) =E= GADJ*qbarg(C);

YGDEF..
  YG =E= SUM(INSDNG, TINS(INSDNG)*YI(INSDNG))
        + SUM(f, tf(F)*YF(F))
        + SUM(A, tva(A)*PVA(A)*QVA(A))
        + SUM(A, ta(A)*PA(A)*QA(A))
        + SUM(C, tm(C)*pwm(C)*QM(C))*EXR
        + SUM(C, te(C)*pwe(C)*QE(C))*EXR
        + SUM(C, tq(C)*PQ(C)*QQ(C))
        + SUM(F, YIF('GOV',F))
        + trnsfr('GOV','ROW')*EXR;

EGDEF..
  EG =E= SUM(C, PQ(C)*QG(C)) + SUM(INSDNG, trnsfr(INSDNG,'GOV'))*CPI;


*System constraint block==================================

 FACEQUIL(F)..  SUM(A, QF(F,A)) =E= QFS(F);

 COMEQUIL(C)..
   QQ(C) =E= SUM(A, QINT(C,A)) + SUM(H, QH(C,H)) + QG(C)
           + QINV(C) + qdst(C) + QT(C);

 CURACCBAL..
  SUM(C, pwm(C)*QM(C)) + SUM(F, trnsfr('ROW',F)) =E=
  SUM(C, pwe(C)*QE(C)) + SUM(INSD, trnsfr(INSD,'ROW')) + FSAV;

 GOVBAL.. YG =E= EG + GSAV;

 TINSDEF(INSDNG)..
  TINS(INSDNG) =E= tinsbar(INSDNG)*(1 + TINSADJ*tins01(INSDNG))
              + DTINS*tins01(INSDNG);

 MPSDEF(INSDNG)..
```

```
 MPS(INSDNG)  =E= mpsbar(INSDNG)*(1 + MPSADJ*mps01(INSDNG))
              + DMPS*mps01(INSDNG);

SAVINVBAL..
  SUM(INSDNG, MPS(INSDNG) * (1 - TINS(INSDNG)) * YI(INSDNG))
   + GSAV + FSAV*EXR =E=
  SUM(C, PQ(C)*QINV(C)) + SUM(C, PQ(C)*qdst(C)) + WALRAS;

TABSEQ..
 TABS =E=
  SUM((C,H), PQ(C)*QH(C,H)) + SUM((A,C,H), PXAC(A,C)*QHA(A,C,H))
 + SUM(C, PQ(C)*QG(C)) + SUM(C, PQ(C)*QINV(C)) + SUM(C, PQ(C)*qdst(C));

INVABEQ.. INVSHR*TABS =E= SUM(C, PQ(C)*QINV(C)) + SUM(C, PQ(C)*qdst(C));

GDABEQ..  GOVSHR*TABS =E= SUM(C, PQ(C)*QG(C));

OBJEQ..   WALRASSQR   =E= WALRAS*WALRAS ;



*9. MODEL DEFINITION ################################################

 MODEL STANDCGE  standard CGE model
 /
*Price block (10)
 PMDEF.PM
 PEDEF.PE
 PQDEF.PQ
 PXDEF.PX
 PDDDEF.PDD
 PADEF.PA
 PINTADEF.PINTA
 PVADEF.PVA
 CPIDEF
 DPIDEF

*Production and trade block (17)
 CESAGGPRD
 CESAGGFOC
 LEOAGGINT
 LEOAGGVA
 CESVAPRD.QVA
 CESVAFOC
 INTDEM.QINT
 COMPRDFN.PXAC
 OUTAGGFN.QX
 OUTAGGFOC.QXAC
 CET
 CET2
 ESUPPLY
 ARMINGTON
 COSTMIN
```

```
 ARMINGTON2
 QTDEM.QT

*Institution block (12)
 YFDEF.YF
 YIFDEF.YIF
 YIDEF.YI
 EHDEF.EH
 TRIIDEF.TRII
 HMDEM.QH
 HADEM.QHA
 EGDEF.EG
 YGDEF.YG
 GOVDEM.QG
 GOVBAL
 INVDEM.QINV

*System-constraint block (9)
 FACEQUIL
 COMEQUIL
 CURACCBAL
 TINSDEF.TINS
 MPSDEF.MPS
 SAVINVBAL.WALRAS
 TABSEQ.TABS
 INVABEQ
 GDABEQ
 /
 ;

*10. FIXING VARIABLES NOT IN MODEL AT ZERO ############################

  PDD.FX(C)$(NOT CD(C)) = 0;
  PDS.FX(C)$(NOT CD(C)) = 0;
  PE.FX(C)$(NOT CE(C)) = 0;
  PM.FX(C)$(NOT CM(C)) = 0;
  PX.FX(C)$(NOT CX(C)) = 0;
  PXAC.FX(A,C)$(NOT SAM(A,C)) = 0;

  QD.FX(C)$(NOT CD(C)) = 0;
  QE.FX(C)$(NOT CE(C)) = 0;
  QF.FX(F,A)$(NOT SAM(F,A)) = 0;
  QG.FX(C)$(NOT SAM(C,'GOV')) = 0;
  QH.FX(C,H)$(NOT SAM(C,H)) = 0;
  QHA.FX(A,C,H)$(NOT BETAH(A,C,H)) = 0;
  QINT.FX(C,A)$(NOT SAM(C,A)) = 0;
  QINV.FX(C)$(NOT CINV(C)) = 0;
  QM.FX(C)$(NOT CM(C)) = 0;
  QQ.FX(C)$(NOT (CD(C) OR CM(C))) = 0;
  QT.FX(C)$(NOT CT(C)) = 0;
  QX.FX(C)$(NOT CX(C)) = 0;
  QXAC.FX(A,C)$(NOT SAM(A,C)) = 0;
  TRII.FX(INSDNG,INSDNGP)$(NOT SAM(INSDNG,INSDNGP)) = 0;
```

```
 YI.FX(INS)$(NOT INSD(INS)) = 0;
 YIF.FX(INS,F)$((NOT INSD(INS)) OR (NOT SAM(INS,F))) = 0;


*11. MODEL CLOSURE ###################################################
$ontext
In the simulation file, SIM.GMS, the user chooses between
alternative closures. Those choices take precedence over the choices
made in this file.

In the following segment, closures is selected for the base model
solution in this file. The clearing variables for micro and macro
constraints are as follows:

FACEQUIL - WF: for each factor, the economywide wage is the
market-clearing variable in a setting with perfect factor mobility across
activities.

CURACCBAL - EXR: a flexible exchange rate clears the current account of
the RoW.

GOVBAL - GSAV: flexible government savings clears the government
account.

SAVINVBAL - SADJ: the savings rates of domestic institutions are scaled
to generate enough savings to finance exogenous investment quantities
(investment-driven savings).

The CPI is the model numeraire.
$offtext

*Factor markets=======

 QFS.FX(F)      = QFS0(F);
 WF.LO(F)       = -inf;
 WF.UP(F)       = +inf;
 WFDIST.FX(F,A)  = WFDIST0(F,A);
* WFDIST.LO(F,A)  = -INF;
* WFDIST.UP(F,A)  = +INF;


*Current account of RoW===========

* EXR.FX      = EXR0;
 FSAV.FX     = FSAV0;

*Import and export prices (in FCU) are fixed. A change in model
*specification is required if these prices are to be endogenous.
 PWM.FX(C)    = PWM0(C) ;
 PWE.FX(C)    = PWE0(C) ;


*Current government balance=======
```

```
*GSAV.FX     = GSAV0 ;
 TINSADJ.FX  = TINSADJ0;
 DTINS.FX    = DTINS0;
 GADJ.FX     = GADJ0;
*GOVSHR.FX   = GOVSHR0 ;


*Savings-investment balance=======

 MPSADJ.FX = MPSADJ0;
 DMPS.FX   = DMPS0;
*IADJ.FX   = IADJ0;
*INVSHR.FX = INVSHR0 ;


*Numeraire price index============

 CPI.FX       = CPI0;
*DPI.FX       = DPI0;


*12. DISPLAY OF MODEL PARAMETERS AND VARIABLES ######################

DISPLAY
*All parameters in this file and include files are displayed in
*alphabetical order.

ALPHAA   , ALPHAVA0  , ALPHAAC  , ALPHAQ   , ALPHAT    , ALPHAVA
BETAH    , BETAM     , BUDSHR   , BUDSHR2  , BUDSHRCHK , CPI0
CUTOFF   , CWTS      , CWTSCHK  , DELTAA   , DELTAAC   , DELTAQ
DELTAT   , DELTAVA   , DPI0     , DMPS0    , DTINS0    , DWTS
DWTSCHK  , EG0       , EH0      , ELASAC   , ELASCHK   , EXR0
FRISCH   , FSAV0     , GADJ0    , GAMMAH   , GAMMAM    , GOVSHR0
GSAV0    , IADJ0     , ICA      , ICD      , ICE       , ICM
INTA     , INVSHR0   , IVA      , LESELAS1 , LESELAS2  , MPS0
MPSADJ0  , MPSBAR    , PA0      , PDD0     , PDS0      , PE0
PINTA0   , PM0       , POP      , PQ0      , PRODELAS  , PRODELAS2
PVA0     , PWE0      , PWM0     , PX0      , PXAC0     , QA0
QBARG    , QBARG0    , QBARINV  , QD0      , QDST      , QDST0
QE0      , QF0       , QF2BASE  , QFBASE   , QFS0      , QFSBASE
QG0      , QH0       , QHA0     , QINT0    , QINTA0    , QINV0
QM0      , QQ0       , QT0      , QVA0     , QX0       , QXAC0
RHOA     , RHOAC     , RHOQ     , RHOT     , RHOVA     , SAM
SAMBALCHK, SHCTD     , SHCTE    , SHCTM    , SHIF      , SHIFCHK
SHII     , SHRHOME   , SUMABSDEV , SUPERNUM , TA       , TA0
TABS0    , TAXPAR    , TE       , TE0      , TF        , TF0
THETA    , TINS0     , TINSADJ0 , TINSBAR  , TM        , TM0
TQ       , TQ0       , TRADELAS , TRII0    , TRNSFR    , TVA
TVA0     , WALRAS0   , WF0      , WFA      , WFDIST0   , YF0
YG0      , YI0       , YIF0
;
```

```
*13. SOLUTION STATEMENT ################################################

OPTIONS ITERLIM = 1000, LIMROW = 3, LIMCOL = 3, SOLPRINT=ON,
      MCP=PATH, NLP=CONOPT2 ;

$ontext
These options are useful for debugging. When checking whether the
initial data represent a solution, set LIMROW to a value greater than
the number of equations and search for three asterisks in the listing
file. SOLPRINT=ON provides a complete listing file. The program also
has a number of display statements, so when running experiments it is
usually not necessary to provide a solution print as well.
$offtext

 STANDCGE.HOLDFIXED   = 1 ;
 STANDCGE.TOLINFREP   = .0001 ;

$ontext
The HOLDFIXED option converts all variables which are fixed (.FX) into
parameters. They are then not solved as part of the model.
The TOLINFREP parameter sets the tolerance for determinining whether
initial values of variables represent a solution of the model
equations. Whether these initial equation values are printed is
determimed by the LIMROW option. Equations which are not satsfied to
the degree TOLINFREP are printed with three asterisks next to their
listing.
$offtext


 SOLVE STANDCGE USING MCP ;



*14. OPTIONAL NLP MODEL DEFINITION AND SOLUTION STATEMENT ############

$ontext
Define a model that can be solved using a nonlinear programming (NLP)
solver. The model includes the equation OBJEQ which defines the
variable WALRASSQR, which is the square of the Walras' Law variable,
which must be zero in equilibrium.
$offtext

 MODEL NLPCGE  standard CGE model for NLP solver
 /
*Price block (10)
 PMDEF
 PEDEF
 PQDEF
 PXDEF
 PDDDEF
 PADEF
 PINTADEF
```

```
PVADEF
CPIDEF
DPIDEF

*Production and trade block (17)
CESAGGPRD
CESAGGFOC
LEOAGGINT
LEOAGGVA
CESVAPRD
CESVAFOC
INTDEM
COMPRDFN
OUTAGGFN
OUTAGGFOC
CET
CET2
ESUPPLY
ARMINGTON
COSTMIN
ARMINGTON2
QTDEM

*Institution block (12)
YFDEF
YIFDEF
YIDEF
EHDEF
TRIIDEF
HMDEM
HADEM
EGDEF
YGDEF
GOVDEM
GOVBAL
INVDEM

*System-constraint block (9)
FACEQUIL
COMEQUIL
CURACCBAL
TINSDEF
MPSDEF
SAVINVBAL
TABSEQ
INVABEQ
GDABEQ
OBJEQ
/
;

NLPCGE.HOLDFIXED   = 1 ;
NLPCGE.TOLINFREP   = .0001 ;
```

```
*SOLVE NLPCGE MINIMIZING WALRASSQR USING NLP ;


*15. SOLUTION REPORTS ##############################################

*Optional include file defining report parameters summarizing economic
*data for the base year.

$INCLUDE REPBASE.INC


$STITLE Input file: MOD101.GMS. Standard CGE modeling system, Version 1.01

 STANDCGE.MODELSTAT = 0;
 STANDCGE.SOLVESTAT = 0;
 STANDCGE.NUMREDEF  = 0;

 NLPCGE.MODELSTAT = 0;
 NLPCGE.SOLVESTAT = 0;
 NLPCGE.NUMREDEF  = 0;


*#*#*#*#*# THE END OF MOD101.GMS #*#*#*#*
```

Appendix C: Full Descriptive Statistics

I. Industrial Output

Table B.1: Industrial Output for the State of Louisiana by Model in Millions

Industry	Ag/For	Mining	Utilities	Construction	Manufacturing	Wholesale	Trans	Retail	Prof Serv	Educ/Health	Non-NAICS
Actual*											
average	1,404	23,932	3,715	8,782	38,208	8,313	6,899	12,110	41,323	14,788	27,362
std dev	180	7,088	338	436	10,954	625	1,150	774	3,015	4,086	1,905
2001 CGE											
average	4,119	19,910	4,090	19,967	81,254	9,881	13,657	12,994	56,476	41,291	46,085
std dev	229	1,109	228	1,061	4,526	730	1,348	724	3,145	2,299	2,566
2002 CGE											
average	3,932	12,955	4,296	17,559	70,882	9,978	12,797	14,092	61,040	38,252	48,260
std dev	221	731	242	986	3,963	564	719	797	4,493	2,152	2,784
2006 CGE											
average	4,117	23,439	4,217	17,513	166,625	11,706	14,422	16,451	72,928	37,868	51,439
std dev	246	1,356	222	1,217	9,713	690	864	966	4,248	2,513	2,911
2007 CGE											
average	4,941	27,204	5,879	26,251	135,318	12,978	27,935	16,138	76,210	43,834	73,969
std dev	275	1,520	346	1,332	7,772	724	1,438	845	4,275	1,895	4,038
2008 CGE											
average	5,389	33,769	6,529	27,855	170,805	13,951	12,981	18,313	88,362	48,926	48,023
std dev	280	3,814	360	1,401	7,337	928	751	1,012	4,572	2,812	2,597
2001 SAM											
average	3,733	17,759	3,861	18,675	68,758	9,921	13,332	13,906	47,548	30,423	39,507
std dev	208	989	215	1,040	3,829	552	742	774	2,648	1,694	2,200
2002 SAM											
average	3,309	10,954	3,612	14,871	59,973	8,403	11,892	11,585	43,072	25,695	32,034
std dev	645	2,134	704	2,898	11,686	1,637	2,317	2,258	8,393	5,007	6,242
2006 SAM											
average	3,532	20,244	3,633	14,983	143,920	10,065	14,156	13,131	51,947	27,077	35,420
std dev	701	4,018	721	2,974	28,562	1,998	2,809	2,606	10,309	5,374	7,030
2007 SAM											
average	4,114	22,751	4,939	22,256	114,423	10,821	24,636	13,820	53,418	29,841	39,178
std dev	920	5,088	1,105	4,977	25,587	2,420	5,509	3,091	11,945	6,673	8,761
2008 SAM											
average	4,670	28,591	5,615	24,583	150,763	12,321	16,294	15,784	63,610	34,797	46,696

Table B.1 Con't

std dev	810	4,958	974	4,263	26,143	2,136	2,825	2,737	11,030	6,034	8,097
2001 TS											
average	1,483	44,249	3,718	9,707	21,457	9,498	6,245	14,650	44,072	11,431	32,361
std dev	120	2,425	144	516	431	939	399	1,681	2,938	297	1,530
2002 TS											
average	1,153	31,610	3,311	8,370	30,092	9,580	5,745	14,678	43,845	12,020	33,063
std dev	9	80	143	211	1,066	614	174	1,081	1,882	267	1,042
2006 TS											
average	1,668	12,793	3,316	10,629	63,167	10,963	10,655	16,940	50,712	12,645	31,024
std dev	137	4,905	355	1,365	14,137	1,925	2,894	3,241	8,698	884	2,749
2007 TS											
average	1,640	13,524	3,431	10,157	44,405	11,680	11,501	16,070	55,359	13,109	29,000
std dev	114	4,372	297	1,012	8,462	1,975	3,175	2,411	10,025	907	2,027
2008 TS											
average	1,524	11,844	3,863	10,968	40,717	11,403	16,213	15,245	61,740	14,744	31,969
std dev	86	4,346	164	1,199	6,881	1,626	5,735	2,238	12,304	1,352	2,926

II. Industrial Employment

Table B.2: Industrial Employment Estimates for the State of Louisiana, by Model and Number of Employees

Industry	Ag/For/Mining	Constr/Util	Manufacturing	Whole/Retail	Trans	Prof Serv	Educ/Health	Non-NAICS
Actual*								
average	81,102	194,595	159,812	345,667	91,531	497,126	339,302	740,403
std dev	7,561	13,591	9,478	4,929	2,077	35,730	19,851	17,833
2001 CGE								
average	67,372	367,022	330,736	838,636	305,910	107,249	1,009,100	112,121
std dev	3,751	19,579	18,424	44,039	30,186	5,972	56,191	6,243
2002 CGE								
average	47,348	327,752	288,517	891,237	286,628	115,915	934,825	117,414
std dev	2,667	18,406	16,130	50,418	16,100	8,532	52,598	6,774
2006 CGE								
average	77,261	326,458	678,225	1,041,888	323,046	138,490	925,444	125,148
std dev	4,491	22,098	39,535	61,239	19,344	8,066	61,409	7,082
2007 CGE								
average	90,124	486,423	550,794	1,058,915	625,711	144,722	1,071,242	179,961
std dev	5,028	24,919	31,637	49,523	32,207	8,118	46,322	9,825
2008 CGE								
average	109,790	518,068	695,242	1,182,504	290,767	167,800	1,195,683	116,837
std dev	11,266	26,153	29,865	65,237	16,815	8,682	68,711	6,318
2001 SAM								
average	60,258	343,930	279,870	881,334	298,630	90,295	743,508	96,118
std dev	3,355	19,151	15,584	49,076	16,629	5,028	41,401	5,352

Table B.2 Con't

2002 SAM								
	39,989	277,406	244,113	737,658	266,376	81,794	627,966	77,936
std dev	7,792	54,055	47,568	143,741	51,906	15,938	122,366	15,187
2006 SAM								
average	66,663	279,450	585,808	849,427	317,080	98,648	661,737	86,175
std dev	13,230	55,460	116,260	168,578	62,928	19,578	131,329	17,102
2007 SAM								
average	75,325	412,099	465,744	899,617	551,808	101,441	729,277	95,317
std dev	16,844	92,154	104,150	201,173	123,396	22,684	163,081	21,315
2008 SAM								
average	93,254	456,232	613,664	1,026,532	364,957	120,796	850,407	113,607
std dev	16,170	79,111	106,410	178,002	63,284	20,946	147,462	19,700
2001 TS								
average	126,276	188,313	87,540	876,135	137,319	82,133	277,478	77,694
std dev	8,426	8,632	1,725	120,763	10,741	6,654	8,425	4,432
2002 TS								
average	94,833	166,230	117,697	874,857	127,357	81,541	289,537	79,012
std dev	7,286	4,807	12,401	101,363	4,845	5,360	11,961	4,214
2006 TS								
average	52,450	183,813	208,314	872,570	180,880	81,772	289,980	70,322
std dev	3,414	6,733	11,155	48,932	17,112	3,978	5,263	626
2007 TS								
average	54,398	180,601	152,475	865,486	186,907	86,081	298,467	66,744
std dev	2,358	3,006	4,602	31,369	13,923	3,914	4,195	759

Table B.2 Con't

2008 TS								
average	51,401	189,349	137,890	814,413	213,167	89,799	321,162	69,615
std dev	1,715	3,464	2,308	18,512	18,644	3,579	5,521	855
2001 CGE Nat								
average	73,670	126,379	397,297	230,426	238,980	203,001	228,865	29,374
std dev	4,102	6,649	22,132	12,398	23,582	11,303	12,744	1,636
2002 CGE Nat								
average	51,774	114,809	346,582	242,447	223,916	219,404	212,019	30,761
std dev	2,917	6,451	19,377	13,714	12,577	16,149	11,929	1,775
2006 CGE Nat								
average	84,483	114,154	814,719	283,623	252,366	262,134	209,892	32,787
std dev	4,911	7,473	47,492	16,678	15,112	15,268	13,928	1,855
2007 CGE Nat								
average	98,550	168,786	661,643	293,277	488,811	273,931	242,959	47,147
std dev	5,498	8,760	38,004	13,524	25,160	15,365	10,506	2,574
2008 CGE Nat								
average	120,054	180,630	835,162	324,992	227,150	317,613	271,182	30,609
std dev	12,319	9,172	35,876	18,271	13,136	16,433	15,584	1,655
2001 SAM Nat								
average	65,891	118,388	336,195	240,002	233,292	170,910	168,628	25,182
std dev	3,669	6,592	18,721	13,364	12,991	9,517	9,390	1,402
2002 SAM Nat								
average	43,727	97,095	293,241	201,341	208,095	154,821	142,423	20,418
std dev	8,521	18,920	57,141	39,233	40,550	30,168	27,753	3,979
2006 SAM Nat								
average	72,895	97,790	703,703	233,653	247,706	186,721	150,083	22,577
std dev	14,467	19,408	139,658	46,371	49,160	37,057	29,786	4,481

Table B.2 Con't

2007 SAM Nat								
average	82,366	142,862	559,476	248,207	431,077	192,009	165,401	24,972
std dev	18,419	31,947	125,110	55,504	96,398	42,937	36,987	5,584
2008 SAM Nat								
average	101,972	158,636	737,165	283,099	285,107	228,643	192,873	29,763
std dev	17,682	27,508	127,825	49,090	49,438	39,647	33,444	5,161
2001 TS Nat								
average	138,081	70,093	105,157	235,642	107,275	155,463	62,932	20,355
std dev	9,213	2,154	2,072	31,831	8,391	12,595	1,911	1,161
2002 TS Nat								
average	99,657	61,653	145,312	236,570	98,952	154,403	65,960	20,767
std dev	1,952	786	6,578	24,750	4,769	10,028	2,126	971
2006 TS Nat								
average	57,353	68,324	250,238	235,045	141,305	154,779	65,768	18,423
std dev	3,733	1,795	13,400	12,930	13,368	7,529	1,194	164
2007 TS Nat								
average	59,484	67,646	183,160	235,559	146,014	162,935	67,693	17,486
std dev	2,579	739	5,528	8,708	10,877	7,408	951	199
2008 TS Nat								
average	56,206	71,394	165,641	223,278	166,528	169,973	72,840	18,238
std dev	1,875	1,003	2,773	5,239	14,565	6,775	1,252	224

III. Occupation Employment

Table B.3: Occupational Employment Estimates for the State of Louisiana, by Model and Number of Employees

Industry	11	13	15	17	19	21	23	25	27	29	31	33	35	37	39	41	43	45	47	49	51	53
Actual*																						
average	94,413	54,820	17,321	31,280	12,429	25,243	13,344	106,111	16,408	107,950	53,713	53,321	157,919	58,309	47,805	193,899	301,243	5,099	119,213	96,801	122,658	156,048
std dev	11,639	4,100	893	2,531	733	1,941	948	4,083	1,246	5,428	3,137	1,965	5,406	2,598	3,997	7,018	6,585	1,224	4,633	4,119	5,038	7,095
2001 CGE																						
average	340,565	1,003,416	5,684	11,637	94,813	5,671	4,523	7,242	113,854	205,116	8,803	17,431	19,855	24,162	36,997	68,129	515,018	6,818	7,284	303,532	26,457	243,764
std dev	18,092	55,875	317	648	5,279	316	252	403	6,342	11,426	490	971	1,106	1,345	2,060	3,794	28,684	380	397	29,952	1,557	18,014
2002 CGE																						
average	299,486	929,559	5,266	12,186	99,288	5,939	3,946	6,317	99,321	178,933	9,515	18,840	21,459	26,115	39,986	73,881	558,503	7,394	7,548	284,400	28,266	246,139
std dev	16,812	52,302	296	703	5,728	343	221	353	5,553	10,004	700	1,387	1,579	1,922	2,943	4,180	31,600	418	426	15,975	1,594	13,918
2006 CGE																						
average	298,715	920,231	5,213	12,989	105,828	6,330	9,276	14,851	233,476	420,622	11,368	22,509	25,638	31,201	47,774	86,252	652,017	8,632	8,722	320,534	27,743	288,776
std dev	20,755	61,063	346	735	5,988	358	541	866	13,610	24,519	662	1,311	1,493	1,817	2,783	5,065	38,287	507	515	19,194	1,463	17,015
2007 CGE																						
average	447,743	1,065,207	6,034	18,678	152,180	9,103	7,533	12,060	189,609	341,592	11,879	23,522	26,792	32,605	49,924	84,609	639,596	8,467	10,957	620,847	38,680	320,150
std dev	22,723	46,061	261	1,020	8,308	497	433	693	10,891	19,620	666	1,319	1,503	1,829	2,800	4,429	33,483	443	518	31,957	2,274	17,856
2008 CGE																						
average	475,110	1,188,948	6,735	12,126	98,800	5,910	9,508	15,223	239,334	431,176	13,774	27,272	31,065	37,804	57,885	96,013	725,805	9,609	9,174	288,507	42,958	344,163
std dev	23,899	68,324	387	656	5,342	320	408	654	10,281	18,522	713	1,411	1,607	1,956	2,995	5,304	40,097	531	502	16,685	2,371	22,902
2001 SAM																						
average	318,525	739,320	4,188	9,976	81,280	4,862	3,828	6,128	96,344	173,570	7,412	14,676	16,716	20,343	31,148	72,909	551,150	7,297	7,572	296,308	25,404	244,728
std dev	17,737	41,168	233	556	4,526	271	213	341	5,365	9,665	413	817	931	1,133	1,734	4,060	30,690	406	422	16,499	1,415	13,627
2002 SAM																						
average	253,641	624,428	3,537	8,089	65,905	3,942	3,339	5,345	84,035	151,394	6,714	13,294	15,142	18,428	28,216	60,741	459,166	6,079	6,445	264,305	23,765	207,299
std dev	49,425	121,676	689	1,576	12,842	768	651	1,042	16,375	29,501	1,308	2,590	2,951	3,591	5,498	11,836	89,473	1,185	1,256	51,503	4,631	40,394

Table B.3 Con't

2006 SAM																					
average	255,551	658,010	3,728	8,944	72,872	4,359	8,012	12,827	201,662	363,307	8,097	16,033	18,263	22,225	34,030	68,846	520,440	6,890	7,423	314,615	23,899	248,293
std dev	50,717	130,589	740	1,775	14,462	865	1,590	2,546	40,022	72,102	1,607	3,182	3,624	4,411	6,754	13,663	103,287	1,367	1,473	62,439	4,743	49,276
2007 SAM																					
average	379,602	725,168	4,108	9,893	80,602	4,821	6,370	10,198	160,331	288,846	8,327	16,487	18,780	22,854	34,994	72,459	547,751	7,252	9,507	547,519	32,497	266,938
std dev	84,887	162,163	919	2,212	18,024	1,078	1,424	2,280	35,853	64,592	1,862	3,687	4,200	5,111	7,825	16,203	122,488	1,622	2,126	122,436	7,267	59,693
2008 SAM																					
average	419,288	845,617	4,790	11,791	96,069	5,747	8,393	13,437	211,251	380,583	9,915	19,633	22,363	27,214	41,670	82,757	625,594	8,282	8,796	362,120	36,944	303,939
std dev	72,705	146,631	831	2,045	16,658	996	1,455	2,330	36,631	65,993	1,719	3,404	3,878	4,719	7,226	14,350	108,479	1,436	1,525	62,792	6,406	52,703
2001 TS																					
average	163,656	275,914	1,563	8,064	65,700	3,930	1,197	1,917	30,135	54,291	6,742	13,349	15,205	18,504	28,333	74,269	561,435	7,433	6,416	136,252	24,657	227,649
std dev	9,662	8,377	47	460	3,748	224	24	38	594	1,070	546	1,082	1,232	1,499	2,295	10,632	80,371	1,064	849	10,658	1,032	27,930
2002 TS																					
average	142,582	289,187	1,638	8,227	67,032	4,010	1,654	2,649	41,642	75,021	6,696	13,258	15,102	18,378	28,140	74,385	562,307	7,444	6,341	125,681	22,217	229,379
std dev	3,356	9,322	53	385	3,134	187	75	120	1,885	3,396	435	861	981	1,194	1,828	8,193	61,936	820	637	6,057	1,371	22,064
2006 TS																					
average	133,839	241,842	1,370	6,146	50,078	2,996	2,387	3,821	60,074	108,226	5,666	11,219	12,779	15,551	23,812	62,821	474,895	6,287	5,704	150,571	20,126	194,215
std dev	58,733	104,096	590	2,577	20,999	1,256	1,043	1,670	26,258	47,305	2,358	4,670	5,319	6,473	9,911	24,742	187,037	2,476	2,301	66,461	8,404	77,700
2007 TS																					
average	136,155	255,513	1,448	6,072	49,473	2,959	1,962	3,141	49,381	88,963	6,048	11,975	13,640	16,600	25,417	62,377	471,538	6,243	5,744	159,808	21,096	205,087
std dev	35,672	70,707	401	1,590	12,955	775	428	686	10,783	19,427	1,747	3,460	3,941	4,796	7,343	17,259	130,470	1,727	1,592	45,080	5,906	59,851
2008 TS																					
average	136,755	264,459	1,498	6,035	49,170	2,941	1,688	2,703	42,491	76,550	6,233	12,342	14,058	17,108	26,195	59,005	446,045	5,905	5,611	173,903	21,830	203,296
std dev	46,269	89,624	508	2,030	16,542	990	568	910	14,299	25,761	2,077	4,112	4,684	5,701	8,729	19,075	144,194	1,909	1,842	61,872	7,403	65,785
2001 CGE Nat																					
average	114,627	110,333	101,318	107,339	66,017	30,565	35,072	33,836	65,979	40,989	19,735	30,129	42,756	42,958	26,166	91,419	124,663	21,299	111,774	114,152	100,218	96,648
std dev	6,100	5,892	5,387	5,662	3,606	1,702	1,912	1,862	3,512	2,171	1,078	1,611	2,244	2,261	1,373	4,969	6,653	1,149	5,928	6,123	5,361	5,220
2002 CGE Nat																					
average	107,259	104,211	96,121	100,344	61,539	30,010	34,232	33,120	63,181	40,452	20,063	29,493	42,870	42,676	26,212	86,500	114,513	20,659	102,579	104,698	92,286	88,694

Table B.3 Con't

std dev	6,178	6,007	5,551	5,787	3,603	1,844	2,078	2,017	3,700	2,427	1,280	1,810	2,564	2,552	1,629	5,005	6,593	1,301	5,903	6,023	5,325	5,124
2006 CGE Nat																						
average	152,326	147,376	138,129	144,202	98,853	31,933	39,630	37,853	102,483	55,890	22,737	40,828	55,367	56,648	31,700	128,972	162,704	30,336	147,832	151,357	143,079	133,924
std dev	8,981	8,694	8,128	8,503	5,793	1,935	2,370	2,268	5,991	3,273	1,337	2,392	3,247	3,326	1,862	7,602	9,679	1,773	8,816	9,001	8,433	7,891
2007 CGE Nat																						
average	174,362	168,621	154,561	161,000	93,760	36,302	43,397	41,117	96,050	56,406	24,922	42,046	58,483	58,812	36,844	146,126	191,463	30,110	170,301	177,444	159,196	153,781
std dev	8,794	8,479	7,768	8,140	4,901	1,743	2,094	1,973	4,938	2,800	1,238	2,162	2,832	2,883	1,849	7,480	9,847	1,599	8,789	9,151	8,266	7,958
2008 CGE Nat																						
average	168,281	161,302	149,556	160,586	109,174	39,419	47,724	45,819	111,695	63,109	27,304	46,250	62,878	64,903	36,350	138,762	188,127	34,743	172,650	173,685	157,211	147,862
std dev	8,154	7,698	7,080	7,728	4,997	2,163	2,506	2,426	5,057	3,051	1,452	2,219	3,157	3,208	1,872	6,452	9,270	1,663	8,382	8,492	7,456	7,255
2001 SAM Nat																						
average	102,013	98,182	89,777	95,063	54,635	23,763	28,179	26,994	58,314	36,245	16,346	25,591	38,877	38,899	22,787	83,560	113,922	18,202	102,430	104,800	91,392	88,519
std dev	5,680	5,467	4,999	5,293	3,042	1,323	1,569	1,503	3,247	2,018	910	1,425	2,165	2,166	1,269	4,653	6,344	1,014	5,704	5,836	5,089	4,929
2002 SAM Nat																						
average	92,333	89,620	82,243	85,896	49,885	21,681	25,718	24,630	53,263	33,042	15,039	23,525	35,364	35,267	20,991	76,534	101,402	16,678	91,122	93,429	82,221	79,171
std dev	26,720	25,935	23,801	24,858	14,436	6,274	7,443	7,128	15,414	9,562	4,352	6,808	10,234	10,206	6,074	22,148	29,345	4,827	26,370	27,037	23,794	22,912
2006 SAM Nat																						
average	128,692	124,431	116,207	121,209	80,451	22,740	29,569	27,930	84,699	44,724	16,588	32,257	44,315	45,463	24,797	110,746	139,616	24,158	126,878	130,588	123,358	115,713
std dev	25,540	24,695	23,063	24,055	15,966	4,513	5,868	5,543	16,810	8,876	3,292	6,402	8,795	9,023	4,921	21,979	27,708	4,794	25,180	25,917	24,482	22,965
2007 SAM Nat																						
average	142,651	137,863	125,886	131,158	73,025	24,439	30,960	28,939	77,376	43,910	17,417	31,923	45,911	46,248	27,937	122,078	159,987	23,332	142,939	149,283	133,785	129,322
std dev	31,900	30,829	28,151	29,330	16,330	5,465	6,923	6,471	17,303	9,819	3,895	7,139	10,267	10,342	6,247	27,299	35,776	5,217	31,964	33,383	29,917	28,919
2008 SAM Nat																						
average	150,243	144,320	133,126	141,765	90,550	28,762	36,300	34,373	94,951	52,146	20,557	37,655	52,476	54,062	29,932	126,997	169,914	28,101	155,021	157,834	142,952	135,222
std dev	26,052	25,025	23,084	24,582	15,701	4,987	6,294	5,960	16,465	9,042	3,565	6,529	9,099	9,374	5,190	22,021	29,463	4,873	26,881	27,369	24,788	23,448
2001 TS Nat																						
average	64,331	56,491	50,384	59,937	27,527	13,874	16,671	16,116	35,505	25,356	11,967	16,323	29,326	29,270	16,122	49,848	82,709	13,382	73,500	75,806	63,799	66,753
std dev	4,875	4,352	4,002	4,454	1,526	823	1,131	1,104	3,056	2,350	964	1,223	2,998	2,847	1,361	4,046	5,826	984	5,112	5,397	4,791	5,099

Table B.3 Con't

2002 TS Nat																					
average	62,497	56,654	51,426	58,475	29,609	13,912	16,824	16,276	37,498	25,698	11,866	16,663	29,258	29,095	15,885	49,982	76,234	13,127	68,131	69,972	60,434	61,615
std dev	4,025	4,213	4,215	3,741	2,211	843	1,177	1,158	3,675	2,422	926	1,317	2,918	2,721	1,255	3,898	3,393	872	3,031	3,185	3,414	3,199

2006 TS Nat																					
average	67,735	63,714	58,489	63,388	34,785	13,523	16,979	16,251	42,509	26,872	11,539	17,816	29,742	29,695	16,170	57,048	79,221	13,638	71,486	73,266	65,010	63,956
std dev	7,181	8,391	8,338	6,631	5,102	662	1,280	1,155	6,397	3,054	790	1,961	3,225	3,065	1,507	8,077	5,401	1,161	5,125	5,336	6,150	4,783
2007 TS Nat																					
average	65,216	61,100	55,890	60,867	32,333	13,677	16,889	16,210	39,902	26,047	11,607	17,261	29,161	28,999	16,075	54,430	76,746	13,238	69,010	70,828	62,192	61,736
std dev	5,158	6,364	6,382	4,684	3,912	808	1,194	1,112	4,782	2,391	854	1,537	2,705	2,457	1,417	6,073	3,230	845	3,048	3,232	3,983	2,885

2008 TS Nat																					
average	65,234	61,164	55,858	60,791	32,110	13,911	17,035	16,347	39,267	25,787	11,673	17,246	28,851	28,687	16,156	54,342	76,778	13,197	68,930	70,819	61,966	61,594
std dev	5,236	6,466	6,402	4,676	3,873	1,097	1,355	1,265	4,557	2,204	939	1,533	2,438	2,201	1,524	6,057	3,356	820	3,051	3,313	3,899	2,866

The Vita

Drew Alexander Varnado, a native of Denham Springs, Louisiana, received his bachelor's degree in Mathematics and Economics at Millsaps College in 2005. After graduating he entered directly into doctoral candidacy at Louisiana State University in Economics. He worked through the university as a consultant to the Louisiana State Board of Regents on their Master Plan efforts. The experience was impactful and in pursuit of more economic application he transferred into the Agricultural Economics Department at Louisiana State University. He will receive his doctoral degree from that department in August 2014 and plans to continue work under Dr. Matt Fannin as a research associate.